

# HIV Drug Resistance Genotyping External Quality Assurance (EQA) for Laboratories in the World Health Organization (WHO) ResNet During 2007-2009.

Neil Parkin<sup>1\*</sup>, Joe Fitzgibbon<sup>2</sup>, James Bremer<sup>3</sup>, and Silvia Bertagnolio<sup>1</sup>  
<sup>1</sup>WHO, Geneva, Switzerland; <sup>2</sup>NIAID, NIH, Bethesda MD, USA; <sup>3</sup>VQA, Rush University, Chicago, IL, USA.  
 \*HIV/AIDS Dept., ATC Unit, World Health Organization, Geneva, Switzerland parkinn@who.int

**XVIII international HIV DRUG RESISTANCE workshop**  
 basic principles & clinical implications  
 June 9-13, 2009 • Fort Myers, Florida

## BACKGROUND

- The WHO has developed a global laboratory network to support HIV drug resistance genotyping in resource-limited countries<sup>1</sup>.
- An external QA program is being implemented to ensure the reliability of genotyping data generated by the various laboratories.
- Three proficiency panels were developed in collaboration with VQA and NIH and sent to 58 network member or candidate laboratories in Europe, North America, Asia, Africa and the Caribbean during 2007-09.
- Appropriate and widely accepted evaluation criteria have not been described previously.

## METHODS

- Each panel was composed of 5 clinical samples:

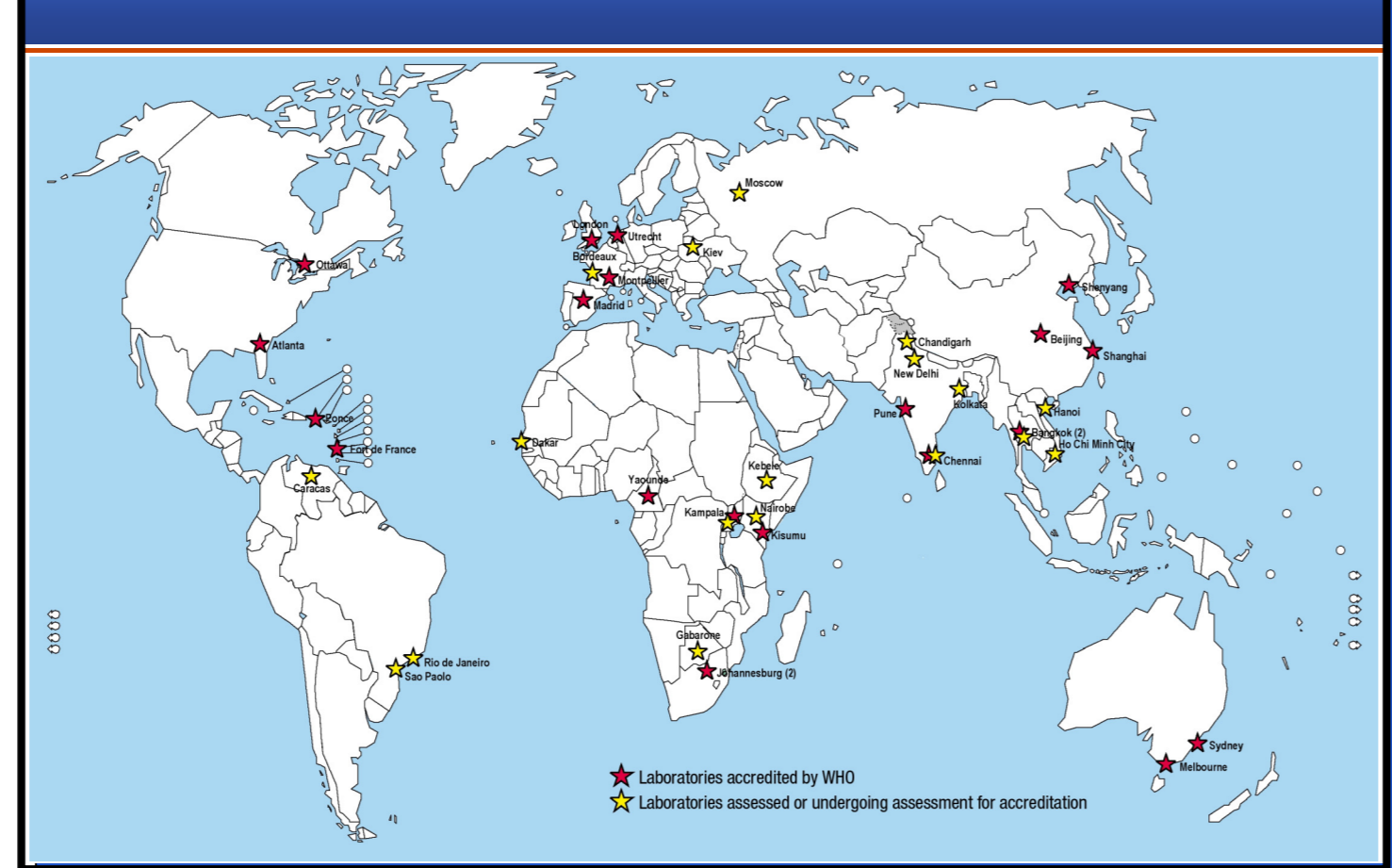
Panel	Sample	Subtype	Viral load	Primary DRM*	Comments
1	1	D	18,547	No	
	2	C	19,139	No	Sample 2 and 3 are the same
	3	C	12,380	No	
	4	C	48,955	No	Sample 4 and 5 are the same
	5	C	16,663	No	
2	1	CRF01_AE	27,810	No	
	2	C	14,282	No	
	3	F	24,238	No	
	4	B	3557	Yes	Multiple NAMS, NNRTI and major PI mutations
	5	CRF02_AG	57,005	No	
3	1	C	7748	Yes	Major NRTI and NNRTI mutation
	2	B	8398	Yes	Major NNRTI mutations
	3	F	13,053	No	
	4	B	10,735	No	
	5	C	10,958	Yes	Multiple NAMS, NNRTI and major PI mutations

