

SCIENTIFIC REPORTS



OPEN

The sexually driven epidemic in youths in China's southwestern border region was caused by dynamic emerging multiple recombinant HIV-1 strains

Received: 17 November 2014

Accepted: 01 May 2015

Published: 02 July 2015

Huamian Wei^{1,*}, Hui Xing^{1,*}, Jenny H. Hsi^{1,*}, Manhong Jia^{2,*}, Yi Feng¹, Song Duan³, Cui He¹, Shitang Yao³, Yuhua Ruan¹, Xiang He⁴, Lingjie Liao¹, Yanling Ma², Yunda Huang⁵, Lin Lu² & Yiming Shao¹

Dehong prefecture, Yunnan province on China's southwestern border was the gateway of the country's AIDS epidemic. Studies on HIV-1 molecular epidemiology will provide key information on virus transmission dynamics and help to inform HIV prevention strategies. HIV-1 infected youths (age 16–25 years) diagnosed in the continuous 3 months in 2009 to 2012 were enrolled. By means of phylogenetic and statistical analyses, it was shown that two thirds (133/205) of youths in Dehong, of which 74.1% were infected sexually, were infected by uncharacterized recombinant HIV-1 strains. Among them about 59.4% (79/131) were unique recombinant forms (URFs) and 40.6% (54/131) formed 11 transmission clusters, termed potential circulating recombinant forms (pCRFs). The emergence of recombinants was statistically significant related with people of low education, residents outside the capital city of Dehong and being Myanmar residents. It was the first report with ongoing HIV-1 recombinant strains in a sexually driven epidemic area in China. Great efforts should be put on reducing multiple risk exposures behavior in local young people, containing the spread of pCRFs to other regions, and preventing the URFs from evolving into future CRFs. Collaborative prevention across border is needed to better control the local AIDS epidemic.

HIV-1 remains a global threat to health and wellbeing, with an estimated 35.3 million people living with HIV in 2012 (<http://www.unaids.org/>). Although the number of newly infected people has declined since 2001, viral diversity is increasing¹. HIV diversity can be attributed to its high rate of mutation and virus replication as well as recombination between different viral strains. The identification of a circulating recombinant form (CRF) is defined through at least three epidemiologically unlinked individuals with full genome sequences displaying the same recombinant structure, while those that do not meet this criterion are categorized as unique recombinant forms (URFs)². Currently, more than 65 CRFs have been included in the HIV Database (<http://www.hiv.lanl.gov/>), and an even greater number of URFs has been

¹State Key Laboratory for Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China. ²Yunnan Center for Disease Control and Prevention, Kunming, Yunnan, China. ³Dehong Center for Disease Control and Prevention, Dehong, Yunnan, China. ⁴Guangdong Provincial Institute of Public Health, Guangdong, China. ⁵Fred Hutchinson Cancer Research Center, Seattle, USA. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Y.S. (email: yshao08@gmail.com)

reported. The increasing complexity of local and global HIV epidemics brings challenges to effective diagnosis, antiretroviral treatment, disease progression monitoring, and virus transmission prevention¹.

The HIV epidemic in China has become increasingly genetically complex, with subtypes B, B', C, CRF01_AE, CRF07_BC, CRF08_BC, and other strains co-circulating³. In particular, Dehong prefecture in Yunnan province, southwest China, is considered the gateway of China's HIV-1 epidemic, from where most of the HIV-1 strains currently circulating in the country had first appeared^{4–9}. The Dehong government has organized an impressive AIDS prevention campaign, and was the first prefecture in China that achieved a stable and then declining cases of HIV infections, a sign of the epidemic containment. Nevertheless, new recombinant strains are continually reported there among key high-risk populations such as injecting drug users (IDUs)^{10,11}. HIV recombination is more commonplace in those infected through parenteral routes as compared with sexual routes, since the mucosal barrier in the latter poses a greater genetic barrier against dual and multiple infection¹². Recombination in IDUs pose fewer concerns for generating a widespread recombinant epidemic in China, as their HIV-1 incidence has been found in decline in most regions of the country^{13,14}. There is seldom report thus far on high proportions of recombinant HIV-1 found in sexual transmission networks, which has become the major driving force in China's current AIDS epidemic. At the same time, existing reports on recombination also often lack appropriate sampling methods aiming at recent HIV infections, and hence may not accurately represent the current HIV epidemic.

This study design is based on the WHO Threshold Survey for transmitted HIV drug resistance surveillance that sought to capture recent infections by sampling from newly reported HIV infections in people less than 25 years of age. It aims to capture the full picture of the current ongoing epidemic in Dehong through characterizing the scale and patterns of recombinant HIV infections as well as factors related to their cause.

Results

Demographic characteristics of study population. The basic characteristics of the sample population were summarized in Table 1. Generally, the number of samples per year, the proportion of sex and the marital status were basically the same during study period of 2009–2012, as expected when an epidemic has reached plateau. Dehong is a multi-ethnic region with over half being non-Han ethnic minorities. Since cross-border migration regulations are relatively permissive, 17.6% of study subjects are Myanmar residents. Noteworthy, heterosexual transmission accounts for the vast majority of sampled subjects (74.1%), which is very similar to the heterosexually infected HIV positive population (74%, 4142/5608) reported during 2009–2012 in Dehong.

