

# Chapter 15

## Socioepidemiology of Injection Drug Users in Miami and HIV-1B Envelope (V1–V5) Genetic Diversity: A Preliminary Study

Paul Shapshak, Brian T. Foley, Seetharaman Balaji, David M. Segal, Clyde McCoy, and J. Bryan Page

**Keywords** Socioepidemiology • Relationship • Network • Risk locale • Risk behavior • Injection drug user (IDU) • Sexual risk • Psychiatric morbidities • HIV-1B infection • Envelope • ENV V1-V5 hypervariable domains • Nucleotide sequence • Cluster • Phylogenetic analysis (phylogeny) • Protein sequence • Molecular similarity • Glycosylation • Entropy

### Core Message

Injection drug use is a major risk behavior associated with transmission of HIV-1B.

Sociocultural and socioepidemiological studies are required to characterize networks of injection drug users (IDUs).

---

P. Shapshak, PhD (✉)

Division of Infectious Disease and International Health, Department of Internal Medicine, Tampa General Hospital, University of South Florida, Morsani College of Medicine, Tampa, FL 33606, USA  
e-mail: [pshapshak@gmail.com](mailto:pshapshak@gmail.com)

B.T. Foley, PhD

Theoretical Division, MS K710, T-10, HIV Databases, Los Alamos National Laboratory, Los Alamos, NM, USA

S. Balaji, MSc, PhD

Department of Biotechnology, Manipal Institute of Technology, Manipal University, Manipal 576104, Karnataka, India

D.M. Segal, PhD

Department of Bioinformatics, College of Health Sciences, Walden University, Minneapolis, MN 55401, USA

C. McCoy, PhD

Department of Epidemiology, University of Miami Miller School of Medicine, Miami, FL 33136, USA

Comprehensive Drug Research Center, University of Miami Miller School of Medicine, Miami, FL 33136, USA

The possibility of links between subject socioepidemiology and viral sequence diversity, phylogenetic relationships, signatures, thermodynamics, and glycosylation patterns is studied.

This chapter addresses whether in risk locales where many people inject together, there are variations in probability of relatedness of HIV-1B env sequences.

## 15.1 Introduction

What became known as HIV-1 was discovered and first isolated in three different laboratories, namely of Drs. J. Levy, L. Montagnier, and R.C. Gallo [1–4]. The complex distribution of human immunodeficiency virus type 1B (HIV-1B) genotypic variants, or quasispecies, within infected individuals has been characterized extensively [5]. There are many potential routes of HIV-1B transmission, including parenteral (injection drug use, injection drug user (IDU), heterosexual and homosexual sexual transmission), blood transfusion, and perinatal (*prepartum* and *postpartum*). Upon transmission, some studies indicate that the major variant found in the donor is transmitted [6], while others report transmission of a minor variant [7–9]. For either alternative, the person newly infected through these routes has an HIV-1B sequence population that is initially homogeneous [7, 10, 11], irrespective of the heterogeneity of the donor's sequences. With time, the HIV-1B high mutation rate and selective pressures generate heterogeneous populations, quasispecies, or sequence variant clouds [8, 12–19].

Sequence heterogeneity in the envelope region occurs primarily in the hypervariable domains, designated V1–V5, which has attracted much attention for characterizing genotypic and phenotypic variants [20]. Because the env gene is particularly prone to mutation [21], it has proven invaluable in molecular epidemiological studies tracing the patterns of disease transmission and progression [9]. Several analytical approaches have been implemented using different variable domains of the HIV-1B ENV gene to investigate viral transmission: genetic distances, phylogenetic trees, and sequence signature patterns. Phylogenetic analysis has been used to establish the likelihood of HIV-1B transmission from infected health care workers to patients [11, 14]. Only recently, has such analysis been used to determine whether different transmission groups possess characteristic variants [22–24]. Risk group-associated variations in sequence signatures have been described between drug user, transfusion, and homo-

---

J.B. Page, PhD

Department of Anthropology, University of Miami, Coral Gables, FL 33133, USA

Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL 33136, USA

Department of Epidemiology, University of Miami Miller School of Medicine, Miami, FL 33136, USA

Comprehensive Drug Research Center, University of Miami Miller School of Medicine, Miami, FL 33136, USA

sexual/hemophilic DNA sequences within the V3 loop domain [25, 26]. Mother-infant transmissions reveal that the V3 sequences are much more heterogeneous in the mother than in the infant after birth. The amino acid sequence signatures correlated with linked vertical transmission between these pairs [27]. Earlier in the epidemic, two predominant genotypes of HIV-1B (Thai A and B) were known to exist in Thailand [28]; interestingly, they showed in studies of the V3 loop region that there had been independent introduction of the virus into two high-risk populations distinguished by mode of transmission (sexual activity vs intravenous drug use). Characterization of recombinant segments in the C2–V3 region demonstrated transmission between both individuals within a sexual risk couple [7]. In addition, dual transmission of HIV-1B and C subtypes was detected in a family, husband to wife to child [29].

Injection drug use is a primary risk behavior associated with transmission of HIV [30–52]. Many sociological and epidemiological characterizations of drug abuse and HIV infection networks include NIDA. There have been many detailed studies characterizing the IDU and sexual transmission of HIV-1B and other epidemiologically linked viruses including HCV. Detailed molecular studies identified and tracked virus strains in well-defined drug abuse networks as part of sociological and epidemiological studies.

One factor that complicates deducing HIV infection-related risk networks is the infectivity potential of virus strains. This involves inferring the probability that various virus strains infect new hosts after each single exposure due to use of contaminated drug injection needle-syringe, cotton, cooker, and washwater. The CDC reports that the average probability of HIV infection is 0.33 per 100 needlestick or cut exposures<sup>1</sup>. Consequently, the establishment of risk networks in which a single V1–V5 HIV-1B strain variant dominates is unlikely. There is also a lowered chance of transmission of any given strain, especially in circumstances of multiple IDUs injecting in the same risk locale. However, long-term intimate relations among socially characterized participants warrant a higher likelihood of exhibiting detectable clustering of HIV strains with increased sequence relatedness than diffusely distributed networks [53–57].

In this study, we characterized and defined the sociodynamics of four networks of seropositive IDUs. In addition, we characterized the viral sequence within specific hypervariable domains (V1–V5) of 37 HIV ENV genes derived from the IDUs.

## 15.2 Materials and Methods

### 15.2.1 Network Epidemiology

Epidemiological methods for assessment and outreach of individuals who inject drugs included several approaches and protocols. Four epidemiological networks linked by sexual interactions and/or IDU behavior were studied (Table 15.1).

---

<sup>1</sup>For comparison, the average risks for HCV and HBV infections are 1.8 per 100 and 6–30 per 100 needlestick or cutaneous cut exposures, respectively [53].

**Table 15.1** Patient HIV-1B IDU demographics

Participant	Gender	Risk locale <sup>a,b</sup>	Risk behavior	Drugs used	Date of blood draw	HIV-1B DNA clones sequenced
1001	Female	Overtown	IDU and sex	Heroin and crack cocaine	8-7-1995 8-29-1996	3
1002	Male	Overtown	IDU and sex	Heroin and crack cocaine	8-7-1995 8-29-1996	5
1003	Female	Overtown	IDU and sex	Heroin and Dilaudid	8-21-1995 8-29-1996	2
1004	Male	Overtown	IDU and sex	Heroin	8-21-1995 8-30-1996	4 5
1005	Female	Overtown	IDU and sex	Heroin	9-20-1995 8-30-1996	4
1006	Male	Overtown	IDU and sex	Heroin and cocaine	10-13-1995 11-15-1996	8 2
1008	Female	Opa-Locka	IDU	Heroin and cocaine	2-8-1996	2
1011	Female	Opa-Locka	IDU	Heroin	3-7-1996	
1012	Male	Liberty City	IDU	Heroin	3-8-1996	
1013	Male	Liberty City	IDU	Heroin	3-5-1996	
1014	Male	Liberty City	IDU	Heroin	3-29-1996	
1015	Female	Opa-Locka	IDU	Heroin	4-9-1996	2
1017	Male	Liberty City	IDU	Heroin	11-15-1996	
1018	Male	Liberty City	IDU	Heroin	11-15-1996	
1019	Male	Liberty City	IDU	Heroin	11-15-1996	

This study was performed in Miami (Dade County, Florida, USA)

<sup>a</sup>Risk locale is the neighborhood where the participants carried out risk activities

<sup>b</sup>It should be noted that IDU risk activity involves the use of cottons, cookers, rinse waters, and injection needle-syringes. Paraphernalia used at shooting galleries have been shown to be HIV-1B contaminated [32, 42, 82, 83]

Fieldwork with questionnaires was used to investigate injection drug use in Miami. Factors including frequency of participation in shooting galleries, high-risk contacts, injection drug use, sex, gender, age, and mobility were characterized to elucidate the dynamic temporal and spatial relationships between risk and network members [30, 36, 37, 42–46, 48, 58, 59]. Confidentiality was strictly maintained throughout this work. The laboratory studies were all done devoid of any personal identifiers. The Internal Review Board (IRB) rules of the University of Miami were all maintained and strictly enforced. Human subject approvals were obtained at the time of these studies from the IRB at the University of Miami.

### 15.2.2 Specimens

Blood was obtained from 15 heterosexual HIV-positive subjects (Tables 15.1 and 15.2). Blood samples were obtained for individuals in networks 1, 2, and 3 at the same initial time, whereas those in network 4 were obtained 14 months later. Follow-up samples were subsequently obtained 12 months afterward for members of the first three networks. Blood was obtained from study individuals using standard EDTA-containing vacutainer tubes. Peripheral blood lymphocyte (PBMNC) pellets were produced using Ficoll–Hypaque density gradient centrifugation and were cryopreserved at  $-85^{\circ}\text{C}$  until needed [15–17, 19].

