

SCIENTIFIC REPORTS



OPEN

Escape from humoral immunity is associated with treatment failure in HIV-1-infected patients receiving long-term antiretroviral therapy

Yabo Ouyang^{1,5}, Qianqian Yin¹, Wei Li^{2,3}, Zhenpeng Li¹, Desheng Kong¹, Yanling Wu², Kunxue Hong¹, Hui Xing¹, Yiming Shao¹, Shibo Jiang^{2,4}, Tianlei Ying^{1,2} & Liying Ma¹

Interindividual heterogeneity in the disease progression of HIV-1-infected patients receiving long-term antiretroviral therapy suggests that some host-related factors may have limited treatment efficacy. To understand the nature of factors contributing to treatment failure, we performed a retrospective cohort study of 45 chronically HIV-1-infected individuals sharing a similar demographics and route of infection, compared the differences between virologically suppressed (VS) and treatment failure (TF) patients with respect to clinical, immunological and virological characteristics. We found that the baseline diversity of HIV-1 env quasispecies was the major difference between VS and TF group, and higher baseline diversity in TF patients. We further predicted TF-related env mutations using a selection pressure-based approach, followed by an analysis of these mutations based on the available three-dimensional structures of gp120/gp41 or their complexes with neutralizing antibodies. Notably, almost all of the identified residues could be mapped to the epitopes of known HIV-1 neutralizing antibodies, especially the epitopes of broadly neutralizing antibodies, and these mutations tended to compromise antibody-antigen interactions. These results indicate that the escape of HIV-1 from host humoral immunity may play a direct role in TF in long-term antiretroviral-experienced patients and that based on env gene sequence of the viruses in the patients.

Rapid replication dynamics and high mutation rate of human immunodeficiency virus type 1 (HIV-1) allows the virus to evolve continuously and quickly during the course of infection¹. With the advent of highly active antiretroviral therapy (ART), a combination of three or more potent antiretroviral drugs, profound suppression of viral replication has been achieved, and HIV-1-associated mortality has been dramatically reduced^{2,3}. Nonetheless, considerable variation of the virus is observed among patients in response to treatment, with a substantial proportion of individuals experiencing treatment failure, including virologic failure, i.e., inability to achieve or maintain suppression of viral replication; immunologic failure, i.e., failure of CD4 count to increase; and subsequent clinical failure^{4,5}. This interindividual heterogeneity in disease progression suggests that some host-related factors may have affected the effectiveness of ART. An understanding of such contributing factors associated with treatment failure is critical to developing an effective long-term HIV treatment strategy.

Host immune response elicited by HIV-1 contributes to the repression of viral replication, but it varies greatly in different individuals⁶. Some HIV-1-infected individuals are able to develop heavily mutated, broadly neutralizing antibodies (bnAbs) capable of neutralizing a broad spectrum of HIV-1 isolates^{7–10}. These bnAbs only

¹State Key Laboratory of Infectious Disease Prevention and Control, National Center for AIDS/STD Control and Prevention (NCAIDS), Chinese Center for Disease Control and Prevention, Beijing, China and Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, China. ²Key Laboratory of Medical Molecular Virology of MOE/MOH, School of Basic Medical Sciences, Fudan University, Shanghai, China. ³Protein Interactions Section, Cancer and Inflammation Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, Maryland, USA. ⁴Lindsley F. Kimball Research Institute, New York Blood Center, New York, USA. ⁵Beijing You'an Hospital, Capital Medical University, Beijing, China and Beijing Institute of Hepatology, Beijing, China. Yabo Ouyang and Qianqian Yin contributed equally to this work. Correspondence and requests for materials should be addressed to S.J. (email: shibojiang@fudan.edu.cn) or T.Y. (email: tlying@fudan.edu.cn) or L.M. (email: mal@chinaaids.cn)

appear after several years of HIV-1 infection, suggesting the continuous viral escape and selection by autologous neutralizing antibodies^{11,12}. It is also intriguing to observe that a small group of HIV-1-infected individuals (<1%), termed elite controllers, spontaneously maintain undetectable levels of viral replication in the absence of antiretroviral therapy by mounting robust immune responses directed at HIV-1^{13,14}. Moreover, the escape mutants in the envelope glycoprotein (Env) during early HIV-1 infection primarily involved changes in N-linked glycosylation that could sterically inhibit the accessibility of neutralizing epitopes^{11,15}. All these findings suggest a critical role of selective pressure exerted by host humoral immune response in driving HIV-1 evolution. However, in ART-treated patients, the selective pressure is predominantly governed by antiretroviral drugs, and the effects of HIV-1 evolution and escape from host immune selective pressure on disease progression under long-term ART are largely unknown.

Therefore, to identify host and viral factors associated with differing clinical outcomes in HIV-1-infected patients receiving long-term antiretroviral treatment, we performed a retrospective cohort study of 45 chronically HIV-1-infected individuals sharing a similar demographic and ethnic background. All these patients were infected with HIV-1 subtype B through a similar route of infection, and they were enrolled in the National Free Antiretroviral Treatment Program in Anhui and Henan provinces, China. These patients were categorized into two groups: virologically suppressed (VS, n = 25) and treatment failure (TF, n = 20), according to disease outcome. The differences between the two groups with respect to clinical, immunological and virological characteristics were then compared. Interestingly, diversity of HIV-1 *env* quasispecies was found to be the major difference between the two groups. More specifically, compared to those in individuals in the VS group, both gp120 and gp41 gene quasispecies of HIV-1 in the TF patients showed higher baseline diversity. Treatment failure-related adaptive mutations were further predicted by using a selection pressure-based analysis of single-genome amplification (SGA)-derived full-length HIV-1 *env*. Notably, most of the identified mutations could be mapped to epitopes of known HIV-1 neutralizing antibodies, especially epitopes of bnAbs. These results demonstrate that the escape of HIV-1 from host humoral immunity contributes to disease progression in patients on long-term ART.

Results

Population characteristics. All 45 enrolled patients, who participated in a multicenter AIDS Cohort Study in Anhui and Henan provinces in China from 2004 to 2011, shared a common ethnic background and a similar route of infection, namely plasma donation. Similarity of immunological and viral genetic backgrounds provided a unique opportunity to explore the relationship between viral evolution and disease progression in different individuals.

Patients were divided into two groups according to disease outcome. The virologic suppression (VS) group was defined as viral load persistently <400 copies/ml after ART for 6 months. The treatment failure (TF) group was defined as repeated viral load >400 copies/ml after 6 months. By the close of the study period, 20 patients were classified as TF, and 25 patients were classified as VS. Baseline demographic and clinical features of all patients are shown in Table 1. The age, gender, high-risk behavior, and initial regimen showed no significant difference between VS and TF groups. Neither viral load nor CD4⁺ cell count differed between the two groups before treatment. Similar to the TF group, drug-resistant mutations in the pol region of the VS group were rarely found (P = 0.309, data not shown). Although there is a significant difference in switching regimens among the treatment failure cases, it occurred after the “treatment failure”.

Most individuals displayed declining trends in viral load and increased CD4⁺ cells after 6 months of ART (Fig. 1a). In the VS group, viral load levels remained at an undetectable level (≤400 copies/ml) after ART for 6 months (median treatment duration of 60 months). However, viral load in the TF group fluctuated widely, and detectable viral load was found during ART (median treatment duration of 63 months).

Env sequences and phylogenetic relationship. For an accurate representation of HIV-1 *env* quasispecies *in vivo*, we used nested PCR single-genome amplification (SGA) to amplify the full-length *env* genes as previously reported¹⁶, whereby PCR products were derived from a single template molecule. Plasma collected at baseline and during ART was used to evaluate *env* clonal genotypes. A total of 1,021 single-genome full-length gp160 amplicons were obtained for cross-sectional and longitudinal analysis by SGA. Among these, 638 sequences derived from plasma of VS (n = 352) and TF (n = 286) at baseline were used for cross-sectional analysis. Analysis of gp160 nucleotide sequences at baseline showed that all sequences clustered phylogenetically with subtype B reference sequences (Fig. 1b). A neighbor-joining phylogenetic tree, including all sequences, showed no evidence of cross-contamination between study subjects or contamination with laboratory strains. As expected, the sequences for each patient were clustered together and formed monophyletic groups, except subjects 2124 and 3594 who were infected with two unrelated HIV-1 variants of subtype B.