Subtypes and recombinant patterns identification. Phylogenetic analysis of 205 HIV-1 *pol* gene fragments showed that the majority of infections (64.9%) in 2009–2012 among young persons in Dehong prefecture consisted of not the conventional subtype group, but the new recombinant forms (Fig. 1, Table 2). This observation was confirmed by combined results of phylogenetic tree analyses, breakpoint analyses and near full-length genome (NFLG) (Fig. 2). The conventional subtypes account for 35.1% of all the samples, in which CRF01_AE was the most prevalent strain at 14.1%, followed by subtype C (10.2%), CRF08_BC (4.9%), B (4.4%) and CRF07_BC (1.5%) (Fig. 1). Among the new recombinant viruses, 12 clusters were identified from 54 subjects with at least three sequences in each cluster with bootstrap values no less than 0.7. The remaining new recombinant viruses from 79 subjects were interspersed widely throughout the phylogenetic tree. Bootscanning analysis was performed by both Simplot and RIP to identify recombinant breakpoints and confirm strain designations. The result was list in Supplemental Table S1. The new B'/C recombinants consisted 40% of the total sample, followed by CRF_01AE/B'/C (19.5%), CRF01_AE/C (4.9%), and CRF_01AE/B' recombinants (0.5%) (Fig. 1).

URFs and potential CRFs. For the 12 high-bootstrap clusters with non-conventional genotypes, NFLG sequencing was performed to clarify their genomic structures, obtaining a total of 15 NFLGs (Fig. 2), with at least one sequence in each cluster. One sequence at the root of cluster 11 was removed due to inconsistent structure with the rest of the cluster; while clusters 3 and 4 were found to have identical structures and merged at the outermost common branch point. Thus, these clusters with identical breakpoints were labeled as potential circulating recombinant forms (pCRFs) with consideration of the criteria for CRF. A total of 11 pCRFs consisting of 54 sequences (26.3% of the sample) were identified. The remaining 38.5% sequences (n = 79) were designated as URFs, since no phylogenetic cluster can be formed among them. For pCRFs, only two recombinant patterns, 01/B'/C (n = 31) and B'/C (n = 23), were found. For URFs, all recombinant patterns observed in the study were found in the URFs, including B'/C (n = 51), 01/B'/C (n = 17), 01/C (n = 10) and 01/B' (n = 1) (Fig. 1). Although recombinant structures vary widely, some sequences, especially B'/C recombinants, showed common breakpoints with each other despite being far apart on the phylogenetic tree; some URFs also shared breakpoints with CRF07_BC, CRF08_BC, or the pCRFs, suggesting common ancestry or related mechanisms of recombinant generation.

Characteristics of potential CRFs. The structures and basic information on the pCRFs are summarized in Fig. 2. pCRFs1, 4, 6, 7, and 10 included fragments from subtypes B', C, and CRF01_AE

	N	% Total
Total	205	100%
Year of diagnosis		
2009	54	26.3%
2010	49	23.9%
2011	43	21.0%
2012	59	28.8%
Sex		
Female	104	50.7%
Male	101	49.3%
Marital status		
Married or co-habiting	102	49.8%
Single, separated, or divorced	103	50.2%
Ethnicity		
Han	83	40.5%
Dai (minority)	47	22.9%
Jingpo (minority)	63	30.7%
Other (minorities)	12	5.9%
Education level		
No schooling	47	22.9%
Primary school	76	37.1%
Junior school	69	33.7%
High school and above	13	6.3%
Occupation		
Farmer	153	74.6%
Not farmer	52	25.4%
Residence		
Luxi County	83	40.5%
Other Dehong county	76	37.1%
Myanmar	36	17.6%
Other region in China	10	4.9%
High-risk contact history		
Heterosexual	152	74.1%
IDU	40	19.5%
MSM	6	2.9%
N/A	7	3.4%

Table 1. Basic characteristics of study population. Newly diagnosed HIV infection cases in young people ≤ 25 years old consecutively reported in approximately 3 months of sampling each year from 2009 to 2012.

($n = 23/205$, 11.2%), while the others contained only B' and C fragments ($n = 31/205$, 15.1%). pCRF2 ($n = 5$) was highly associated with being male, IDU, Jingpo ethnicity, and permanent residence in Longchuan county within Dehong. The largest cluster, pCRF3, was strongly associated with permanent residence in Luxi county, the prefecture's capital ($n = 11/14$), as well as being of Jingpo ethnicity and history of extramarital sex ($n = 10/14$). pCRF4 was found in those permanently residing in four Dehong counties and Myanmar, and was associated with extramarital sex ($n = 8/9$) and being female ($n = 6/9$), which may be related to commercial sex work. pCRF6 was exclusively found in those with Burmese permanent residence ($n = 3$), and pCRFs 1 and 6, both CRF01_AE/B'/C recombinants, were found exclusively in Han Chinese ($n = 6$). It is interesting that all pCRFs can be transmitted through heterosexual contact. The 01/B'/C recombinants were highly associated with sexual transmission ($n = 22/23$) as opposed to intravenous drug use. In contrast, nearly all B'C clusters in pCRFs were both circulating in IDU and heterosexual population.

The pCRFs sequences were blast for similarities to all known sequences in the Los Alamos HIV Sequence Database from Dehong, elsewhere in China, and the neighboring countries India, Myanmar,

