

# SCIENTIFIC REPORTS



OPEN

## Transmission of Multiple HIV-1 Subtype C Transmitted/founder Viruses into the Same Recipients Was not Determined by Modest Phenotypic Differences

Hongshuo Song<sup>1</sup>, Bhavna Hora<sup>1</sup>, Elena E. Giorgi<sup>2</sup>, Amit Kumar<sup>1</sup>, Fangping Cai<sup>1</sup>,  
Tanmoy Bhattacharya<sup>2</sup>, Alan S. Perelson<sup>2</sup> & Feng Gao<sup>1,3</sup>

Received: 04 July 2016  
Accepted: 04 November 2016  
Published: 02 December 2016

A severe bottleneck exists during HIV-1 mucosal transmission. However, viral properties that determine HIV-1 transmissibility are not fully elucidated. We identified multiple transmitted/founder (T/F) viruses in six HIV-1-infected subjects by analyzing whole genome sequences. Comparison of biological phenotypes of different T/F viruses from the same individual allowed us to more precisely identify critical determinants for viral transmissibility since they were transmitted under similar conditions. All T/F viruses used coreceptor CCR5, while no T/F viruses used CXCR4 or GPR15. However, the efficiency for different T/F viruses from the same individual to use CCR5 was significantly variable, and the differences were even more significant for usage of coreceptors FPRL1, CCR3 and APJ. Resistance to IFN- $\alpha$  was also different between T/F viruses in 2 of 3 individuals. The relative fitness between T/F viruses from the same subject was highly variable (2–6%). Importantly, the levels of coreceptor usage efficiency, resistance to IFN- $\alpha$  and viral fitness were not associated with proportions of T/F viruses in each individual during acute infection. Our results show that the modest but significant differences in coreceptor usage efficiency, IFN- $\alpha$  sensitivity and viral fitness each alone may not play a critical role in HIV-1 transmission.

HIV-1 mucosal transmission is associated with a stringent population bottleneck that remains poorly understood. Despite the highly diverse quasispecies population in HIV-1-infected donors, about 60–80% of HIV-1 infections through mucosal transmission are established by a single transmitted/founder (T/F) virus<sup>1–5</sup>. Some evidence suggests that HIV-1 transmission is not entirely a stochastic process in which all replication competent viruses have an equal probability to be transmitted, but rather that certain genetic and phenotypic traits, which can facilitate viral transmission, are selected<sup>6</sup>. Genetic comparisons of the *env* sequences of T/F viruses showed shorter variable loops and fewer potential N-linked glycosylation sites for subtype A and C T/F viruses<sup>3,7–9</sup>, but not for subtype B T/F viruses compared to the donor or chronic viruses<sup>8,10,11</sup>. Comparison of the subtype B T/F viruses and chronic viruses identified some signatures in T/F *env* sequences<sup>12,13</sup>. Biological studies with infectious molecular clones (IMCs) representing unambiguously inferred full-length T/F HIV-1 genomes showed that T/F viruses replicate with high efficiency in primary CD4+ T cells, but less efficiently in monocyte-derived macrophages<sup>2,14,15</sup>. Recent comprehensive phenotypic comparisons between the T/F and chronic viruses demonstrated that T/F viruses have enhanced infectivity, incorporate higher Env content per viral particle, bind dendritic cells more efficiently and are relatively more resistant to inhibition by IFN- $\alpha$  than chronic viruses<sup>16,17</sup>.

Despite the severe population bottleneck, around 20–40% of mucosal HIV-1 transmissions are established by multiple HIV-1 T/F viruses<sup>1–5</sup>. The mechanisms for transmission of multiple HIV-1 variants into the same subjects are still unknown. Statistical analysis of cases of multiple HIV-1 transmissions through the heterosexual

<sup>1</sup>Department of Medicine, Duke University Medical Center, Durham, NC 27710, USA. <sup>2</sup>Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. <sup>3</sup>National Engineering Laboratory for AIDS Vaccine, School of Life Sciences, Jilin University, Changchun 130012, China. Correspondence and requests for materials should be addressed to F.G. (email: fgao@duke.edu)

Subject	Fiebig stage	Country	Year	Days from infection (95% CI)*	No. of T/F viruses	
					5' half	3' half
CH0200	I/II	Malawi	2007	11 (8, 13)	3	5
CH0228	III	Malawi	2007	19 (14, 24)	2	3
CH0078	I/II	South Africa	2007	11 (8, 14)	1	2
CH0047	III	Malawi	2007	19 (13, 24)	2	2
CH0010	I/II	Malawi	2006	9 (7, 12)	3	6
CH1754	III	Malawi	2010	14 (12, 15)	6	7

**Table 1. Demographic characteristics of subjects infected with multiple T/F viruses.** Days from infection were estimated taking the harmonic average of the estimates obtained using the two genome halves (see Supplementary Table 1).

route revealed that the number of transmitted viruses does not follow a Poisson distribution, suggesting that the transmission of multiple T/F viruses are not independent events with low probability<sup>4</sup>. Therefore, it has been hypothesized that, during transmission of multiple HIV-1 variants, either the transmission rate is transiently increased due to certain cofactors, such as inflammatory genital infections that may reduce the transmission barrier by destroying the intact mucosa or increasing the availability of activated CD4 + T cells, or that transmission of multiple variants are linked events, such as through a multiply-infected cell or viral aggregates<sup>3,4</sup>. Regardless of the mechanisms involved, the fact that different HIV-1 variants can simultaneously cross the mucosal barrier in the same host and establish productive infection imply that these T/F viruses may have similar transmission efficiency. However, little is known to date regarding whether multiple T/F viruses in the same HIV-1-infected individual share the same phenotypic properties. Addressing this question would provide unique insight into critical viral properties associated with HIV-1 transmissibility and help to better understand potential mechanisms of HIV-1 transmission.

We have identified six individuals who were infected with two or more subtype C T/F viruses. To investigate if multiple T/F viruses in each individual have distinct biological phenotypes, we compared their coreceptor tropism, sensitivity to IFN- $\alpha$  inhibition and replication fitness. Our results demonstrated that T/F viruses in the same infected individual have different coreceptor usage efficiencies, resisted IFN- $\alpha$  at different levels and had different fitness. These observed differences indicate that those phenotypic properties cannot solely predict the transmissibility of HIV-1.