Panel 1 (2007): 21 labs reported results from a total of 25 assays: 8 ViroSeq, 6 TruGene, 11 In-house  
 Panel 2 (2008): 26 labs reported results from a total of 33 assays: 10 ViroSeq, 7 TruGene, 16 In-house; data from 2 labs not analyzed due to amplification failures.  
 Panel 3 (2009): 23 labs reported results from a total of 33 assays: 13 ViroSeq, 6 TruGene, 14 In-house; data from 2 labs not analyzed due to amplification failures.  
 \* minor PR polymorphisms present and considered in DRM site scoring

- Consensus sequences (PR codons 10 to 99 and RT 38-240) for each sample were generated based on >80% concordance of results from >30 labs and all platforms; individual test results were compared to that sample's consensus sequence. If 80% concordance is not achieved at any given position, an "N" is inserted in the consensus sequence and is not considered in the scoring.
- An overall sequence identity score as well as concordance at Drug Resistance Mutation (DRM) codons (as defined by IAS-USA) were used to assess lab performance.
  - Mixed vs. unmixed bases were counted as discrepancies for nucleotide sequence alignment score
  - For the DRM site score, if a mutant codon is present in the consensus and included in the reported sequence, regardless of being mixed or not, it was not counted as a discrepancy (assessed at the nucleotide level for 2007-2008, changed to amino acid level in 2009)
- These evaluation methods and criteria differ from those used by the VQA

<sup>1</sup>Bertagnolio et al., Antiviral Therapy 13 Suppl 2:49-57.

## 1. HIVDR Laboratory Network, March 2009



## 2. Drug Resistance Mutation Site Scores

PANEL	SAMPLE	N*	N>99%	N>95%	min	max	mean
1	1	25	17	22	88.9	100	98.6
	2	25	25	25	100	100	100
	3	25	25	25	100	100	100
	4	25	18	24	96.7	100	99.5
	5	25	18	21	96.7	100	99.3
overall		22	23	96.8	99.7	99.5	
2	1	31	29	30	96.9	100	99.9
	2	30	25	25	95.5	100	99.4
	3	31	28	30	97.0	100	99.8
	4	30	18	26	95.5	100	99.0
	5	30	25	27	72.7	100	98.7
overall		22	28	94.5	100	99.4	
3	1	32	29	31	94.1	100	99.7
	2	33	31	31	83.1	100	99.4
	3	33	31	31	55.9	100	98.6
	4	33	31	32	90.6	100	99.7
	5	23	18	22	64.2	100	98.2
overall		33	30	30	74.3	100	98.3

