GEMS
GLOBAL EVALUATION OF MICROBICIDE SENSITIVITY

COMPARISON OF HIV DRUG-RESISTANT MUTANT DETECTION BY NGS WITH AND WITHOUT UNIQUE MOLECULAR IDENTIFIERS (UMI)

Kevin D McCormick; Kerri J Penrose; Rahil Sethi; Jacob Waldman; William Schwarzmann; Uma Chandran; John W Mellors and Urvi M Parikh

27th International Workshop on HIV Drug Resistance and Treatment Strategies

October 22, 2018

University of Pittsburgh, Pittsburgh, PA, USA

MTN microbicide trials network

Sensitive, Scalable and Affordable Assays are needed for HIV Drug Resistance Testing

	Sanger Sequencing	Next Generation Sequencing (NGS)
Sensitivity	Cannot detect minor variants (>20%)	High sensitivity for minor variant detection (**?**<20%)
Scalability	Low throughput Limited potential for automation	Sample multiplexing - potential for automation
Cost	High cost per sample (>100 USD)	Reduced cost per sample (<50 USD)
Validation	Widely Validated; adapted into kits	Sensitivity, specificity and reproducibility unknown

(Table adapted from Inzaule et al. The Lancet. 2016)

The effects of sequencing error and variation of sampling depth on the accuracy of mutant detection by NGS are not well defined.

Unique Molecular Identifiers (UMIs) may be used to reduce sequencing artifacts and quantify the true sampling depth.