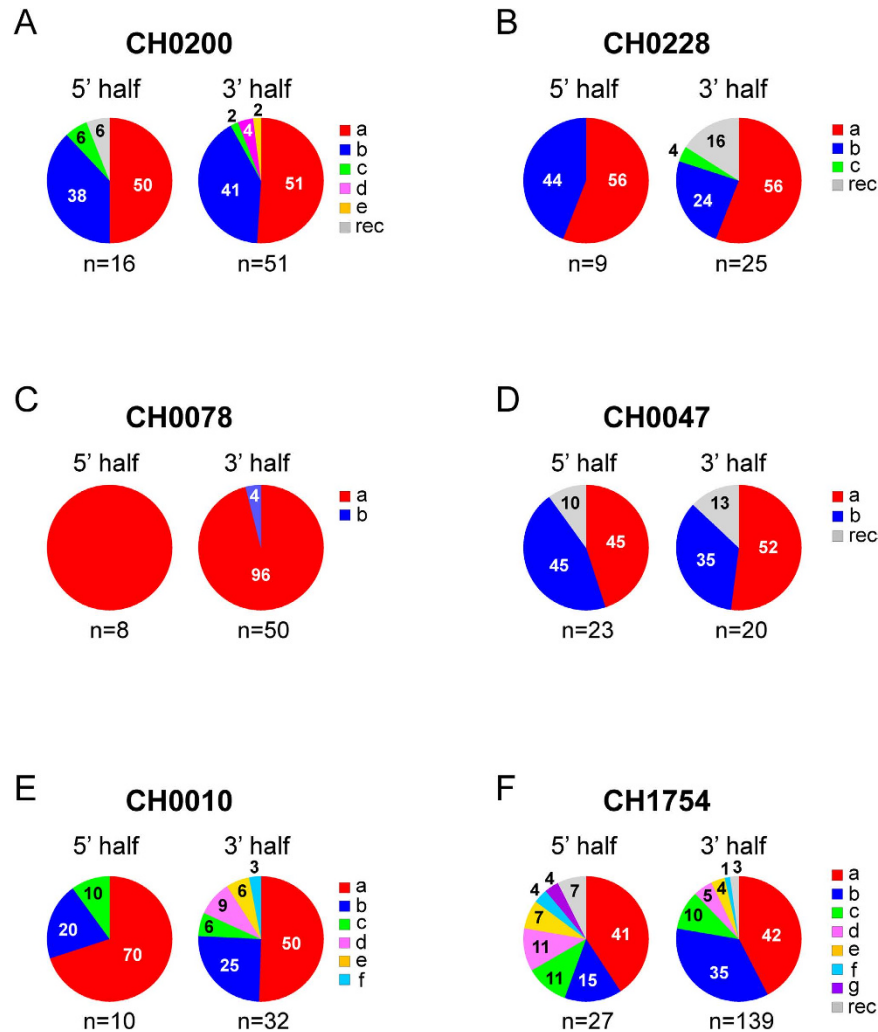
## Results

**Identification of multiple T/F whole genome sequences from the same individuals.** Two or more T/F viruses were identified in six subjects by analyzing nearly full-length genome sequences (in two overlapping halves) generated by SGA from the first HIV-1 RNA positive plasma samples (at or before Fiebig stage III) in the CHAVI001 acute infection cohort (Table 1). The analysis of both halves of viral genomes using the Poisson model<sup>18</sup> on each lineage and then combining the estimates as described in the methods showed that the first HIV-1 RNA positive samples were collected between 9 and 19 days after transmission for these six individuals (Table 1). In all of our time estimates, whether the infection was caused by single or multiple T/F viruses, we stated a 95% confidence interval whose width depended on the sample size (the number of sequences for analysis) (Supplementary Table 1). Within that time window, our model cannot confidently distinguish whether multiple genetically distinct T/F viruses infected the host at the same time or a few days apart. In most cases, estimates and confidence intervals calculated for different T/F virus lineages within the same individual overlapped. Goodness of fit p-values, which measured the probability that the overall Hamming distance (HD) distribution differed from the expected Poisson form, were non-significant after multiple-testing correction. Both elements indicated that different T/F virus lineages could have evolved at the same time. The one exception was the CH1754a lineage in the 5'-half genome, which seemed much more homogeneous than the other T/F lineages, indicating a possibility that CH1754a T/F virus was a later infection than the other T/F virus lineages causing the overall HD distribution to diverge from a Poisson form (Supplementary Table 1).

Between two and seven T/F viruses were identified in each of the six subjects (Table 1 and Fig. 1). The proportions of T/F viruses in each individual were highly variable, with major T/F viruses as high as 96% in CH0078 and minor T/F viruses as low as 1% in CH1754 (Fig. 1). The highly unequal proportions of T/F viruses at acute infection suggest different replication advantages among T/F viruses in the same HIV-1-infected individuals.

### Different efficiencies in use of various coreceptors by T/F viruses from the same individual.

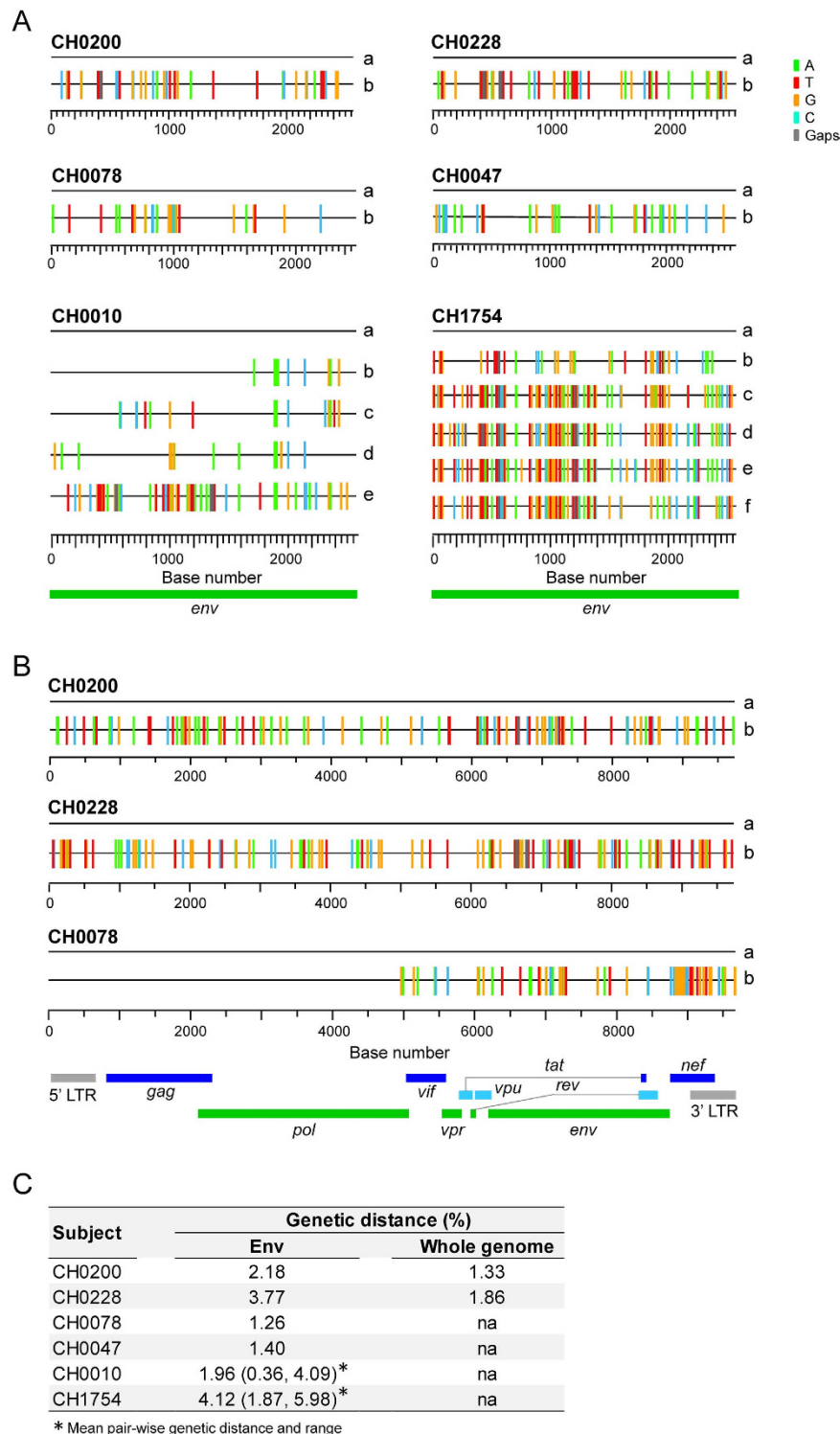
Since proportions of different T/F viruses were highly variable in the same HIV-1 infected individual, one important question was whether they had different efficiencies in coreceptor usage that led to the disproportional expansion of different T/F viruses early after transmission. To address this question, we cloned all the *env* genes for which the T/F sequences could be reliably inferred with sufficient number of sequences. Two T/F *env* clones (a and b) were generated for CH0200, CH0228, CH0078 and CH0047, while five (a through e) and six (a through f) *env* clones were made for CH0010 and CH1754, respectively (Fig. 2A). The genetic diversity among the T/F *env* genes in each individual ranged from 1.26% (between two T/F *env* genes in CH0078) to 4.12% (among six T/F *env* genes in CH1754) (Fig. 2C). Env pseudoviruses were generated by cotransfection with the pNL43- $\Delta$ Env-vpr<sup>+</sup>-luc<sup>+</sup> clone that expresses luciferase upon infection and were used to infect a panel of NP-2



**Figure 1. Percentage of T/F virus populations in each HIV-1-infected individual.** The number of T/F viruses and their percentage were determined based on the 5'- and 3'-half viral genomes at the screening time point in subjects CH0200 (A), CH0228 (B), CH0078 (C), CH0047 (D), CH0010 (E) and CH1754 (F). Different T/F sequences were indicated by different colors and recombinants were shown as gray. The numbers on the pie charts indicate the percentage of each T/F virus as well as the recombinants. The number of SGA sequences (n) for 5'- and 3'-half genomes in each individual is indicated under the pie chart. The number included those from both half genome sequences and NFLG sequences for CH0200, and both long and short 3'-half genome sequences for CH0228.