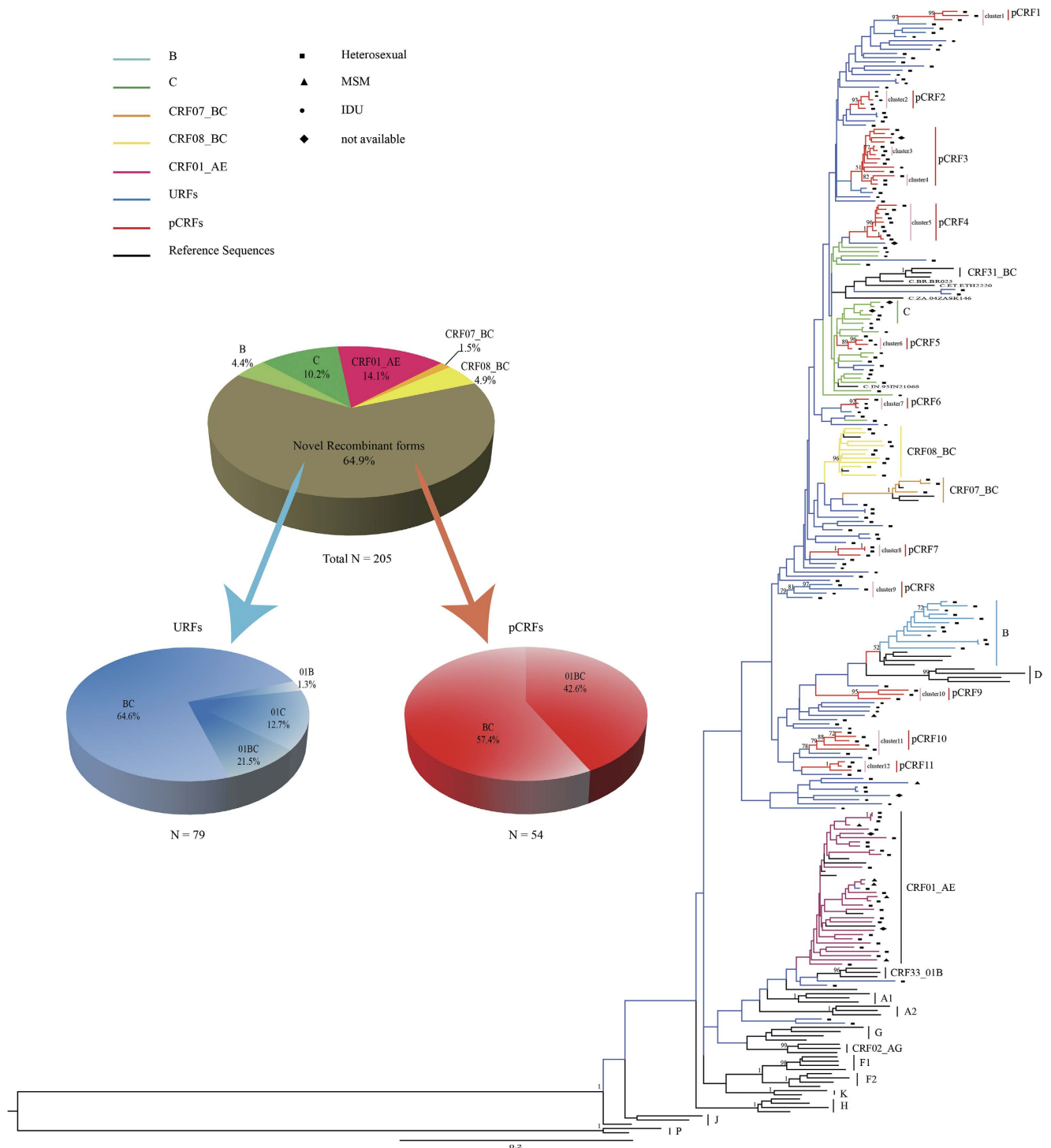


Figure 1. Phylogenetic analysis of the 1.2 kb *pol* gene fragments of 205 samples. Sequences were aligned and phylogenetic tree constructed using the maximum-likelihood method in PhyML. Strain designations are marked with different colors. Clusters with high bootstrap values and potential circulating recombinant clades (pCRFs, $n = 11$) are marked with vertical lines and labeled accordingly. Markers on the individual sequences indicate the route of transmission: ■, heterosexual sex; ▲, MSM sex; ●, intravenous drug use; and ◆, not available. The proportion of each subtype is also shown in the figure.

Thailand, and Vietnam. Phylogenetic analyses indicated few links between the current Dehong recombinant strains and historical transmissions elsewhere (Supplemental Figure S1). All pCRFs, except pCRF6, clustered closely with numerous sequences sampled from Dehong in 2005, 2009, and 2010^{11,15,16}. pCRF6, which found solely in Burmese permanent residents, co-clustered only with sequences sampled from

