Blinded Comparison of an *in vitro* Diagnostic NGS Assay with Sanger Sequencing for HIV Genotypic Resistance Testing.

V Varghese[1], P Ariyaratne[2], PL Tzou[1], E Rakhmanaliev[2], C Lee[2], C Villy[2], M Yee[2], K Tan[2], BA Pinsky[1], G Michel[2], RW Shafer[1]

[1]Stanford University School of Medicine, Stanford, CA, [2]Vela Diagnostics, Singapore Science Park II, Singapore

Introduction

- Next generation sequencing (NGS) platforms are increasingly used for HIV genotypic resistance testing in research settings. However, there are few studies in which an *in vitro* diagnostic NGS product has been blindly compared with standard genotypic resistance testing (SGRT) using dideoxynucleotide Sanger sequencing.
- We identified cryopreserved samples that had previously undergone SGRT, and were found to have HIV drug-resistance mutations (DRMs), for a blinded comparison with the the Vela Diagnostics *Sentosa*® SQ HIV Genotyping NGS Assay.
- We determined the extent of concordance between SGRT and NGS for the identification of amino acid mutations, drug-resistance mutations (DRMs), and predicted levels of reduced susceptibility to the most commonly used antiretroviral (ARV) drugs.

Methods

Patients and samples: We selected 49 cryopreserved samples from 45 HIV-1-infected individuals that had undergone SGRT at the Stanford Healthcare Diagnostic Virology Lab for a blinded comparison with the Vela Diagnostics *Sentosa*® SQ HIV Genotyping NGS Assay. We selected samples that by SGRT had large numbers of the most commonly occurring and clinically relevant nucleoside RT inhibitor (NRTI), non-nucleoside RT inhibitor (NNRTI), and protease inhibitor (PI)-resistance mutations. The study was approved by the Stanford University IRB.

Standard genotypic resistance testing (SGRT): SGRT using direct PCR Sanger sequencing had been performed as part of routine patient management. Plasma samples had been ultracentrifuged and the resulting pellet was subjected to RNA extraction and SuperScript® III One-Step RT-PCR with Platinum® *Taq* DNA Polymerase (ThermoFisher Scientific) followed by a second round of PCR. Bidirectional-sequencing encompassing HIV protease (PR) and the first 300 codons of RT was performed using BigDye® Terminators (ThermoFisher Scientific) with products resolved electrophoretically on an ABI 3730 sequencer.

Next-generation sequencing (NGS): The *Sentosa*® SQ HIV Genotyping Assay is an NGS-based integrated workflow, comprising (i) A robotic liquid handling system for RNA extraction and NGS library preparation; (ii) Ion Torrent's instruments for deep sequencing; (iii) Kits for RNA extraction, HIV NGS library preparation and sequencing, and (iv) Data analysis and reporting software. The *Sentosa*® HIV NGS workflow is highly automated and requires about 3.5 hrs. hands-on time with total turn-around time about 27 hrs. The assay processes up to 15 plasma samples simultaneously. The system sequences the entire PR, first 385 bp of RT, and entire IN. For the purposes of the current study, nucleotide sequences were exported as FASTA files in which positions containing nucleotide mixtures with variants present at a level of ≥5% were represented using IUPAC ambiguities.

Data analysis: We submitted the SGRT and NGS FASTA files containing PR and the first 250 codons of RT to the HIVDB genotypic resistance test (GRT) interpretation program and selected the "Sequence quality" and "Drug resistance" summary spreadsheet output options (https://hivdb.stanford.edu/hivdb/by-sequences/). The "Sequence quality" spreadsheet has one row per sequence with (i) Start and stop position for each gene, (ii) HIV-1 subtype; (iii) Proportion of nucleotides with ambiguities; (iii) Lists of PR mutations grouped into major DRMs, accessory DRMs, and other mutations; (iv) Lists of RT mutations grouped into NRTI DRMs, NNRTI DRMs, and other mutations; and (v) QC metrics. The "Drug resistance summary" spreadsheet has one row per sequence with numeric drug resistance scores and categorical drug resistance levels for each of the NRTIs, NNRTIs, and PIs.