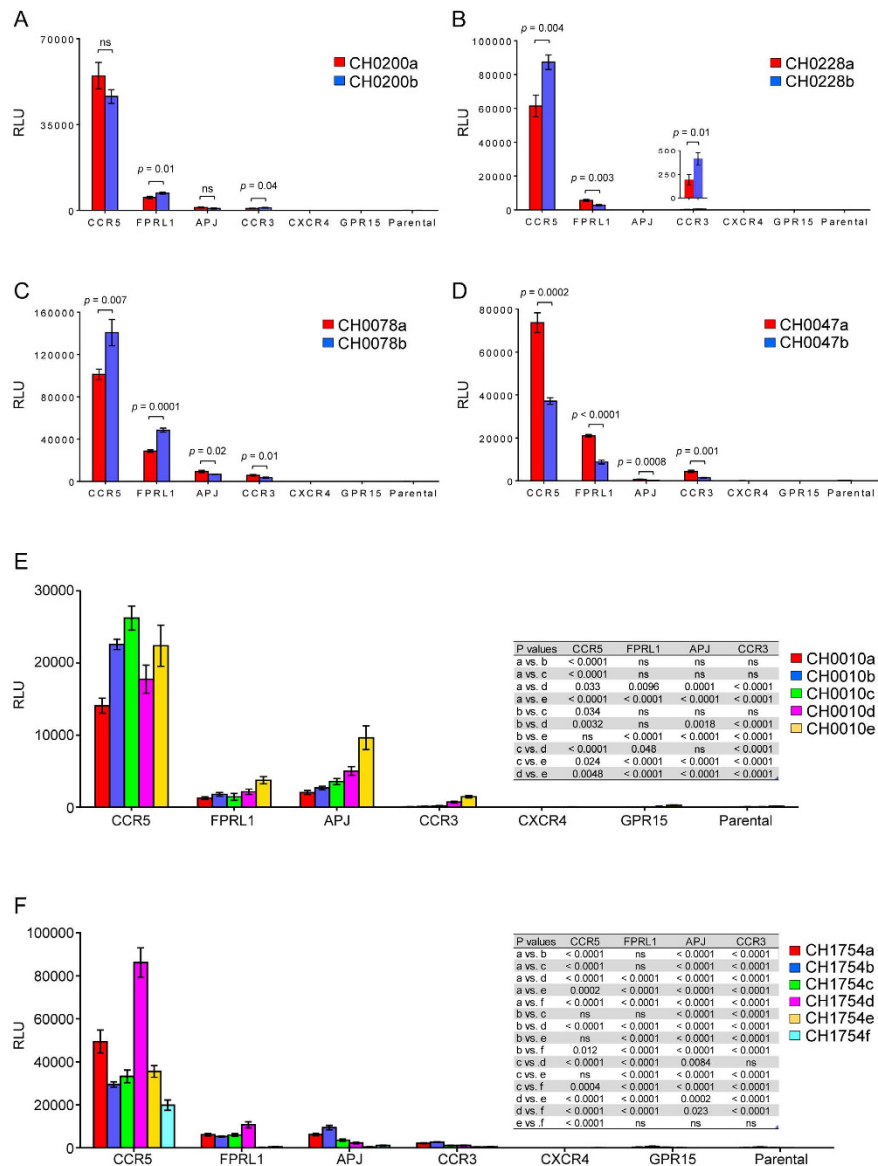
cell lines that express CD4 and one of two major coreceptors (CCR5 or CXCR4) or other commonly used coreceptors (FPRL1, APJ, CCR3 or GPR15)<sup>19–22</sup>. Consistent with previous reports<sup>1</sup>, all 19 T/F Env pseudoviruses were CCR5 tropic, while none of them used CXCR4 as coreceptor (Fig. 3). We also found that they all used coreceptors FPRL1 and CCR3, and 17 of 19 Env pseudoviruses, i.e., all except the two T/F viruses from CH0228, used APJ for entry, although the efficiency of these alternative coreceptors usage is much lower than of CCR5 usage (Fig. 3). None of the Env pseudoviruses infected GPR15+ cells.

Next, we investigated whether there was a difference in coreceptor usage efficiency between the T/F viruses from the same individual. Although all Env pseudoviruses used CCR5, the levels of infection among different T/F viruses were significantly different in 5 out of 6 individuals. Significant differences in use of the CCR5 coreceptor were found between the *env* genes of the two T/F from CH0228 ( $p = 0.004$ , two-tailed t-test), CH0078 ( $p = 0.007$ , two-tailed t-test) and CH0047 ( $p = 0.0002$ , two-tailed t-test), while no difference was observed between the two T/F *env* genes from CH0200 (Fig. 3A–D). The five T/F Env pseudoviruses from CH0010 significantly differed in infectivity ( $p < 0.0001$ , One-way ANOVA test), except for the comparison between CH0010b and CH0010e (Fig. 3E). When the six T/F Env pseudoviruses from CH1754 were compared, no differences in infectivity were observed among CH1754b, CH1754c and CH1754e, while all other T/F viruses were significantly different ( $p \leq 0.012$ , One-way ANOVA test; Fig. 3F). The highest difference in infectivity in CCR5 + NP2 cells was observed between CH1754d and CH1754f (4.3 folds;  $p < 0.0001$ , One-way ANOVA test). These results showed that significant, but moderate, differences (<5 folds) in CCR5 usage were often observed among T/F viruses in the same infected subject.



**Figure 2. Genetic diversity among different T/F virus sequences in the same individuals.** Highlighter plots ([http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter\\_top.html?choice=mismatches](http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html?choice=mismatches)) showed the locations of the base differences in the *env* gene (A) and the whole genome (B), compared to the most dominant T/Fa virus in each individual. Pairwise genetic distances among different T/F viruses in the *env* gene and the whole genome in each individual are shown (C).

Although T/F Env pseudoviruses also used other alternative coreceptors at much lower levels compared to CCR5, the infectivity of the various T/F from the same subjects varied much more in cells expressing those coreceptors than in CCR5 expressing cells (Fig. 3). CH1754e used CCR5 less efficiently than CH1754a and CH1754d, more efficiently than CH1754f, and similar to CH1754b and CH1754c (Fig. 3F). However, on average it was

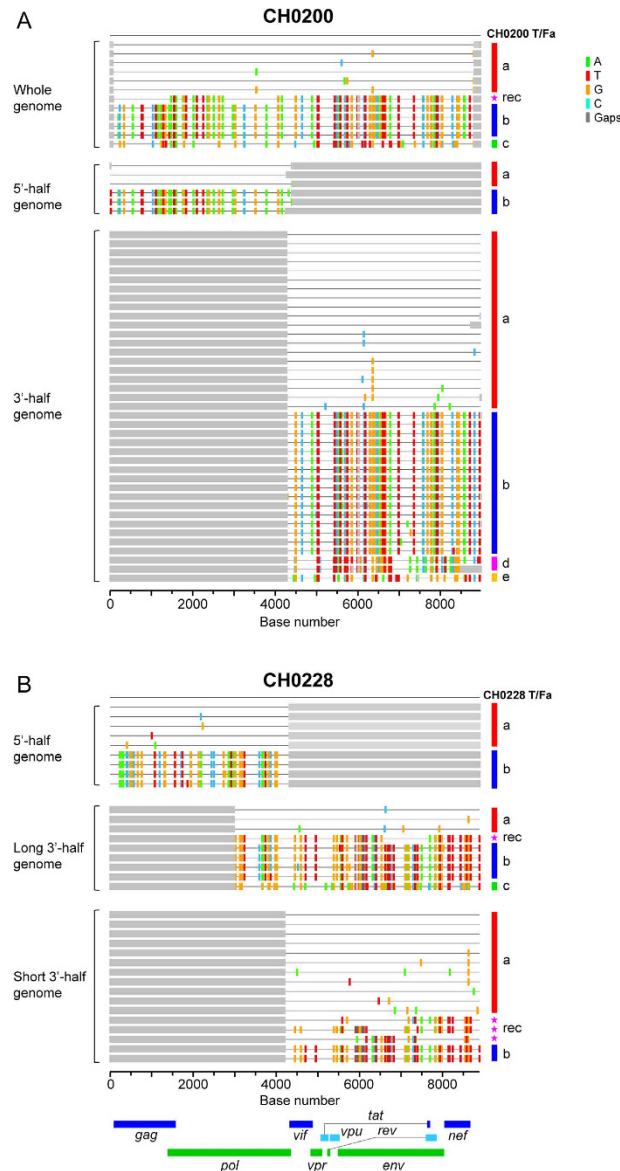


**Figure 3. Comparison of coreceptor usage among different T/F viruses from the same HIV-1-infected individual.** Same amount of T/F Env-pseudoviruses (200 TCID<sub>50</sub>) was used to infect NP-2 cell lines expressing CD4 together with one of other coreceptors (CCR5, CXCR4, FPRL1, APJ, CCR3 or GPR5). The parental NP-2 cell line expressing CD4 but not any coreceptor was used as a control. The luciferase activity was measured 3 days post infection. The positive virus replication was defined if a virus gave luciferase activity three times above its background luciferase activity in the NP-2 parental cell line. All infections were performed in triplicates.

66.5-fold less efficient in infecting FPRL1+ cells than CH1754d ( $p < 0.0001$ , One-way ANOVA test), 21.9-fold less efficient than CH1754b to infect APJ+ cells ( $p < 0.0001$ , One-way ANOVA test), and 6.9-fold less efficient than CH1754b to infect CCR3+ cells ( $p < 0.0001$ , One-way ANOVA test). CH0010a used CCR5 1.6-fold less efficiently than CH0010e, but was 19.8-fold less efficient in infecting CCR3+ cells than CH0010e ( $p < 0.0001$ , One-way ANOVA test) (Fig. 3E). Various levels of differences in alternative coreceptor usage were also observed for different T/F viruses in four other individuals, although the differences were comparatively smaller (Fig. 3A–D). Taken together, these results showed that different T/F Envs from the same subject use the same coreceptor at differing efficiencies, and the efficiency in using CCR5 was not indicative for usage efficiencies of other coreceptors.