Total, n (%)	Conventional		pCRFs		URFs		pCRFs		URFs	
	72 (35.1%)	54 (26.3%)	79 (38.5%)	Crude OR (95% CI)	Adjusted OR (95% CI)	Crude OR (95% CI)	Adjusted OR (95% CI)			
Year of diagnosis										
2009	22 (40.7%)	20 (37.0%)	12 (2.2%)	1		1				
2010	20 (40.8%)	12 (24.5%)	17 (34.7%)	0.66 (0.26–1.69)		1.56 (0.60–4.05)				
2011	12 (27.9%)	7 (16.3%)	24 (55.8%)	0.64 (0.21–1.95)		3.67 (1.37–9.84)*				
2012	18 (30.5%)	15 (25.4%)	26 (44.1%)	0.92 (0.37–2.29)		2.69 (1.05–6.68)*				
Sex										
Female	41 (39.4%)	28 (26.9%)	35 (33.7%)	1		1				
Male	31 (30.7%)	26 (25.7%)	44 (43.6%)	1.23 (0.60–2.50)		1.66 (0.87–3.17)				
Ethnicity										
Han	35 (42.2%)	24 (28.9%)	24 (28.9%)	1		1				
Jingpo (minority)	16 (25.4%)	19 (30.2%)	28 (44.4%)	1.73 (0.75–4.03)		2.55 (1.14–5.71)*				
Other minority	21 (35.6%)	11 (18.6%)	27 (45.8%)	0.76 (0.31–1.87)		1.87 (0.87–4.06)				
Education level										
Junior and above	36 (43.9%)	28 (34.1%)	18 (22.0%)	1	1	1	1			
No schooling and Primary	36 (29.3%)	26 (21.1%)	61 (49.6%)	0.93 (0.46–1.88)	0.81 (0.38–1.74)	3.39 (1.68–6.82)*	2.541 (1.09–5.36)*			
Occupation										
Not farmer	48 (31.4%)	40 (26.1%)	65 (42.5%)	1		1				
Farmer	24 (46.2%)	14 (26.9%)	14 (26.9%)	1.43 (0.65–3.12)		2.32 (1.09–4.95)*				
Residence										
Luxi County	37 (44.6%)	26 (31.3%)	20 (24.1%)	1	1	1	1			
Other Dehong county	21 (27.6%)	23 (30.3%)	32 (42.1%)	1.56 (0.72–3.39)	1.56 (0.70–3.49)	2.82 (1.30–6.11)*	2.90 (1.28–6.60)*			
Myanmar	10 (27.8%)	2 (5.6%)	24 (66.7%)	0.29 (0.06–1.41)	0.28 (0.54–1.41)	4.44 (1.78–11.1)*	3.17 (1.20–8.39)*			
Other region in China	4 (40.0%)	3 (30.0%)	3 (30.0%)	1.07 (0.22–5.18)	1.18 (0.21–6.51)	1.39 (0.28–6.82)	1.79 (0.26–12.4)			
High-risk contact history										
Heterosexual	54 (35.5%)	43 (28.3%)	55 (36.2%)	1		1				
Injecting drug use	10 (25.0%)	10 (25.0%)	20 (50.0%)	1.26 (0.48–3.29)		1.96 (0.84–4.58)				
MSM ^e	4 (66.7%)	0 (0%)	2 (33.3%)	–		–				
Not available ^e	4 (57.1%)	2 (28.6%)	1 (14.3%)	–		–				

Table 2. Multiple Logistic regression analysis of factors associated with infection with pCRFs and URFs, as compared to conventional subtypes (B, C, CRF01_AE, CRF07_BC, and CRF08_BC). *OR greater than one for both upper and lower bound of 95% CI. ^eThe high-risk contact category “MSM”, as well as the “Not available” category was excluded from analysis due to its small number of cases. Cases were designated missing “-”.

northern Myanmar in 2009¹⁰. In addition, pCRF4 also co-clustered with one sequence from Anhui province¹⁷, while pCRF9 with one sequence in Yunnan elsewhere¹⁸ and another one from neighboring Guangxi province¹⁹.

Factors related with the generation of new HIV-1 recombinant strains. Chi-square tests between each subtype and the demographic risk factors were performed (Supplemental Table S2). To study the relationship between HIV subtype and demographic risk factors, both univariate and multivariate logistic regression analyses were performed (Table 2). Viral strains were grouped as conventional subtypes, pCRFs, and URFs, with consideration of the epidemic time and scope. The percentage of each category with particular factors is also summarized in Table 2. Univariate regression analysis showed that sampling year (2011: OR = 3.66, 95% CI = 1.37–9.84; 2012: 2.65, 1.05–6.68), ethnicity (Jingpo: 2.55, 1.14–5.71), education (primary school and no schooling: 3.39, 1.68–6.82), occupation (farmer: 2.32, 1.09–4.95) and residence (other Dehong county: 2.82, 1.30–6.11; Myanmar: 4.44, 1.78–11.1) were related with URFs category, which was mostly the same as chi-square analysis (Supplemental Table S3). Upon adjustment for multivariate logistic regression, only education (AOR = 2.41, 95% CI = 1.09–5.36) and residence (other Dehong county: 2.90, 1.28–6.60; Myanmar: 3.17, 1.20–8.39) remained significant

Cluster	Recombinant map	Subtype	Patient ID ^a	Risk history	Ethnicity	Residence	Sex	Prevalence (%)
pCRF1		01/B/C	DH10JJX01 DH10JJX32 DH12JJX43	Married sex(2/3) Extramarital sex(1/3)	Han(3/3)	Longchuan(1/3) Yingjiang(1/3) Luxi(1/3)	Female(2/3) Male(1/3)	1.46
pCRF2		B/C	DH09JJX01 DH09JJX18 DH10JJX34 DH12JJX27 DH12JJX41	IDU(4/5) Extramarital sex(1/5)	Jingpo(4/5) Dai(1/5)	Longchuan(5/5)	Male(5/5)	2.44
pCRF3		B/C	DH09JJX07 DH09JJX13 DH09JJX24 DH10JJX24 DH09JJX40 DH12JJX14 DH09JJX42 DH10JJX26 DH09JJX43 DH12JJX18 DH09JJX46 DH12JJX50	Extramarital sex(10/14) Married sex(1/14) IDU(2/14) NA(1/14)	Jingpo(10/14) Han(3/14) Dai(1/14)	Luxi(11/14) Longchuan(2/14) Lianghe(1/14)	Female(7/14) Male(7/14)	6.83
pCRF4		01/B/C	DH09JJX19 DH10JJX18 DH09JJX20 DH10JJX29 DH09JJX35 DH11JJX33 DH09JJX50 DH12JJX24 DH12JJX42	Extramarital sex(8/9) Married sex(1/9)	Han(6/9) Dai(3/9)	Luxi(5/9) Myanmar(1/9) Lianghe(1/9) Tengchong(1/9) Ruili(1/9)	Female(6/9) Male(3/9)	4.39
pCRF5		B/C	DH09JJX17 DH09JJX31 DH11JJX02	IDU(2/3) Extramarital sex(1/3)	Han(1/3) Dai(1/3) Jingpo(1/3)	Longchuan(2/3) Luxi(1/3)	Male(2/3) Female(1/3)	1.46
pCRF6		01/B/C	DH11JJX04 DH12JJX23 DH12JJX38	IDU(1/3) Married sex(1/3) Extramarital sex(1/3)	Han(3/3)	Myanmar(2/3) Longchuan(1/3)	Female(2/3) Male(1/3)	1.46
pCRF7		01/B/C	DH09JJX02 DH10JJX43 DH11JJX39	Extramarital sex(2/3) Married sex(1/3)	Han(1/3) Jingpo(1/3) other minority(1/3)	Longchuan(2/3) Ruili(1/3)	Female(2/3) Male(1/3)	1.46
pCRF8		B/C	DH09JJX53 DH12JJX56 DH12JJX05	Extramarital sex(2/3) IDU(1/3)	Han(2/3) Jingpo(1/3)	Myanmar(1/3) Luxi(2/3)	Male(2/3) Female(1/3)	1.95
pCRF9		B/C	DH10JJX48 DH12JJX09 DH12JJX47	Extramarital sex(2/3) Married sex(1/3)	Han(2/3) Dai(1/3)	Longchuan(1/3) Longling(1/3) Luxi(1/3)	Male(2/3) Female(1/3)	1.46
pCRF10		01/B/C	DH09JJX30 DH10JJX08 DH11JJX18 DH12JJX12 DH11JJX42	Extramarital sex(6/6)	Jingpo(3/6) Han(1/6) Dai(1/6) other minority(1/6)	Luxi(5/6) Yingjiang(1/6)	Female(4/6) Male(2/6)	2.44
pCRF11		B/C	DH09JJX04 DH10JJX45 DH11JJX23	Extramarital sex(3/3)	Han(2/3) other minority(1/3)	Longchuan(1/3) Baoshan(1/3) Yingjiang(1/3)	Female(2/3) Male(1/3)	1.46