### 15.2.3 Polymerase Chain Reaction

Genomic DNA was extracted from the cryopreserved PBMC pellets using DNA isolation kits (United States Biochemicals, Cleveland, OH) and subsequently precipitated with ethanol and resuspended in Tris-EDTA (THE) buffer (pH 7.5).

**Table 15.2** Patient networks (15 participants, 4 networks)

Network	Risk locale	Patient	Duration in network <sup>a</sup>	Additional information
1	Overtown	1001, 1002	4y IDU and sex	Recent couple
2	Overtown	1003, 1004, 1005, 1006	10y IDU	Couples with greater than 8 years of relations
		1003, 1004	13y IDU and sex	
		1005, 1006	8y IDU and sex	
3	Opa-Locka	1008, 1011	20y IDU	Female support group
		1015, 1011	20y IDU	
4	Liberty City	1017, 1018	20y IDU	Brothers
		1017, 1019	20y IDU	Brothers
		1012, 1013, 1014	15y IDU	Running partners

<sup>a</sup>y years, duration of risk activity

The V1–V5 region of the HIV-1B ENV gene was amplified using two rounds of PCR with a set of nested primers as previously described [19]. For the first PCR, we used approximately 1 mg genomic DNA in a 100  $\mu$ l reaction containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 20 nmoles (200 mM) each of dNTPs, 2.5 U AmpliTaq DNA polymerase (Perkin Elmer, Foster City, CA), and primers XPR1 (GGGATCAAAGCCTAAAGCCA, sense, nucleotide positions 6558–6577 in HXB2) and XPR7 (ACTTCTCCAATTGTCCCTCA, antisense, positions 7647–7666). Amplification consisted of a hot start (95 °C, 5'), then 30 cycles of 94 °C, 0'35"; 60 °C, 1'35"; and 72 °C, 2'35", followed by a final extension (72 °C, 10'). For the nested PCR reaction, primers XPR2 (GAATTCACCCCACTCTGTGTTA, sense, positions 6590–6605) and XPR6 (AAGCTTCTCCTCCAGGTCTGA, antisense, positions 7625–7639) were used to amplify 1 ml of the first PCR reaction under the same conditions as above. The size of the PCR amplicon was determined by analysis on 1.2% agarose gel-1X TAE in 0.5-mg/ml ethidium bromide. Typically, a single PCR band was observed using only 10  $\mu$ l of the nested PCR reaction.

Negative controls consisted of a water sample (instead of peripheral blood mononuclear cells), RT reaction mixtures without added reverse transcriptase, and reagent controls run in parallel with the tested samples [15–17, 19].

To minimize the possibility of carryover contamination, separate rooms (neither connected by airflow nor by air conditioning) were always used for the addition of reagents and sample DNA template vs performance of PCR and the handling of PCR products. Only one subject's sample was ever handled at any time during processing as well as subsequent amplification and cloning steps. In addition, sentinel tubes were evaluated weekly for contamination. Several subjects had blood samples redrawn after 12 months so that cloned sequences could be compared to verify phylogenetic relationships and to rule out sample mix-up and/or contamination [15–17, 19]. Sequence integrity was analyzed – Rodrigo and Learn [60] produced a review of several methods.

### ***15.2.4 Molecular Cloning and DNA Sequencing***

The PCR product was purified and concentrated using the High Pure PCR Product Purification Kit (Boehringer Mannheim, Mannheim, Germany) and quantified by fluorescence (Hoefler Instruments, San Francisco, CA). The amplified product (40 ng) was subcloned into pCR2.1 vector using the TA cloning kit (Invitrogen, San Diego, CA). An aliquot of the ligation reaction (110 ng) was then used to transform INVaF' cells. Plasmid DNA was isolated from positive clones, which were selected by kanamycin resistance and lacZa complementation (blue/white), using the Wizard Miniprep (Promega Corporation, Madison, WI). After digestion with EcoRI, the size of the cloned insert DNA was verified by agarose gel electrophoresis. Glycerol stocks of positive clones were prepared for long-term storage. A variety of sequencing primers both internal and external to the cloned fragment were used to sequence the primary and complementary strands (ACGT Inc., Northbrook, IL) [15–17, 19]. Clone designation: number designates subject ID; "L" designates blood, followed

by clone number; and then “D” or “R” designates DNA or RNA nucleic acid source, respectively. The addition of “+12”, e.g., 1004 + 12L4D, designates follow-up samples, obtained 12 months after the initial samples.

## 15.2.5 Sequence Analysis

### 15.2.5.1 DNA and Protein Alignments

The *env* sequences of HIV-1B IDUs from two different risk locales such as Overtown and Opa-Locka, Dade County, Florida (Table 15.2), made a total of 37 nucleotide sequences, and their corresponding proteins were also used for sequence alignments. The alignment files were generated using Clustal X 1.83 [61] with default parameters. Both intra- and intermolecular sequence variations were analyzed using Clustal X. The sequence variations and consensus patterns (signatures) were displayed using Weblogo 2.8.2 <http://weblogo.berkeley.edu/>, [62].

### 15.2.5.2 Entropy Analysis

Shannon entropy [63] is a simple quantitative measure in uncertainty units in a dataset and was used to measure the variation in DNA and protein sequence alignments. Entropy calculation was done for each position of the input sequence set. It can be used as a measure of the relative variation in the regions of an aligned gene or protein. The online web server available at the Los Alamos National Laboratories (LANL) (Los Alamos, NM) (<http://www.lanl.gov>) and ([https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy\\_one.html](https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy_one.html)) were used.

### 15.2.5.3 Neutrality Hypothesis

The null hypothesis assumes that in most protein-coding genes, the number of synonymous nucleotide substitutions per site ( $d_S$ ) is equal to the number of non-synonymous nucleotide substitutions per site ( $d_N$ ), i.e., ( $H_0: d_N = d_S$ ), and the alternate hypothesis is ( $H_1: d_N \neq d_S$ ). The probability ( $P$ ) of rejecting the null hypothesis of strict neutrality ( $d_N = d_S$ ) is less than 0.05, considered significant at the 5% level. This may favor the alternative hypothesis. The variance of the difference was computed using the bootstrap method with 500 replicates [64]. Analyses were conducted using the Nei-Gojobori method [65] in MEGA5 [66]. All positions containing gaps and missing data were eliminated from the dataset. Tajima’s test of neutrality [67] was conducted using MEGA5 to compare the number of segregating sites per site with the nucleotide diversity. An essential parameter in the theory of neutral evolution is “ $4Nu$ ,” where “ $N$ ” is the effective population size and “ $u$ ” is the mutation rate per site. The difference in the estimate obtained provides an indication of non-neutral evolution. Codon-based Z-test was also conducted using MEGA5 to

test whether positive selection is operating on a gene by comparing the relative abundance of synonymous and non-synonymous substitutions that have occurred in the gene sequences.

#### **15.2.5.4 Estimates of Transition/Transversion Bias**

Evolutionary analyses were conducted in MEGA5 [66]. Two statistical methods such as maximum likelihood (ML) and maximum composite likelihood (MCL) were used. Substitution pattern and rates were estimated for nucleotide sequences using Tamura-Nei model (+G) [68, 69]. A discrete gamma distribution was used to model evolutionary rate differences among sites (5 categories, [+G], parameter = 0.6115). All positions containing gaps and missing data were eliminated. The analysis involved 37 nucleotide sequences. Codon positions included were 1st + 2nd + 3rd + Noncoding.

#### **15.2.5.5 Disparity Index Test**

A simple statistic (disparity index test) measures the homogeneity of substitution patterns between a pair of sequences [70]. It works by randomly comparing the nucleotide (or amino acid) frequencies of the two descendent sequences and using the number of observed differences between them. MEGA5 computed the disparity index per site, which is given by the total disparity index between two sequences divided by the number of positions compared, excluding gaps and missing data. It is more powerful than a chi-square test of the equality of base frequencies between sequences (<http://www.megasoftware.net/manual.pdf>). The test was performed to infer substitution pattern homogeneity on pairwise nucleotide sequence comparisons. Monte Carlo simulations with 500 replicates were computed. All positions containing gaps and missing data were eliminated.

#### **15.2.5.6 Molecular Relatedness and Phylogenetic Analysis**

Neighbor-joining (NJ) method was used to generate phylogenetic trees. MCL method assumed substitutions included transitions and transversions. Substitutions among lineages were treated as uniform rates, and the pattern among lineages was set homogeneous. All positions containing gaps and missing data were deleted. Codon positions included were 1st + 2nd + 3rd + Noncoding [69].

The generated sequence alignment files were imported from Clustal X to SplitsTree4 version 4.13.1, built 16 Apr 2013 [71], to compute splits as well as to infer variations among the sequences. The nucleotide dataset (37 sequences, each 1086 bp) had 578 constant sites, 277 non-parsimony-informative sites, and 196 gapped sites (no missing data). Whereas the translated proteins of the nucleotide dataset (37 sequences, each 367 residues) had 157 constant sites, 127 non-



parsimony-informative sites, and 71 gapped sites (no missing data). The distances were computed from characters (nucleotides) with default parameters, and then splits were generated from distances using neighbor-net [72] approach to produce a set of circular splits on the network as defined in split decomposition [73], and the amino acid distances used a neighbor-net variance approach performed using ordinary least squares.

### 15.2.5.7 Confidence in the Resulting Phylogenetic Trees

Confidence in the resulting phylogenetic trees was assessed using bootstrap analysis [64]. One thousand bootstrap replicates were generated to assess the reliability of each edge in the tree as well as the network.

### 15.2.5.8 Phi Test for Recombination

The characters were analyzed using Phi test [74] and found informative sites using window size of 100.