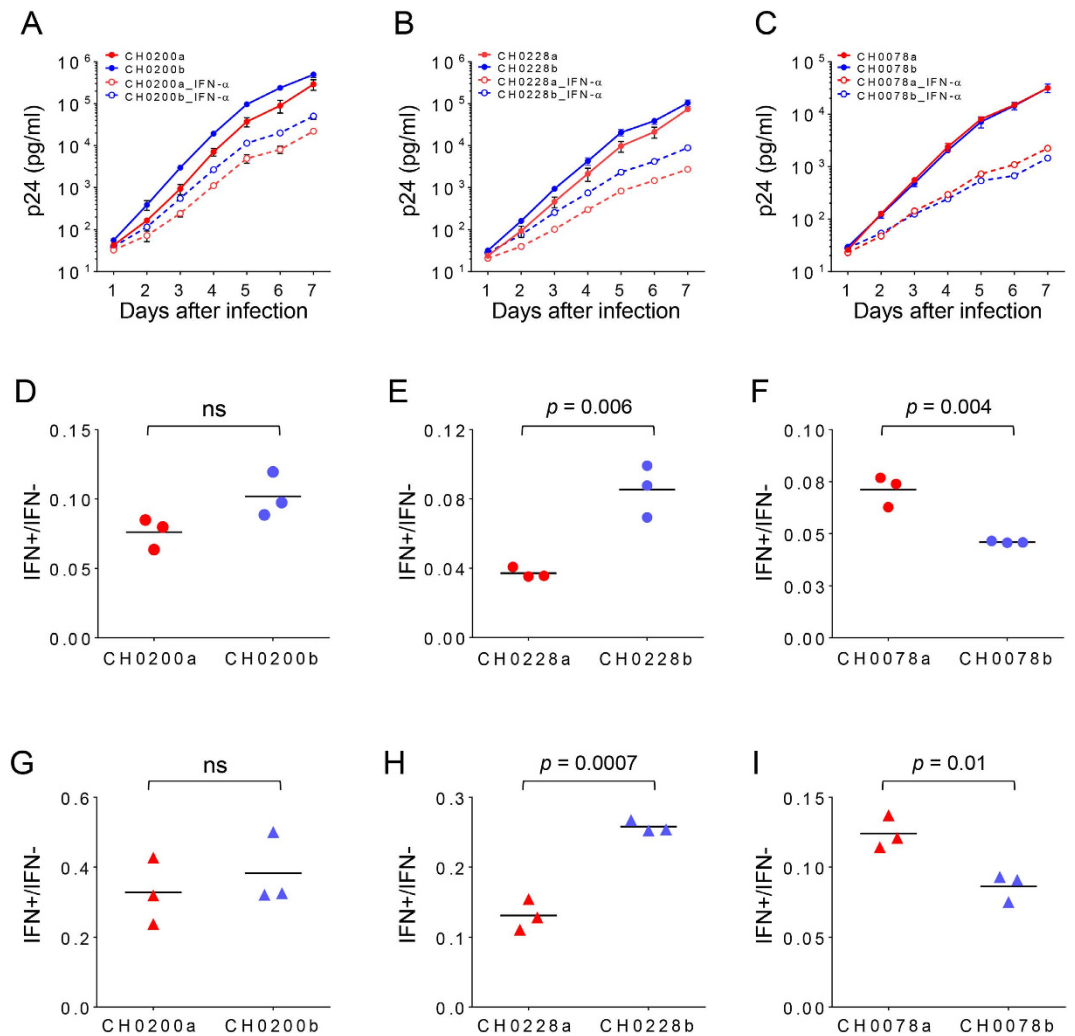
**Various sensitivities to IFN- $\alpha$  inhibition among T/F viruses from the same subject.** Recent studies showed that subtype B T/F viruses were relatively more resistant to IFN- $\alpha$  than chronic viruses or matched 6-month viruses from the same infected individuals, while generally no significant differences were observed for subtype C viruses<sup>17,23,24</sup>. To determine whether T/F viruses from the same infected individual resist IFN- $\alpha$  at different levels, we synthesized T/F genomes and constructed infectious molecular clones (IMCs) to represent the T/F viruses for which the T/F whole genome sequences could be reliably inferred (Fig. 2B). To accurately match the 5'- and 3'-half genome sequences for each whole T/F genome in CH0200, we successfully





**Figure 4. Identification of matching 5'- and 3'- half genomes from the same T/F virus in each individual.** Highlighter plots identified the matching 5'- and 3'- half genomes from the same T/F virus by using NFLG genome sequences in CH0200 (A) and long overlapping sequences (~1000 bp) in CH0228 (B). The T/Fa and T/Fb sequences that were identical or highly similar in NFLG and both half genome sequences (CH0200) or in the overlapping region in the middle of HIV-1 genomes (CH0228) are indicated by red and blue bars, respectively.

obtained 12 near full length genome (NFLG) sequences by SGA. Highlighter plot analysis of all those sequences together showed that CH0200a and CH0200b sequences from 5'- or 3'-half genome were identical or highly similar to CH0200a and CH0200b NFLG sequences (Fig. 4A). Phylogenetic tree analysis of both half genome sequences confirmed that CH0200a and CH0200b from NFLG or half genome sequences were indistinguishable (Supplementary Fig. 1). Thus, both CH0200a and CH0200b 5'- and 3' half genome sequences could be unequivocally linked together based on the NFLG sequences. We were unable to obtain NFLG sequences for the viruses in CH0228. Instead, we successfully amplified nine long 3'-half genome SGA sequences, which overlapped with the 5'-half genome sequences by ~1000 bp. Highlighter plot analysis of the overlapping region showed that both CH0228a and CH0228b from 5'- and 3'-half genome sequences were identical or highly similar at this overlapping region and the sequences from both long and short 3'-half genome sequences are indistinguishable (Fig. 4B). Phylogenetic tree analysis of the overlapping sequences confirmed these relationships (Supplementary Fig. 2). Thus, both analyses allowed for unambiguous identification of the correct 5'- and 3'-half genome sequences derived from the same T/F genome in CH0228 based on the overlapping region sequences. The missing LTR region for each T/F virus in the two overlapping half genome sequences



**Figure 5. Sensitivity of different T/F viruses to IFN- $\alpha$  inhibition.** The replication kinetics of each multiple T/F virus in the absence (solid lines) and presence of 500 U/ml IFN- $\alpha$  (dashed lines) from subjects CH0200 (A), CH0228 (B), CH0078 (C) were determined in primary CD4 T cells from a health donor C24. Activated primary CD4 + T cells from donor C24 were infected with equal amount (500 TCID<sub>50</sub>) of each virus. Viral replication was monitored by measuring the p24 concentration in the culture supernatant daily for 7 days. All infections were performed in triplicates. The resistance to IFN- $\alpha$  inhibition by different T/F viruses from CH0200 (D), CH0228 (E), CH0078 (F) was determined by comparing the ratio of viral p24 production in the presence and absence of IFN- $\alpha$  at day 7. The resistance to IFN- $\alpha$  inhibition by different T/F viruses from CH0200 (G), CH0228 (H) and CH0078 (I) was also determined in primary CD4 + T cells from another donor C29 at day 7.

was amplified independently as previously described<sup>14,16,25</sup>. Two T/F IMCs (a and b) were generated for subjects CH0200, CH0228 and CH0078 (Fig. 2B).

The screening samples were generally limited in volume since they were the first samples collected from all subjects. In addition, the 5'-half genome was more difficult to amplify than the 3'-half genome. Since we performed SGA analysis of 3'-half genome first, samples were often exhausted before obtaining a large number of SGAs for the 5'-half genome. This was why we only had relatively small numbers of SGAs for CH0078, CH0010 and CH0228. The minority T/F (CH0078b) viruses were only present at 4% of 3'-half genome sequences (50), we were not able to get enough SGAs to obtain sequences for this minority T/F in the 5'-half genome sequences before we exhausted the screening sample (Fig. 1C) and only one 5'-half genome derived from the CH0078b T/F virus was detected 4 weeks later<sup>26</sup>. To determine if the 3'-half genome could have any impact on biological phenotypes, two chimeras that shared the same 5'-half genome were generated and tested in this study since the full-length T/F sequences for the CH0078b T/F virus could not be reliably inferred (Fig. 2B).

Different T/F viruses from subjects CH0200, CH0228 and CH0078 had very similar replication kinetics in activated primary CD4 + T cells when each was cultured independently (Fig. 5A–C). All T/F viruses replicated at significantly lower levels in the presence of 500 U/ml IFN- $\alpha$  than in the absence of IFN- $\alpha$  (dashed lines in Fig. 5A–C). To compare the level of resistance to IFN- $\alpha$  for each virus, we determined the ratio of p24 concentrations in the culture with or without IFN- $\alpha$  at day 7 as previously reported<sup>16</sup>. In CH0228, the major T/F

virus CH0228a was on average 2.3-fold more sensitive to IFN- $\alpha$  than a less predominant T/F virus CH0228b ( $p = 0.006$ , two-tailed t-test, Fig. 5E). But in CH0078, the major T/F virus CH0078a was on average 1.5-fold more resistant to IFN- $\alpha$  than the minor T/F virus CH0078b ( $p = 0.004$ , Fig. 5F). No significant difference was observed between the most predominant virus CH0200a and a less predominant virus CH0200b in CH0200 (Fig. 5D). Similar results were obtained with primary CD4 + T cells from a different donor (Fig. 5G–I).