A total of 552 single-genome full-length gp160 amplicons, including 169 sequences from baseline, derived from plasma of 12 TF patients, who had never switched regimens and had good adherence until the last follow-up point, were used for longitudinal analysis (Fig. 1c). TF patients were included to compare *env* diversity between baseline (T0) and during ART. Sequence diversity was calculated using the Kimura 2-parameter model option in MEGA 5. The number of sequences of 12 patients included in the analysis was shown in Table S3. During ART, gp160 nucleotide sequences were crossed with each other or formed a separate cluster according to treatment time.

Genetic characteristics of HIV-1 *env*. Demographic characteristics and baseline drug-resistant mutations were the same between the VS and TF groups; therefore, we sought to investigate the difference between the two groups based on the diversity of HIV-1 *env* quasispecies. Quasispecies heterogeneity was quantitated by the mean genetic distance (d) of the sequences in each group. As shown in Fig. 2a, a relatively high level of heterogeneity occurred within V1-V2, V4 and V5 regions of gp120. The TF group exhibited higher baseline diversity in both

	VS (n = 25)	TF (n = 20)	P Value	All Patients (n = 45)
Age, median years (Range) ^a	37 (26–64)	40 (30–56)	0.272	38 (26–64)
Male, No. (%) ^b	7 (28.00)	9 (45.00)	0.236	16 (35.56)
High-risk behavior, No. (%) ^c			1.000	
Former plasma donation	24 (96.00)	20 (100.00)		44 (97.78)
Sex	1 (4.00)	0 (0.00)		1 (2.22)
Initial Regimen, No. (%) ^c			0.314	
AZT + NVP + 3TC	13 (52.00)	12 (60.00)		25 (55.56)
D4T + NVP + 3TC	12 (48.00)	6 (30.00)		18 (40.00)
D4T + EFV + 3TC	0 (0.00)	1 (5.00)		1 (2.22)
AZT + EFV + 3TC	0 (0.00)	1 (5.00)		1 (2.22)
Viral Load at baseline, median log ₁₀ copies/ml (Range) ^a	4.84 (3.21–6.98)	5.20 (4.07–5.98)	0.568	5.07 (3.21–6.98)
CD4 ⁺ count at baseline, median cells/ μ l (Range) ^a	166 (15–515)	247 (83–459)	0.110	219 (15–515)
Treatment time, median months (Range) ^a	60 (16–72)	63 (26–72)	0.670	63 (16–72)
Have switched regimens, No. (%) ^c	0 (0.00)	5 (25.00)	0.013	5 (11.11)
Have missed ART dose in last 7 days, No. (%) ^c	0 (0.00)	3 (15.00)	0.080	3 (6.67)
Have related drug-resistant mutation at baseline, No. (%) ^c	1 (4.00)	3 (15.00)	0.309	4 (8.89)

Table 1. Demographic and clinical characteristics of HIV-1 subtype B'-infected patients. ^aMann-Whitney Test; ^bChi-Square Test; ^cFisher's Exact Test; Abbreviations: 3TC, lamivudine; AZT, zidovudine; d4T, stavudine; EFV, efavirenz; NVP, nevirapin.

gp120 ($p = 0.019$) and gp41 ($p = 0.011$) compared to that of the VS group. Interestingly, the most significant difference between the diversity of VS and TF groups occurred in the gp120 C2 region ($p = 0.005$).

In the longitudinal dataset, TF patients who never switched regimens were included to compare *env* diversity between baseline (T_0) and during ART. Following baseline, two time points were selected based on viral load level: 1) when viral load reached a level >400 copies/ml (T_1) and 2) maximum value of viral load (T_2). A total of 536 gp160 sequences derived from 12 TF patients were analyzed and displayed no significant difference between baseline, T_1 and T_2 in gp160, or its subregion diversity (Fig. 2b). Although the TF group had a higher viral diversity at baseline, these results indicate that the viral *env* quasispecies did not further increase in these patients during ART.

We also analyzed the length of amino acid sequences and the number of potential N-linked glycosylation sites (PNGs) of HIV-1 *env* at baseline. No evident difference was seen in either length of amino acid sequence or number of PNGs in gp120 between the VS and TF groups (Table S1). Although the TF group exhibited relatively shorter length of amino acid sequence in gp41, the number of PNGs was no different from that of the VS group. Overall, no substantive difference was observed in gp160 between the two groups with respect to either length of amino acid sequence or number of PNGs.

The gp120 V3 loop has been suggested as a major viral determinant for co-receptor preference¹⁷. Thus, we also compared the V3 loops of the VS and TF groups. As shown in Fig. 2c, no evident difference was detected in the charge and amino acid usage of V3 loops between the two groups. The two groups also exhibited similar tetramer compositions of the tip of V3 loops, typically GPGR and GPGQ.

To evaluate variables that have contributed to treatment outcome at baseline, a logistic regression analysis was performed. Variables, including CD4⁺ T cell count, viral load, gp160 diversity, drug-resistant mutations and co-receptor usage, were all pooled for analysis. Interestingly, gp160 diversity was the only variable that remained significant in the final model, and it was the best single variable for predicting TF for 75.6% individuals in this cohort (Fig. 2d). An area under the curve (AUC) of 0.794 (95% confidence interval 0.657–0.931) was obtained in the ROC curve, with a sensitivity and positive predictive value of 75.0% at a 0.50 cutoff and a specificity and negative predictive value of 80.0%. Taken together, these results suggest that higher gp160 diversity significantly associated to the risk of treatment failure.

Determination of treatment failure-related HIV-1 *env* adaptive mutation. To gain more insight into the difference in *env* quasispecies diversity between VS and TF groups, a selection pressure-based analysis was performed to investigate patterns of *env* evolution. The ratios of non-synonymous (K_a) and synonymous (K_s) nucleotide substitutions were calculated, as described in the Methods section, and used as an indicator of selective pressure acting on *Env*-coding genes. Treatment failure-related adaptive mutations were screened by the K_a/K_s ratio of individual residues as follows: (a) $K_a/K_s > 1$ and sequence frequency > 0.01 in the TF group, while $K_a/K_s < 1$, or $K_a/K_s > 1$, but $LOD < 2$, in the VS group; (b) sequence frequency of TF group higher than that of VS group ($p < 0.05$), while the corresponding population frequency of the TF group was over 0.15.

In total, 36 and 10 putative adaptive mutations related to treatment failure from gp120 and gp41 regions were obtained, respectively (Table 2). Among them, 4 critical mutations (G145N, E150S, I371V and N392S) of gp120 (Fig. 2e) and 5 critical mutations (A607N, E662K, N677K, E734G and V832L) of gp41 (Fig. 2f) were further predicted to be directly related to treatment failure using Bayesian network. The G145N, I371V, A607N, E662K and N677K mutations still existed with high frequency during ART (Table S2). No mutations were observed in the

