#234

Searching for Rare HIV Strains in Rural Democratic Republic of Congo (2001-2003)

Vallari A[1]*, McArthur C[2], Sthreshley L[3], Brennan CA[1]

[1]Infectious Disease Research, Abbott Diagnostic, Abbott Park IL, [2]School of Dentistry, University of Missouri-Kansas City, Kansas City, MO, [3]Presbyterian Church (USA), Kinshasa, DRC

Contact Information:
Ana Vallari
Abbott Diagnostics
100 Abbott Park Rd
Dept 09NG/Building AP20
Abbott Park IL 60064-6015
Email: ana.vallari@abbott.com

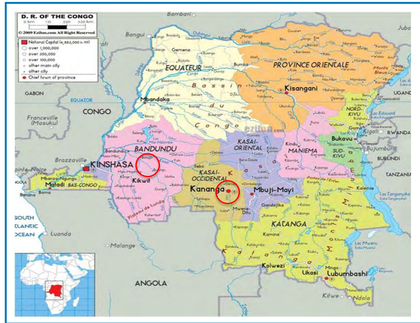
Abstract

Background: Democratic Republic of Congo (DRC) is thought to be the epicenter of the HIV pandemic. Not only were the oldest documented HIV-1 infections found in Kinshasa (circa 1959-1960) but the many diverse HIV-1 strains circulating in the DRC show high levels of intrasubtype diversity and intersubtype recombination indicative of an old epidemic. In this study, we characterized HIV strains in two rural areas of the DRC to identify and obtain sequences for rare subtypes and recombinants found only in the DRC.
Methodology: Specimens were collected, between 2001-2003, at the Vanga Hospital, Bandundu Province and The Good Shepard Hospital located 12 kilometers from Kananga, Kasia-Occidental Province in the DRC. A total of 278 HIV-infected specimens from voluntary testing and pregnant women participating in a PMTCT program were characterized. HIV serotype was determined based on antibody reactivity to HIV type and group specific peptides derived from *env* gp120 V3 loop and gp41 immunodominant region using a multiplex immunoassay. Strain classification was determined by RT-PCR amplification and phylogenetic analysis of the *env* gp41 sequence; for rare subtypes, *gag* and *pol* sequences were also obtained and evaluated.
Results: The peptide serotyping assay classified all infections as HIV-1 group M except for 12 that were nonreactive. Phylogenetic analysis of 172 *env* sequences showed a high level of strain diversity. Subtype A predominated (43.6%) but 8 additional subtypes and 5 CRFs were found: subtype C (2.9%), D (9.3%), F1 (4.6%), G (11.0%), H (2.9%), J (1.2%), K (0.6%), L (0.6%), CRF01 (5.8%), CRF02 (9.9%), CRF11 (0.6%), CRF25 (1.7%), CRF27 (0.6%) plus unclassified (4.7%). Three subtype H, and the subtype J, K and L strains had concordant classification based on *gag*, *pol*, and *env* sequences.
Conclusions: Molecular characterization of HIV-infected specimens collected in the DRC show many different strains are circulating in the population and identified rare HIV subtypes H, J, K, and L for which very limited sequences are currently available. Subtype H, J and K sequences are most frequently found within CRFs thus the potentially pure subtype H, K and L strains identified here require confirmation by full genome sequencing. Continued surveillance of HIV strain diversity is important and essential to address the challenge posed by ongoing evolution of HIV and to monitor the rapidly changing HIV pandemic.

Study Population

Specimens were collected at the Vanga Hospital, Bandundu Province and The Good Shepard Hospital located 12 kilometers from Kananga, Kasia-Occidental Province in the DRC between 2001-2003. The specimens came from persons seeking voluntary HIV testing and pregnant women participating in a Prevention of Mother To Child Transmission (PMTCT) program. Of a total of 341 specimens, 278 were determined to be HIV-infected and selected for further characterization to evaluate HIV strain diversity in the population.

Figure 1: Map of the Democratic Republic of Congo and Collection Sites



Circles show regions where the specimens were collected for this study

Serological Testing

Serotype of HIV-infected specimens:

A research use only peptide immunoassay, utilizing *env* V3 loop and immunodominant region (IDR) peptides, was used to serotype HIV-infected specimens (N=278) into HIV-1 groups (M, N, O and P) and HIV-2.

The assay was run on a Luminex instrument to simultaneously evaluate antibody reactivity to 12 different peptides derived from HIV-1, HIV-2 and SIV (chimpanzee and gorilla). A schematic of the assay is shown in Figure 2.

Figure 2: Peptide Immunoassay

1. Coat 12 different beads with peptides
   env gp41 IDR
   - HIV-1 group M, N, O, P
   - HIV-2
   - SIVgor, SIVcpz

   env gp120 V3
   - HIV-1 group M, N, O, P
   - SIVcpz
2. Add sample, wash
3. Add goat anti-human IgG-biotin conjugate, wash
4. Add streptavidin-phycoerythrin (SAPE) label, wash
5. Read on FlexMAP system

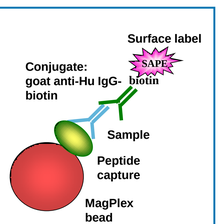


Table 1: Peptide Immunoassay Result: Controls only

	IDR Peptide							V3 Peptide				
	M	HIV-2	O	N	P	CPZ	Gor	M	O	N	P	CPZ
NC	43	24	79	29	25	34	23	57	21	24	71	21
PC M	16514	578	6255	9194	44	9403	4451	5677	38	51	58	1054
PC O	2207	26	11159	118	25	26	500	40	2134	23	56	17
PC N	3582	53	1139	9183	30	9078	797	192	27	510	52	898
PC HIV-2	44	4787	82	27	23	26	20	44	21	23	43	20

Shown in Table 1 is an example of assay results for control specimens; values are mean fluorescense intensity (MFI). Classification of serotype is assigned based on the pattern of reactivity. For example, group N infected specimen (PC N) has highest reactivity (red) to group N and SIV$_{CPZ}$ IDR and V3 peptides.