**Various fitness differences among T/F viruses from the same subject.** We next sought to investigate whether different T/F viruses in the same individual had similar replication fitness. To more sensitively compare the fitness of T/F viruses, we performed competition PASS fitness assays, as previously described<sup>27–30</sup>. Two T/F viruses from CH0200, CH0228 or CH0078 were cultured together and the proportion of each virus in the culture was determined at days 1, 3 and 5 by PASS. The less predominant T/F virus CH0200b was nearly as fit as the major T/F virus CH0200a (relative fitness  $3\% \pm 2\%$  [mean of 3 replicates  $\pm$  SE, see methods],  $p = 0.13$ ) (Fig. 6A). The less predominant T/F virus CH0228b was 7% less fit than the major T/F virus CH0228a ( $p = 0.006$ ) (Fig. 6B). Although the minor T/F virus in CH0078 accounted for only 4% of the viral population at the screening time point<sup>26</sup>, the major T/F virus was only a little more fit than the minor T/F virus in CH0078 (relative fitness  $4\% \pm 2\%$ ,  $p = 0.009$ ) (Fig. 6C). Overall, major T/F viruses were more fit than the less predominant or minor T/F viruses in three viruses but only in the cases of CH0228 and CH0078 was the difference statistically significant.

**No association between predominance of T/F viruses and phenotypes.** We next determined whether the predominance of a T/F virus was associated with more efficient use of coreceptors, more resistance to IFN- $\alpha$  inhibition and better fitness in the same subject during acute infection. We found that the most predominant T/F viruses used CCR5 less efficiently than the less predominant or minor T/F viruses in CH0228, CH0078, and CH0010 (Fig. 3B,C,E and F) while the most predominant T/F virus used CCR5 more efficiently in CH0047 (Fig. 3D). In CH1754, the most predominant virus CH1754a used CCR5 significantly less efficiently than the minor T/F virus CH1754d ( $p < 0.0001$ , One-way ANOVA test), although with significant higher efficiency than all other four viruses (Fig. 3F). The predominant T/F virus was more resistant to IFN- $\alpha$  in CH0078, but less resistant to IFN- $\alpha$  than the minor T/F virus in CH0228. In the third subject CH0200, no significant difference in resistance to IFN- $\alpha$  was observed between the most predominant T/F virus and a less predominant one. The predominant T/F virus was generally more fit than the minor T/F virus in all three subjects (CH0200, CH0228 and CH0078). However, the major T/F virus was only 4% more fit in CH0078 compared to a 7% difference in fitness in CH0028, but the major CH0078a T/F virus dominated the minor CH0078b T/F virus (96% vs. 4%), while the CH0228a was a little more prevalent than the CH0228b (56% vs 24%). These results indicated that the proportions of T/F viruses in each HIV-1-infected individual were not determined by observed differences in efficiency in coreceptor usage, sensitivity to IFN- $\alpha$  or viral fitness alone at the acute infection stage.

## Discussion

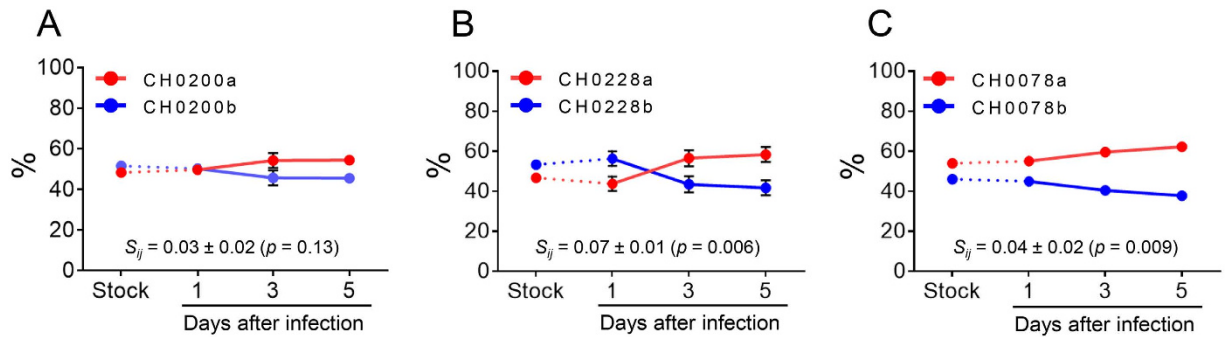
Understanding the properties of T/F viruses will have important implications for HIV-1 vaccine development. Comparing the biological phenotypes among different T/F viruses from the same individuals infected with HIV-1 through mucosal transmission can reveal if some biological phenotypes determine HIV-1 transmission since these T/F viruses are likely transmitted under similar conditions. We compared biological phenotypes of different T/F viruses from each of six subjects and found significant differences in coreceptor usage efficiency, resistance to IFN- $\alpha$  and viral replication fitness among T/F viruses from the same individuals. These results indicate that transmission of different T/F viruses from a donor into a new recipient was possibly not determined solely by observed differences in CCR5 usage, alternative coreceptor usage, resistance to IFN- $\alpha$ , and viral replication fitness.

We found that the ability of different T/F Envs from the same individual to infect CCR5 + cells varied up to 4.3 fold, indicating that viruses with these levels of difference in CCR5 usage efficiencies could still all be efficiently transmitted into a new host. Previous studies have examined the usage of alternative coreceptors by HIV-1 isolates from acute/early infections<sup>31,32</sup>. However, little is known about the utilization of alternative coreceptors by T/F viruses and their potential roles in HIV-1 transmission. We found that the majority of the T/F Envs did not use CXCR4 or GPR15, but could use FPRL1, APJ and CCR3 for entry, although at reduced efficiencies compared to the use of CCR5. The usage of FPRL1, APJ and CCR3 differed by 20–67 folds between different T/F Envs from the same individuals (CH1754 and CH0010). These results showed that T/F Envs with significant differences in use of alternative coreceptors could again all be transmitted into the same recipient, suggesting that efficiency in use of these alternative coreceptors may also not be an important factor for HIV-1 transmission.

HIV/SIV can lead an interferon storm during acute infection<sup>33–35</sup>. Previous studies showed that subtype B T/F viruses were more resistant to IFN- $\alpha$  inhibition than chronic HIV-1 viruses from different individuals or matched 6-month viruses from the same infected individuals<sup>16,17</sup>. However, no statistical differences in IFN- $\alpha$  resistance were observed for subtype C T/F viruses compared to chronic virus or linked viruses from the donors quasispecies<sup>16,24</sup>, although two IMC-derived subtype C T/F viruses showed enhanced resistance than matched 6-month viruses after infection<sup>17</sup>. We now showed that modest differences (<3 folds) in IFN- $\alpha$  resistance were present among different T/F viruses from the same individuals infected with subtype C viruses. These data suggest that the observed differences in resistance to IFN- $\alpha$  among subtype C viruses, similar to those observed between acute and chronic viruses, may not be sufficient to explain the differential success of viruses after HIV-1 transmission.

The relationship between fitness and transmission of HIV-1 has been investigated based on partial viral genomes in previous studies<sup>36–38</sup>. Results were not conclusive since other parts of the viral genome can also affect viral fitness and play a critical role in transmission. Thus, the role of viral fitness in HIV-1 transmission remains unclear. A recent study on epidemiologically linked heterosexual transmission pairs showed that T/F virus did not exhibit higher replication capacity than the corresponding non-transmitted variants in the donors,





**Figure 6. Comparison of viral fitness of different T/F viruses from the same HIV-1-infected individual.** The relative fitness of different T/F viruses in the same individual was determined using the competition PASS fitness assay<sup>27,29</sup>. Activated primary CD4 + T cells were infected with viral stock mixture containing equal amount (5 ng p24) of each compared virus. The proportions of compared viruses in the culture supernatant were measured on days 1, 3 and 5. The major (red) and the minor or less predominant (blue) T/F viruses from CH0200 (A), CH0228 (B) and CH0078 (C) were compared.

which argued against the importance of viral replication fitness during transmission<sup>24</sup>. In our study, the complete genome sequence for each T/F viruses was unequivocally determined and free of RT-PCR artifacts since each T/F whole genome sequence was generated by the consensus of multiple SGA sequences. We found that the fitness differences were 3–7% between different T/F viruses in each of three individuals. All those individuals were infected through mucosal transmission based on the patient infection history. Thus, our results indicate that differences in viral replication fitness of this magnitude, as measured *in vitro*, do not affect HIV-1 transmission through the mucosal barrier. However, the study of additional cases is warranted to better characterize the role of replication fitness in HIV-1 transmission.