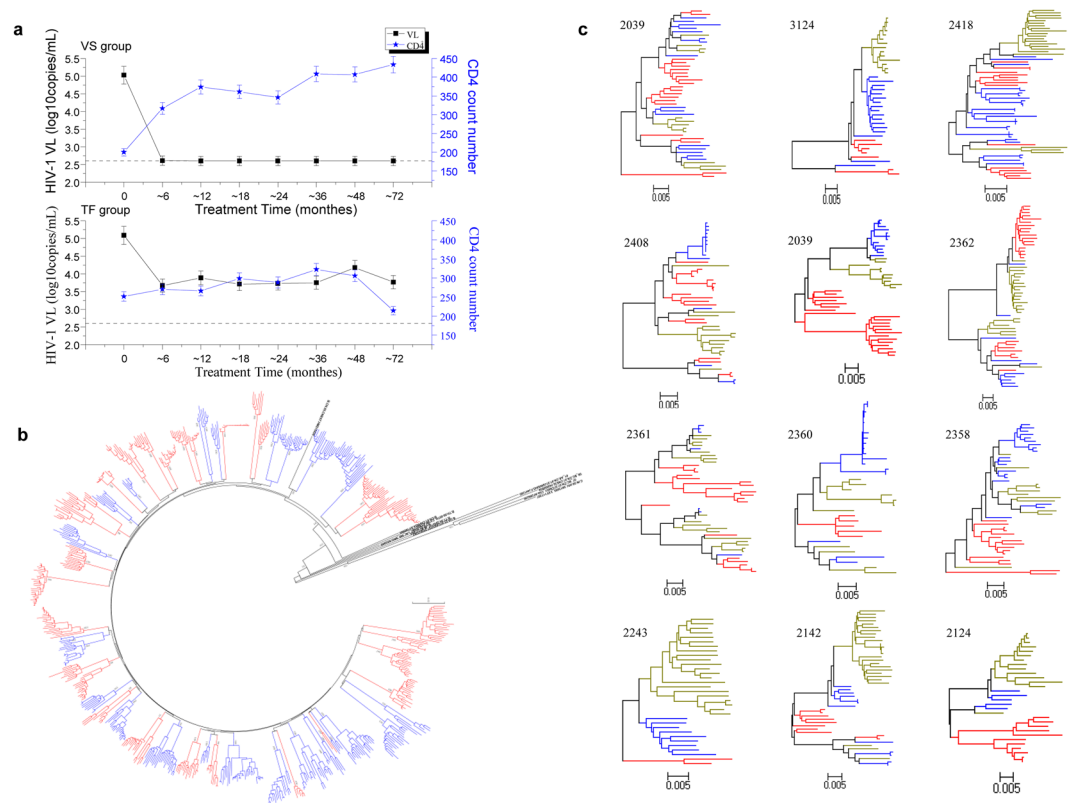


Figure 1. (a) Longitudinal changes in viral load (square) and HIV-1 CD4⁺ T cell count (pentacle) for VS (upper panel) and treatment failure TF (upper panel) during ART. HIV-1 vial load level of all VS was undetectable during ART, while the viral load levels of most TF fluctuated widely during ART, albeit lower than baseline viral load. Data are presented in mean \pm standard deviation (SD) from three replicates. (b) Phylogenetic trees of 638 gp160 sequences at baseline between VS and TF groups. Each color represents one group (red: VS; blue: TF). Bootstrap values >85% are shown. No sequences showed evidence of cross-contamination between study subjects or contamination with laboratory strains. As expected, all sequences for each patient were clustered together. (c) Phylogenetic analysis of 552 gp160 amplicons, including 169 sequences from baseline, derived from 12 TF patients (red: baseline; blue: when viral load reached a level of >400 copies/ml; yellow: the maximum value of viral load).

N-terminal heptad repeat (NHR) region (residues 542–592) of gp41, the target site of the HIV-1 fusion inhibitors, such as enfuvirtide (also known as T20) derived from the gp41 C-terminal heptad repeat (CHR) region¹⁸.

Afterwards, using a flow cytometry-based multiple-cycle growth competition assay, as described previously¹⁹, we also measured the relative replication fitness of several treatment failure-related Env adaptive mutations, including G145N, I371V, A607N, E662K and N677K mutant viruses. Although the $1 + s$ fitness values of A607N and N677K increased slightly, the $1 + s$ fitness values of G145N, I371V and E662K viruses were not significantly different from those of the wild-type, suggesting that the change in viral fitness of these variants may not be a major determinant of treatment outcome (Fig. 2g).

Mapping of treatment failure-related mutations on HIV-1 Env structure. To better understand the mutations identified by bioinformatic prediction, we mapped these residues on the high-resolution structures of HIV-1 gp120 (PDB ID 4TVP) and gp41 (PDB ID 2PV6). To our surprise, almost all of these residues (28 of 31 in gp120 exterior Env and 2 of 3 in the outer membrane part of gp41) appear to be involved with possible epitopes of previously reported neutralizing antibodies (Fig. 3). In the gp120 region, residues 134, 135, 137, 138, 145, 149 and 150 are located in a small flexible loop of the V1/V2 region which is accessible to a number of V1/V2-directed antibodies^{20–23}. For example, the bnAb PGT122 recognizes glycan and protein elements in both V3 and V1/V2 region, including residues in the 134–150 loop²². Residues 166 and 170 are located in another area of the V1/V2 region and can be targeted by another panel of neutralizing antibodies. The bnAbs PG9 and PG16 have direct interaction with the gp120 residue 170^{24,25}. The surface area where residues 179, 183 and 194 are located is part of the epitopes of mAb 830A²⁶. Residues 232, 240, 242, 267 and 269 form a surface that can be targeted by the 35O22-like bnAbs, and 35O22 has direct interaction with gp120 residue 240²⁷. The surface comprised of residues 292, 293 and 295 can also be accessed by neutralizing antibodies, such as mAb b12²⁸. Residue 371 is located in the CD4-binding site and has direct interaction with a number of bnAbs, including VRC-class antibodies (e.g., VRC01, VRC06, VRC07 and VRC23)^{29–31}, CH103³², b12 and F105³³, NIH45-46³⁴, 12a21³⁵, and 3bnc117³⁶. The surface area where residues 392, 396, 397, 402, 405 and 407 are located can be accessed by some glycan-dependent

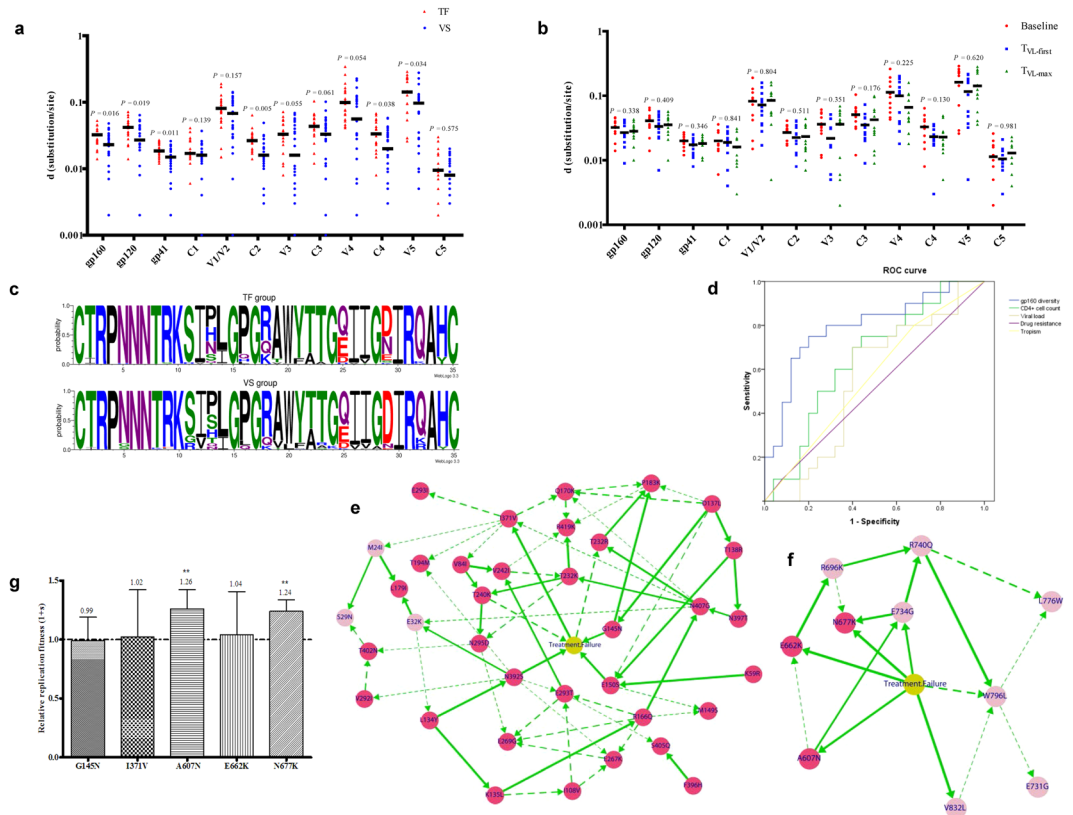


Figure 2. Comparison of *env* nucleotide diversity between patients with virologic suppression (VS) and treatment failure (TF) at baseline (**a**) and comparison of *env* nucleotide diversity from 12 patients with TF at three time points (**b**) (T₀: baseline; T₁: when viral load reached a level of >400 copies/ml; T₂: the maximum value of viral load). The short lines represent the median value. P values were calculated by nonparametric Mann-Whitney test. (**c**) Consensus V3 sequences of VS group and TF group. The values on the y-axis indicate the percentage of each site, and relative height of each amino acid reflects its percentage at that site. (**d**) Logistic regression and receiver operating characteristic (ROC) analysis were performed to develop a predictive model and predictive value of HIV-1 *env* diversity for therapy outcome. Among the covariates including CD4+ T cell count, viral load, gp160 diversity, drug-resistant mutations and co-receptor usage, diversity of gp160 was the best single variable for predicting TF for 75.6% of individuals in this cohort, and an area under the curve (AUC) of 0.794 (CI95%, 0.657–0.931) was obtained in the ROC curve. Bayesian network analysis was used to screen therapy outcome-related putative adaptive mutations. Four mutation sites (G145N, E150S, I371V and N392S) of gp120 (**e**) and five mutation sites (A607N, E662K, N677K, E734G and V832L) of the gp41 region (**f**) were further predicted using Bayesian network. The width of edge was proportional to bootstrap support. The solid edges were considered to be robust and had bootstrap support over 65%, while dashed edges had bootstrap support between 35% and 65%. (**g**) Relative replication fitness of outcome-related putative adaptive mutants. Each mutant virus was competed against the wild-type in PM1 cells. A fitness value of 1.0 indicated that fitness was equal to that of wild type. Each bar represented the mean “1 + s value” ± SD from three replicates. Significant differences between mutant virus and wild-type virus were marked with asterisks (**p < 0.01).

bnAbs. For example, bnAb PGT135 recognizes the glycosylation site Asn392³⁷. Residue 419 also interacts with some neutralizing antibodies, including b12, 17b³⁷ and X5³⁸.