Figure 2. Characteristics of potential circulating recombinant forms. The recombinant patterns of the 11 pCRFs were listed, with brief summary of demographic information of each cluster. Patient IDs in bold indicate samples with NFLG sequences, and all other IDs indicate 1.2 kb *pol* gene fragments.

factors. Generally, the URFs category which represents newly emerged recombined HIV strains was highly related with people with lower education levels and who were living in Myanmar or other regions beyond the capital city of Dehong. However, no relationship was found between pCRFs category and factors refer to the conventional subtype's category. And the distribution of new recombinant HIV-1 strains showed no significant differences among populations with different high-risk contact history.

Discussion

Since the 1980s, Dehong prefecture in China's southwest border near the drug producing "the Golden Triangle" has been a hotspot of HIV infection and recombination, with great impact on China's overall HIV epidemic^{5,6}. Numerous studies have described the prevalence of recombinant forms and their transmissions between neighboring regions^{10,11,15,16,20–23}. However, these studies either constrained to a single group of population^{10,11,20} or failed to identify the current transmitting strains^{20,21} as well as explore the correlates of emerging recombinants^{15,16,22,23}. Numerous studies reported high rates of recombinant HIV-1 found in IDUs^{10,11,24}, since the parental strains with greater genetic similarity are more likely to recombine²⁵. Infection caused by multiple founder viruses happened more often in IDUs compared heterosexually infected population¹². Without the mucosal barrier in IDUs during transmission, it may also lead more easily to recombinant HIV-1 strains. There is rarely report for high proportions of recombinant HIV-1 strains in sexual driving epidemic in Asia. Two studies reported comparatively high proportions of recombinant HIV-1 strains in sexually driven epidemics in sub-Saharan Africa^{26,27}, one with 25% found in a group of antiretroviral treated patients in Ghana, the other with 47% found in Democratic Republic of Congo(DRC). However, these studies did not distinguish long term and recent infections, and thus the recombinants observed may not reflect the current epidemic.

In this study, it characterized the currently occurring epidemic of HIV-1 infections among young people in Dehong Prefecture in Yunnan Province of China. The sampling methodology was based on

WHO's threshold survey of HIV drug resistance strategy with strong supports by China's public health infrastructure. All young people newly diagnosed in the first Quarter of each year by all county CDCs in Dehong Prefecture were enrolled into the study consecutively for 4 years, from 2009 to 2012. Our study found that two third (64%) of the newly diagnosed HIV-1 infections in Dehong was caused by new forms of recombinant HIV-1 strains. To our knowledge, this is reported for the first time very high proportion of new recombinant HIV-1 in predominantly heterosexually infected populations.

Among the newly described new recombinant forms, phylogenetic and breakpoint analyses showed that there were at least 11 recombinant clusters that have been formed with capability to transmit locally, in addition to the conventional subtype clusters. The newly formed clusters were considered as pCRFs to distinguish them with conventional CRFs that circulating in China, while other recombinants referred to URFs. The emergence of the pCRFs and URFs were uneven and statistically correlated with residency outside the prefecture capital, migration from Myanmar, and lower education levels.

Within pCRFs clusters, nearly all structures were composed of a subtype C genomic backbone, except for one whose genome consisted of roughly one third subtype B', C, and 01_AE. Previous studies found that CRF01_AE strains were mainly associated with sexual transmission in China, while the B'/C recombinant subtypes were mainly associated with intravenous drug use³. Hence, it is likely that the pCRFs observed in the predominantly sexually transmitted populations in our study were related to HIV-1 strains originated from earlier IDU epidemics. Further studies are needed to clarify this hypothesis.