\*number of lab-assay combinations (some labs used more than one assay); amplification failures excluded

## 3. Nucleotide Sequence Alignment Scores

PANEL	SAMPLE	N*	N>99%	N>95%	min	max	mean
1	1	25	8	23	97.3	99.0	98.8
	2	25	25	25	99.6	100	100
	3	25	25	25	99.8	100	100
	4	25	0	23	97.2	98.8	98.5
	5	25	8	24	97.6	99.1	98.8
overall		22	25	98.7	99.4	99.2	
2	1	31	30	31	98.4	99.3	99.3
	2	30	28	30	98.3	99.9	99.6
	3	31	30	31	99.0	100	99.9
	4	30	18	29	98.1	99.5	99.1
	5	30	26	27	90.3	99.9	99.3
overall		27	30	97.8	99.7	99.4	
3	1	32	31	31	97.9	100	99.9
	2	33	30	32	91.0	100	99.6
	3	33	31	31	71.2	100	99.0
	4	33	32	32	93.9	100	99.7
	5	23	22	22	92.6	100	98.6
overall		33	30	31	74.9	100	98.8

\*number of lab-assay combinations (some labs used more than one assay); amplification failures excluded

## 4. DRM Site Discrepancies (Panel 1)

Sample	PR or RT	Position	Number of labs	Consensus	Reported Variants
1	PR	11	7	GTG (V)	GTC (V)
4	PR	20	5	AAR (K)	AAG (K)
5	PR	13	3	RTA (IV)	ATA (I)
5	PR	20	2	AAR (K)	AAG (K)

Also 12 sites with 1 discrepancy each red: differences vs. consensus

This table, and the adjacent 2 (Tables 5 & 6) shows that the majority of discrepancies can be accounted for by a small number of specific sites in a few samples, and that most of these sites contain mixtures in the consensus or reported sequence

## 5. DRM Site Discrepancies (Panel 2)

Sample	PR or RT	Position	# labs	Consensus	Reported Variants
4	RT	184	7	CTA (L)	CTR (L), MTR (I/L/M), MTA (I/L), ATA (I), ATG (M)
2	PR	93	6	CTK (L)	CTT (L), CT G (L), CT S (L), GGG (G)
2	PR	89	4	ATR (IM)	ATG (M)
4	PR	71	3	AYT (IT)	ACT (T)
4	PR	32	3	GTA (V)	GTR (V)
2	RT	65	2	AAG (K)	AAR (K), CCA (P)
2	PR	58	2	CAG (Q)	CAR (Q), MAG(K/Q)
3	RT	151	2	CAA (Q)	CAG (Q)
3	RT	225	2	CCC (P)	MCC (P/T), CC T (P)
4	PR	62	2	ATA (I)	GTA (V), RTA (IV)
4	RT	62	2	GTT (V)	GAA (E), GYT (AV)
5	PR	36	2	ATA (I)	ATG (M)

Also 27 sites with 1 discrepancy each red: differences vs. consensus

## 6. DRM Site Discrepancies (Panel 3)

Sample	PR or RT	Position	# labs	Consensus	Reported Variants
2	RT	116	8	TTC (F)	TTY (F), TT S (F/L), TT T (F)
4	PR	62	6	RTA (IV)	GTA (V)
4	PR	69	6	CAY (H)	CAT (H)
1	RT	181	6	TAT (Y)	TRT (Y/C)
2	RT	188	6	TGT (C)	TRT (Y/C)
4	PR	71	5	GCW (A)	GCA (A)
2	RT	62	5	GCC (A)	GCY (A)
4	PR	10	3	CTC (L)	CTY (L), CTT (L)
4	PR	16	3	GGG (G)	GGR (G)
1	RT	184	3	GTG (V)	RTG (MV)
5	PR	36	3	ATG (M)	ATA (I), AT R (M/I), AT S (M/I)
2	PR	20	3	AAR (K)	AAA (K)

