World Health Organization

List of mutations for surveillance of transmitted drug resistant HIV: 2009 update

Diane Bennett
HIV/AIDS Dept., ATC Unit
World Health Organization
Geneva, Switzerland
bennettd@who.int

DE Bennett[1], RJ Camacho[2], D Otelea[3], DR Kuritzkes[4], H Fleury[5], M Kiuchi[6], W Heneine[7], R Kantor[8], MR Jordan[9], JM Schapiro[6], A-M Vandamme[10], P Sandstrom[11], CAB Boucher[12], D van de Vijver[13], S-Y Rhee[6], TF Liu[6], D Pillay[14], RW Shafer[6]

[1]World Health Organization, Geneva, Switzerland; [2]Centro Hospitalar de Lisboa Ocidental, Lisbon, Portugal; [3"]Prof. Dr. Matei Bals" Institute for Infectious Diseases, Bucharest, Romania; [4]Brigham and Women's Hospital Harvard Medical School, Boston MA, USA; [5]Laboratoire de Virologie EA 2968, Université de Bordeaux, France; [6]Stanford University, Stanford, CA; [7]Centers for Disease Control and Prevention, Atlanta, GA, USA; [8]Brown University, RI, USA; [9]Tufts University School of Medicine, Boston, MA, USA; [10]Rega Institute for Medical Research, Katholieke Universiteit Leuven, Belgium, [11]Public Health Agency of Canada, Ottowa, Ontario, Canada; [12]University Medical Center Rotterdam, The Netherlands; [13]University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands [14]University College London and Centre for Infections, Health Protection Agency, London UK.

BACKGROUND

To compare transmitted HIV drug resistance rates across geographic regions and times, the World Health Organization recommends the adoption of a consensus genotypic definition of transmitted HIV-1 drug resistance. In January 2007, we outlined criteria for developing a list of mutations for drug-resistance surveillance and compiled a list of 80 RT and protease mutations meeting these criteria (surveillance drug resistance mutations; SDRMs). Since January 2007, several new drugs have been approved and several new drug-resistance mutations have been identified. Based on the same criteria, we analyzed publicly available sequences to update the list of mutations for surveillance of transmitted drug-resistant HIV.

Criteria

1. Mutations should be commonly recognized as causing or contributing to resistance.
2. Mutations should be non-polymorphic in untreated persons and should not occur at highly polymorphic positions.
3. The mutation list should be applicable to the 8 most common HIV-1 subtypes.
4. The list should be short and unambiguous.

METHODS

Mutations causing or contributing to drug resistance

Mutations present on three or more of the following five expert lists were considered to be recognized as causing or contributing to drug resistance:
1. ANRS drug resistance interpretation algorithm (2008.07),
2. HIVdb drug resistance interpretation algorithm (4.3.7),
3. IAS-USA Mutations Associated With Drug Resistance (March/April 2008),
4. Los Alamos National Laboratories HIV Sequence database (2007), or
5. Rega Institute Drug Resistance Interpretation Algorithm (7.1.1)

Identification of nonpolymorphic mutations

Some mutations occur commonly in the absence of drug selective pressure; these could lead to falsely elevated estimates of transmitted resistance and were excluded from our list. We defined nonpolymorphic mutations as mutations present at ≤0.5% in ARV-naive individuals infected with subtypes for which >1,000 sequences were available in our dataset and at levels >0.5% in no more than one subtype for which fewer than 1,000 sequences were available. Nonpolymorphic mutations occurring at polymorphic positions, defined as positions with mutations occurring at >1% in among naive individuals infected with any subtype, were generally excluded. Exceptions were made for major mutations that directly cause resistance.

HIV-1 Subtype Identification

A set of 100 reference sequences was compiled by combining 65 representative group M sequences curated by the Los Alamos Sequence Database. An additional 35 samples were added to the dataset so that it would include three or more divergent reference sequences for each pure subtype and many of the most common CRFs. Neighbor joining trees were created from an alignment of each sequence with the 100 reference sequences. Sequences clustering within clades formed by subtypes A, B, C, D, F, G, H, J, and K, and CRF01_AE and CRF02_AG sequences were assigned to that clade. Sequences grouping within clades CRF_03 to CRF_19 were assigned to that clade unless the region spanned by the CRF mapped onto one of the pure subtypes or CRF01_AE or CRF02_AG, in which case the sequence was assigned to one of these. Sequences that were not within a clade were assigned to the subtype or CRF of the closest node. For the purposes of this study, we analyzed mutation prevalence rates only within subtypes A, B, C, D, F, G, CRF01_AE, and CRF02_AG. CRF sequences and non-CRF recombinants that clustered with one of these eight subtypes within protease and/or RT were also included

Exclusion of rare mutations

We excluded rare drug-resistance mutations, defined as those mutations present at a frequency below 0.5% among treated individuals in the subtype having the highest prevalence of that mutation. Because the number of isolates from treated persons for some subtypes was low, we also required that the mutations be present in sequences from at least two different persons with the subtype having the highest prevalence of that mutation.