Dehong is China's first HIV-1 epidemic region and past studies indicated that subtype B', C, CRF07_BC and CRF08_BC were all introduced to China through Dehong⁴⁻⁹. In this study, eleven pCRFs are found, indicating that the recombinant strains are already obtaining the ability to spread in this region. It is reported that recombination is more evolutionarily advantageous than point mutations, since it introduces a large number of mutations while maintaining the capability of transmission²⁸. In addition to the new recombinants' transmission ability, other facilitating factors will further determine its transmission impact. The drug trafficking routes through Dehong facilitated the spread of the CRF07_BC and CRF08_BC from Thailand, Myanmar, and Laos to other areas of China in the 1990s. With great effort put on crackdown drug trafficking and methadone treatment to IDUs, the HIV incidence in IDUs of Dehong has declined²⁹. However, the identification of multiple circulating recombinant HIV-1 strains in local sexually active youths is worrisome, considering the gateway position of Dehong in China's HIV epidemic history. Meanwhile, recent promotion of border tourism and trades may introduce millions of people from other parts of China (<http://www.dehong.gov.cn/>), increasing the transmission risk of novel recombinant HIV-1 strains. The situation therefore deserves attention from local and national public authorities.

In this study, there are more Myanmar that sampled in 2011 and 2012 than the two previews years. This may be related to concurrent conflicts in northern Myanmar, which increased the outgoing refugee count from northern Myanmar to Dehong (<http://www.unhcr.org/5049e42b6.html>). For historical reasons, people in Dehong and Myanmar along the border share close relationships. Our study found that the newly formed URFs were significantly related to the Myanmar residents near the border, in which more than half are of Jingpo ethnicity, an ethnic minority group that inhabits both sides of the China-Myanmar border. In statistical analyses, high risk contact history showed no effect on URF generation, but a close association is observed between IDU status and Jingpo ethnicity (Supplemental Table S3, $p = 0.005$). In consideration of the fact of higher probability of recombination in IDUs, it is speculated that the Jingpo ethnicity in Myanmar may play an important role in the formation of recombinant strains. Meanwhile, significant association was also found between URFs generation and lower education levels, suggesting that those youths are most at risk for multiple infections and recombinant infections.

Since 1980s, the local governments have paid much effort on HIV prevention and control. Dehong is one of the few places in China where the number of reported HIV infection cases have experienced a plateau or slight decline in recent years²⁹. However, the high proportion of new recombinant HIV-1 strains among recently infected youths, as well as increasing cross-border migration, indicate that the fight against AIDS in Dehong may still be far from over. The challenges for further controlling the HIV epidemic in Dehong include: 1) heterosexual and IDU transmissions are intertwined in overlapping populations, which facilitated the recombinant strain formation and transmission; 2) cross-border migrations of both Burmese and Chinese residents; 3) difficulties in reaching people with lower education levels and those living in remote regions beyond the prefecture capital. The findings of the study also indicated that monitoring the prevalence of HIV recombinants has important implications for assessing the overall risk exposure to HIV infection and for serving as an early warning indicator for potential outbreaks by new viral recombinants.

As the low hanging fruits in AIDS prevention had been picked up, redoubled efforts are needed to implement in-depth interventions and leave no untouched stones. More specific prevention measures should be targeted to reduce the multiple risk exposures in both the IDUs and heterosexuals populations in Dehong. Such efforts include continued crackdown on drug trafficking and improve harm reduction services, such as methadone treatment clinics, needle exchanges programs and condom promotion, to the last mile reaching people with low education and living in remote areas. As the spread of infectious disease recognizes no border, so do the control efforts. The Dehong government should extend regular information sharing with the neighboring Burmese authorities to coordinate synergistic and joint HIV/

AIDS control and intervention programs. The success of Dehong's AIDS control campaign will rely on collaborative prevention and control efforts on both sides of the border.

The study has some limitations. Since the epidemic of HIV-1 in Dehong is quite unique, and many recombinant strains have been reported here, using of short gene segments (approximately 1200 bps) to define the genotype may underestimate the variety of and frequency of recombinants. We have attempted amplification of *pol* and *gag* fragments and near full-length genomes for as many samples as possible; however, they were not obtained for all strains due to limited blood plasma stock and long term cryo-preservation.

Methods

Ethics Statement. The use of Human Blood Plasma in the current study has been approved by the institutional review boards of the National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention. All study participants gave their written informed consent. The methods were carried out in "accordance" with the approved guidelines. All experimental protocols were approved by the institutional review boards of the National Center for AIDS/STD Control and Prevention, China CDC.

Study population. The study population was selected according to the WHO Threshold Survey method for transmitted HIV drug resistance surveillance in 2009 to 2012³⁰. Per national regulations, newly diagnosed cases of HIV infection are reported to the Centers for Disease Control and Prevention (CDC) system at the county, prefecture, province, and national level within 24 hours (http://english.gov.cn/laws/2005-10/10/content_75718.htm). All HIV-1 infected persons of less than 25 year old reported to the county CDCs in Dehong prefecture were consecutively enrolled. If less than 34 eligible participants were enrolled by the end of a three-month period, the sampling period was extended until the minimum sample size was reached. Participants were excluded if their CD4⁺ cell counts < 500 cells/ml and, if female, had previous pregnancy histories. Basic demographic information, risk behavior histories, and a blood sample were acquired from each participant following written informed consent. Participants infected through mother-to-child transmission were further excluded from this study.

Genotyping. Blood plasma samples were subjected to viral RNA extraction and cDNA synthesis as previously described^{31,32}. A 1208 base-pair (bp) segment of the HIV-1 *pol* gene (HXB2 coordinates: 2242–3450 nt) was amplified and sequenced using an in-house drug resistance genotyping method³¹. Briefly, viral RNA was extracted from patient blood plasma using the QIAamp Viral RNA Mini Kit (Qiagen, Germany). cDNA transcription and first-round polymerase chain reaction (PCR) was performed by one-step RT-PCR using primers RT21 (CTGTATTTCAGCTATCAAGTCTTTTGATGGG) and MAW26 (TGGAAATGTGGAAAA GAAGGAC). The primers of second-round PCR were PRO-1 (CAGAGCCAACAGCCCCACCA) and RT4R (CTGCCAATTCTAATTCTGCTTC). PCR products were purified and sequenced with an ABI PRISM 3100 DNA sequencer (Applied Biosystems, USA). For further confirmation of the sequence subtype, a segment of *gag* gene (790–1862) was also amplified.