In gp41, 2 of the 3 identified residues are located in the highly conserved membrane-proximal external (MPER) region in the outer membrane part of gp41. Residue 662 can directly interact with bnAbs 2F5 and m66^{39,40}, while residue 677 can interact with bnAbs 4E10 and 10E8^{41,42}. Taken together, these results indicate a significant correlation between treatment failure-related mutations and the epitopes of HIV-1 neutralizing antibodies.

Molecular docking. We next employed molecular docking to investigate if the treatment failure-related adaptive mutations of HIV-1 Env would compromise the binding capability of bnAbs. By using the Z-DOCK program⁴³, three relatively high-frequency mutations, including I371V in gp120 CD4-binding loop within the epitope of VRC01; Q170K in gp120 V2 loop within the epitope of CH59; and T240K in gp120 seven-stranded sheets within the epitope of 35O22, were evaluated to determine how extensively these mutations could affect antibody-gp120 binding. For the docked VRC01/gp120 binding³⁶, a Z-rank was obtained with a top pose score of −120.767 for indicating to wild-type gp120, in contrast to the score of −112.569 for binding to mutant gp120 I371V, indicating that the binding of VRC01/gp120 encounters a more energetically favorable docking

Mutation ^a	bnAbs	Population frequency		Region ^a	TF group		VS group		
		TF group	VS group		Ka/Ks	Sequence frequency	Ka/Ks	Sequence frequency	
K59R	—	0.15	0.04	C1	24	0.08	0.75	0.04	
V84I	—	0.2	0.04	C1	30.95	0.06	0.25	0	
I108V	—	0.2	0.08	C1	17.27	0.08	2.18	0.01	
L134Y	PGT122	0.2	0.04	V1-V2	7.2	0.13	0.82	0.03	
K135L		0.15	0.04	V1-V2	7.67	0.08	0.35	0.03	
D137L		0.15	0	V1-V2	27	0.09	—	—	
T138R		0.15	0	V1-V2	21	0.07	—	—	
G145N		0.5	0.08	V1-V2	58	0.2	0.67	0.01	
M149S		0.15	0	V1-V2	4	0.01	—	—	
E150S		0.15	0	V1-V2	40	0.14	—	—	
R166Q		PG9, PG16	0.15	0	V1-V2	25	0.09	—	—
Q170K	0.15		0	V1-V2	72.26	0.09	—	—	
L179I	830A	0.15	0	V1-V2	138.6	0.1	—	—	
P183K		0.15	0	V1-V2	30	0.1	—	—	
T194M		0.2	0.04	V1-V2	7	0.02	2	0.01	
T232K	35O22	0.3	0.16	C2	3.87	0.18	0.97	0.08	
T232R		0.15	0	C2	2.43	0.11	—	—	
T240K		0.15	0.04	C2	71.69	0.1	2.38	0	
V242I		0.15	0.04	C2	1.99	0.08	0.04	0	
E267K		0.15	0.04	C2	1.7	0.06	0.5	0	
E269G		0.25	0.12	C2	2.27	0.09	0.69	0.05	
V292I		B12	0.25	0.08	C2	2.46	0.04	1.04	0.01
E293T			0.15	0.04	C2	23	0.08	0.5	0
E293I	0.2		0.04	C2	7	0.02	0.5	0	
N295D	0.3		0.04	C2	3.5	0.07	0.33	0	
I371V	VRC01	0.55	0.04	C3	27.65	0.21	1.36	0	
N392S	PGT135	0.2	0.04	V4	1.18	0.14	0.03	0	
F396H		0.15	0	V4	14	0.05	—	—	
N397T		0.2	0	V4	63.93	0.08	—	—	
T402N		0.2	0.04	V4	47.79	0.04	14.27	0.01	
S405Q		0.15	0.04	V4	1.18	0.05	0.1	0.01	
N407G		0.5	0.28	V4	1.57	0.25	0.98	0.14	
R419K		B12	0.3	0.08	C4	1.14	0.13	0.55	0.03
A607N	—	0.15	0	gp41	46	0.16	—	—	
E662K	2F5	0.35	0	gp41	61	0.21	—	—	
N677K	4E10	0.8	0.44	gp41	3.33	0.61	0.94	0.27	

Table 2. Therapy failure-related adaptive mutations. ^aHXB2 Env was used as a reference sequence.—: no associated bnAbs.

compared to the binding of VRC01/gp120-I371V. In detail, VRC01 VH moved closer to the gp120 CD4-binding loop by ~ 1 Å shift in response to the I371 to V371 mutation (Fig. 4a), resulting in a conformation of VRC01 VH more like that of CD4, and the VRC01 VL disengaged from V5 loop to release steric clash with gp120 inner domain core. Consequently, the clasped V5 loop at the VRC01 VH/VL interface, which constitutes an important factor for V5 sequence variation tolerance of VRC01 neutralization breadth²⁹, became more solvent-exposed in the gp120 I371V mutant. Thus, it is possible that the breadth and potency of VRC01 may be compromised in neutralizing HIV-1 bearing the gp120 I371V mutation. In addition, we found that I371_{gp120} is close to A56_{VRC01} with a distance of 3.6 Å for VRC01/gp120, indicating an σ - σ super-conjugation interaction which is compromised in VRC01/gp120-I371V with a longer distance of 4.3 Å. Similarly, for the docked CH59/gp120 complex⁴⁴, CH59 binds more favorably to wild-type gp120 than to gp120 with Q170K mutation (Z-score of -109.932 for wild type vs. -86.583 for Q170K mutant). Notably, the strong H-bond of Q170_{gp120} NE2-N31_{CH59} OD1, which exists at the CH59/gp120 interface with a distance of 2.1 Å and is the major non-bonding interaction for associating CH59 and gp120 V2 peptide, was found to be broken in the CH59/gp120-Q170K complex (Fig. 4b). For the docked 35O22/gp160 binding²⁷, the H-bond of T240_{gp120}-E72_{35O22} in 35O22/gp160 complex was evidently perturbed by T240 to K240 mutation, in which the long side chain of lysine elicits substantial steric clash with the glutamine carbonate group of E72_{35O22}. Consequently, to facilitate 35O22 binding, the gp120 T240K mutant needs to expel the protruding VH CDR2 loop of 35O22 by a ~ 2.1 Å shift from the binding cleft formed by the gp120 seven-stranded sheets group and gp41 $\alpha 8$ helix (Fig. 4c). Such change would compromise the hydrophobic interactions, consisting of F72H_{35O22} and W45_{gp120}, L86_{gp120}, and V89_{gp120}, as observed in the 35O22/gp160

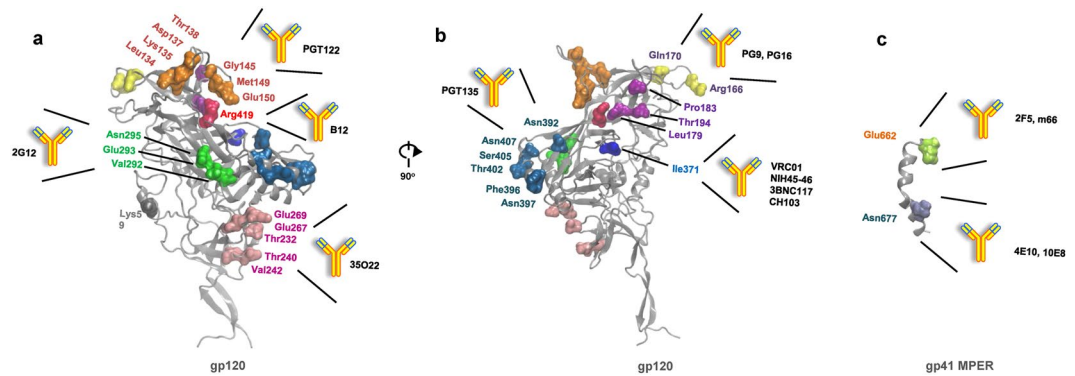


Figure 3. Mapping of treatment failure-related mutations on the high-resolution structures of HIV-1 gp120 (a) and 90° views (b) and gp41 MPER (c). Most treatment failure-related mutations overlapped with the neutralization epitopes of different kinds of bnAbs.