Similar differences in Env content, cell-free infectivity, dendritic cell interaction and IFN- $\alpha$  resistance were also observed between genetically much more divergent T/F and chronic viruses from different individuals<sup>23</sup>. In addition, no differences in CCR5 usage were found between T/F and chronic viruses from different individuals<sup>23,39</sup>. Thus, the subtle but significant differences in fitness and coreceptor usage between different T/F viruses in the same subject should represent the real differences among different T/F viruses. However, such differences were not likely to play an important role in transmission since different T/F viruses were transmitted at the same time. HIV-1 transmission may not be determined by any one of these biological characteristics alone, instead it is possible that it is governed by these viral properties in concert<sup>23</sup>. Of note, two CH0078 IMCs shared the same 5'-half CH0078a T/F genome because that 5'-half CH0078b T/F genome were not available. Thus, the differences in sensitivity to IFN- $\alpha$  inhibition and replication fitness in this case were only determined by 3'-half genome sequences. One limitation of this study is its relative small number of subjects. Future studies with a large sample size including subjects infected with other HIV-1 subtypes are warranted to confirm these observations.

Taken together, coreceptor usage efficiency, resistance to IFN- $\alpha$ , and replication fitness were different among multiple T/F viruses in the same individuals, suggesting that these observed differences each alone may not be sufficient to determine HIV-1 transmissibility. Moreover, the proportion of each T/F virus during acute infection was not associated to any of these biological phenotypes investigated in this study. This indicates that replication advantages of a T/F virus after transmission into a recipient may be determined by other host or viral factors. Whether higher levels of differences in these factors or other biological phenotypes are important to determine HIV-1 transmission remain to be studied.

## Methods

**Analysis of near full-length HIV-1 genome sequences.** Six subjects who were infected with multiple T/F viruses were identified from the CHAVI001 acute infection cohort. All subjects were male and none were on antiretroviral therapy. All subjects were infected with subtype C viruses through heterosexual transmission. Written informed consent was obtained from all subjects and the study was approved by the Duke University Institutional Review Board. All methods were performed in accordance with the relevant guidelines and regulations. Viral RNA was extracted from plasma samples using the Qiagen EZ1 Virus mini kit on the BioRobot EZ1 (Qiagen, Valencia, CA) and used for cDNA synthesis using Superscript III Reverse Transcriptase (Life Technologies; Carlsbad, CA) and primers 1R3.B3R (5'-ACTACTTGAAGCACTCAAGGCAAGCTTTATTG-3'; nt 9611-9642 in HXB2) and 07Rev9 (5'-CTTCCTGCCATAGGAGATGCCTAA-3'; nt 5957-5980) for 3'- and 5'-half HIV-1 genome, respectively. Single genome amplification (SGA) was performed to obtain two overlapping half genomes as described<sup>16,40</sup>. For amplification of a long 3'-half genome (~6000 bp) of CH0228, the first round PCR was carried out with primers 07for5 (5'-GGDSTGCCACACTAATGAT-3'; nt 3622-3641) and 2.R3.B6R (5'-TGAAGCACTCAAGGCAAGCTTTATTGAGGC-3'; nt 9636-9607); the second round PCR was done with primers IMC\_For3 (5'-GAAAGCATAGTAATATGGGGAAAGACTCC-3'; nt 3681-3709) and Low2c (5'-TGAGGCTTAAGCAGTGGGTTCC-3'; nt 9591-9612). The PCR thermocycling conditions were as follows: one cycle at 94°C for 2 min; 35 cycles of a denaturing step at 94°C for 15 sec, an annealing step at 55°C for 30 sec, an extension step at 68°C for 6 min and 15 sec; and one cycle of an additional extension at 68°C for 10 min.

For amplification of the near full length genome of CH0200, the first round PCR was carried out using primers Upper1A (5'-AGTGGCGCCCGAACAGG-3'; nt 634-650) and 2.R3.B6R, and the second round PCR was done with primers Upper2 (5'-CTCTCTCGACGCAGGACTCGGCTT-3'; nt 681-704) and LTRD (5'-CTGGAAAGTCCCCAGCGGAAAGTC-3'; nt 9460-9437). Two microliters of the first round PCR products were used for the second round PCR. The PCR thermocycling conditions were as follows: one cycle at 94 °C for 2 min; 10 cycles of a denaturing step at 94 °C for 15 sec, an annealing step at 55 °C for 30 sec, an extension step at 68 °C for 8 min; 20 cycles of a denaturing step at 94 °C for 15 sec, an annealing step at 55 °C for 30 sec, an extension step at 68 °C for 8 min with an incremental 20 sec for each successive cycle; and one cycle of an additional extension at 68 °C for 10 min.

SGA amplicons were sequenced directly by cycle sequencing and dye terminator methods using an ABI 3730xl genetic analyzer (Applied Biosystems, Foster City, CA). Individual sequences were assembled and edited using the Sequencher program 4.10 (Gene Codes, Ann Arbor, MI).

**Estimating days since infection.** The days from infection were estimated using the Poisson model<sup>18</sup> based on pairwise SGA sequence Hamming distances at screening calculated within each distinct lineage and then combined using the algorithm described below. Specifically, we divided each first-time point alignment into single-founder lineages. Sequences significantly enriched for hypermutation ( $p < 0.1$  by Fisher exact test), detected recombinants, and sequences with ambiguous nucleotides were excluded from the analysis. Lineages with less than 3 sequences remaining were also excluded to avoid ambiguities in inferring the lineage consensus. For each of the remaining lineages, we calculated the Hamming distance (HD) of each sequence from the lineage consensus, which was taken to be the lineage T/F, and then merged all the HDs from the same alignment into one distribution. If all T/Fs evolved from the same time point, they were all equally fit and accumulated random mutations at the same rate, then the combined HDs should follow a Poisson distribution. We estimated the parameter of this Poisson distribution and used it to obtain the time since infection<sup>18</sup>. For each subject, time estimates from the two genome-halves were then averaged using a harmonic mean, which minimizes the asymptotic sampling variance.

**Generation of infectious molecular clones.** The full-length genomes for T/F viruses were inferred as described in previous studies<sup>14,25</sup>. Infectious molecular clones (IMCs) were chemically synthesized and constructed for different T/F viruses from three subjects (CH0200, CH0228 and CH0078) as previously described<sup>2,14</sup>. Among them, CH0200a, CH0228a and CH0228b were reported in a previous study<sup>16</sup>. The viral stocks were generated by transfecting the IMCs into 293T cells as described previously<sup>25</sup>. The p24 concentrations of viral stocks were determined with the p24 ELISA kit (PerkinElmer, Waltham, MA) and TCID<sub>50</sub> of viral stocks was determined on TZM-bl cells.

**Generation of Env pseudoviruses.** The *env* gene of each T/F virus was amplified from the SGA-derived PCR amplicon of the 3'-half viral genome using the *Taq* High Fidelity polymerase (Invitrogen Life Technologies, Carlsbad, CA). The *env* amplicons were then gel-purified and ligated into the pcDNA3.1 vector using the pcDNA3.1 Directional TOPO Expression Kit (Invitrogen Life Technologies, Carlsbad, CA) as described previously<sup>41</sup>. All *env* clones were confirmed by sequencing. To produce Env-pseudoviruses, 6 µg of each *env* clone was co-transfected with 10 µg of pNL43-ΔEnv-vpr<sup>+</sup>-luc<sup>+</sup> into 293T cells using the FuGENE6 transfection reagent (Roche Diagnostics; Indianapolis, IN). The culture supernatants containing the pseudoviruses were harvested 72 hours post transfection, aliquoted and stored at -80 °C until use. Infectious titers (TCID<sub>50</sub>) of pseudovirus stocks were determined on TZM-bl cells.

**Determination of coreceptor usage.** The usage of coreceptors was determined on a panel of NP-2 cell lines expressing CD4 together with CCR5, CXCR4 or other G protein-coupled receptors (FPRL-1, APJ, CCR3 and GPR15)<sup>31,42</sup>. The parental NP-2 cell line expressing CD4 but not coreceptors was used as a control. NP-2 cell lines expressing different coreceptors were seeded into a 24-well plate one day before infection at a density of  $1 \times 10^5$  cells per well. On the next day, NP-2 cells were infected with same amount (200 TCID<sub>50</sub>) of each pseudovirus. After 4 hours of incubation at 37 °C, the infected cells were washed twice with complete culture medium and cultured with DMEM with 10% FBS at 37 °C for three days. Infected cells were lysed 72 hours post infection and the infectivity was assessed by measuring relative luciferase units (RLU) of lysed cells from each well. All experiments were performed in triplicates. Viral replication in each NP-2 coreceptor cell line was considered positive when the RLU value was three times above the background RLU value in the parental NP-2 cell line.