Sequences analyzed

We analyzed publicly available RT and PI sequences from drug-naive individuals, with two exclusions:
1. Sequences from primarily infected persons in regions with high rates of transmitted resistance.
2. Sequences with two or more mutations from the 2007 SDRM list based on the premise that the mutations in such sequences could have resulted from previous selective drug pressure.

Analysis

1. We identified mutations present on ≥ 3 lists and analyzed the sequences from drug-naive individuals, within subtypes A, B, C, D, F, G, CRF01_AE, and CRF02_AG, for the frequency of those mutations by subtype.
2. For each of the mutations from the previous analysis that met our criteria for non-polymorphism or for non-occurrence at a highly polymorphic position, we examined publicly available sequences for the frequency of each mutation among individuals reported to be *treated* with the relevant drug class. Mutations that met the criterion for rarity were excluded from the list.
3. We reached the final list through review of the results of the analysis by a panel comprising the authors of this paper. Some mutations that occurred as low-level polymorphisms among several subtypes were further excluded for parsimony.
4. Finally, because many mutations on the resulting 2009 SDRM list occurred among sequences of untreated individuals in the database at a frequency greater than zero, we examined the frequencies of sequences with one or more mutations on the 2009 SDRM list among the sequences in the dataset from untreated persons, in order to characterize potential over-estimation of transmitted resistance that could occur using our list.

RESULTS

Table 1. Numbers of Reverse Transcriptase Inhibitor (RTI) and Protease Inhibitor (PI)-Naïve Persons from Whom Publicly Available Sequences Were Available for Analysis in 2006 and 2008

	RTI-Naïve Persons			PI-Naïve Persons		
Subtype	2006	2008	% ▲	2006	2008	% ▲
B	2,240	5,672	+153	3,704	7,439	+101
Non-B Subtypes						
A	499	1,305	+162	686	1,528	+123
AE	635	770	+21	684	902	+32
AG	484	1,035	+114	811	1,437	+77
C	915	2,020	+121	1025	2,182	+113
D	192	320	+69	355	515	+45
F	137	265	+93	180	598	+232
G	145	403	+178	270	619	+129
Non-B Total	3,007	6,118	+105	4011	7,781	+94

RESULTS (CONT.)

Table 2. Nucleoside RT Inhibitor Surveillance Drug Resistance Mutation (SDRM) List: 34 Mutations at 15 Positions

Position	AA	Lists	New	A (%)	AE (%)	AG (%)	B (%)	C (%)	D (%)	F (%)	G (%)	Maximum frequency among untreated	Maximum Frequency among treated
Number of individuals:				1,305	770	1,035	5,672	2,020	324	265	403	11,586	14,621
41	L	5		0	0.3	0.1	0.3	0	1.2	0	0	1.2	39
65	R	5		0.1	0	0	0	0.2	0	0	0	0.2	6.5
67	N	5		0	0	0	0	0	0	0	0	0	44
	G	4		0	0	0.1	0.2	0.1	0	0	0	0.2	1.6
	E	3	√	0	0	0	0	0	0	0	0	0	1.7
69	D	4		0.1	0	0	0	0	0	0	0	0.1	7.1
	Ins	5		0	0	0	0	0	0	0	0	0	1.2
70	R	5		0	0.3	0	0.2	0	0	0	0	0.3	29
	E	5	√	0	0	0	0	0	0	0	0	0	1.6
74	V	5		0	0	0	0	0	0	0	0.2	0.2	7.5
	I	4	√	0	0	0	0	0	0	0	0	0	3.3
75	M	4		0	0.5	0	0	0	0	0	0	0.5	3.8
	T	4		0	0	0	0	0	0	0	0	0	2.7
	A	3		0	0	0	0	0.1	0	0	0	0.1	0.8
	S	3		0	0	0	0	0	0	0	0	0	1.4
77	L	3		0	0.3	0	0.1	0	0	0	0	0.3	2.0
115	F	5		0	0	0	0	0	0	0	0	0	1.4
116	Y	3		0	0.1	0	0	0	0	0	0	0.1	3.3
151	M	5		0	0	0	0	0	0	0	0	0	5.2
184	V	5		0.2	0.1	0.2	0.2	0.1	0.6	0	0	0.6	61
	I	5		0	0.3	0.1	0	0	0.3	0	0	0.3	2.8
210	W	5		0	0	0	0	0	0	0	0	0	25
215	Y	5		0.2	0	0	0	0	0	0	0	0.2	38
	F	5		0.1	0	0	0	0	0	0	0	0.1	22
	I	3		0.2	0.1	0	0	0	0	0	0	0.2	5.7
	S	3		0	0.1	0.1	0.4	0	0.4	0	0	0.4	0.6
	C	3		0	0	0	0.2	0	0	0	0	0.2	3.1
	D	3		0	0	0	0.4	0	0.3	0	0	0.4	0.8
	V	3		0	0	0	0	0	0	0	0	0	1.0
	E	3		0	0	0	0.1	0	0	0	0	0.1	2.1
219	Q	5		0.2	0.4	0.4	0	0	0	0	0	0.4	25
	E	5		0	0	0	0	0	0	0	0	0	9.4
	N	4	√	0.1	0.4	0.2	0.1	0	0	0	0	0.4	2.5
	R	4		0.2	0	0.1	0.1	0.1	0	0	0	0.1	2.2