For near full-length genome (NFLG) amplification, RNA was reverse transcribed and cDNA with near-end point dilution was amplified followed by positive PCR products purifying and sequencing³². Phylogenetic analysis was used to identify the subtype of sequences with the standard reference sequences downloading from Los Alamos database. Phylogenetic analysis was performed using the maximum likelihood (ML) method in PhyML³³. The general time reversible (GTR) plus gamma model of nucleotide substitution was selected by FindModel (<http://www.hiv.lanl.gov>) to construct ML trees. Branch significance was analyzed by bootstrap with 200 replicates. Recombinants including CRF07_BC, CRF08_BC and sequences that cannot be classified into B, C and CRF01_AE were verified by bootscan analysis using SimPlot v 3.5.1 with window size of 120 bp and step size of 10 bp³⁴ and Recombinant Identification Program (RIP)³⁵.

Statistical analysis. Samples were divided into three categories according to the genotypes of their recombinant forms: conventional subtypes, potential circulating recombinant forms (pCRFs), and URFs. The conventional subtype groups include subtypes B, C, CRF01_AE, CRF07_BC, and CRF08_BC, which are widely prevalent in China³. The URFs group refers to newly formed or non-propagating recombinant strains, while the pCRFs group consists of strains with three or more samples clustered in phylogenetic clusters with high bootstrap values (≥ 0.7), indicating strains that are currently in local circulation with limited range of spread. The associations between demographic risk factors and infection by the three categories were analyzed using Chi-square tests and univariate and multivariate logistic regression with both the forward selection and backward elimination stepwise methods. In multivariate regression analysis, risk behavior history categories with very few cases such as men-who-have-sex-with-men and those with unknown routes of infection, were excluded from analysis. A two-sided *p* value of 0.05 or less was regarded as significant. All statistical analysis was performed using SPSS Statistics version 19.

References

1. Hemelaar, J. *et al.* Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS* **25**, 679–689, doi:10.1097/QAD.0b013e328342ff93 (2011).