The DRC specimens were classified as HIV-1 group M (266 of 278; 95.7%) and 12 specimens (4.3%) were nonreactive to all peptides and therefore unclassified.

Molecular Characterization

Molecular characterization of DRC specimens:

RT-PCR was performed using HIV-1 specific primers that amplify a 676 nt region of *env* gp41 including the IDR. If the *env* RT-PCR failed, alternative primers in *env* were used.

For rare subtypes, RT-PCR was also performed for *gag* p24 (468 nt) and *pol* integrase (864 nt).

Viral sequences were aligned with reference strains and phylogenetic analysis was performed using PHYLIP 3.5c (J. Felsenstein, University of Washington, Seattle).

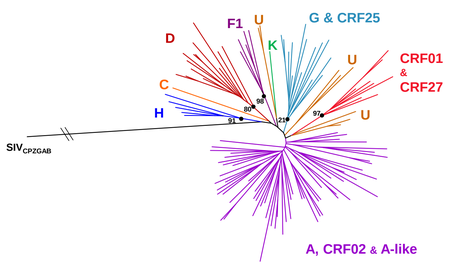
Table 2: Classification Based on *env* Sequence

Subtype/CRF	N=	%
A	75	43.6
C	5	2.9
D	16	9.3
F1	8	4.6
G	19	11.0
H	5	2.9
J	2	1.2
K	1	0.6
L	1	0.6
CRF01	10	5.8
CRF02	17	9.9
CRF11	1	0.6
CRF25	3	1.7
CRF27	1	0.6
Unclassified	8	4.7
Total	172	

Nine different subtypes and 5 CRFs were identified in the population. Subtype A was the most prevalent strain. Rare subtypes H, J, K and L were present at a low frequency.

Note that of the 12 peptide immunoassay nonreactive specimens, only 3 were RT-PCR positive; 1 subtype A, 1 D and 1 unclassified.

Figure 3: *env* Phylogenetic Tree

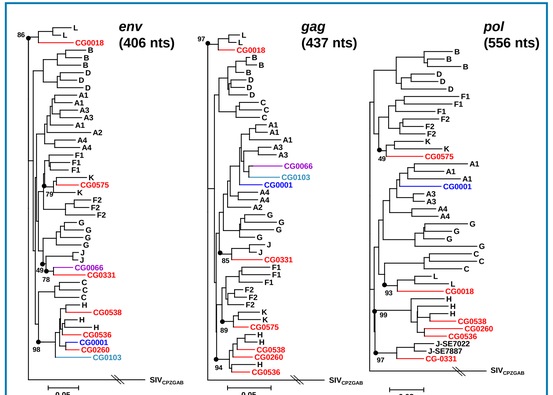


To illustrate the high level of strain diversity present in the DRC population this tree includes a subset of 117 *env* sequences (664 nts) obtained in the study.

The branches for subtypes D, F1, and H, and CRF01 are defined and supported by bootstrap values ≥80%. In contrast, subtypes A, G and CRF derivatives of these subtypes, form clusters with basal branching, long branch lengths and low bootstrap values indicative of long term evolution and recombination.

The majority of unclassified strains (U) appear to be recombinants, however the sequence lengths are too short for definitive classification of composition.

Figure 4: Phylogenetic Trees of Rare HIV Subtypes



Phylogenetic trees were derived from alignments of *env, gag and pol* sequences from specimens that were classified as subtypes H, J, K, and L based on *env*. Strains that have concordant subtype classification in all 3 regions are shown in red; discordant subtype classification are shown in blue and purple. Reference sequences are denoted by black letters. Bootstrap values (%) are shown for key branches.

Table 3: Molecular Characterization of Rare Subtypes

Sample ID	gag	pol	env	Overall subtype
CG0260	H	H	H	H
CG0538	H	H	H	H
CG0536	H	H	H	H
CG0001	A	A	H	URF
CG0103	A	not done	H	URF
CG0331	J	J	J	J
CG0066	A	Neg	J	URF
CG0575	K	K	K	K
CG0018	L	L	L	L

For the rare subtypes H, J, K and L, RT-PCR was done to obtain *gag* and *pol* sequences. Phylogenetic analysis shows concordant classification in all 3 regions for 3 subtype H, 1 subtype J, 1 subtype K and 1 subtype L. Two strains are unique subtype A and H recombinants; these are not the same URFs (blue in Figure 4). A URF of subtypes A and J was also identified (purple in Figure 4).

Conclusions

In the Democratic Republic of Congo HIV-1 group M strain diversity is high. In this study of 172 HIV-infected specimens, 9 subtypes, 5 CRFs, and several URFs were identified. Subtype A is the most prevalent strain accounting for 43.6% of infections, followed by subtype G comprising11% of infections.

Rare HIV subtypes H, J, K and L were identified at a low frequency. Groups O, N and P, and subtype B were not found.

Full genome sequences of the diverse viruses in this population are needed to determine the true level of virus complexity.

Continued surveillance of HIV strain diversity is important and essential to address the challenge posed by ongoing evolution of HIV and to monitor the rapidly changing HIV pandemic.