## 15.2.6 Nucleotide Sequence Accession Numbers

Sequences were submitted to GenBank and the Los Alamos National Laboratory HIV sequence database, and the accession numbers obtained are KT984127–KT984163 (Table 15.3).

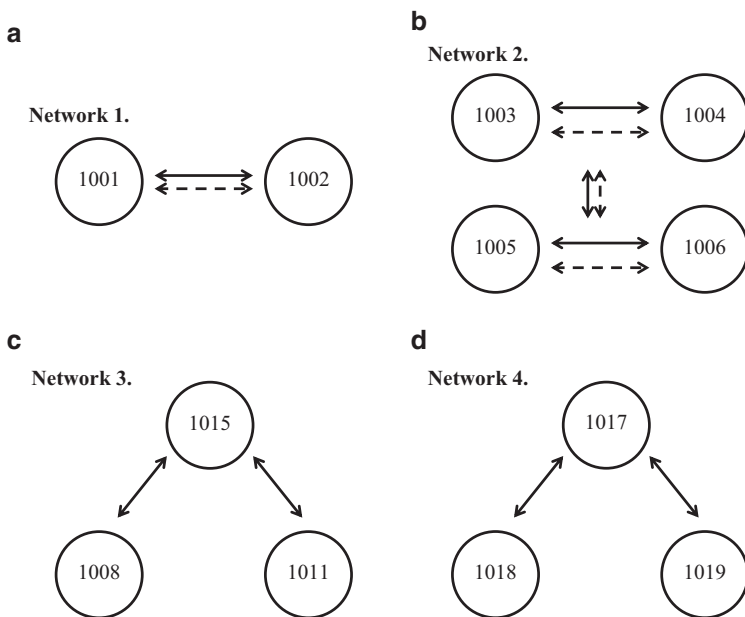
## 15.3 Results

### 15.3.1 Socioepidemiology

The four IDU networks consisted of (i) a male/female dyad (subjects 1002/1001) with sexual relations and shared IDU habits for 4 years; (ii) two separate male/female dyads each of whom maintained sexual relationships for 13 years (subjects 1004/1003) and 8 years (subjects 1006/1005) and all of whom have shared IDU habits for 10 years; (iii) a female triad (subjects 1008, 1011, and 1015) with shared IDU habits for over 20 years; and (iv) a familial triad (subjects 1017, 1018, and 1019) with shared IDU habits. Shared IDU habits would include common use of needle-syringes, cookers, cottons, and rinse water. All individuals in these networks were located in Dade County, Florida. Subjects in networks 1 and 2 resided in Overtown (Dade County, Florida) that was a separate locale from those in network 3 who lived in Opa-Locka (Dade County, Florida) (Tables 15.1 and 15.2, Fig. 15.1).

**Table 15.3** HIV-1B ENV clones (37 sequences)

Patients	Clone identification	NCBI accession numbers
1001	1001L6D	KT984127
	1001L7D	KT984128
	1001L8D	KT984129
1002	1002L1D	KT984130
	1002L2D	KT984131
	1002L3D	KT984132
	1002L5D	KT984133
	1002L7D	KT984134
1003	1003L4D	KT984135
	1003L8D	KT984136
1004	1004L1D	KT984137
	1004L2D	KT984138
	1004L4D	KT984139
	1004L5D	KT984140
	1004_12L1D	KT984142
	1004_12L4D	KT984143
	1004_12L8D	KT984144
	1004_12L9D	KT984145
	1004_12L11D	KT984141
	1005	1005L1D
1005L2D		KT984147
1005L3D		KT984148
1005L4D		KT984149
1006	1006L1R	KT984152
	1006L2R	KT984153
	1006L3R	KT984154
	1006L4R	KT984155
	1006L5R	KT984156
	1006L6R	KT984157
	1006L7R	KT984158
	1006L8R	KT984159
	1006_L12R	KT984150
1006_L13R	KT984151	
1008	1008L4D	KT984160
	1008L5D	KT984161
1015	1015_12L6D	KT984163
	1015_12L12	KT984162



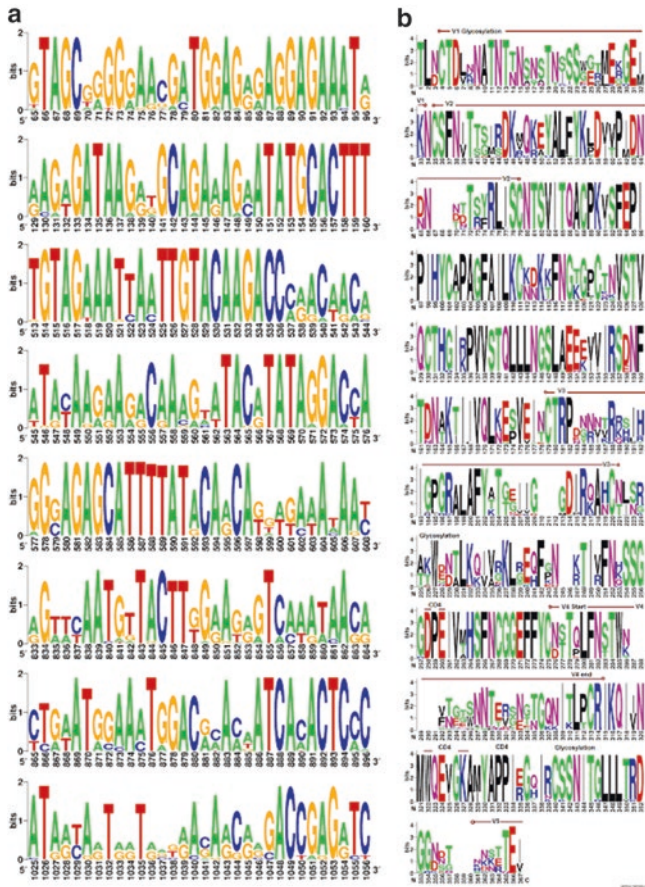
**Fig. 15.1 (a–d)** Four socioepidemiological networks. Refer to Tables 15.1 and 15.2 for additional information. IDU = injection drug use with sharing needle-syringes, cottons, cookers, and washwaters. Solid arrow = IDU. Dotted arrow = sex

### 15.3.2 DNA and Protein Alignments

Nucleotide variations were analyzed based on sequence alignments and are presented in Fig. 15.2a. The transition/transversion bias was also estimated (Table 15.4). The protein translations of the variable regions (V1–V5) showed lesser variation in the V1 loop in comparison with the published global sequence variations of HIV-1 B isolates of blood and brain [75], whereas V2 loop is variable but not hypervariable. However, V3 is hypervariable especially between 182 and 190 and after the starting residues CTRP (Fig. 15.2b). The variable region (V4) is less variable compared to blood env proteins [75].

The overall amino acid variation in V1–V4 is lesser in comparison with blood isolates, and the variations are moderate in brain isolates. The differences in the variability of V1–V4 may be due to the accessibility of loops that may reflect mutations to escape from the immune system as well as in vivo variation in biological properties, such as tropism for macrophages or other cell types or ability to form syncytia [76, 77]. The analyzed dataset had a partial V5 start region. Hence, gp41 start, fusion peptide, and immunodominant regions were not available for further comparison.

It was observed that the residues in the start and end of the variable regions (V1–V4) are well conserved (Fig. 15.2b). There were no differences observed in the



**Fig. 15.2** (a) Env coding genes V1–V4 of HIV-1B IDU patients; (b) The translated env proteins show sequence conservations (*tall characters*), semi-conserved substitutions (*stacked with similar colored characters*), and variations (*stacked with different colors*). The annotation is based on the benchmark HXB2 sequence (accession K03455) provided in HIV Sequence Compendium [90]

**Table 15.4** Maximum composite likelihood estimate of the pattern of nucleotide substitution

	A	T	C	G
A	–	<i>3.43</i>	<i>2.34</i>	<b>14.53</b>
T	<i>5.37</i>	–	<b>12.11</b>	<i>2.81</i>
C	<i>5.37</i>	<b>17.73</b>	–	<i>2.81</i>
G	<b>27.73</b>	<i>3.43</i>	<i>2.34</i>	–

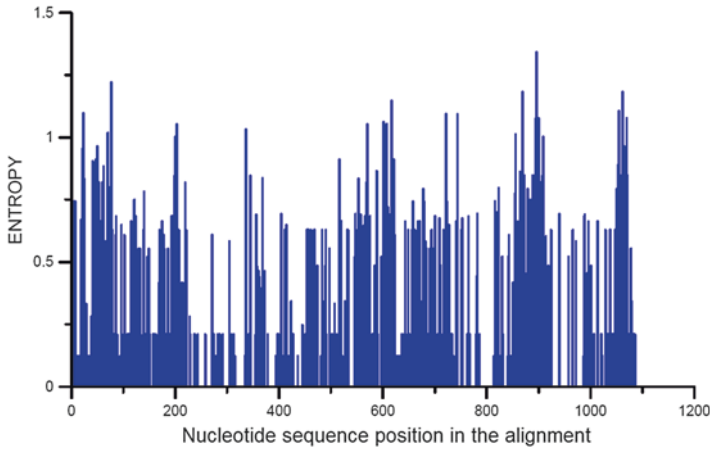
Note: Each entry shows the probability of substitution ( $r$ ) from one base (row) to another base (column [69]). For simplicity, the sum of  $r$  values is made equal to 100. Rates of different transitional substitutions are shown in **bold**, and those of transversional substitutions are shown in *italics*. The nucleotide frequencies are 38.48% (A), 24.57% (T/U), 20.17% (C), and 16.78% (G). The transition/transversion rate ratios are  $k_1 = 5.165$  (purines) and  $k_2 = 5.172$  (pyrimidines). The overall transition/transversion bias is  $R = 2.532$ , where  $R = [A * G * k_1 + T * C * k_2] / [(A + G) * (T + C)]$

glycosylation patterns and CD4 regions of the analyzed dataset as in blood- and brain-derived Env sequences of HIV-1, as it recognizes the same CD4 receptor in all strains of HIV. The CD4 located between V3 and V4 is sequentially flanked by well-conserved residues from positions ranging between 254 and 274, except a single residue variation at 263. The most conserved residues throughout the sequences (V1–V4) are A, D, E, F, G, K, L, M, N, P, R, and W.