**IFN-α inhibition assay.** Peripheral blood mononuclear cells (PBMCs) were obtained by leukopheresis from healthy donors as described previously<sup>27</sup>. Written consent was obtained from the donors and the study was approved by the Duke University Institutional Review Board. All methods were performed in accordance with the relevant guidelines and regulations. The CD4+ T cells were negatively selected from PBMCs with an auto MACS Pro Separator using the CD4+ T cell Isolation Kit II (Miltenyi Biotec, Auburn, CA). Before infection, purified CD4+ T cells were stimulated for 3 days in RPMI 1640 medium containing 10% fetal bovine serum (FBS), 32 U/ml interleukin2 (IL-2) (Advanced Biotechnologies, Columbia, MD), 0.2 mg/ml soluble anti-CD3 (eBioscience, San Diego, CA) and 0.2 mg/ml anti-CD28 (BD Bioscience, San Diego, CA). The stimulated CD4+ T cells were seeded into a 96-well U bottom plate ( $5 \times 10^5$  cells per well) and infected with 500 TCID<sub>50</sub> of each virus (m.o.i of 0.001). After incubation at 37 °C for 4 hours, the cells were washed 3 times with RPMI 1640 and then cultured in a 48-well plate with 500 µl of RPMI 1640 containing 10% FBS and 32 U/ml IL-2. Each virus was cultured in triplicates. Viral replication kinetics was monitored daily for 7 days by measuring p24 concentrations in the culture supernatant.

To determine virus sensitivity to IFN- $\alpha$  inhibition, stimulated CD4 + T cells were infected with each virus as described above. After washing, the infected cells were cultured in 500  $\mu$ l of RPMI 1640 containing 10% FBS, 32 U/ml IL-2 and 500 U/ml IFN- $\alpha$  as previously described<sup>16</sup>. The viral replication in the presence of IFN- $\alpha$  was monitored daily for 7 days by measuring p24 concentrations in culture supernatants. On day 3, half of the culture medium in each well was exchanged with fresh medium containing 500 U/ml IFN- $\alpha$ . To compare the sensitivity to IFN- $\alpha$  inhibition, the ratio of viral p24 production in the presence and absence of IFN- $\alpha$  was calculated for each virus on day 7. All experiments were performed in triplicates.

**PASS fitness assay.** Viral fitness was compared using a single-round competitive PASS fitness assay as we described previously<sup>27</sup>. In brief, purified CD4 + T cells were stimulated with soluble anti-CD3 and anti-CD28 for three days. After stimulation, 50  $\mu$ l of cell suspension ( $1 \times 10^6$ ) was seeded into each well of a 96-well U bottom plate and infected with the vial stock mixture containing equal amount of each compared virus (5 ng p24 per virus;  $\sim 0.001$  m.o.i). After absorption at 37 °C for 4 hours, the cells were washed 3 times with RPMI 1640. The infected cells were then transferred into a 24-well plate and cultured with 600  $\mu$ l of RPMI 1640 containing 10% FBS and 32 U/ml IL-2 for 5 days. The culture supernatant was harvested and replaced with fresh medium daily. The kinetics of virus replication was monitored by measuring the p24 concentration in the culture supernatant. All experiments were performed in triplicates.

The harvested culture supernatants were treated with RNase-free DNase to eliminate the residual plasmid DNA in viral stocks used for transfection. Viral RNA was then extracted from 200  $\mu$ l of the treated supernatant using the PureLink Viral RNA/DNA Mini Kit (Invitrogen, Carlsbad, CA). RNA was eluted with 20  $\mu$ l of RNase-free water. A total of 17  $\mu$ l viral RNA was used for cDNA synthesis using the SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA) with viral specific primers. The cDNA was stored at  $-20$  °C until use.

The parallel allele-specific sequencing (PASS) assay was performed as described previously<sup>27,29</sup>. The in-gel PCR reaction was performed in a PTC-200 Thermal Cycler using amplification primers specifically designed for viruses from each subject (CH0200 forward: 5'-TAGATCCCCAAAATATAACAGACAATGCCAAGAC-3' (nt 7040–7072 in HXB2) and CH0200 reverse: 5'-Acry-CCCAAGAACCCAAGGAACACAGC-3' (nt 7773–7795) for subject CH0200; CH0228 forward: 5'-TGCAATACCTCAGCCATAACACAA-3' (nt 6810–6833) and CH0228 reverse: 5'-Acry-TTCCAGAGCAGCCCCAAATC-3' (nt 8006–8025) for subject CH0228; CH0078 forward: 5'-ATTTTGTGCTCCAGCTGGTTATG-3' (nt 6871–6894) and CH0078 reverse: 5'-Acry-TTTCCTCCATCACGTGTCAATAGTAATCC-3' (nt 7575–7603) for subject CH0078). The proportion of each virus in the culture supernatant was determined by annealing the sequencing primers to the PCR amplicons for single base extension (SBE) with fluorescent-labeled nucleotides specific for each compared virus (CH0200 seq: 5'-ACAGACAATGCCAAGACAATAATAGTACATCTT-3' (nt 7056–7088) for subject CH0200; CH0228 seq: 5'-TGTCCAAAGGTCTCTTTTGACCCAAT-3' (nt 6837–6862) for subject CH0228; CH0078 seq: 5'-GACAATGTCAAACAATAATAGTCCA-3' (nt 7059–7084) for subject CH0078<sup>29</sup>). The relative fitness ( $s_{ij}$ ) was determined as we previously described<sup>27</sup>.

**Statistical analysis.** Values obtained from different T/F viruses were compared using a two-tailed t-test or one-way ANOVA test followed by multiple comparisons using the Prism 6.0 software. P values less than 0.05 were considered significant. Relative fitness differences from each replicate were analyzed as previously described<sup>27</sup>. In particular, these fitness differences were analyzed using a two-stage procedure<sup>43</sup>, assuming they were subject to two independent random measurement errors: a standard error ( $\sigma_i$  for the  $i^{\text{th}}$  replicate) in the determination of the fitness from each of the  $n$  replicates, arising both from Poisson fluctuations in the PASS assay and fluctuations in the viral growth over time [24], and a replicate-to-replicate variance ( $\sigma_R$ ) in the experimental setup. These two variances were separately estimated and added in quadrature to obtain a conservative<sup>44</sup> estimate of the final stand-

ard error  $\sigma = \sqrt{\frac{\sigma_R^2}{n} + \frac{\sum_{i=1}^n \sigma_i^2}{n^2}}$ . Assuming that the  $\sigma_R$  are determined more accurately than  $\sigma_i$ , we use the mean fitness difference along with  $\sigma$  in a t-test to calculate the significance of the fitness difference, with the degrees of freedom given by the one less than the number of replicates. We prefer this as being more conservative estimate than the Satterthwaite approximation which would multiply the degrees of freedom by  $\sigma^4/\sigma_R^4$  (ref. 45). To correct for the number of fitness comparisons being made in this study, we followed Bonferroni<sup>46</sup> and reduced our threshold of significance from the nominal  $p = 0.05$  to  $p = 0.05/n_{\text{comp}}$ , where  $n_{\text{comp}} = 3$  denotes the number of comparisons.

All the calculations of fitness differences were done using the statistical package R<sup>47</sup> and the nonlinear fits used the R package subplex<sup>48</sup> and ucminf<sup>49</sup>.

## References

- Keele, B. F. *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 7552–7557 (2008).
- Li, H. *et al.* High Multiplicity Infection by HIV-1 in Men Who Have Sex with Men. *PLoS pathogens* **6**, e1000890 (2010).
- Haaland, R. E. *et al.* Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1. *PLoS pathogens* **5**, e1000274, doi: 10.1371/journal.ppat.1000274 (2009).
- Abrahams, M. R. *et al.* Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *Journal of virology* **83**, 3556–3567, doi: 10.1128/JVI.02132-08 (2009).
- Chen, Y. *et al.* Comprehensive Characterization of the Transmitted/Founder env Genes From a Single MSM Cohort in China. *J Acquir Immune Defic Syndr* **69**, 403–412, doi: 10.1097/QAI.0000000000000649 (2015).
- Shaw, G. M. & Hunter, E. HIV transmission. *Cold Spring Harbor perspectives in medicine* **2**, doi: 10.1101/cshperspect.a006965 (2012).
- Derdeyn, C. A. *et al.* Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science (New York, N.Y.)* **303**, 2019–2022 (2004).