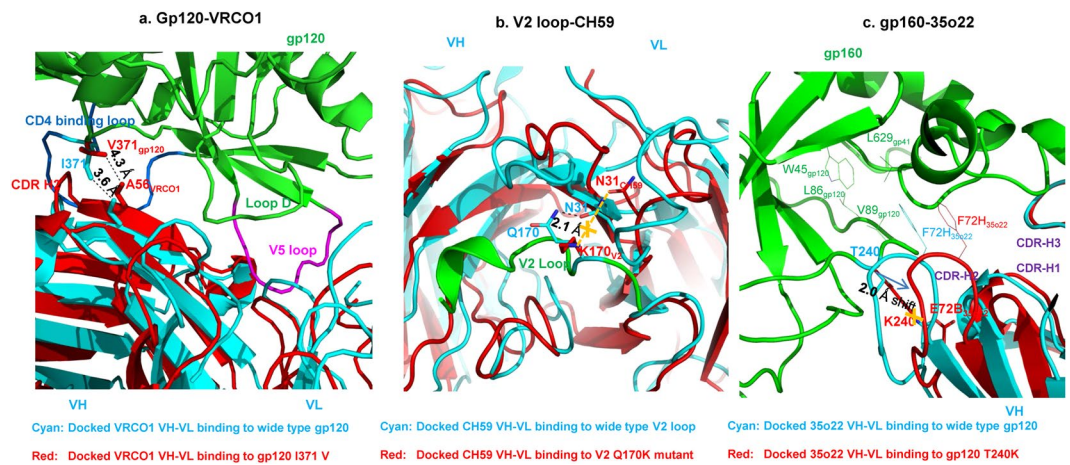


Figure 4. Molecular docking results showed the impact of adaptive mutations of HIV-1 Env on the binding ability of bnAbs. (a) VRC01 was colored in cyan in the docked VRC01 VH-VL binding to wild-type gp120 and colored in red in the docked VRC01 VH-VL binding to gp120 I371 V. (b) CH59 was colored in cyan in the docked CH59 VH-VL binding to wild-type gp120 V2 loop and colored in red in the docked CH59 VH-VL binding to gp120 V2 Q170K mutant. (c) 35O22 was colored in cyan in the docked 35O22 VH-VL binding to wild-type gp120 and colored in red in the docked 35O22 VH-VL binding to gp120 T240K mutant.

complex. Thus, the binding of 35O22/gp160 encounters more energetically favorable docking than the binding of 35O22/gp160-T240K (Z-score of -52.893 vs. -36.811). Based on the docked structural analysis of these three treatment failure-related mutations in Env proteins, antibody binding ability was compromised, irrespective of epitope location.

Virus sensitivity to the bnAbs. The neutralization assay was performed to determine whether the adaptive mutation would affect the sensitivity of mutant-viruses to the bnAbs. The IC_{50} of the bnAb VRC01 against the mutant viruses (I371V) was assessed and compared with that against the wide-type viruses (AT1). Indeed, we found that the I371V mutant virus gained 4.9 times enhancement of resistance to VRC01 ($IC_{50} = 0.0724 \pm 0.0012 \mu\text{g/ml}$ for I371V, and $IC_{50} = 0.0148 \pm 0.0009 \mu\text{g/ml}$ for AT1). This result suggests that some mutant viruses in the treatment-failure patients may have acquired resistance to some bnAbs.

Discussion

A thorough understanding of humoral immune responses to HIV-1 has been a major research focus in recent years⁴⁵. HIV-1 has been reported to escape neutralizing antibodies during early infection via a “glycan shield” mechanism whereby escape mutations in the Env become involved in N-linked glycosylation; such mutations were very sparse and, generally, did not map to known neutralization epitopes¹¹. Humoral responses also evolve in response to these Env escape mutations, but the kinetics of evolution does not correlate with the rate of disease progression⁶. Recently identified broadly neutralizing antibodies (bnAbs) against some conserved epitopes have rekindled interest in exploring antibody-virus co-evolution pathways *in vivo*, but these antibodies only naturally arise in a fraction of individuals after years of HIV-1 infection. Thus, despite intensive research, important gaps remain in our understanding of humoral immune responses relative to HIV-1 disease progression. In

particular, considering that selection pressure exerted on viruses has been largely replaced by antiretroviral drugs in ART-treated patients, it is important to ascertain whether and how humoral immune responses affect disease progression under this circumstance.

In the present study, gene sequencing and selection pressure-based analysis were employed to uncover the underlying relationships between treatment failure and putative adaptive mutations occurring in the HIV-1 *env* gene. A study population was very carefully selected for homogeneity in terms of demographics, ethnic makeup, treatment modality and, importantly, route of infection (infected with the same HIV-1 subtype B' via former plasma donation). This unique feature enabled a comparative study across subpopulations that exhibited different treatment outcomes. Interestingly, although different baseline variables of the TF and VS groups were compared, including, for example, CD4⁺ T cell count, viral load, drug-resistant mutations, and co-receptor usage, the only variable that showed a significant difference between the two groups was gp160 diversity. The gp120/gp41 diversity did not increase or decrease in these TF patients during long-term ART, suggesting that the viruses may have evolved to escape host immune responses before the initiation of ART via quasispecies diversification. This also suggests that the level of gp160 diversity may be used as an indicator to predict disease prognosis under certain circumstances.

The residues that underwent positive selection pressures were determined by calculation of Ka/Ks ratio for individual amino acids in gp120 or gp41. Based on the crystal structures of gp120 and gp41 proteins, we mapped these residues onto the structures of Env. In contrast to the untreated patients in which Env escape mutations were sparse and could not be mapped to known neutralization epitopes, we found that most of these mutation sites overlapped with the neutralization epitopes of different kinds of bnAbs: R419K and I371V locating in the epitope of CD4bs antibody b12 and VRC01; R166Q and Q170K in the epitope of V1/V2 targeting antibodies PG9, PG16, and CH59; T232R, T240K, V242I, E267K, E269G in the epitope of gp120/gp41 interface antibody 35O22; E662K and N677K in the epitope of MPER antibody 2F5 and 4E10 (Fig. 3). These results demonstrate, for the first time, that, to the best of our knowledge, HIV-1 escape from host humoral immunity contributes to disease progression in patients on long-term ART.

Our analysis may have important implications concerning the proposition of an Ab effect on the virological outcome after HAART treatment. For instance, we did not find differences in the number of PNG and length of V3 loop sequence between TF and VS patients, suggesting that the glycosylation sites were not triggered by the immune response in these patients, and also that the V3 loop epitope may not be relevant for immune escape. The highest level of mutation was located in the C2 region of Env, suggesting that this region may be one of the most interesting epitopes to target. Notably, our analysis was based on the published crystal structures of bnAbs complexed with gp120. Due to the limited complex structures available, it is probably that some of the mutations are located within epitopes that have not been structurally characterized, or the Abs targeting them have not been identified. For example, the 134–138 loop (Leu134, Lys135, Asp137, Thr138) are just located on the Env surfaces and could be a dominant epitope by a number of Abs, but we only found one bnAb, PGT122, that mostly recognizes V3 but also has some interactions with the 134–138 loop. We believe that further studies, e.g., by using B cells from the patients described in this study, could lead to the isolation of Abs targeting this loop.

To elucidate the mechanism that governs the development of these treatment failure-related adaptive mutations, we selected three bnAbs (VRC01, 35O22, and CH59) that have available crystal structures in complex with gp120/gp160 for computational modeling analysis to evaluate whether Env adaptive mutations have an impact on antibody binding. The docking results suggest that these mutations tend to compromise antibody-antigen interactions. While the accumulation of HIV-1 escape mutations is a very complex process, this finding suggests that escape of antibody binding could be one of the routes via which the viruses escape from host humoral immunity in long-term antiretroviral-experienced patients. In a proof-of-concept study, we generated a mutant HIV-1 virus with a single I371V mutation, and compared side-by-side the neutralization activity of VRC01, the most advanced clinically developed HIV-1 bnAbs, against the wild-type and the mutant virus. The HIV-1 I371V mutant gained 4.9 times enhancement of resistance to VRC01, suggesting that some viruses from TF patients may have acquired resistance to some bnAbs. Nevertheless, further study is warranted to investigate how the emergences of these HIV-1 mutants cause the ART failure. For instance, we will isolate these viral mutants and test their sensitivity to the antiretroviral drugs used by the TF patients, even though these drugs are not expected to target the HIV-1 Env, because we cannot exclude the possibility that the mutations of viral Env may cause the change of viral replication kinetics and viral sensitivity to some of the antiretroviral drugs.