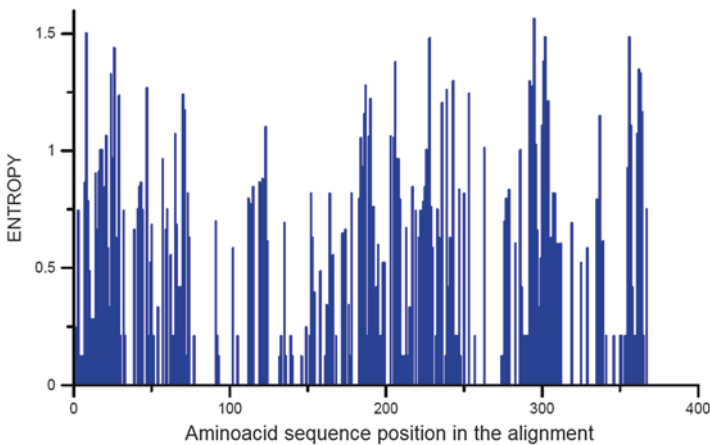
Intra- and inter-sample variations were compared using protein sequence alignment. The intra-clonal variations of patient (1005) shows a variation at the 93rd position for the clone 1005L2D (instead of “F” it is “S”). This variation occurs only in this clone. The clones (L4D and L5D) of the subject 1008, at position 120, “S” were found instead of “K,” whereas in the following patients 1004 (L1D, L2D, 12L1D, and 12L11D) and 1006 (L5R, L6R, L7R, and L8R), it is “T;” and the rest it is “K.” In the subject (1004) at position 125, “T” is replaced with “Q” in L1D, L2D, 12L4D, and 12L9D, whereas in L4D, L5D, 12L1D, and 12L8D, it is replaced by “R.” At position 155, the subjects 1001, 1002, and 1005 had “K,” whereas the rest (1003, 1004, 1006, 1008, and 1015) it was replaced by “E,” except two clones of 1004 (L4D and L5D), where “E” is replaced with “G.” At position 186, the subject 1004 (all clones) had “Y,” whereas the subjects 1001, 1002, and 1005 had “G,” and in others 1003, 1006, 1008, and 1015, it is replaced by “N.” At positions 200–210, there were some insertions and deletions only for 1004, and for all others it is absent. There are some sequence-specific signatures between 290 and 310; for 1001, 1002, and 1005, it is “FNGTWNNTERSNT”; for 1003 it is “NNNTWNSPNRLNS”; for 1008 it is “ST SINANNT EGNE”; for 1004 it is “VTGESNNTVGNNG” except for 12L1D and 12L11D (“GTEMSVENDT” and “FTRESNNTVGNNGT”); for 1006 it is “VTEGSNNTEGN”; for 1015 it is “WSLNGTNTTTNTNE.” These unique subject-specific signatures can drive diversity at the molecular level.

### 15.3.3 Entropy Analysis

The sequence variations and conservations are analyzed using entropy plots for DNA (Fig. 15.3) and protein (Fig. 15.4). The entropy values  $>1$  were observed at the following positions 23, 64, 70, 77, 200, 203, 336, 570, 601, 607, 608, 617, 721, 722, 744, 856, 869, 895–897, 900, 909, 1055, 1062, 1063, and 1070. The entropy values at positions 23, 200, and 203 were due to gapped alignment instead of variations. Similarly for the protein sequence alignment, the entropy values  $>1$  were observed at following positions: 8 (N), 17 (N), 18 (S), 21 (I), 24 (W), 26 (R), 29 (K), 47 (M), 65, 70 (N), 71 (D), 123 (T), 184 (G), 186 (V), 187 (V), 189 (R), 190 (H), 203 (A), 205 (T), 206 (G), 226 (T), 228 (E), 236 (G), 239 (G), 243 (P), 253 (K), 263 (M), 286 (K), 292 (F), 294 (G), 295 (T), 296 (W), 300 (E), 301 (R), 302 (S), 304 (T), 337 (H), 356 (D), 357 (T), 361, 362 (N), 363 (K), and 364 (T). The entropy values at positions 65 and 361 were due to gapped alignment [63].



**Fig. 15.3** Entropy plot of Env *nucleotide* sequence alignment. The plot is generated by comparing residue positions of the first sequence of the input (L6D\_1001) with the rest



**Fig. 15.4** Entropy plot of Env *protein* sequence alignment. The plot is generated by comparing residue positions of the first sequence of the input (L6D\_1001) with the rest

### 15.3.4 Tajima's Neutrality Test

The analysis involved 37 nucleotide sequences ( $m$ ), and the number of segregating sites ( $S$ ) is 32, the nucleotide diversity ( $\pi = 0.097$ ), and the Tajima's test statistic,  $D = 0.593$ . The positive test statistic reflects intermediate-frequency mutations, suggesting diversifying selection [78]. The HIV-1 sequences have not revealed evidence for natural selection in *env* [79]. Although the test assumes that nucleotides are equally mutable, it is not true for coding regions because the polymorphism is

not same for first, second, and third codon positions, and codon usage biases may further complicate the mutation pattern. The difference in the estimate obtained provides an indication of non-neutral evolution. Since Tajima's test is not very powerful and DNA polymorphisms are largely synonymous, it should be verified further with experimental work [80].

### ***15.3.5 Codon-Based Z-Test of Selection***

Only two random pair of clones (L2D 1005 vs L5D 1002) and (L2D 1004 vs L1D 1004) had probability (P)-value less than 0.05, i.e., 0.04 and 0.01, respectively, agreeing the null hypothesis. The rest of the sequence pairs suggest rejecting the null hypothesis of neutrality and preferring to accept the alternate hypothesis.

### ***15.3.6 Estimates of Transition/Transversion Bias***

There was a total of 890 transition/transversion positions in 37 nucleotide sequences. The estimated value of the shape parameter for the discrete gamma distribution is 0.6115. Mean evolutionary rates in these categories were 0.04, 0.21, 0.54, 1.16, and 3.05 substitutions per site. The nucleotide frequencies are  $A = 38.48\%$ ,  $T/U = 24.57\%$ ,  $C = 16.78\%$ , and  $G = 20.17\%$ . The maximum log likelihood for this computation was  $-3618.371$ . The ML-estimated transition/transversion bias ( $R$ ) is 1.97, and MCL-estimated transition/transversion bias is  $R = 2.532$  (Table 15.4).

### ***15.3.7 Disparity Index Analysis***

The substitution pattern between lineages was calculated by assuming that the sequences have evolved with the same evolutionary pattern of nucleotide substitution. The following sequences had disparity index ( $ID > 0$ ) indicated evolutionary divergence between sequences based on composition bias. Therefore, we can reject the null hypothesis at the 5% level. The sequences of patients 1004 and 1006 (the clones are L1D 1004, L2D 1004, 12L1D 1004, 12L11D 1004, L5R 1006, L6R 1006, L3R 1006, L13R 1006, L4R 1006, L12R 1006, and L1R 1006.) had composition bias ( $ID = 0.1$ ) in random comparison with patient 1005 (clone: L2D 1005), no bias with others. The sequences of patients 1001, 1002, and 1005 (their corresponding clones L8D 1001, L5D 1002, L3D 1002, L7D 1002, L2D 1002, L1D 1002, L7D 1001, L6D 1001, L4D 1005, L1D 1005, and L2D 1005) showed composition bias ( $ID = 0.1$ ) with the patient 1005 (clone 12L6D 1015). A different clone of patient 1005 (12L12D 1015) showed composition bias ( $ID = 0.1$ ) with the following sequences of patients 1001, 1002, and 1005 (their corresponding clones are L8D

1001, L5D 1002, L3D 1002, L7D 1002, L2D 1002, L1D 1002, L7D 1001, L6D 1001, L4D 1005, and L3D 1005). It is to be noted that for a couple of clonal sequences (L1D 1005 and L2D 1005)  $ID = 0.2$ , when compared with a patient clone 1015 (12L12D 1015). Thus, the test identified lineages and genes that are evolving with substantially different evolutionary processes as reflected in the atypical patterns of change [70].