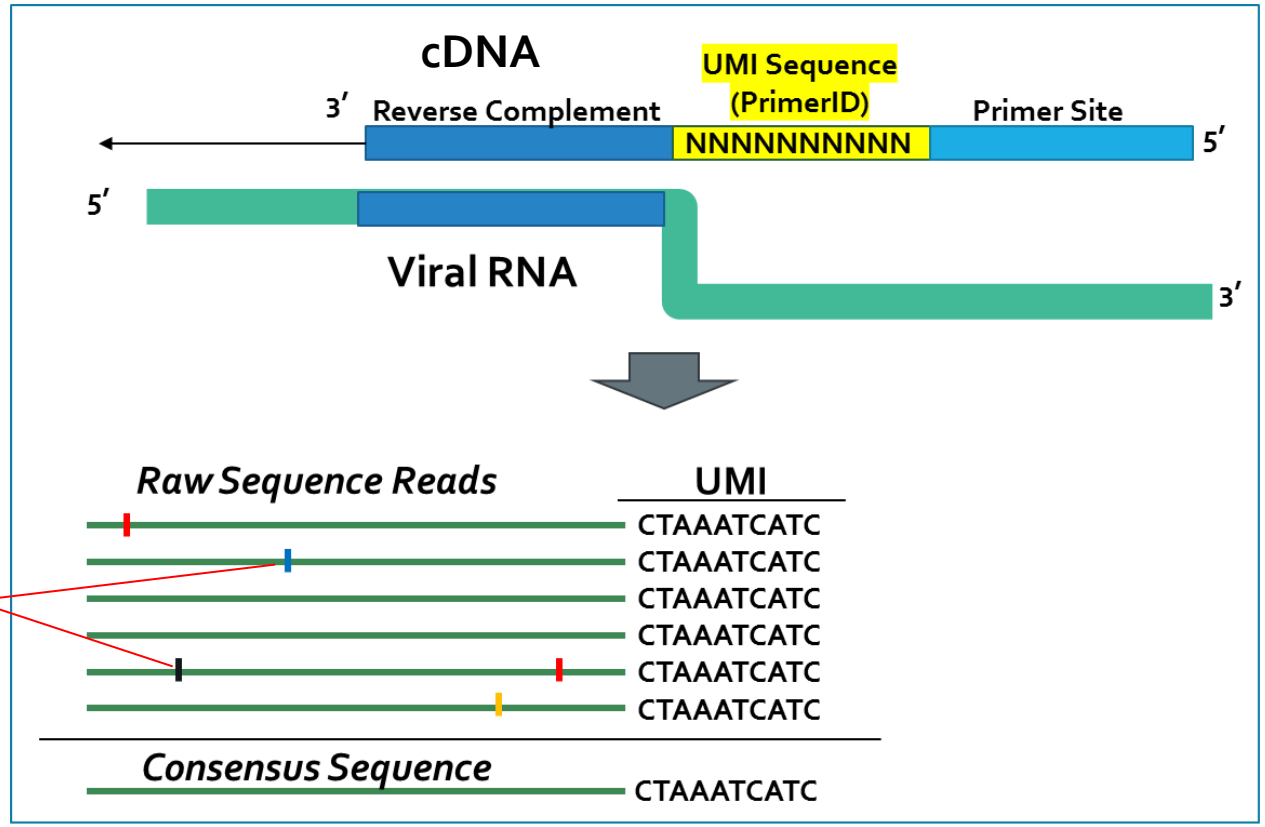


Figure adapted from https://tcs-dr-dept-tcs.cloudapps.unc.edu/ (Jabara, et al. PNAS. 2011 and Zhou et al. JVI 2015)

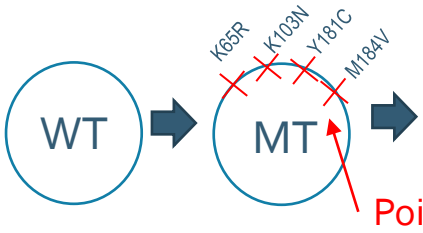
OBJECTIVE

Perform a comparative assessment of UMI- and non-UMI-based NGS assays for mutant detection and confident analysis of low frequency HIV DR.

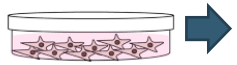
1) **The sensitivity and specificity**
2) **The limits of detection**
3) **The accuracy and precision with clinical samples**

METHODS: Design of the Recombinant HIV Drug Resistant Mixture Panel

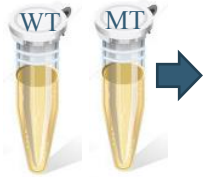
xxLAI-
X
RT-Subtype C
Plasmids



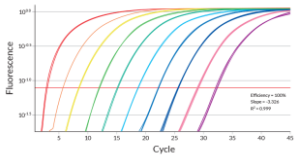
Transfect
293T



Harvest
Virions



Sybr Green
qPCR
Quantification



MT:WT Mixtures

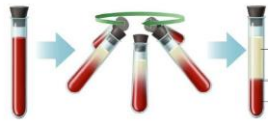
100%
20%
10%
5%
2.5%
1.0%
0.5%
0%

1% and 5%
Serial Dilution
(copies/ml)

1E6
1E5
1E4
1E3
1E2
1E1

Resuspend mixtures with
HIV-1 Negative
Donated Blood

Isolate Plasma
by Centrifuge



1) Virion Production

2) Virion Quant
(10^9 copies/ml)

3) Mix MT:WT
(50,000 copies)

4) Serial
Dilution

5) Blood reconstitution
Plasma Isolation

METHODS: NGS Library Prep and Data Analysis

All NGS libraries were constructed from the RT region of HIV-1 using the ultrasensitive single-genome sequencing method (Boltz et al. 2016).

non-UMI-based NGS Analysis

- **PASeq v1.4 (https://www.paseq.org)**
- **10,000 raw read coverage with >30 Miseq quality filter**
- **Minor allele frequencies were reported from the "Amino acid Variants .CSV" file**

UMI-based NGS Analysis

- **Zhou method was used for consensus building. (Zhou et al. JVI 2015)**
- **80% homologous UMI binned reads required for consensus; allowing for 1 mismatched codon**
- **298 and 54 UMI consensus sequences were required to call 1% and 5% minor variants respectively (95% confidence)**

Comparable HIV-1 drug resistance mutation detection in both UMI and non-UMI NGS (Sensitivity)