These findings have considerable importance for developing an effective long-term antiviral strategy against HIV-1. Although highly-active ART (HAART) is very effective in suppressing viral replication, it can result in toxicity with unpredictable effects and resistance to therapy, especially with long-term use, resulting in treatment failure in a substantial proportion of individuals^{2,4,5}. Of note, advances in antibody-based therapeutics are coming at an ever-increasing pace, and, as a result HIV-1 bnAbs, hold promise in the ongoing effort to eradicate HIV^{20,46–49}. In the present study, it was shown that host humoral immunity may play a role in treatment failure in long-term ART-treated patients. Some mutations in HIV-1 gp120/41 may have resulted in escape from some neutralizing antibodies and thus contributed to disease progression. This implies that a combination of conventional antiretroviral drugs with potent bnAbs, especially a cocktail of bnAbs targeting different epitopes, could be beneficial in patients who have shown poor response to ART. Moreover, given that there is no mutation in the HIV-1 gp41 NHR region, the target of HIV-1 fusion inhibitors, the combination of T20 with antiretroviral drugs may be provided to the patients immediately after their viruses are identified to carry the mutations in the sites overlapped with the neutralization epitopes of bnAbs, to avoid the treatment failure.

In conclusion, our study provides evidence that escape from host humoral immunity may play a role in treatment failure in long-term antiretroviral-experienced patients. Treatment failure-related *env* mutations identified by SGA sequencing and selection pressure-based analysis may provide an immunological basis for understanding the co-evolution of viruses and bnAbs in a particular patient. Based on *env* gene information of the viruses in the

patients, we may be able to predict the treatment outcome, based on which individualized therapeutic regimen, such as the combinations of antiretroviral drugs with HIV-1 entry inhibitors targeting the coreceptor (e.g., maraviroc) or gp41 (e.g., enfuvirtide), rather than gp120 where the mutations are located, may be designed in advance for the patients with TF potential.

Methods

Ethics statement. This study was approved by the Institutional Research Ethics Community of the Chinese Center for Disease Control and Prevention, and all subjects signed informed consent to participate in the research study prior to blood and data collection. All experiments were performed in accordance with relevant guidelines and regulations.

Study patients. The study population consisted of 45 chronically HIV-infected patients on a stable ART regimen who participated in a multicenter AIDS Cohort Study in Anhui and Henan provinces in China from 2004 to 2011. Patients who initially received therapy with a first-line regimen of lamivudine (3TC), stavudine (d4T), or zidovudine (AZT) and nevirapine (NVP) or efavirenz (EFV) were selected. Patients in this study met the following inclusion criteria: be naïve-ART before enrollment and be 18–65 years old with uninterrupted HAART for more than 12 months. 3 patients in TF group and 2 patients in VS group who were lost to follow-up at the last sampling point were excluded. Blood samples collected at each follow-up point were analyzed for viral load, CD4⁺ T cell count and drug-resistance mutations. Plasma viral loads were determined with real-time Nucleic Acid Sequence Based Amplification (NASBA) (NucliSense Easy Q, BioMerieux, France) according to the manufacturer's instructions. CD4⁺ T cells in whole blood were quantitated by flow cytometry using Tritest CD3/CD4/CD45 TruCount Tubes (BD Biosciences, San Jose, CA). For detection of drug-resistance mutations, viral RNA was first extracted from plasma using the QIAamp[®] Viral RNA Mini Kit (Qiagen, Chatsworth, CA) according to the manufacturer's protocol and then used to synthesize first-strand cDNA using the SuperScript[™] III First-Strand Synthesis System (Invitrogen). The HIV-1 pol gene was amplified with specific primers that annealed to conserved regions of the pol region, and sequences of the pol gene were compared to the consensus B reference sequence using HIVdb software (Stanford HIV Drug Resistance Database, <http://hivdb.stanford.edu>) to detect the drug-resistance mutations.

cDNA synthesis, single-genome amplification and HIV-1 env sequencing. HIV-1 RNA extraction, cDNA synthesis, and SGA were performed as previously reported⁵⁰. For SGA, according to a Poisson distribution, a positive reaction rate of 30% or lower, or cDNA diluted to ensure fewer than 29 PCR products yielded from a total of 96 independent PCRs thus ensuring that amplicons were derived from a single template. Sequences of the *env* region were amplified by a nested polymerase chain reaction using Premix Ex Taq version 2.0 (Takara, Japan). PCR products derived from cDNA dilutions yielding less than 30% PCR positivity were sequenced by an ABI 3730XL Sequencer using BigDye terminators (Applied Biosystems, Foster City, California, USA). The chromatogram data were cleaned and assembled using Sequencher v4.9 (Gene Codes, Ann Arbor, Michigan, USA).

Sequence analysis. Multiple nucleotide sequence alignments of gene sequences encoding for the Env CDS, V1V2, V3, V4 and V5 of HIV-1 from each subject were conducted using Clustal W, input in Gene Cutter tool from the Los Alamos HIV sequence database (http://www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html) and then adjusted manually with BioEdit software (version 7.1.3). Phylogenetic analysis was performed using the Kimura-2 model of substitution within neighbor-joining method in MEGA 5 at the nucleotide level, and the reliability of the branching orders was tested by bootstrap analysis of 1,000 replicates⁵¹. HIV-1 co-receptor usage (CCR5 and CXCR4) was predicted using WEBPSSM (<http://indra.mullins.microbiol.washington.edu/webpssm/>). Sequence diversity was calculated using the Kimura 2-parameter model option in MEGA 5. The number and location of potential N-linked glycosylation sites (PNGSs) for each sequence were estimated using the N-GLYCOSITE web tool from the Los Alamos HIV database (<http://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html>).

Identification of putative adaptive mutations directly related to treatment failure. A total of 638 *env* sequences, including 352 sequences in the VS group and 286 sequences in TF group, acquired from patients samples before treatment (baseline) were analyzed to screen treatment failure in relation to adaptive mutation by a selection pressure-based analysis. The Ka/Ks ratio could be used as an indicator of selective pressure acting on a protein-coding gene. The computation of Ka/Ks for an individual site or specific amino acid was based on the model of Chen *et al.*⁵². A Ka/Ks value of 1, >1 or <1 indicated a neutral selection, a positive selection or a negative selection, respectively. Log odds (LOD) confidence score was used to measure the significance of selection pressure, and LOD > 2 indicated a significance of the positive selection. Ka/Ks and LOD were calculated by the CorMut R package⁵³. When performing the computation, HXB2 served as a reference.

Bayesian network was applied to further analyze the correlation between predicted mutations and treatment outcome. Generally, the Bayesian network is a graphical model that encodes a joint probability distribution of a set of random variables, where nodes represent random variables, and arrows represent probabilistic dependencies between them⁵⁴. The Bayesian network was constructed by the bnlearn R package, and the results were revised by the nonparametric bootstrap method. More than 65% of bootstrap support was considered robust, while 35–65% was considered thin⁵⁵. The mutations having direct connection with treatment failure with bootstrap value over 65% were identified.

Virus growth competition assay. A multiple-cycle recombinant virus growth competition assay was applied to determine the impact of quasispecies variation on viral fitness. Mutations were introduced to the

wild-type plasmid pAT1 by PCR to construct the mutant plasmids which contained the critical adaptive mutation sites. Then HEK-293T cells were transfected with mutant plasmids to rescue the viruses with adaptive mutations. TCID₅₀ of viruses with adaptive mutations was quantitated to detect viral infection ability. Mutant virus and wild-type virus AT2 co-infected PM1 cells with equal P24 antigen levels. Cells were collected from 3 to 6 days post-infection. Cells were stained by anti-Thy1.1 antibody and anti-Thy1.2 antibody, respectively, and subsequently both antibodies. The ratio of the two viruses was detected by FACS. Relative fitness was calculated using the online tool VFitness (<http://bis.urmc.rochester.edu/VFitness/FitnessTwo.aspx>).

Molecular docking. The Z-DOCK program⁴³ implanted in the Accelrys Discovery Studio (<http://accelrys.com>) was used to fit bnAbs VRC01, CH59 and 35O22 onto gp120. The output Z-scores for the top poses was used to determine the effects of treatment failure-related residues of Env proteins on antibody binding. The three-dimensional structures for these Fab and HIV Env were extracted from the resolved crystal structure of VRC01/gp120 (PDB ID 4LST)³⁶, CH59/gp120 (PDB ID 4HPY)⁴⁴, and 35O22/gp160 complexes (PDB ID 4TVP)²⁷. For calculation convenience, the constant regions of Fab were removed with only consideration of the V regions (VH-VL). Env mutants, including I371V for VRC01, Q170K for CH59, T240K for 35O22, were modeled by the PyMOL molecular graphics system (version 1.5.0.4; Schrödinger, LLC) without atomic clashes.