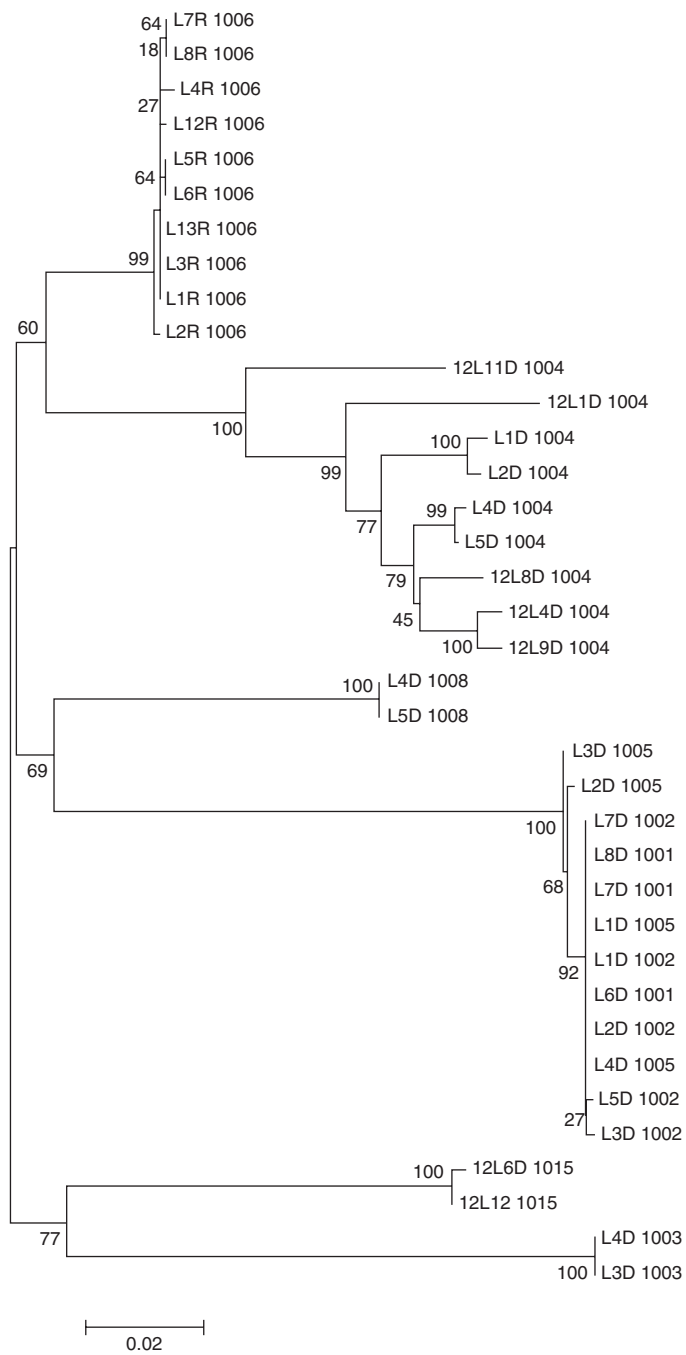
### 15.3.8 *Molecular Relatedness and Dendrogram*

Dendrograms are suitable to display and infer evolutionary model assuming mutations and speciation events. A consensus tree is displayed after 1000 bootstrap replications (Fig. 15.5). However, it is well known that for some complex evolutionary scenarios involving gene loss, duplication, hybridization, horizontal gene transfer, or recombination, a dendrogram is not suitable for an appropriate representation of evolutionary events. Hence, the incompatible and ambiguous signals in the dataset (such as socioepidemiological data) were represented by split networks that provide only an “implicit” representation of evolutionary history [71]. The estimated proportion of invariant sites of nucleotides is 0.334 and for proteins is 0.427 [81].

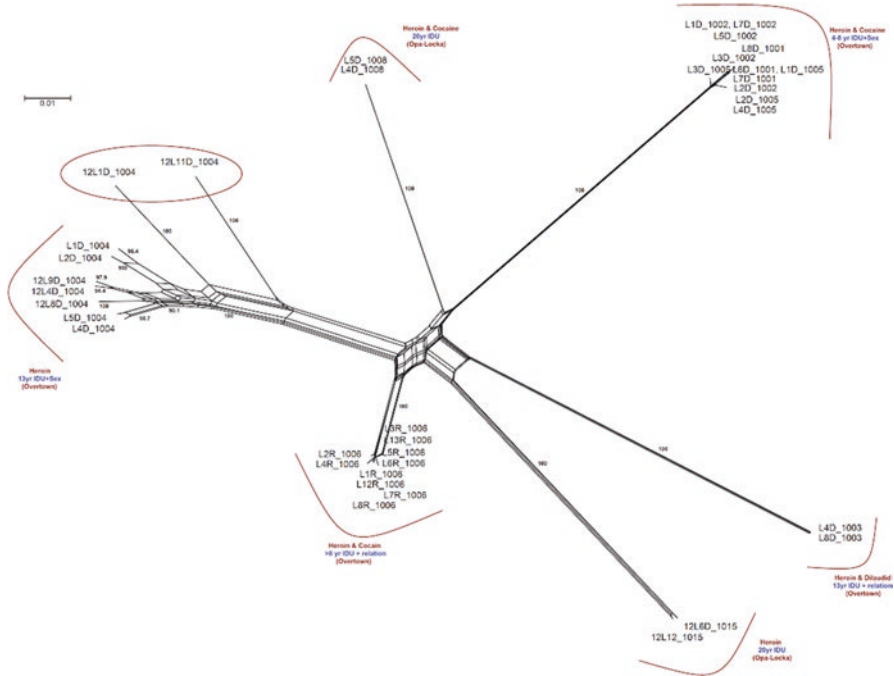
As depicted in Fig. 15.1 (network 1), the IDU and sexual relationship between 1001 and 1002 agrees with the molecular connections as represented in the dendrogram with 92% confidence on the branch. The network 2 of the same figure is inconsistent with the molecular data, having 1003 and 1004 in two separate clusters with 100% confidence on the branch. Whereas the other subjects (1005 and 1006) were also distributed into two different clusters, 1005 is clustered along with 1001 at a confidence of 92%, and 1006 is entirely a unique cluster with 99% confidence. The network 3 is also in disagreement with the molecular data of 1015 and 1008. However, the sequence details of 1011 are required to confirm with 1015. (No molecular data is available to compare with network 4.)

The resulting nucleotide-based split network is showed in Fig. 15.6, and the protein-based split network is showed in Fig. 15.7. This pattern suggests that the dataset contains conflicting evolutionary signals (such as duration of IDUs, sexual relationship, hypervariable regions, recombination, risk locale, random genetic drift, etc.) and is consistent with the hypothesis of recombination events (refer, vi. Phi test) among the major lineages. Two of clones belonging to IDU patient (1004) appear isolated from the clusters, whereas few others clones of 1004 show sequence admixtures, preferably due to recombination events. Compared to all other subjects, 1004 cluster is more diverse and some clones converge. It is to be noted that split networks provide only an “implicit” picture of an evolutionary relationship and “nodes” in the “split network” do not represent ancestral species.





**Fig. 15.5** Evolutionary relationships of HIV-1 B Env based on nucleotide alignment (37 sequences). The branching pattern was generated by the neighbor-joining (NJ) method, and the confidence of the clades was assessed by bootstrap values ( $n = 1000$  replicates)



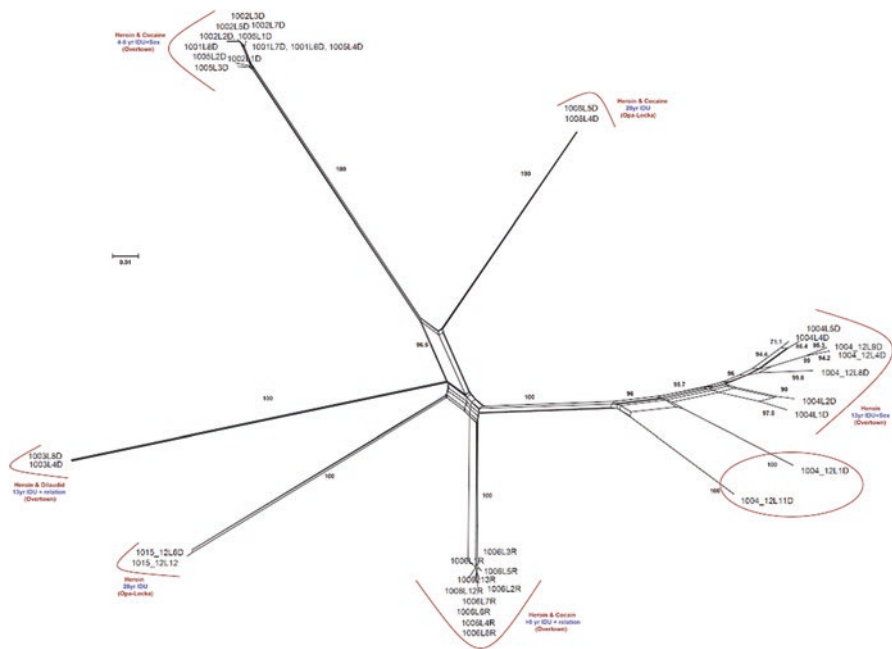
**Fig. 15.6** A split decomposition network for 37 Env sequences of IDU HIV-1 B patients based on *nucleotide* sequence alignment, computed using the neighbor-net method with bootstrap values ( $n = 1000$  replicates). The computed splits are displayed as a network with equal angle. Two of the IDU patient sequences appear isolated from the clusters and are encircled

### 15.3.9 Phi Test for Recombination

The nucleotide and amino acid characters were analyzed using Phi test [74]. The results are summarized in Table 15.5. Although the split network based on nucleotides and proteins had a similar tree topology (Figs. 15.6 and 15.7), only the network based on nucleotides showed significant evidence for recombination in the Phi test, which was not reflected in proteins.

## 15.4 Discussion

We address the question as to whether clustering observed in IDU networks reported in this chapter may reflect relatively concentrated shared co-injection behaviors. The dyads that we characterized injected with each other feasibly hundreds of times between 1982 and 1988, by which time all had seroconverted. The participants in the double dyad, additionally, had injected with each member of the group many



**Fig. 15.7** A split decomposition network for 37 Env sequences of IDU HIV-1 B patients based on protein sequence alignment, computed using the neighbor-net method with bootstrap values ( $n = 1000$  replicates). The computed splits are displayed as a network with equal angle. Two of the IDU patient sequences appear isolated from the clusters and are encircled

times. There is no evidence of direct interpersonal connections between participants 1005 and 1002 and 1001. It is relevant to note that the carrier of a founding HIV-1B infection need not be personally present when the next victim is exposed. That individual may no longer be on the premises; just his/her paraphernalia need be present for additional individuals to become HIV-1B infected with the strains at hand. These paraphernalia include contaminated needle-syringes, cottons, cookers, and washwaters, all generally used in the IDU venue. IDU paraphernalia have been extensively characterized in Miami-Dade County and shown to be contaminated with HIV-1B [42, 82, 83].