HIV RNA Input = 50,000 copies/ sample

Expected Allele Frequency (% MT)	non-UMI-based NGS detected allele frequency (%)				UMI-based NGS detected allele frequency(%)			
	K65R	K103N	Y181C	M184V	K65R	K103N	Y181C	M184V
100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20.0	28.0	28.0	27.0	27.0	20.3	20.3	20.3	20.3
10.0	14.6	14.6	14.6	14.6	8.3	8.3	8.3	8.3
5.0	7.5	7.6	7.8	7.8	3.7	3.7	3.7	3.7
2.5	3.4	3.4	3.5	3.5	1.6	1.6	1.6	1.6
1.0	1.3	1.3	1.5	1.5	0.5	0.5	0.5	0.5
0.5	0.7	0.6	0.6	0.6	0.3	0.3	0.3	0.3
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Similar false positive mutation rates in UMI and non-UMI-based Wildtype samples (Specificity)

False positive Mutation Frequencies	non-UMI-based NGS (n=8)	UMI-based NGS (n=17)
Ave Number	482	68
Mean	0.06%	0.08%
Median	0.02%	0.05%
Min	0.01%	0.01%
Max	0.46%	0.46%

Higher false negatives result with non-UMI NGS at lower viral RNA input and UMI NGS at high RNA input

Expected (% MT)	Viral Load (cp/ml)	non-UMI-based NGS % Detected			UMI-based NGS % Detected			# UMI
		K103N	Y181C	M184V	K103N	Y181C	M184V	
5%	1.00E+06	3.8	3.8	3.8	0.6	1.5	1.5	4308
	1.00E+05	4.0	4.0	4.0	1.3	3.2	3.2	24991
	1.00E+04	3.9	4.0	4.0	4.0	4.7	4.7	2149
	1.00E+03	4.8	4.8	4.8	3.1	3.1	3.1	261
	1.00E+02	0.8	0.9	0.9	**<BCO**	**<BCO**	**<BCO**	**16**
	1.00E+01	**0[a]**	**0[a]**	**0[a]**	**<BCO**	**<BCO**	**<BCO**	**0**
1%	1.00E+06	0.9	0.9	0.9	0[a]	0[a]	0[a]	5010
	1.00E+05	0.9	0.8	0.8	0[a]	0.7	0.7	20848
	1.00E+04	0.8	0.8	0.8	0.6	0.7	0.7	2325
	1.00E+03	1.6	1.6	1.6	0.8	0.8	0.8	266
	1.00E+02	**0[a]**	**0[a]**	**0[a]**	**<BCO**	**<BCO**	**<BCO**	**19**
	1.00E+01	**0[a]**	**0[a]**	**0[a]**	**<BCO**	**<BCO**	**<BCO**	**3**

Below Biological Cutoff (<BCO)

0[a] = false negative detection; **<BCO** = below biological cutoff (insufficient UMI counts)

non-UMI NGS false negative detection with inefficient PCR from polymorphic primer binding sites