Table 3. Non-Nucleoside RT Inhibitor Surveillance Drug Resistance Mutation (SDRM) List: 19 Mutations at 10 Positions

Position	AA	Lists	New	A (%)	AE (%)	AG (%)	B (%)	C (%)	D (%)	F (%)	G (%)	Maximum frequency among untreated	Maximum Frequency among treated
Number of individuals:				1,305	770	1,035	5,672	2,020	324	265	403	11,784	14,621
100	I	5		0	0	0	0	0	0	0	0	0	5.4
101	E	5		0	0	0	0.2	0	0	0	0	0.2	6.4
	P	5	√	0	0	0	0	0	0	0	0	0	1.4
103	N	5		0.1	0	0.2	0.3	0.2	0	0	0	0.3	40
	S	4		0	0	0	0	0	0	0	0	0	0.6
106	M	5		0	0	0	0	0	0	0	0	0	12
	A	5		0	0	0.1	0	0	0	0	0	0.1	1.1
179	F	5	√	0	0	0	0	0	0	0	0	0	0.2
181	C	5		0.1	0.1	0.1	0	0.1	0	0	0	0.1	14
	I	5		0.1	0	0	0	0	0	0	0	0.1	1.1
	V	5		0	0	0	0	0	0	0	0	0	1.1
188	L	5		0	0.3	0	0	0	0	0	0	0.3	2.4
	H	5		0	0	0	0.1	0	0.3	0	0	0.3	0.6
	C	5		0	0	0.1	0	0	0	0	0	0.1	0.9
190	A	5		0.1	0	0	0	0.1	0.3	0	0	0.3	10
	S	5		0	0	0	0	0	0	0	0	0	1.6
	E	4		0	0	0.1	0	0	0	0	0	0.1	0.5
225	H	5		0	0	0	0	0	0	0	0	0	7.6
230	L	3		0	0	0	0	0	0	0	0	0	3.6

Table 4. Protease Inhibitor (PI) Surveillance Drug Resistance Mutation (SDRM) List: 40 Mutations at 18 Positions*

Position	AA	New	Lists	A (%)	AE (%)	AG (%)	B (%)	C (%)	D (%)	F (%)	G (%)	Maximum frequency among untreated	Maximum Frequency among treated
Number of individuals:				1,528	902	1,437	7,439	2,182	515	598	619	15,220	7,886
23	I	√	3	0.1	0.1	0	0	0	0	0	0	0.1	2.1
24	I		5	0	0	0.1	0	0	0	0.2	0	0.2	11
30	N		5	0	0	0	0	0	0	0	0	0	9.0
32	I		5	0	0	0.1	0	0	0	0	0	0.1	7.3
46	I		5	0.3	0.4	0.1	0.4	0.2	0	0.2	0	0.6	29
	L	√	5	0.3	0.2	0.3	0.3	0.1	0.2	0.2	0.5	0.5	13
47	V		5	0	0	0	0	0.1	0.2	0	0.3	0.3	6.8
	A		5	0	0	0	0	0	0	0	0	0	0.5
48	V		5	0	0	0	0	0	0	0	0	0	5.7
	M	√	3	0	0	0	0	0	0	0	0	0	2.3
50	V		5	0	0	0	0	0	0	0	0	0	2.5
	L		5	0	0	0	0	0	0	0	0	0	1.8
53	L		5	0	0.1	0	0	0	0	0.2	0.2	0.2	7.8
	Y	√	4	0.1	0	0.1	0	0	0	0	0	0.1	0.6
54	V		5	0	0	0	0	0	0	0	0	0	43
	L		5	0	0	0	0	0	0	0	0	0	4.6
	M		5	0	0	0	0	0	0	0	0	0	4.3
	A		5	0	0	0	0	0	0	0	0	0	2.6
	T		5	0.1	0	0	0	0	0	0	0	0.1	1.7
	S		5	0	0	0	0	0	0	0	0	0	1.7
73	S		5	0	0	0.1	0	0	0	0	0	0.1	13
	T		4	0	0	0	0	0	0	0	0	0	4.3
	C		3	0	0	0	0	0	0	0	0	0	2.0
	A		4	0	0	0	0	0	0	0	0	0	0.8
76	V	√	5	0	0	0	0	0	0.2	0	0	0.2	4.3
82	A		5	0	0	0	0	0.1	0	0.2	0	0.2	30
	T		5	0	0	0	0	0	0	0	0	0	13
	F		5	0	0	0	0	0	0	0	0	0	2.1
	S		5	0	0	0	0	0	0	0	0	0	1.6
	C	√	3	0	0	0	0	0	0	0	0	0	0.8
	M		4	0	0	0	0	0	0	0	0	0	2.5
	L	√	4	0	0	0	0	0	0	0	0	0	0.5
83	D	√	3	0	0	0.1	0	0.1	0	0	0	0.3	9.3
84	V		5	0	0.1	0	0	0	0	0	0	0.1	21
	A		4	0	0	0	0	0	0	0	0	0	0.8
	C		3	0	0	0	0	0	0	0	0	0	0.8
85	V	√	4	0.1	0	0.1	0.1	0.2	0	0.2	0	0.2	6.4
88	D		5	0	0.1	0.1	0.1	0	0	0	0.2	0.2	6.4
	S		5	0	0	0	0	0	0	0	0	0	3.4
90	M		5	0.1	0.3	0	0.1	0.1	0	0	0.5	0.5	45