One observed condition in which this kind of exposure takes place is when a group of injectors arrives at a risk locale and uses needle-syringes and paraphernalia contaminated by the people who just finished ‘shooting up’ (injecting drugs). They use the contaminated available paraphernalia and syringes, not necessarily sharing them, but often not adequately cleansing them of prior contaminants. The primary hypothesis to be tested by sequencing ENV variable loops is, if we can expect sequence characteristics of strains of virus will reflect interpersonal risk behaviors characterized by sociocultural studies. In cases where people who run together (socialize) also inject together under private circumstances, we would expect clustering to reflect that fact. In cases where people who inject together do so in risk locales

**Table 15.5** Phi test for recombination in the split network

Sl. No	Subject ID	Informative sites	$k$	Mean	Variance	Observed	$P$ -value	Statistical significance
1.	1001–1006, 1008, and 1015 (genes)	360	33	0.899	6.703	0.080	1.190	Yes
2.	1003–1006 (genes)	286	27	0.692	7.459	0.056	7.143	Yes
3.	1003 and 1004 (genes)	192	18	0.053	1.160	0.039	1.695	Yes
4.	1005 and 1006 (genes)	116	11	–	–	–	–	No
5.	1008 and 1015 (genes)	131	13	–	–	–	–	No
6.	1001–1006, 1008, and 1015 (proteins)	168	46	0.090	2.008	0.090	0.497	No
7.	1001 and 1002 (proteins)	Non-informative	–	–	–	–	–	–
8.	1003 and 1004 (proteins)	105	30	0.029	7.989	0.030	0.595	No
9.	1005–1006 (proteins)	72	21	–	–	–	–	No
10.	1008 and 1015 (proteins)	81	23	–	–	–	–	No

where many inject, we would expect lower levels of clustering of virus strains. Additional study in more localities would aid in understanding the importance of shooting companions and shooting venues in the spread of HIV-1B (as well as additional viruses including HCV and HBV). Moreover, the extent to which individuals participate in out-of-network HIV-1B risk activity needs socioepidemiological comparisons and sequence relatedness characterization as well [42, 47, 48, 49, 84–89].

The full V1–V5 ENV domains are included in this study to characterize social effects and molecular changes. The structure and genomes of HIV strains are well known and on a firm basis for such studies [90]. For example, previous studies suggest that the overall tertiary conformation of the entire env protein may be important in tropism-determining activity. In particular, domains within the V1/V2 [91, 92] and/or V3 are critical determinants of macrophage cellular tropism [12, 25, 85, 91, 93–95]. A 94-amino acid domain, including the V3 loop, is involved with HIV-1B infection of macrophages being both necessary and sufficient for virus

entry [91, 93]. A mutation of residue 287 from a lysine to a glutamic acid converts a non-macrophagetropic isolate of HIV-1B to one capable of replicating in macrophages. In addition, the V3 domain is a 35–37-amino acid loop bounded by a pair of disulfide-bonded cysteine residues [96]. It forms two antiparallel beta turns and a short C-terminal alpha helix [97]. It is an epitope for neutralizing antibodies as well as cytotoxic T lymphocytes [98, 99].

Based on sequence analysis, specific signatures, transition/transversion bias, statistical test of neutrality, and molecular diversity as reflected in dendrograms, we conclude that there are genetic variations in V1–V4 region of HIV-1 B env. For example, disparity index confirmed composition bias in a couple of clones (12L1D\_1004 and 12L11D\_1004) belonging to 1004 with patient 1005 (clone: L2D 1005). This is very well reflected as isolated branches of split network. Similarly, the sequences of patients 1004 and 1006 had composition bias with patient 1005, and the sequences of patients 1001, 1002, and 1005 showed composition bias with the clone 12L6D of 1015 as reflected in the split network (Figs. 15.6 and 15.7), which confirms non-neutral evolution as indicated by Tajima's test. The unsystematic variations introduced by recombination may set an evidence for non-neutral evolution. It is also known that the natural selection is frequently masked by recombination and the natural selection over the env V1–V4 region had a minor role in driving diversity [76, 100, 101].

Finally, there is a non-intuitive and unexpected relationship between needle-syringe and paraphernalia sharing and psychiatric morbidities associated with IDU risk behavior. In the current volume of *Global Virology II*, Thames and Jones (Chap. 12) indicate that IDU needle-syringe and paraphernalia-sharing as well as reduced needle-syringe cleansing behaviors are associated with psychiatric comorbidities due to HIV-1 infection. These comorbidities include antisocial personality disorder (ASPD) and exhibit such as remorseless, impulsive, and irresponsible behaviors. Thames and Jones further report and discuss additional comorbidities associated with such behaviors including DSM-III axis II diagnosis as well as opioid and cocaine consumption. [102]. Additionally, marijuana use is also associated with such set of behaviors and conditions. [103].

## 15.5 Conclusions

The results described here may lead to new directions of understanding natural selection, random genetic drift, and recombination in the HIV-1B env protein as well as diversity during HIV-1 infections in a defined socioepidemiological context. Additional work is needed to characterize in detail the effects of differing risk activities including contemporaneously in the post-HAART, cART era as the world progresses into the next era of more advanced molecular and immunological therapies. Moreover, the contextual application of studies of IDU risk behaviors and molecular epidemiology should also include characterization of the associated psychiatric morbidities as well as the possible role of brain-related HIV-1 infections.

**Acknowledgments** Support for this work included NIH Grants from NIDA (NIH, Bethesda, MD) (PS, DS, CMcC, and JBP).

**Conflict of interest** The authors report no conflicts of interest.

## References

1. Levy JA, Shimobokuro J (1985) Recovery of AIDS associated retroviruses from patients with AIDS or ARC and from clinically healthy individuals. *J Infect Dis* 152:734–738
2. Montagnier L, Gruest J, Chamaret S, Dauguet C, Axler C, Guetard D, Nugeyre MT, Barre-Sinoussi F, Chermann JC, Brunet JB, Klatzman D, Gluckman JC (1984) Adaptation of lymphadenopathy associated virus (LAV) to replication in EBV-transformed b lymphoblastoid cell lines. *Science* 225(4657):63–66
3. Goedert JJ, Gallo RC (1985) Epidemiological evidence that HTLV-III is the AIDS agent. *Eur J Epidemiol* 1(3):155–159
4. Barre-Sinoussi F, Mathur-Wagh U, Rey F, Brun-Vezinet F, Yancovitz DR, Rouzioux C, Montagnier L, Mildvan D, Chermann JC (1985) Isolation of lymphadenopathy-associated virus (LAV) and detection of LAV antibodies from U.S. patients with AIDS. *JAMA* 253(12):1737–1739
5. Junqueira DM, de Medeiros RM, Matte MC, Araújo LA, Chies JA, Ashton-Prolla P, Almeida SE (2011) Reviewing the history of HIV-1: spread of subtype B in the Americas. *PLoS One* 6(11):e27489. doi:10.1371/journal.pone.0027489. PMID: 22132104; PubMed Central PMCID: PMC3223166
6. Wolfs TFW, Zwart G, Bakker M, Goudsmit J (1992) HIV-1 genomic RNA diversification following sexual and parental virus transmission. *Virology* 189:103–110
7. Sabino EC, Shpaer EG, Morgado MG, Korber BT, Diaz RS, Bongertz V, Cavalcante S, Galvão-Castro B, Mullins JI, Mayer A (1994) Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals from Brazil. *J Virol* 68(10):6340–6346
8. Wolinsky SM, Wike CM, Korber B, Hutto C, Parks WP, Rosenblum LL, Kunstman KJ, Furtado MR, Munoz J (1992) Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants. *Science* 255:1134–1137
9. Zhu T, Mo H, Wang N, Nam DS, Cao Y, Koup RA, Ho DD (1993) Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. *Science* 261:1179–1181
10. Holmes EC, Zhang LQ, Simmonds P, Ludlam CA, Brown AJL (1992) Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. *Proc Natl Acad Sci U S A* 89:4835–4839
11. Holmes EC, Zhang LQ, Simmonds P, Rogers AS, Brown AJL (1993) Molecular investigation of human immunodeficiency virus (HIV) infection in a patient of an HIV-infected surgeon. *J Infect Dis* 167:1411–1414
12. Chesebro B, Wehrly K, Nishio J, Perryman S (1992) Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell tropic isolates: definition of critical amino acids involved in cell tropism. *J Virol* 66:6547–6554
13. Myers G, Berzofsky JA, Korber B, Smith RF, Pavlakis GN (1992) Human retroviruses and AIDS: a compilation of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos
14. Ou C-Y, Cielsielski CA, Myers G, Bandea C, Luo C, Korber BTM, Mullins JI, Schochetman G, Berkelman RL, Economou AN, Witte JJ, Furman LJ, Satten GA, MacInnes KA, Curran IW, Jaffe HW (1992) Molecular epidemiology of HIV transmission in a dental practice. *Science* 256:1156–1171