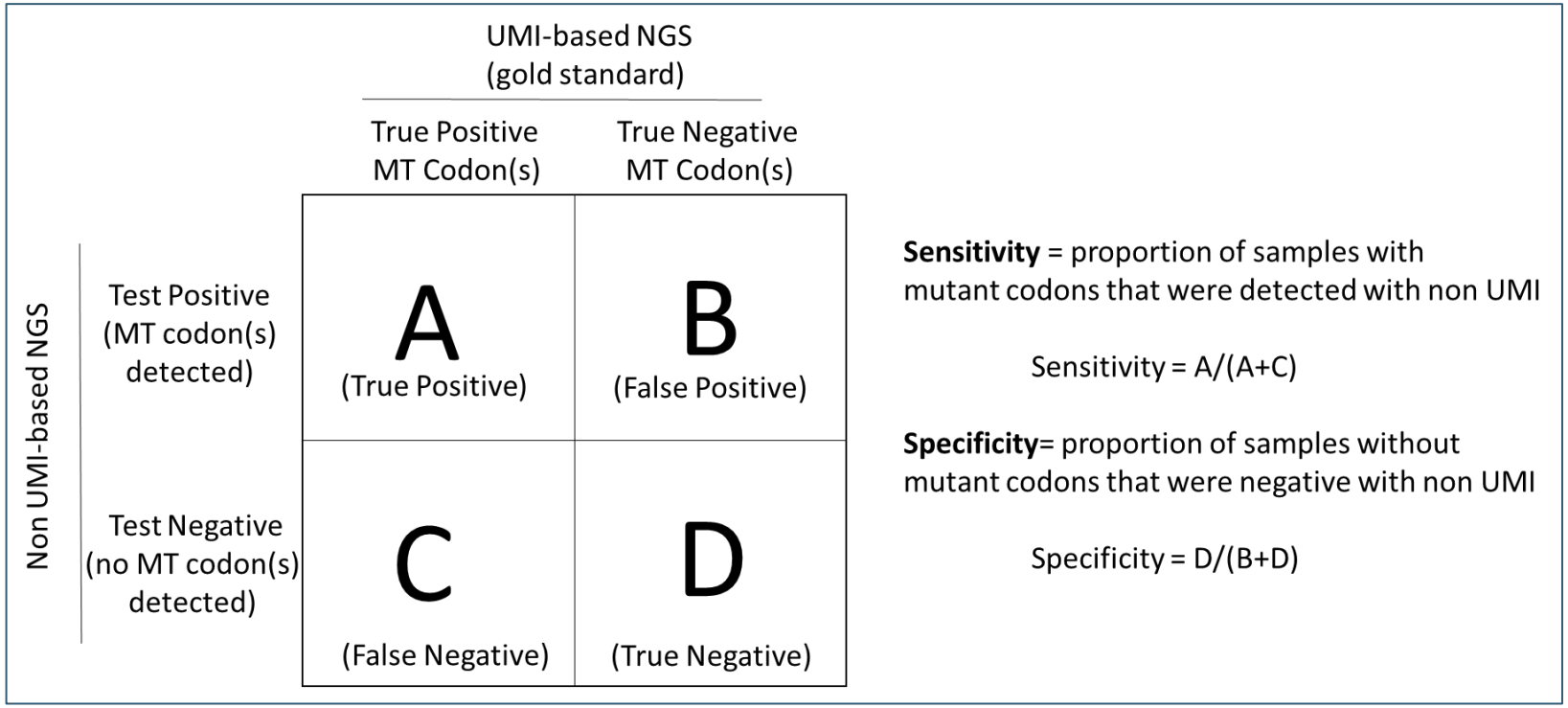
HIV RNA Input = 10,000 copies/ sample

# of PCR FWD primer mismatches	Expected Allele Frequency (% MT)	non-UMI-based NGS detected allele frequency (%)	UMI-based NGS detected allele frequency (%)	# UMI
0	5.0	7.0	4.0	1213
	1.0	1.5	0.9	1447
2	5.0	7.7	10.7	215
	1.0	0[a]	**<BCO**	248