8. Chohan, B. *et al.* Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1-V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. *Journal of virology* **79**, 6528–6531 (2005).
9. Sagar, M. *et al.* Selection of HIV variants with signature genotypic characteristics during heterosexual transmission. *The Journal of infectious diseases* **199**, 580–589, doi: 10.1086/596557 (2009).
10. Frost, S. D. *et al.* Characterization of human immunodeficiency virus type 1 (HIV-1) envelope variation and neutralizing antibody responses during transmission of HIV-1 subtype B. *Journal of virology* **79**, 6523–6527 (2005).
11. Wilen, C. B. *et al.* Phenotypic and immunologic comparison of clade B transmitted/founder and chronic HIV-1 envelope glycoproteins. *Journal of virology* **85**, 8514–8527, doi: 10.1128/JVI.00736-11 (2011).
12. Gnanakaran, S. *et al.* Recurrent signature patterns in HIV-1 B clade envelope glycoproteins associated with either early or chronic infections. *PLoS pathogens* **7**, e1002209, doi: 10.1371/journal.ppat.1002209 (2011).
13. Asmal, M. *et al.* A signature in HIV-1 envelope leader peptide associated with transition from acute to chronic infection impacts envelope processing and infectivity. *PLoS one* **6**, e23673, doi: 10.1371/journal.pone.0023673 (2011).
14. Salazar-Gonzalez, J. F. *et al.* Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *The Journal of experimental medicine* **206**, 1273–1289 (2009).
15. Ochsenbauer, C. *et al.* Generation of transmitted/founder HIV-1 infectious molecular clones and characterization of their replication capacity in CD4 T lymphocytes and monocyte-derived macrophages. *Journal of virology* **86**, 2715–2728, doi: 10.1128/JVI.06157-11 (2012).
16. Parrish, N. F. *et al.* Phenotypic properties of transmitted founder HIV-1. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6626–6633, doi: 10.1073/pnas.1304288110 (2013).
17. Fenton-May, A. E. *et al.* Relative resistance of HIV-1 founder viruses to control by interferon-alpha. *Retrovirology* **10**, 146, doi: 10.1186/1742-4690-10-146 (2013).
18. Giorgi, E. E. *et al.* Estimating time since infection in early homogeneous HIV-1 samples using a poisson model. *BMC bioinformatics* **11**, 532, doi: 10.1186/1471-2105-11-532 (2010).
19. Shimizu, N. *et al.* A formylpeptide receptor, FPRL1, acts as an efficient coreceptor for primary isolates of human immunodeficiency virus. *Retrovirology* **5**, 52, doi: 10.1186/1742-4690-5-52 (2008).
20. Edinger, A. L. *et al.* An orphan seven-transmembrane domain receptor expressed widely in the brain functions as a coreceptor for human immunodeficiency virus type 1 and simian immunodeficiency virus. *Journal of virology* **72**, 7934–7940 (1998).
21. He, J. *et al.* CCR3 and CCR5 are co-receptors for HIV-1 infection of microglia. *Nature* **385**, 645–649, doi: 10.1038/385645a0 (1997).
22. Choe, H. *et al.* The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* **85**, 1135–1148 (1996).
23. Parrish, N. F. *et al.* Phenotypic properties of transmitted founder HIV-1. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6626–6633, doi: 10.1073/pnas.1304288110 (2013).
24. Deymier, M. J. *et al.* Heterosexual Transmission of Subtype C HIV-1 Selects Consensus-Like Variants without Increased Replicative Capacity or Interferon-alpha Resistance. *PLoS pathogens* **11**, e1005154, doi: 10.1371/journal.ppat.1005154 (2015).
25. Jiang, C. *et al.* Primary infection by a human immunodeficiency virus with atypical coreceptor tropism. *Journal of virology* **85**, 10669–10681, doi: 10.1128/JVI.05249-11 (2011).
26. Ritchie, A. J. *et al.* Recombination-mediated escape from primary CD8 + T cells in acute HIV-1 infection. *Retrovirology* **11**, 69, doi: 10.1186/s12977-014-0069-9 (2014).
27. Song, H. *et al.* Impact of immune escape mutations on HIV-1 fitness in the context of the cognate transmitted/founder genome. *Retrovirology* **9**, 89, doi: 10.1186/1742-4690-9-89 (2012).
28. Song, H. *et al.* Reversion and T cell escape mutations compensate the fitness loss of a CD8 + T cell escape mutant in their cognate transmitted/founder virus. *PLoS one* **9**, e102734, doi: 10.1371/journal.pone.0102734 (2014).
29. Cai, F. *et al.* Detection of minor drug-resistant populations by parallel allele-specific sequencing. *Nature methods* **4**, 123–125, doi: 10.1038/nmeth995 (2007).
30. Liu, D. *et al.* Preexisting compensatory amino acids compromise fitness costs of a HIV-1 T cell escape mutation. *Retrovirology* **11**, 101, doi: 10.1186/s12977-014-0101-0 (2014).
31. Nedellec, R. *et al.* Virus entry via the alternative coreceptors CCR3 and FPRL1 differs by human immunodeficiency virus type 1 subtype. *Journal of virology* **83**, 8353–8363 (2009).
32. Isaacman-Beck, J. *et al.* Heterosexual transmission of human immunodeficiency virus type 1 subtype C: Macrophage tropism, alternative coreceptor use, and the molecular anatomy of CCR5 utilization. *Journal of virology* **83**, 8208–8220 (2009).
33. Sandler, N. G. *et al.* Type I interferon responses in rhesus macaques prevent SIV infection and slow disease progression. *Nature* **511**, 601–605, doi: 10.1038/nature13554 (2014).
34. Abel, K., Roche, D. M., Chohan, B., Fritts, L. & Miller, C. J. Temporal and anatomic relationship between virus replication and cytokine gene expression after vaginal simian immunodeficiency virus infection. *Journal of virology* **79**, 12164–12172, doi: 10.1128/JVI.79.19.12164-12172.2005 (2005).
35. Stacey, A. R. *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. *Journal of virology* **83**, 3719–3733, doi: 10.1128/JVI.01844-08 (2009).
36. Arnott, A. *et al.* High viral fitness during acute HIV-1 infection. *PLoS one* **5**, doi: 10.1371/journal.pone.0012631 (2010).
37. Kong, X. *et al.* The human immunodeficiency virus type 1 envelope confers higher rates of replicative fitness to perinatally transmitted viruses than to nontransmitted viruses. *Journal of virology* **82**, 11609–11618 (2008).
38. Quinones-Mateu, M. E. *et al.* A dual infection/competition assay shows a correlation between *ex vivo* human immunodeficiency virus type 1 fitness and disease progression. *Journal of virology* **74**, 9222–9233 (2000).
39. Wilen, C. B. *et al.* Phenotypic and immunologic comparison of clade B transmitted/founder and chronic HIV-1 envelope glycoproteins. *Journal of virology* **85**, 8514–8527, doi: 10.1128/JVI.00736-11 (2011).
40. Sanchez, A. M. *et al.* Development of a contemporary globally diverse HIV viral panel by the EQAPOL program. *Journal of immunological methods* **409**, 117–130, doi: 10.1016/j.jim.2014.01.004 (2014).
41. Kirchherr, J. L. *et al.* High throughput functional analysis of HIV-1 env genes without cloning. *J Virol Methods* **143**, 104–111 (2007).
42. Shimizu, N. *et al.* Broad usage spectrum of G protein-coupled receptors as coreceptors by primary isolates of HIV. *Aids* **23**, 761–769 (2009).
43. Steimer, J. L., Mallet, A., Golmard, J. L. & Boisvieux, J. F. Alternative approaches to estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed-effect model. *Drug metabolism reviews* **15**, 265–292, doi: 10.3109/03602538409015066 (1984).
44. Davidian, M. & Giltinan, D. M. *Nonlinear models for repeated measurement data*. (Chapman & Hall/CRC, 1998).
45. Satterthwaite, F. E. An approximate distribution of estimates of variance components. *Biometrics* **2**, 110–114 (1946).
46. Bonferroni, C. E. *Teoria statistica delle classi e calcolo delle probabilità*. (Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 1936).
47. R: A Language and Environment for Statistical Computing, R Core Team, R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org> (2016).
48. subplex: Subplex optimization algorithm. R package version 1.1-3. <http://CRAN.R-project.org/package=subplex> (2008).
49. ucminf: General-purpose unconstrained non-linear optimization. R package version 1.1-3. <http://CRAN.R-project.org/package=ucminf> (2012).



## Acknowledgements

This work was supported by the Center for HIV/AIDS Vaccine Immunology (AI067854), the Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (AI100645) and RO1 grant (AI028433) from National Institutes of Health. We thank Anna Berg for technical assistance.

## Author Contributions

H.S. and F.G. conceived and designed the experiments. H.S., B.H., A.K. and F.C. performed experiments. T.B., E.E.G. and A.S.P. performed mathematical modeling and statistical analysis of the fitness results. H.S., T.B., E.E.G., A.S.P. and F.G. analyzed the data. H.S., T.B., E.E.G., A.S.P. and F.G. wrote and edited the manuscript. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Song, H. *et al.* Transmission of Multiple HIV-1 Subtype C Transmitted/founder Viruses into the Same Recipients Was not Determined by Modest Phenotypic Differences. *Sci. Rep.* **6**, 38130; doi: 10.1038/srep38130 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016