15. Segal DM, Shapshak P, Zhang BT, Crandall KA, Page JB, Fujimura RF, Goodkin K, Douyon R, McCoy CB (1997) Novel tetrameric tip motifs (APGK and VPGK) in the V3 loop of HIV type 1 envelope sequences in blood and brain from two injection drug users in Miami, Florida. *AIDS Res Hum Retrov* 13:1643–1646
16. Shapshak P, Segal DM, Crandall KA, Fujimura RK, Zhang BT, Xin KQ, Okuda K, Petito C, Eisdorfer C, Goodkin K (1999) Independent evolution of HIV-1 in different brain regions. *AIDS Res Hum Retr* 15:811–820
17. Shapshak P, Minagar A, Kanguene P, Frost S, Kosakovsky-Pond SL, Zarate S, Singer E, Commins D, De Groot A, Quinones-Mateu M, Terwilliger E (2007) Blood and brain HIV strain variation. In: Goodkin K, Shapshak P (eds) *The spectrum of neuroAIDS disorders: pathophysiology, diagnosis, and treatment*. ASM Press, Washington, DC
18. Viscidi RP (1999) HIV evolution and disease progression in longitudinal studies. In: Crandall KA (ed) *The evolution of HIV*. Johns Hopkins University Press, Baltimore, pp 346–389
19. Xin K-Q, Shapshak P, Kawamoto S, Nagano I, McCoy CB, Okuda K (1995) Highly divergent env sequences of HIV-1 B subtype with two novel V3 loop motifs detected in an AIDS patient in Miami. *Florida AIDS Res Human Retrov* 11:1139–1141
20. Modrow S, Hahn BH, Shaw GM, Gallo RC, Wong-Staal F, Wolf H (1987) Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. *J Virol* 61:570–578
21. Saag MS, Hahn BH, Gibbons J (1988) Extensive variation of human immunodeficiency virus type 1 in vivo. *Nature* 334:440–444
22. Lorenzo E, Herrera RJ, Lai S, Fischl MA, Hill MD (1996) The tat and C2-V3 envelope genes in the molecular epidemiology of human immunodeficiency virus-1. *Virology* 221:310–317
23. Crandall KA (1995) Intraspecific phylogenetics: support for the dental transmission of human immunodeficiency virus. *J Virol* 69:2351–2356
24. Salminen M, Nykanen A, Brummer-Korvenkontio H, Kantanen ML, Liitsola K, Leinikki P (1993) Molecular epidemiology of HIV-1 based on phylogenetic analysis of in vivo gag p7/p9 direct sequences. *Virology* 195:185–194
25. Kuiken CL, Zwart G, Baan E, Couthino RA, van den Hoek JAR, Goudsmit J (1993) Increasing antigenic and genetic diversity of the V3 variable domain of the human immunodeficiency virus envelope protein in the course of the AIDS epidemic. *Proc Natl Acad Sci U S A* 90:9061–9065
26. Wolfs TFW, Nara PL, Goudsmit J (1993) Genotypic and phenotypic variation of HIV-1: impact on AIDS pathogenesis and vaccination. *Chem Immunol* 56:1–33
27. Sundaravaradan V, Hahn T, Ahmad N (2005) Conservation of functional domains and limited heterogeneity of HIV-1 reverse transcriptase gene following vertical transmission. *Retrovirology* 2:36. doi:[10.1186/1742-4690-2-36](https://doi.org/10.1186/1742-4690-2-36)
28. Ou CY, Takebe Y, Luo CC, Kalish M, Auwanit W, Bandea C, de la Torre N, Moore JL, Schochetman G, Yamazaki S (1992) Wide distribution of two subtypes of HIV-1 in Thailand. *AIDS Res Hum Retroviruses* 8(8):1471–1472
29. Janini LM, Tanuri A, Schechter M, Peralta JM, Vicente ACP, Dela Torre N, Pieniazek NJ, Luo C-c, Ramos A, Soriano V, Schochetman G, Rayfield MA, Pieniazek D (1998) Horizontal and vertical transmission of human immunodeficiency virus type 1 dual infections caused by viruses of subtypes B and C. *JID* 177:227–231
30. Chitwood DD, Griffin DK, Comerford M, Page JB, Trapido EJ, Lai S, McCoy CB (1995) Risk factors for HIV-1 seroconversion among injection drug users: a case-control study. *Am J Public Health* 85(11):1538–1542
31. Allen DM, Onorato IM, Green TA (1992) HIV infection in intravenous drug users entering drug treatment, United States, 1988 to 1989. *Am J Public Health* 82:541–546
32. McCoy CB, Metsch LR, Chitwood DD, Shapshak P, Comerford ST (1998) Parenteral transmission of HIV among injection drug users: assessing the frequency of multi-person use of needles, syringes, cookers, cotton, and water. *J Acquir Immune Defic Syndr Hum Retrovirol* 18:S25–S29

33. McCoy CB, Shapshak P, Rivers JE, McCoy HV, Weatherby NL, Chitwood DD, Shah SM, Inciardi JA, McBride DC, Watters JK (1994) The risks of injection practices and compliance to risk reduction protocols in the prevention of HIV-1 among injecting drug users. *J AIDS* 7:773–776
34. McCoy CB, Shapshak P, Shah SM, McCoy HV, Rivers JE, Page JB, Chitwood DD, Weatherby NL, Inciardi JA, McBride DC, Mash DC, Watters JK (1994) HIV-1 prevention: interdisciplinary studies and review on efficacy of bleach and compliance to bleach prevention protocols. National Research Council and Institute of Medicine, National Academy Press, Washington, DC, pp 255–283
35. McCoy Clyde B, Metsch LR, Collado-Mesa F, Arheart KL, Messiah SE, Katz D, Shapshak P (2004) The prevalence of HCV and HIV among injection drug users who use high risk inner-city locales in Miami, Florida, vol 99. *Memórias do Instituto Oswaldo Cruz, Rio de Janeiro*, pp 789–793
36. McCoy HV, Chitwood DD, Page JB, McCoy CB (1997) Skills for HIV risk reduction: evaluation of recall and performance in injecting drug users. *Subst Use Misuse* 32(3):229–247
37. McCoy CB, Metsch LR, Page JB, McBride DC, Comerford S (1997) Injection drug users' practices and attitudes toward intervention and potential for reducing the transmission of HIV. *Med Anthropol* 18:35–60
38. McCoy C, Inciardi J (1995) Sex, drugs, and the continuing spread of AIDS. Roxbury Publishing Company, Los Angeles
39. McCoy C, Metsch L, Inciardi J, Anwyl R, Wingerd J, Bletzer K (1996) Sex, drugs, and the spread of HIV/AIDS in belle glade, florida. *Med Anthropol Q* 10:83–93
40. McCoy CB, Shapshak P, Metsch LR, McCoy HV, Rivers JE, Weatherby NL, Shah SM, Chitwood DD (1995) HIV-1 prevention: interdisciplinary studies on the efficacy of bleach and development of prevention protocols. *Archivum Immunologie et Therapiae Experimentalis* 43:1–9
41. Needle RH, Coyle SL, Genser SG, Trotter RT (1995) Social networks, drug abuse, and HIV transmission. National Institute on Drug Abuse, NIDA Research Monograph, Rockville, p 151
42. Page JB, Shapshak P, Duran EM, Even G, Moleon-Borodowski I, Llanusa-Cestero R (2006) Detection of HIV-1 in injection paraphernalia: risk in an era of heightened awareness. *AIDS Patient Care STDS* 20(8):576–585
43. Page JB, Smith PC, Kane N (1990) Shooting galleries, their proprietors, and implications for prevention of AIDS. *J Drug Issues* 5(1/2):69–85
44. Page JB, Chitwood DD, Smith PC, Kane N, McBride DC (1990) Intravenous drug abuse and HIV infection in Miami. *Med Anthropol Q* 4(1):56–71
45. Page JB, Salazar JM (1999) Needle use in Miami and Valencia: observations of high and low availability. *Med Anthropol Q* 13(4):413–435
46. Page JB (1990) Shooting scenarios and risk of HIV-1 infection. *Am Behav Sci* 33(4):478–490
47. Page JB (2002) Networks of informal social relations and HIV: problems in characterization and analysis. *Bioinformatics and cybernetics: from gene expression to social interaction*. Sixth World conference on systemics, cybernetics, and informatics, Orlando, 14–18 July 2002
48. Page JB, Singer M (2010) *Comprehending drug abuse: ethnographic research at the social margins*. Rutgers University Press, New Brunswick
49. Rothenberg R (2009) HIV transmission networks. *Epidemiology*: edited by Tim Mastro and Quarraisha Abdool-Karim. *Current Opinion in HIV & AIDS* 4(4):260–265. doi:[10.1097/COH.0b013e32832c7cfc](https://doi.org/10.1097/COH.0b013e32832c7cfc)
50. Salemi M, de Oliveira T, Ciccozzi M, Rezza G, Goodenow MM (2008) High-resolution molecular epidemiology and evolutionary history of HIV-1 subtypes in Albania. *PLoS One* 3(1):e1390. doi:[10.1371/journal.pone.0001390](https://doi.org/10.1371/journal.pone.0001390)
51. Beyrer C, Patel Z, Stachowiak JA, Tishkova FK, Stibich MA, Eyzaguirre LM, Carr JK, Mogilnii V, Peryshkina A, Latypov A, Strathdee SA (2009) Characterization of the emerging