Abbreviations: Pos - amino acid position; AA - amino acid difference from consensus B; Lists - Number of mutation lists with the mutation; New - mutations not present on the 2007 SDRM list; Maximum frequency among untreated- highest prevalence in untreated persons in any of the 8 listed subtypes; Maximum frequency among treated - Prevalence of the mutation in the subtype with the highest prevalence of the mutation among treated individuals, provided the mutation is present in viruses from two or more individuals. Underlined bold mutations are those with a prevalence >0.5% that were included because the >0.5% prevalence occurs in only one subtype with fewer than 1000 sequences for analysis.

RESULTS (CONT.)

- RT sequences from 11,586 RT inhibitor-naïve individuals and protease sequences from 15,220 PI-naïve individuals were publicly available and met the criteria to be included the analysis dataset (Table 1). More than 50% of both the RT and protease sequences were from non-subtype B viruses.
- RT sequences from 14622 RT inhibitor-treated individuals and protease sequences from 7819 PI-treated individuals were publicly available and met the criteria to be included the analysis dataset. Non-B sequences made up 25% of the analyzable RT sequences and 15% of the analyzable PI sequences from treated persons
- Table 2 shows the updated list of 34 NRTI SDRMs at 15 RT positions. The proportion of reportedly drug-naive individuals in the dataset having one or more NRTI-associated SDRMs for each of the eight subtypes is as follows: 0.2% in subtype G, 0.4% in subtype F, 0.6% in subtype C, 1.3% in CRF02_AG, 1.4% in subtype A, 2.3% in subtype B, B,2.4% in subtype D, and 2.9% in CRF01_AE.
- Table 3 contains the updated list of 19 NNRTI SDRMs at 10 RT positions. The proportion of reportedly drug-naive individuals in the dataset with one or more NNRTI-associated SDRM for each of the eight subtypes would be as follows: 0% in subtypes F and G, 0.4% in subtype A and CRF01_AE, 0.5% in subtype C, 0.5% in CRF02_AG, 0.6% in subtype D, and 0.8% in subtype B.
- Table 4 contains the updated list of 40 PI SDRMs at 18 protease positions. The proportion of reportedly drug-naive persons in the dataset having one or more PI-associated SDRMs for each of the eight subtypes would be as follows: 0.4% in subtype D, 0.9% in subtype C, 1.1% in subtype A, 1.1% in subtype B and CRF01_AG, 1.1% in subtype G, 1.2% in subtype F, and 1.5% in CRF01_AE.

DISCUSSION AND CONCLUSIONS

- The 2007 WHO list of SDRMs was used in several published surveys and has facilitated comparability of results and minimized misclassification of transmitted resistance. The new list, based on substantially more sequences, includes several new mutations. Nearly twice the numbers of B and non-B sequences were analyzed for the new list compared to the 2007 list analysis.
- Among the 93 mutations on the 2009 SDRM list, three occurred at a frequency >0.5% (in subtypes with fewer than 1,000 sequences) and 46 occurred at a frequency between 0.1% and 0.5% in at least one subtype. All lists of mutations used to define transmitted resistance include mutations that are polymorphic to some degree; their contribution to probable overestimates of transmitted resistance should be acknowledge. Our list attempts to minimize this contribution and other elements that could lead to inaccurate estimations of transmitted resistance.