Also 5 sites with 2 discrepancies and 29 sites with 1 discrepancy each red: differences vs. consensus

## 7. Distribution of Errors at DRM sites

Panel	Sample	PR	RT	Total
1	1	10	1	11
	2	0	0	0
	3	0	0	0
	4	9	0	9
	5	9	0	9
2	1	1	4	5
	2	14	8	22
	3	0	4	4
	4	14	12	26
	5	3	4	7
3	1	2	12	14
	2	3	22	25
	3	0	2	2
	4	24	2	26
	5	16	13	29*

Discrepancies are not evenly spread out across the samples, suggesting that one or more features specific to some samples are a source of difficulty for several labs

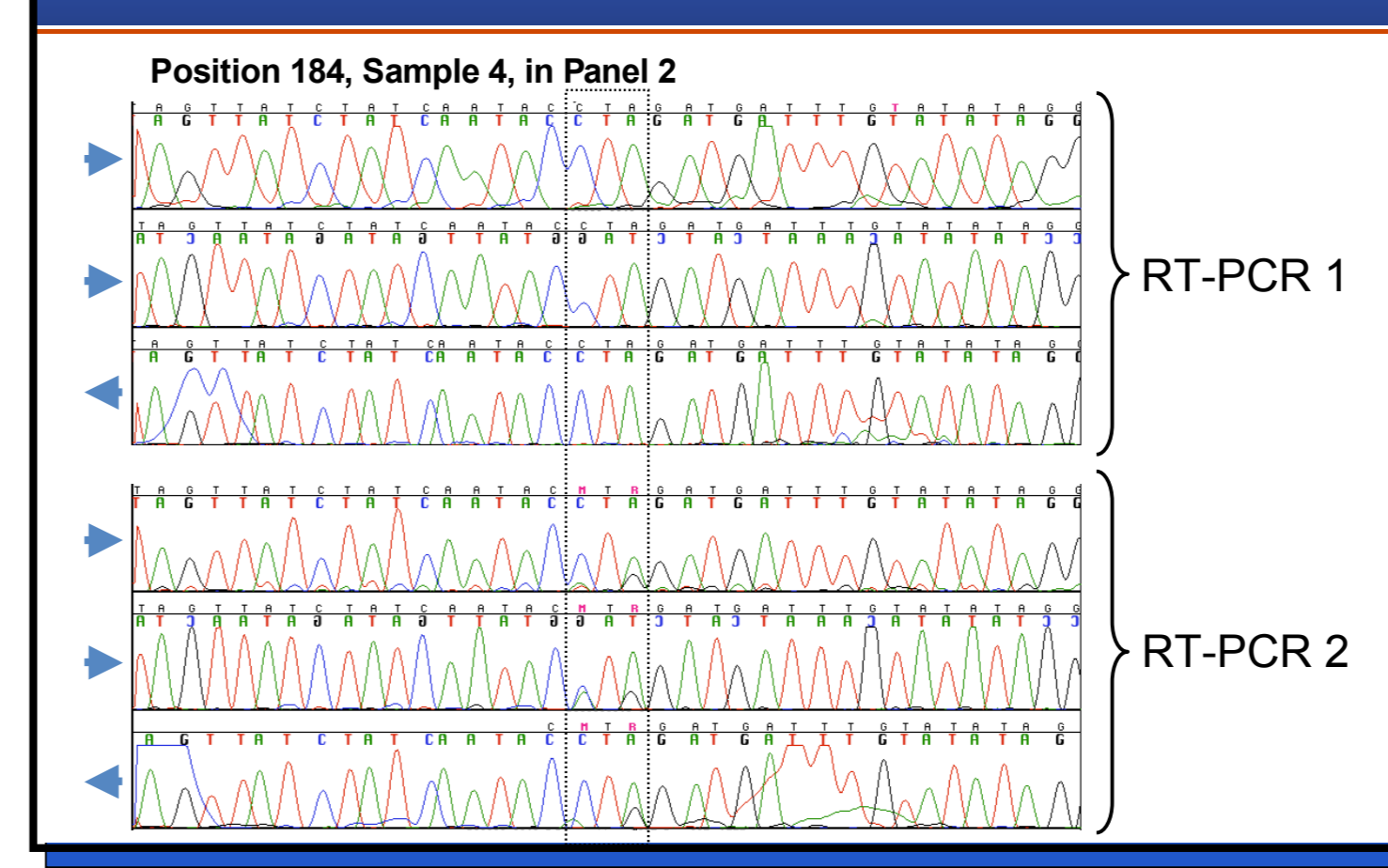
\*24 of these discrepancies were a result of the wrong sequence being submitted (possible contamination)