Before docking, the structures of bnAbs and Env were processed as follows. 1) water molecules were deleted and hydrogen atoms were added at pH7.4 and an ionic strength of 0.145 and in a dielectric environment of 10; 2) energy was minimized based on CHARMM⁵⁶ with a cutoff of 0.9; 3) loop regions were rebuilt according to SEQRES⁵⁷ data; and 4) cysteine bridges in these proteins were defined as blocked regions. Then Z-DOCK performed a systematic search of a uniform sample of docked protein poses and used an internal scoring algorithm to predict optimal interactions. A total of 54000 docked poses were initially produced, which were further filtered and re-ranked to obtain the top 200 poses based on Z-Rank score using electrostatic and desolvation energy and non-deterministic FFT optimization. These 200 poses were then visually scrutinized and judged by Z-Score. The most reasonable poses were chosen as the models for the VH-VL-Env complexes by combined consideration of general features of antibody-antigen interaction, such as hydrogen bonds, salt bridges, hydrophobic packing and other interactions at the interface, without any highly unusual features and steric clashes. The Z-scores for VH-VL docking into wild-type and mutant Envs were evaluated and compared.

Neutralization assay in TZM-bl cells. A sensitive HIV-1 neutralization assay was performed to test the neutralizing activities of the bnAb VRC01 against HIV-1 carrying Env with I371V mutation as previously described⁵⁸. Briefly, the bnAb was at serial dilution was incubated with the mutant virus (200 TCID₅₀) for 1 h at 37 °C. The virus with wild Env sequence (AT1) was included as control. TZM-bl cells (1×10^4) with DEAE-dextran (15 µg/ml) were added to each well. After 48 h incubation, The RLU were determined using the luciferase reporter gene assay system (Bright-Glo; Promega) and the luminometer (PerkinElmer Life Sciences), as described previously⁵⁸. Mock infected cells were used to determine background luminescence. IC₅₀ (the concentration of an inhibitor causing 50% inhibition of infection) value was then calculated⁵⁹.

Statistical analyses. Comparisons of differences between VS and TF groups were calculated using Mann-Whitney, Chi-Square or Fisher's exact test. The relative fitness comparison between mutant viruses and wild-type viruses was done using the Wilcoxon rank sum test. Statistical significance was set at P value < 0.05 and for two-sided tests. Logistic regression and receiver operating characteristic (ROC) analysis were performed to develop a predictive model. Covariates, including CD4⁺ T cell count, viral load, gp160 diversity, drug-resistant mutations and co-receptor usage, were included for logistic analysis. The statistical analyses were performed using the SPSS software package (version 17.0).

References

- Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582–1586 (1996).
- Drobniewski, F. & Eltringham, I. Highly active antiretroviral therapy. *Lancet* **351**, 1059–1060, doi:10.1016/s0140-6736(05)79027-9 (1998).
- Palella, F. J. Jr. *et al.* Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* **338**, 853–860, doi:10.1056/nejm199803263381301 (1998).
- Bartlett, J. A. & Shao, J. F. Successes, challenges, and limitations of current antiretroviral therapy in low-income and middle-income countries. *Lancet Infect Dis* **9**, 637–649, doi:10.1016/s1473-3099(09)70227-0 (2009).
- Aldous, J. L. & Haubrich, R. H. Defining treatment failure in resource-rich settings. *Curr Opin HIV AIDS* **4**, 459–466, doi:10.1097/COH.0b013e328331dea5 (2009).
- Alter, G. & Moody, M. A. The humoral response to HIV-1: new insights, renewed focus. *J Infect Dis* **202**(Suppl 2), S315–322, doi:10.1086/655654 (2010).
- Stamatatos, L., Morris, L., Burton, D. R. & Mascola, J. R. Neutralizing antibodies generated during natural HIV-1 infection: good news for an HIV-1 vaccine? *Nat Med* **15**, 866–870, doi:10.1038/nm.1949 (2009).
- Cao, Y., Qin, L., Zhang, L., Safrit, J. & Ho, D. D. Virologic and immunologic characterization of long-term survivors of human immunodeficiency virus type 1 infection. *N Engl J Med* **332**, 201–208, doi:10.1056/nejm199501263320401 (1995).
- Cecilia, D., Kleeberger, C., Munoz, A., Giorgi, J. V. & Zolla-Pazner, S. A longitudinal study of neutralizing antibodies and disease progression in HIV-1-infected subjects. *J Infect Dis* **179**, 1365–1374, doi:10.1086/314773 (1999).
- Zhang, Y. J. *et al.* Augmented serum neutralizing activity against primary human immunodeficiency virus type 1 (HIV-1) isolates in two groups of HIV-1-infected long-term nonprogressors. *J Infect Dis* **176**, 1180–1187 (1997).
- Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312, doi:10.1038/nature01470 (2003).
- Bunnik, E. M., Pisas, L., van Nuenen, A. C. & Schuitemaker, H. Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection. *J Virol* **82**, 7932–7941, doi:10.1128/jvi.00757-08 (2008).
- Saag, M. & Deeks, S. G. How do HIV elite controllers do what they do? *Clin Infect Dis* **51**, 239–241, doi:10.1086/653678 (2010).