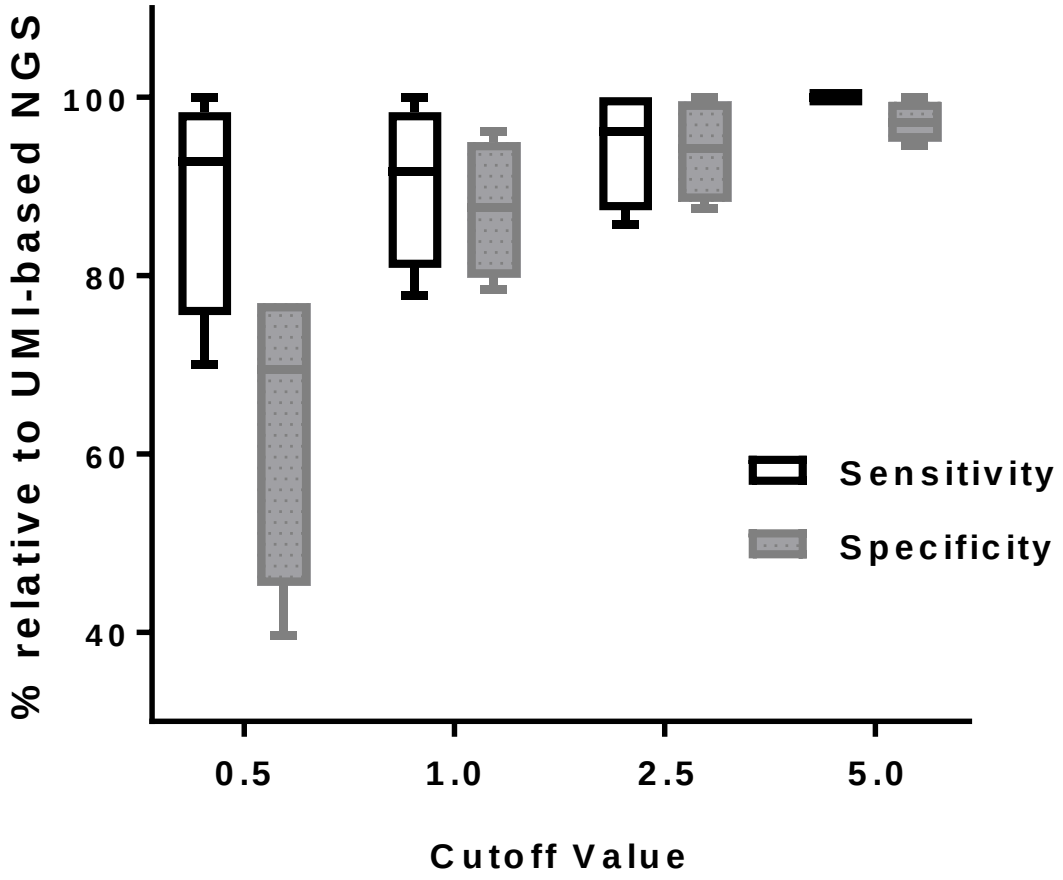
0[a] = false negative detection; **<BCO** = below biological cutoff (insufficient UMI counts)

METHOD: Sensitivity and specificity of non-UMI NGS analysis vs UMI-based consensus building.