2. Robertson, D. L. *et al.* HIV-1 nomenclature proposal. *Science* **288**, 55–56 (2000).
3. He, X. *et al.* A comprehensive mapping of HIV-1 genotypes in various risk groups and regions across China based on a nationwide molecular epidemiologic survey. *PLoS One* **7**, e47289, doi:10.1371/journal.pone.0047289 (2012).
4. MA, Y., Li, Z. & Zhao, S. HIV infected people were first identified in intravenous drug users in China. *Chin J Epidemiol* **11**, 184–185 (1990).
5. Shao, Y. *et al.* Sequence analysis of HIV env genes among HIV infected drug injecting users in Dehong epidemic area of Yunnan province, China. *Chin J Virol* **10**, 291–299 (1994).
6. Luo, C. C. *et al.* HIV-1 subtype C in China. *Lancet* **345**, 1051–1052 (1995).
7. Su, L. *et al.* Characterization of a virtually full-length human immunodeficiency virus type 1 genome of a prevalent intersubtype (C/B') recombinant strain in China. *J Virol* **74**, 11367–11376 (2000).
8. Tee, K. K. *et al.* Temporal and spatial dynamics of human immunodeficiency virus type 1 circulating recombinant forms 08_BC and 07_BC in Asia. *J Virol* **82**, 9206–9215, doi:10.1128/JVI.00399-08 (2008).
9. Li, Z. *et al.* Tracing the origin and history of HIV-1 subtype B' epidemic by near full-length genome analyses. *AIDS* **26**, 877–884, doi:10.1097/QAD.0b013e328351430d (2012).
10. Pang, W. *et al.* Extensive and complex HIV-1 recombination between B, C and CRF01_AE among IDUs in south-east Asia. *AIDS* **26**, 1121–1129, doi:10.1097/QAD.0b013e3283522c97 (2012).
11. Han, X. *et al.* High Prevalence of HIV-1 Intersubtype B'/C Recombinants among Injecting Drug Users in Dehong, China. *PLoS one* **8**, e65337 (2013).
12. Bar, K. J. *et al.* Wide variation in the multiplicity of HIV-1 infection among injection drug users. *J Virol* **84**, 6241–6247, doi:10.1128/jvi.00077-10 (2010).
13. Ruan, Y. *et al.* Incidence of HIV, hepatitis C and hepatitis B viruses among injection drug users in southwestern China: a 3-year follow-up study. *AIDS* **21** Suppl **8**, S39–46, doi:10.1097/01.aids.0000304695.54884.4f (2007).
14. Zhang, L. *et al.* HIV prevalence in China: integration of surveillance data and a systematic review. *The Lancet infectious diseases* **13**, 955–963, doi:10.1016/s1473-3099(13)70245-7 (2013).
15. Li, L. *et al.* Recombination form and epidemiology of HIV-1 unique recombinant strains identified in Yunnan, China. *PLoS One* **7**, e46777, doi:10.1371/journal.pone.0046777 (2012).
16. Chen, M. *et al.* Emerging Variability in HIV-1 Genetics among Recently Infected Individuals in Yunnan, China. *PLoS One* **8**, e60101, doi:10.1371/journal.pone.0060101 (2013).
17. Wu, J. *et al.* New emerging recombinant HIV-1 strains and close transmission linkage of HIV-1 strains in the Chinese MSM population indicate a new epidemic risk. *PLoS One* **8**, e54322, doi:10.1371/journal.pone.0054322 (2013).
18. Bao, L. *et al.* Molecular tracing of sexual HIV Type 1 transmission in the southwest border of China. *AIDS Res Hum Retroviruses* **24**, 733–742, doi:10.1089/aid.2007.0269 (2008).
19. Su, Q., Liang, H., Cen, P., Bi, Z. & Zhou, P. HIV type 1 subtypes based on the pol gene and drug resistance mutations among antiretroviral-naïve patients from Guangxi, Southern China. *AIDS Res Hum Retroviruses* **28**, 725–728, doi:10.1089/AID.2011.0246 (2012).
20. Yang, R. *et al.* On-going generation of multiple forms of HIV-1 intersubtype recombinants in the Yunnan Province of China. *AIDS* **16**, 1401–1407 (2002).
21. Zhang, Y. *et al.* Dominance of HIV-1 subtype CRF01_AE in sexually acquired cases leads to a new epidemic in Yunnan province of China. *PLoS medicine* **3**, e443, doi:10.1371/journal.pmed.0030443 (2006).
22. Chen, M. *et al.* Genetic diversity and drug resistance among newly diagnosed and antiretroviral treatment-naïve HIV-infected individuals in western Yunnan: a hot area of viral recombination in China. *BMC Infect Dis* **12**, 382, doi:10.1186/1471-2334-12-382 (2012).
23. Su, Y. Z. *et al.* Update on diversity and distribution of HIV-1 subtypes in Yunnan province. *Epidemiology and Infection* **141**, 2418–2427, doi:10.1017/S0950268812002713 (2013).
24. Tee, K. K., Saw, T. L., Pon, C. K., Kamarulzaman, A. & Ng, K. P. The evolving molecular epidemiology of HIV type 1 among injecting drug users (IDUs) in Malaysia. *AIDS Res Hum Retroviruses* **21**, 1046–1050, doi:10.1089/aid.2005.21.1046 (2005).
25. Delviks-Frankenberry, K. *et al.* Mechanisms and factors that influence high frequency retroviral recombination. *Viruses* **3**, 1650–1680, doi:10.3390/v3091650 (2011).
26. Delgado, E. *et al.* High prevalence of unique recombinant forms of HIV-1 in Ghana: molecular epidemiology from an antiretroviral resistance study. *J Acquir Immune Defic Syndr* **48**, 599–606, doi:10.1097/QAI.0b013e3181806c0e (2008).
27. Kita, K. *et al.* Genetic diversity of HIV type 1 in Likasi, southeast of the Democratic Republic of Congo. *AIDS Res Hum Retroviruses* **20**, 1352–1357, doi:10.1089/aid.2004.20.1352 (2004).
28. Tebit, D. M., Nankya, L., Arts, E. J. & Gao, Y. HIV diversity, recombination and disease progression: how does fitness “fit” into the puzzle. *AIDS reviews* **9**, 75–87 (2007).
29. Yang, L. *et al.* Analysis on AIDS epidemic posture and control effect in Dehong in 2011. *J Med Pest Control* **29**, 263–265 (2013).
30. Bennett, D. E., Myatt, M., Bertagnolio, S., Sutherland, D. & Gilks, C. F. Recommendations for surveillance of transmitted HIV drug resistance in countries scaling up antiretroviral treatment. *Antiviral therapy* **13** Suppl **2**, 25–36 (2008).
31. Liao, L. *et al.* Genotypic analysis of the protease and reverse transcriptase of HIV type 1 isolates from recently infected injecting drug users in western China. *AIDS Res Hum Retroviruses* **23**, 1062–1065, doi:10.1089/aid.2007.0050 (2007).
32. Wei, H. *et al.* Near full-length genomic characterization of a novel HIV type 1 CRF07_BC/01_AE recombinant in men who have sex with men from Sichuan, China. *AIDS Res Hum Retroviruses* **29**, 1173–1176, doi:10.1089/AID.2013.0104 (2013).
33. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology* **59**, 307–321, doi:10.1093/sysbio/syq010 (2010).
34. Lole, K. S. *et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* **73**, 152–160 (1999).
35. Siepel, A. C., Halpern, A. L., Macken, C. & Korber, B. T. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. *AIDS Res Hum Retroviruses* **11**, 1413–1416 (1995).

Acknowledgements

We would like to thank Prof. Oliver Pybus and Dr. Nuno Faria from the Department of Zoology, University of Oxford and Prof. Yutaka Takebe from AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan for their suggestions on data analysis. This work was supported by grants from National Major Projects for Infectious Diseases Control and Prevention [grant numbers 2012ZX10001-002, 2012ZX10001-008, 2009DFB30420], National Natural Science Foundation of China [grant numbers 81020108030, 81261120393], the International Development Research Center of Canada [grant number 104519-010], and the Chinese State Key Laboratory for Infectious Disease Development Grant [grant number 2012SKLID103]

Author Contributions

YS and HX were responsible for study design. LL, LLu, SY, SD, YM and MJ contributed to data collection; HW, JH, YF, CH, YR, YH, XH and YS performed experiments and analyzed the data; HW, JH, HX and YS contributed to interpretation of results; and HW, JH, and YS wrote the paper with contributions from all authors. All authors have reviewed and approved 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: Wei, H. *et al.* The sexually driven epidemic in youths in China's southwestern border region was caused by dynamic emerging multiple recombinant HIV-1 strains. *Sci. Rep.* **5**, 11323; doi: 10.1038/srep11323 (2015).



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/>