14. Walker, B. D. & Yu, X. G. Unravelling the mechanisms of durable control of HIV-1. *Nat Rev Immunol* **13**, 487–498, doi:[10.1038/nri3478](https://doi.org/10.1038/nri3478) (2013).
15. Derdeyn, C. A., Moore, P. L. & Morris, L. Development of broadly neutralizing antibodies from autologous neutralizing antibody responses in HIV infection. *Current opinion in HIV and AIDS* **9**, 210–216, doi:[10.1097/COH.000000000000057](https://doi.org/10.1097/COH.000000000000057) (2014).
16. Chong, H. *et al.* Genetic and neutralization properties of HIV-1 env clones from subtype B/BC/AE infections in China. *J Acquir Immune Defic Syndr* **47**, 535–543, doi:[10.1097/QAI.0b013e3181663967](https://doi.org/10.1097/QAI.0b013e3181663967) (2008).
17. Jensen, M. A. & van 't Wout, A. B. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS Rev* **5**, 104–112 (2003).
18. Liu, S. *et al.* Different from the HIV fusion inhibitor C34, the anti-HIV drug Fuzeon (T-20) inhibits HIV-1 entry by targeting multiple sites in gp41 and gp120. *J Biol Chem* **280**, 11259–11273, doi:[10.1074/jbc.M411141200](https://doi.org/10.1074/jbc.M411141200) (2005).
19. Dykes, C. *et al.* Evaluation of a multiple-cycle, recombinant virus, growth competition assay that uses flow cytometry to measure replication efficiency of human immunodeficiency virus type 1 in cell culture. *J Clin Microbiol* **44**, 1930–1943, doi:[10.1128/jcm.02415-05](https://doi.org/10.1128/jcm.02415-05) (2006).
20. Kwong, P. D., Mascola, J. R. & Nabel, G. J. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat Rev Immunol* **13**, 693–701, doi:[10.1038/nri3516](https://doi.org/10.1038/nri3516) (2013).
21. Julien, J. P. *et al.* Broadly neutralizing antibody PGT121 allosterically modulates CD4 binding via recognition of the HIV-1 gp120 V3 base and multiple surrounding glycans. *PLoS Pathog* **9**, e1003342, doi:[10.1371/journal.ppat.1003342](https://doi.org/10.1371/journal.ppat.1003342) (2013).
22. Julien, J. P. *et al.* Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1477–1483, doi:[10.1126/science.1245625](https://doi.org/10.1126/science.1245625) (2013).
23. Pollara, J. *et al.* HIV-1 vaccine-induced C1 and V2 Env-specific antibodies synergize for increased antiviral activities. *J Virol* **88**, 7715–7726, doi:[10.1128/JVI.00156-14](https://doi.org/10.1128/JVI.00156-14) (2014).
24. McLellan, J. S. *et al.* Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336–343, doi:[10.1038/nature10696](https://doi.org/10.1038/nature10696) (2011).
25. Pancera, M. *et al.* Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nat Struct Mol Biol* **20**, 804–813, doi:[10.1038/nsmb.2600](https://doi.org/10.1038/nsmb.2600) (2013).
26. Pan, R., Gorny, M. K., Zolla-Pazner, S. & Kong, X. P. The V1V2 Region of HIV-1 gp120 Forms a Five-Stranded Beta Barrel. *J Virol* **89**, 8003–8010, doi:[10.1128/jvi.00754-15](https://doi.org/10.1128/jvi.00754-15) (2015).
27. Pancera, M. *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514**, 455–461, doi:[10.1038/nature13808](https://doi.org/10.1038/nature13808) (2014).
28. Zhou, T. *et al.* Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* **445**, 732–737, doi:[10.1038/nature05580](https://doi.org/10.1038/nature05580) (2007).
29. Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329**, 811–817, doi:[10.1126/science.1192819](https://doi.org/10.1126/science.1192819) (2010).
30. Georgiev, I. S. *et al.* Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization. *Science* **340**, 751–756, doi:[10.1126/science.1233989](https://doi.org/10.1126/science.1233989) (2013).
31. Rudicell, R. S. *et al.* Enhanced potency of a broadly neutralizing HIV-1 antibody *in vitro* improves protection against lentiviral infection *in vivo*. *J Virol* **88**, 12669–12682, doi:[10.1128/jvi.02213-14](https://doi.org/10.1128/jvi.02213-14) (2014).
32. Liao, H. X. *et al.* Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469–476, doi:[10.1038/nature12053](https://doi.org/10.1038/nature12053) (2013).
33. Chen, L. *et al.* Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120. *Science* **326**, 1123–1127, doi:[10.1126/science.1175868](https://doi.org/10.1126/science.1175868) (2009).
34. Diskin, R. *et al.* Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science* **334**, 1289–1293, doi:[10.1126/science.1213782](https://doi.org/10.1126/science.1213782) (2011).
35. Klein, F. *et al.* Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* **153**, 126–138, doi:[10.1016/j.cell.2013.03.018](https://doi.org/10.1016/j.cell.2013.03.018) (2013).
36. Zhou, T. *et al.* Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* **39**, 245–258, doi:[10.1016/j.immuni.2013.04.012](https://doi.org/10.1016/j.immuni.2013.04.012) (2013).
37. Grover, R. K. *et al.* A structurally distinct human mycoplasma protein that generically blocks antigen-antibody union. *Science* **343**, 656–661, doi:[10.1126/science.1246135](https://doi.org/10.1126/science.1246135) (2014).
38. Huang, C. C. *et al.* Structure of a V3-containing HIV-1 gp120 core. *Science* **310**, 1025–1028, doi:[10.1126/science.1118398](https://doi.org/10.1126/science.1118398) (2005).
39. de la Arada, I. *et al.* Structural constraints imposed by the conserved fusion peptide on the HIV-1 gp41 epitope recognized by the broadly neutralizing antibody 2F5. *J Phys Chem B* **113**, 13626–13637, doi:[10.1021/jp905965h](https://doi.org/10.1021/jp905965h) (2009).
40. Ofek, G. *et al.* Structural basis for HIV-1 neutralization by 2F5-like antibodies m66 and m66.6. *J Virol* **88**, 2426–2441, doi:[10.1128/jvi.02837-13](https://doi.org/10.1128/jvi.02837-13) (2014).
41. Bird, G. H. *et al.* Stapled HIV-1 peptides recapitulate antigenic structures and engage broadly neutralizing antibodies. *Nat Struct Mol Biol* **21**, 1058–1067, doi:[10.1038/nsmb.2922](https://doi.org/10.1038/nsmb.2922) (2014).
42. Huang, J. *et al.* Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* **491**, 406–412, doi:[10.1038/nature11544](https://doi.org/10.1038/nature11544) (2012).
43. Chen, R., Li, L. & Weng, Z. ZDOCK: an initial-stage protein-docking algorithm. *Proteins* **52**, 80–87, doi:[10.1002/prot.10389](https://doi.org/10.1002/prot.10389) (2003).
44. Liao, H. X. *et al.* Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2. *Immunity* **38**, 176–186, doi:[10.1016/j.immuni.2012.11.011](https://doi.org/10.1016/j.immuni.2012.11.011) (2013).
45. Ying, T., Prabakaran, P. & Dimitrov, D. S. A systems approach to HIV-1 vaccines. *Nat Biotechnol* **34**, 44–46, doi:[10.1038/nbt.3456](https://doi.org/10.1038/nbt.3456) (2016).
46. Chen, W., Ying, T. & Dimitrov, D. S. Antibody-based candidate therapeutics against HIV-1: implications for virus eradication and vaccine design. *Expert Opin Biol Ther* **13**, 657–671, doi:[10.1517/14712598.2013.761969](https://doi.org/10.1517/14712598.2013.761969) (2013).
47. Klein, F. *et al.* Antibodies in HIV-1 vaccine development and therapy. *Science* **341**, 1199–1204, doi:[10.1126/science.1241144](https://doi.org/10.1126/science.1241144) (2013).
48. van Gils, M. J. & Sanders, R. W. *In vivo* protection by broadly neutralizing HIV antibodies. *Trends Microbiol* **22**, 550–551, doi:[10.1016/j.tim.2014.08.006](https://doi.org/10.1016/j.tim.2014.08.006) (2014).
49. Shingai, M. *et al.* Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature* **503**, 277–280, doi:[10.1038/nature12746](https://doi.org/10.1038/nature12746) (2013).
50. Salazar-Gonzalez, J. F. *et al.* Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J Virol* **82**, 3952–3970, doi:[10.1128/jvi.02660-07](https://doi.org/10.1128/jvi.02660-07) (2008).
51. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739, doi:[10.1093/molbev/msr121](https://doi.org/10.1093/molbev/msr121) (2011).
52. Chen, L., Perlina, A. & Lee, C. J. Positive selection detection in 40,000 human immunodeficiency virus (HIV) type 1 sequences automatically identifies drug resistance and positive fitness mutations in HIV protease and reverse transcriptase. *J Virol* **78**, 3722–3732 (2004).
53. Li, Z. *et al.* CorMut: an R/Bioconductor package for computing correlated mutations based on selection pressure. *Bioinformatics* **30**, 2073–2075, doi:[10.1093/bioinformatics/btu154](https://doi.org/10.1093/bioinformatics/btu154) (2014).
54. Julia Flores, M., Nicholson, A. E., Brunskill, A., Korb, K. B. & Mascaro, S. Incorporating expert knowledge when learning Bayesian network structure: a medical case study. *Artif Intell Med* **53**, 181–204, doi:[10.1016/j.artmed.2011.08.004](https://doi.org/10.1016/j.artmed.2011.08.004) (2011).

55. Deforche, K. *et al.* Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors. *Infect Genet Evol* **7**, 382–390, doi:10.1016/j.meegid.2006.09.004 (2007).
56. Sapay, N. & Tieleman, D. P. Combination of the CHARMM27 force field with united-atom lipid force fields. *J Comput Chem* **32**, 1400–1410, doi:10.1002/jcc.21726 (2011).
57. Ordog, R., Szabadka, Z. & Grolmusz, V. DECOMP: a PDB decomposition tool on the web. *Bioinformatics* **3**, 413–414 (2009).
58. OuYang, Y. *et al.* Neutralization sensitivity of HIV-1 subtype B' clinical isolates from former plasma donors in China. *Virology* **10**, 10, doi:10.1186/1743-422X-10-10 (2013).
59. Chen, W., Zhu, Z., Feng, Y. & Dimitrov, D. S. Human domain antibodies to conserved sterically restricted regions on gp120 as exceptionally potent cross-reactive HIV-1 neutralizers. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17121–17126, doi:10.1073/pnas.0805297105 (2008).

Acknowledgements

We would like to thank the staff of Anhui and Henan Province Center for Disease Control and Prevention for collecting samples, and all subjects who volunteeringly participated in this study. We are grateful to Dr. Carrie Dykes from University of Rochester School of Medicine and Dentistry for providing technical assistance in virus growth competition assay for this study. This work was supported by grants from the State Key Laboratory of Infectious Disease Prevention and Control (2011SKLID102), the National Nature Science Foundation of China (81172733, 81601796 and 81561128006), the 12th Five-Year National Science & Technology Major Project (2013ZX10001-006), and Beijing Natural Science Foundation (7174313).

Author Contributions

L.M. and T.Y. conceived the study. Y.O., Q.Y., W.L., Z.L., D.K., Y.W., K.H., H.X., Y.S., S.J., T.Y., and L.M. performed the experiments and/or participated in the experimental design. Y.S., S.J., T.Y., and L.M. wrote and edited the manuscript. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-05594-5

Competing Interests: The authors declare that they have no competing interests.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017