We compared the rows in the SGRT spreadsheets with the corresponding rows in the NGS spreadsheets to determine the extent of concordance between SGRT and NGS for the detection of (i) Amino acid mutations, defined as differences from the subtype B consensus sequence; (ii) DRMs, defined as mutations assigned a mutation penalty score by the HIVDB GRT interpretation system; and (iii) Categorical drug resistance interpretations for the most commonly used ARVs: the NRTIs 3TC/FTC, which were counted as one drug, abacavir (ABC), AZT, and tenofovir (TDF); the NNRTIs efavirenz (EFV), etravirine (ETR), and rilpivirine (RPV); and the PIs atazanavir (ATV), darunavir (DRV), and lopinavir (LPV). There were 5 drug resistance interpretation levels: susceptible, potential low-level resistance, low-level resistance, intermediate resistance, and high-level resistance.

Results

Patients and samples: ARV treatment histories were available for 38 of the 45 individuals from whom cryopreserved samples had been selected. Among those with known treatment histories, the median number of NRTIs received was 4 (range: 2 to 7), the median number of NNRTIs was 1 (range: 0 to 3), and the median number of PIs was 2 (range: 0 to 6). The median plasma HIV-1 RNA level was 4.1 log copies/ml (range: 2.8 to 5.7). The median sample collection year was 2011 (range: 2002 to 2014). Forty-eight of the samples had subtype B viruses; 1 had a subtype G virus. Table 1 shows the numbers of NRTI, NNRTI, and PI DRMs present by SGRT in the 49 cryopreserved plasma samples used for the SGRT/NGS comparison.

Table 1. NRTI, NNRTI, and PI DRMs Present by SGRT in the 49 Cryopreserved Plasma Samples used for the SGRT/NGS Comparison

NRTI DRMs			NNRTI DRMs			PI DRMs		
DRM	No.	%HIVDB*	DRM	No.	%HIVDB*	DRM	No.	%HIVDB*
M184V	27	54.4%	K103N	22	36.1%	M46I	11	20.8%
M41L	20	27.8%	Y181C	17	20.7%	L90M	11	27.1%
T215Y	14	25.0%	K101E	9	8.2%	L10F	7	7.8%
K65R	13	6.1%	P225H	9	4.7%	I84V	6	11.0%
D67N	11	26.1%	H221Y	7	6.0%	I54V	6	23.1%
L210W	10	16.8%	V108I	7	9.1%	L33F	6	12.1%
L74V	9	7.8%	G190A	6	16.2%	V82A	5	21.7%
K70R	9	17.6%	L100I	5	4.1%	I47V	4	4.4%
T215F	8	9.0%	G190S	3	2.4%	V32I	4	4.7%
L74I	5	3.7%	V179F	3	0.2%	I54M	3	1.9%
A62V	4	4.5%	K103S	2	1.2%	D30N	3	5.5%
K219R	3	2.4%	K101P	2	1.2%	N88D	3	5.0%
M184I	3	2.0%	E138K	1	0.4%	M46L	3	8.5%
T69Ins	2	0.7%	A98G	1	7.3%	I54L	2	2.7%
Q151M	1	2.6%	Y188L	1	4.3%	V82T	2	2.2%
F116Y	1	2.0%	M230L	1	1.9%	V82S	2	0.9%
F77L	1	1.7%	Y188C	1	0.9%	T74P	2	1.9%
K70Q	1	0.2%	G190E	1	0.4%	F53L	2	5.9%
			G190Q	1	0.3%	L89V	2	3.4%
			E138Q	1	1.3%	G48V	1	2.9%
						I50V	1	1.8%

* Percentage of each DRM among NRTI, NNRTI, and PI treated individuals in HIVDB.

SGRT/NGS comparison of nucleic acid mixtures, amino acid mutations, and DRMs:

- A mean 23.2 mutations per sequence were detected by both NGS and SGRT; 3.9 mutations were detected solely by NGS; and 0.5 mutations were detected solely by SGRT.
- A mean 7.5 DRMs per sequence were detected by both NGS and SGRT; 0.8 DRMs were detected solely by NGS; and 0.1 DRMs were detected solely by SGRT.
- NGS detected an increased number of mixed bases per sequence 1.6% (IQR: 1.1% to 2.3%) compared with SGRT 1.2% (IQR: 0.5% to 1.7%; p=1e-6 paired Wilcoxon Rank Sum Test).
- Table 2 lists the DRMs and drug resistance interpretations for the sequences with NNRTI (n=9), NRTI (n=5), and PI (n=4) DRM discordances.