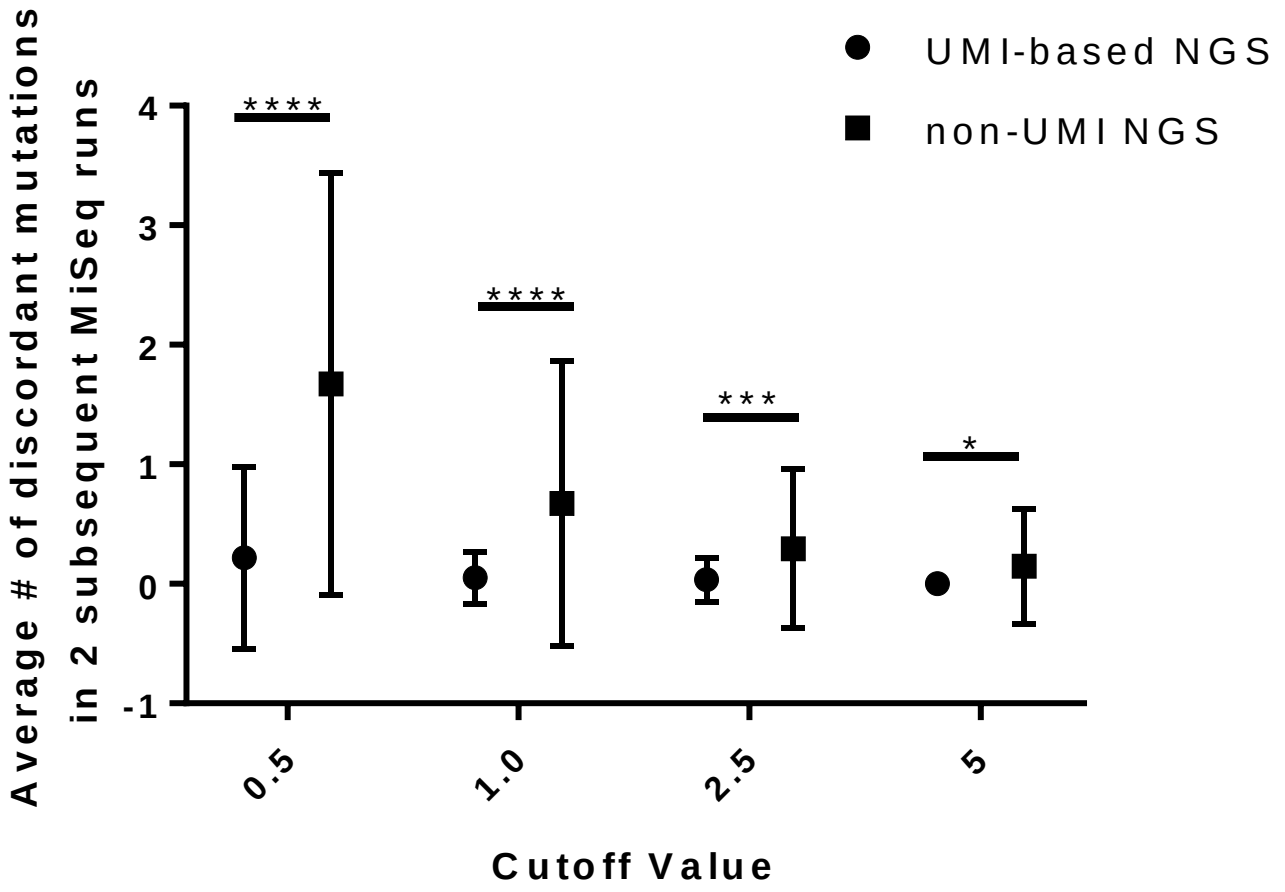
- We re-analyzed <u>without</u> UMIs, a UMI generated clinical dataset.
  - 62 plasma samples from viremic donors with HIV acute infection.



non-UMI-based NGS had similar sensitivity but reduced specificity in the re-analyzed clinical dataset.



Lower reproducibility of non-UMI-based NGS analysis



Note: The same samples were run on MiSeq twice

SUMMARY

- Both UMI and non-UMI methods detected mutants at 0.5% in an isogenic background at high template input (50,000 copies), indicating that both methods are generally robust.
- False negative error rates are higher in non-UMI-based NGS with limited sampling depth (<1000 cp/ml or PCR bias).
- Non-UMI mutant detection at <5% was not reproducible for clinical samples.

IMPLICATIONS

- UMI-based consensus building should be used if calling mutations at frequencies below 5%.
- This is particularly true for samples that are likely to have low template input or inefficient PCR.
  - Low plasma HIV RNA.
  - Dried blood spots!
  - Polymorphic primer binding sites (always a concern).

ACKNOWLEDGMENTS -- THANK YOU!

- Urvi M. Parikh, PhD
- Kerri Penrose, MS
- John W. Mellors, MD
- Amy Heaps, MS
- Breanna Goetz, BS

- Uma Chandran, PhD
- Rahil Sethi, MS
- Jacob Waldman, MS
- William Schwarzmann, BS

- Carole Wallis, PhD
- Raquel Viana, PhD

- Marc Noguera-Julian, PhD

This presentation was made possible by the generous support of the American people through the United States Agency for International Development (USAID) and the U.S. President's Emergency Plan for AIDS Relief (PEPFAR). The contents are the responsibility of the authors and do not necessarily reflect the views of USAID, PEPFAR or the U.S. Government.

We would like to acknowledge the Microbicide Trials Network for the use of the UMI-NGS dataset that was derived from plasma samples from viremic donors with HIV acute infection.