## 8. Sequence Heterogeneity

Panel	Sample	Number of bases				PR DRM mix*	RT DRM mix*
		MIXED	UNMIXED	%MIXED	%MIXED_no N's*		
1	1	10	908	1.1%	0.1%	1	0
	2	1	917	0.1%	0.1%	0	0
	3	1	917	0.1%	0.1%	0	0
	4	19	899	2.1%	0.9%	2	0
	5	18	900	2.0%	1.1%	3	0
2	1	6	891	0.7%	0.0%	0	0
	2	8	888	0.9%	0.8%	2	0
	3	0	897	0.0%	0.0%	0	0
	4	6	891	0.7%	0.2%	2	0
	5	1	896	0.1%	0.0%	0	0
3	1	1	878	0.1%	0.0%	0	0
	2	15	864	1.7%	0.8%	1	0
	3	0	879	0.0%	0.0%	0	0
	4	16	863	1.8%	0.9%	3	0
	5	2	877	0.2%	0.0%	0	0

\*not counting N's since these are ignored in scoring

## 9. Evidence for Low-Level Mixtures



## RESULTS

- Overall, 72 of 91 submissions passed (79%)
  - Success rates were higher with panel 1 (80%) than panel 2 (64%)
  - Scoring criteria were adjusted for panel 3 (94% success rate)
- Specific reasons for failure included:
  - editing errors leading to frameshifts
  - missing sequence (primer design, subtype issues)
  - amplification failure (primer design, subtype issues)
  - cross-contamination or post-processing file naming errors
  - low concordance rate at positions with mixed bases in the consensus.
- Samples and sites with naturally-occurring mixtures were more challenging than others (Tables 7 & 8).
  - no lab obtained a nucleotide alignment score >99% for sample 4 in panel 1 (this sample's consensus sequence had the most mixtures) (Tables 3 and 8)
  - for sample 4 in panel 2, only 18 out of 31 labs reported sequence with >99% concordance, compared to 26 to 30 out of 31 labs for the other 4 samples (Table 3).
- Based on comparison of chromatogram data from several labs (Figures 9 and data not shown), subjectivity in base-calling and PCR amplification bias ("founder effect") can contribute to lower reproducibility when mixtures are present.

## CONCLUSIONS

- The use of a single and stringent criteria (e.g. >99%) for evaluation of sequence-based assays may be unrealistic when using clinical samples containing mixed bases.
- Acceptance criteria may need to be relaxed for such samples, or be flexible and based on the number of mixed positions in each sample.
- These observations led to revised assessment criteria for 2009 proficiency panels:

- Discrepancies not to be counted in the DRM site score:
  - Mixed WT+mutant in consensus, lab reports mutant
    - consensus = AYT (WT = ATT), test sequence = ACT
  - Mixed in consensus, lab reports a compatible base, no change in amino acid
    - consensus = GTR, test sequence = GTG
  - Unmixed in consensus, lab reports a mixture containing the consensus base, no change in amino acid
    - consensus = ACG, test sequence = ACR
- Acceptance criteria: DRM and sequence alignment score ≥98%; considering only positions unmixed in consensus, ≥ 99%

## ACKNOWLEDGEMENTS

- All participating laboratories
- Anir Pandit and William Mackay, QCMD, Glasgow, Scotland
- Cheryl Jennings, VQA, Chicago IL
- Steve Hart (Frontier Science Foundation, Amherst NY), Don Brambilla (RTI Rockville, MD), Suzanne Granger (New England Research Institutes, Inc., Watertown, MA)