SGRT/NGS comparison of drug resistance interpretations:

- Of the 490 drug resistance interpretations (49 sequences x 10 ARVs), 471 (96.1%) were concordant and 19 (3.9%) were discordant.
- The 19 discordances occurred in 9 of the 49 sequences. For 18 discordances, NGS was more resistant than SGRT. For 1 discordance, SGRT was more resistant than NGS.
- For 12 discordances, NGS was 1 level more resistant than SGRT; for 5 discordances, NGS was 2 levels more resistant than SGRT; and for 1 discordance, NGS was 4 levels more resistant than SGRT. For 1 discordance, SGRT was 1 level more resistant than NGS.
- The ARVs with the most discordances were the NNRTIs RPV (n=7), ETR (n=5), and EFV (n=3). The NRTIs AZT and TDF and the PIs LPV and DRV each had one discordance.

Table 2. DRMs and Predicted Drug Resistance Levels for Sequences with an SGRT/NGS DRM Discordance

SID	Drug Class	SGRT and NGS	SGRT only	NGS Only	SGRT => NGS*		
					EFV	ETR	RPV
4	NNRTI	100I,103N,225H		108I	5	4	5
16	NNRTI	103N,181C,188C		106A,190A	5	4	4 => 5
33	NNRTI	103N,238N		179L	5	5	5
35	NNRTI	103N,181C		221Y	5	4	4 => 5
37	NNRTI	181C	101E,103N	221Y	5 => 4	4	5
43	NNRTI	190S		101E	5	2 => 4	3 => 5
44	NNRTI	103N,108I,225H		221Y	5	1 => 2	1 => 3
49	NNRTI	101E		190S	3 => 5	3 => 4	4 => 5
51	NNRTI	181C		98G,101E	4 => 5	4 => 5	4 => 5

					3TC	ABC	AZT	TDF
10	NRTI	41L,44D,184V,210W,215Y		67N	5	5	5	4 => 5
18	NRTI	74V,184V		219R	5	5	1	1
42	NRTI	41L,62V,184V,210W,215Y		219R	5	5	5	4
44	NRTI	65R,74I,184V,219R			5	5	1	5
51	NRTI	65R,67N,184V,219E	70Q	70R,215I	5	5	1 => 5	5

					ATV	DRV	LPV
7	PI	10F,32I,46I,54M,73A,84V,90M		47V	5	5	5
10	PI	46L,82A,84V,90M		54M	5	3 => 4	5
42	PI	46I,90M		10F,58E	4	1	3 => 4
49	PI	10F,23I,46I,47V,54V,82T		82A	5	3	5

* Susceptible (1), Potential low-level resistance (2), Low-level resistance (3), Intermediate resistance (4), High-level resistance (5)

Conclusions

- The overall concordance between the *Sentosa*® SQ HIV Genotyping NGS Assay and SGRT was high for the detection of amino acid mutations, DRMs, and predicted levels of reduced susceptibility to the 10 most commonly used NRTIs, NNRTIs, and PIs.
- Whereas SGRT usually detects nucleotide variants present at levels of ≥20% of the virus population within a sample, the *Sentosa*® SQ HIV Genotyping Assay was calibrated to detect variants at levels ≥5%.
- As a result, the NGS assay appeared more sensitive for the detection of low-abundance amino acid variants, detecting a mean 3.4 additional mutations and 0.7 additional DRMs per sequence.
- The SGRT and NGS assays used different sample preparation methods and primers for PCR amplification, factors that may also contribute to differences in the detection of low-abundance variants.
- The more sensitive detection of low-level variants by NGS resulted in a higher predicted level of drug resistance in 9 (18.4%) of 49 samples and in 18 (3.8) of 490 drug resistance interpretations. The most commonly affected ARV class was the NNRTIs.
- Additional blinded comparisons of samples with non-subtype B viruses and integrase inhibitor resistance mutations are planned.