- HIV type 1 and HCV epidemics among injecting drug users in Dushanbe, Tajikistan. *AIDS Res Hum Retroviruses* 25(9). doi:[10.1089/aid.2008.0206](https://doi.org/10.1089/aid.2008.0206)
52. Pilon R, Leonard L, Kim J, Vallee D, De Rubeis E et al (2011) Transmission patterns of HIV and hepatitis C virus among networks of people who inject drugs. *PLoS One* 6(7):e22245. doi:[10.1371/journal.pone.0022245](https://doi.org/10.1371/journal.pone.0022245)
  53. CDC, Exposure to blood. July, 2003. [http://www.cdc.gov/hai/pdfs/bbp/exp\\_to\\_blood.pdf](http://www.cdc.gov/hai/pdfs/bbp/exp_to_blood.pdf)
  54. Ho DD, Sarngadharan MG, Resnick L, Dimara-Varinese F, Rota TR, Hirsch MS (1985) Primary human t-lymphotropic virus type III infection. *Ann Intern Med* 103:880–883
  55. Neisson-Vernant C, Arfi S, Mathez D, Leibowitch J, Monplaisir N (1986) Needlestick HIV seroconversion in a nurse. *Lancet* 2:814
  56. Oksenhendler E, Horzic M, LeRoux J, Rabian C, Cauvel J (1986) HIV infection with seroconversion after a superficial needlestick injury to the finger. *N Engl J Med* 315:582
  57. Stricof RL, Morse DL (1986) HTLV-III/LAV seroconversion following a deep intramuscular needlestick injury. *N Engl J Med* 314:115
  58. Chitwood DD, Page JB, Comerford M, Inciardi JA, McCoy CB, Trapido EJ, Lai S (1991) The donation and sale of blood by intravenous drug users. *Am J Public Health* 81(5):631–633
  59. Singer M, Jia Z, Schensul JJ, Weeks M, Page JB (1992) AIDS and the IV drug user: the local context in prevention efforts. *Med Anthropol* 14:285–306
  60. Rodrigo AG, Learn GH Jr (2001) Computational and evolutionary analysis of HIV molecular sequences. Kluwer Academic Publishers. ISBN 0-306-46900-6, pp 11–14, 38–40
  61. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25(24):4876–4882
  62. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190
  63. Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana, IL
  64. Efron B, Halloran E, Holmes S (1996) Bootstrap confidence levels for phylogenetic trees. *Proc Natl Acad Sci USA* 93(23):13429–13434
  65. Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3:418–426
  66. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary Genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731–2739
  67. Tajima F (1989) Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595
  68. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512–526
  69. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci* 101:11030–11035
  70. Kumar S, Gadagkar SR (2001) Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. *Genetics* 158(3):1321–1327
  71. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23(2):254–267
  72. Bryant D, Moulton V (2004) NeighborNet: an agglomerative algorithm for the construction of planar phylogenetic networks. *Mol Biol Evol* 21:255–265
  73. Bandelt HJ, Dress AWM (1992) A canonical decomposition theory for metrics on a finite set. *Adv Math* 92:47–105
  74. Bruen T, Philippe H, Bryant D (2005) A quick and robust statistical test to detect the presence of recombination. *Genetics* 172(4):2665–2681
  75. Balaji S, Sneha P, Rama M, Shapshak P (2015) Global protein sequence variation in HIV-1-B isolates derived from human blood and brain. In: Shapshak P et al (eds) *Global virology I – identifying and investigating viral diseases*. Springer Science Business Media New York 2015, New York, pp 613–666

76. Coffin JM (1995) HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* 267:483–489
77. Coffin JM, Hughes SH, Varmus HE (eds) (1997) *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
78. Shriner D, Shankarappa R, Jensen MA, Nickle DC, Mittler JE, Margolick JB, Mullins JI (2004) Influence of random genetic drift on human immunodeficiency virus type 1 env evolution during chronic infection. *Genetics* 166(3):1155–1164
79. Leigh Brown AJ (1997) Analysis of HIV-1 env gene sequences reveals evidence for a low effective number in the viral population. *Proc Natl Acad Sci* 94:1862–1865
80. Nei M, Kumar S (2000) *Molecular evolution and Phylogenetics*. Oxford University Press, New York
81. Lockhart PJ, Steel MA, Huson DH (2000) Invariable site models and their uses in phylogeny reconstruction. *Syst Biol* 49(2):225–232
82. Shah SM, Shapshak P, Rivers JE, Stewart RV, Weatherby NL, Xin K-Q, Page JB, Chitwood DD, Mash DC, Vlahov D, McCoy CB (1996) Detection of HIV-1 DNA in needle/syringes, paraphernalia, and washes from shooting galleries in Miami: a preliminary laboratory report. *JAIDS* 11:301–306
83. Shapshak P, Fujimura RK, Page JB, Segal D, Yang J, Shah SM, Graham G, Rivers JE, Chitwood DD, McCoy CB (2000) HIV-1 RNA load in needle/syringes from shooting galleries in Miami: a preliminary laboratory report. *Drug Alcohol Depend* 58:153–157
84. Amirkhanian Y, Kelly J, Kuznetsova A, Meylakhs A, Yakovlev A, Musatov V, Chaika N (2014) Using social network methods to reach out-of-care or ART-nonadherent HIV+ injection drug users in Russia: addressing a gap in the treatment cascade. *J Int AIDS Soc* 17(4 Suppl 3):19594. doi:[10.7448/IAS.17.4.19594](https://doi.org/10.7448/IAS.17.4.19594)
85. Flaer PJ, Cistone PJ, Younis MZ, Parkash J (2013) A connectivity model for assessment of HIV transmission risk in injection drug users (IDUs). *Eval Program Plann* 39:23–27. doi:[10.1016/j.evalprogplan.2013.03.004](https://doi.org/10.1016/j.evalprogplan.2013.03.004). Epub 2013 Mar 18
86. Harris M, Rhodes T (2013) Injecting practices in sexual partnerships: hepatitis C transmission potentials in a ‘risk equivalence’ framework. *Drug Alcohol Depend* 132(3):617–623. doi:[10.1016/j.drugalcdep.2013.04.012](https://doi.org/10.1016/j.drugalcdep.2013.04.012)
87. Mateu-Gelabert P, Gwadz MV, Guarino H, Sandoval M, Cleland CM, Jordan A, Hagan H, Lune H, Friedman SR (2014) The staying safe intervention: training people who inject drugs in strategies to avoid injection-related HCV and HIV infection. *AIDS Educ Prev* 26(2):144–157. doi:[10.1521/aeap.2014.26.2.144](https://doi.org/10.1521/aeap.2014.26.2.144)
88. West BS, Choo M, El-Bassel N, Gilbert L, Wu E, Kamarulzaman A (2014) Safe havens and rough waters: networks, place, and the navigation of risk among injection drug-using Malaysian fishermen. *Int J Drug Policy* 25(3):575–582. doi:[10.1016/j.drugpo.2013.11.007](https://doi.org/10.1016/j.drugpo.2013.11.007)
89. Young AM, Jonas AB, Mullins UL, Halgin DS, Havens JR (2013) Network structure and the risk for HIV transmission among rural drug users. *AIDS Behav* 17(7):2341–2351. doi:[10.1007/s10461-012-0371-2](https://doi.org/10.1007/s10461-012-0371-2)
90. Foley B, Leitner T, Apetrei C, Hahn B, Mizrachi I, Mullins J, Rambaut A, Wolinsky S, Korber B (eds) *HIV sequence compendium 2013*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, NM, LA-UR 13-26007
91. Westervelt P, Gendelman HE, Ratner L (1991) Identification of a determinant within the HIV-1 surface envelope glycoprotein critical for productive infection of cultured primary monocytes. *Proc Natl Acad Sci USA* 88:3097–3101
92. Westervelt P, Trowbridge DB, Epstein LG, Blumberg BM, Li Y, Hahn BH, Shaw GM, Price RW, Ratner L (1992) Macrophage tropism of determinants of human immunodeficiency virus type 1 in vivo. *J Virol* 66:2577–2582
93. Kuiken CL, De Jong J-J, Baan E, Keulen W, Termsette M, Goudsmit J (1992) Evolution of the V3 envelope domain in proviral sequences and isolates of human immunodeficiency virus type 1 during transition of the viral biological phenotype. *J Virol* 66:4622–4627

94. Wolfs TFW, De Jong J, Van Den Berg H, Tunagel JMGH, Krone WJA, Goudsmit J (1990) Evolution of sequences encoding the principal neutralization epitope of human immunodeficiency virus type 1 is host-dependent, rapid, and continuous. *Proc Natl Acad Sci USA* 87:9938–9942
95. O'Brien WA, Koyanagi Y, Namazie A, Zhao J-Q, Diagne A, Idler K, Zack JA, Chen ISY (1990) HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain. *Nature* 348:69–73
96. Leonard CK, Spellman MW, Riddle L, Harris RJ, Thomas JN, Gregory TJ (1990) Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J Biol Chem* 265:10373–10382
97. LaRosa GJ, Davide JP, Weinhold K, Waterbury JA, Profy AT, Lewis JA, Langlois AJ, Dreesman GR, Boswell BN, Shaddock P, Holley LH, Karplus M, Bolognesi DP, Matthews TJ, Emini EA, Putney SD (1990) Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant. *Science* 149:932–935
98. Jahaverian K, Langlois A, McDanal C, Ross K, Eckler L, Jellis C, Profy A, Rusche J, Bolognesi D, Putney S, Matthews M (1989) Principal neutralizing domain of the human immunodeficiency virus type 1 envelope glycoprotein. *Proc Natl Acad Sci USA* 86:6768–6772
99. Takahashi H, Merli S, Putney SD, Houghten R, Moss B, Germain RN, Berzofsky JA (1989) A single amino acid interchange yields reciprocal CTL specificities for HIV-1 gp160. *Science* 246:118–121
100. Nowak MA, May RM, Phillips RE, Rowland-Jones S, Lalloo DG et al (1995) Antigenic oscillations and shifting immuno-dominance in HIV-1 infections. *Nature* 375:606–611
101. Nowak MA, Anderson RM, Boerlijst MC, Bonhoeffer S, May RM et al (1996) HIV-1 evolution and disease progression. *Science* 274:1008–1011
102. Thames A, Jones J (2017) In: Shapshak et al (eds) *Psychiatric comorbidities in HIV-infection, in HIV and NeuroAIDS*. Springer, New York
103. Haase JM (2009) Co-occurring antisocial personality disorder and substance use disorder: treatment interventions. *Grad J Couns Psychol* 1(2):55–64. <http://epublications.marquette.edu/gjcp/vol1/iss2/6>