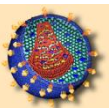


Sequencing Error Assessment and Reduced PCR-based Recombination for HIV-1 Drug Resistance Mutation Linkage Determination by 454 Pyrosequencing

Wei Shao

Advanced Biomedical Computing Center
SAIC Frederick, Inc
NCI-Frederick
Frederick, Maryland, USA



Question 1: PCR mediated recombination

- 1. Does PCR-based recombination affect 454 sequencing?**
- 2. What percent of recombinants are produced by the PCR step?**
- 3. Method to reduce PCR mediated recombination**

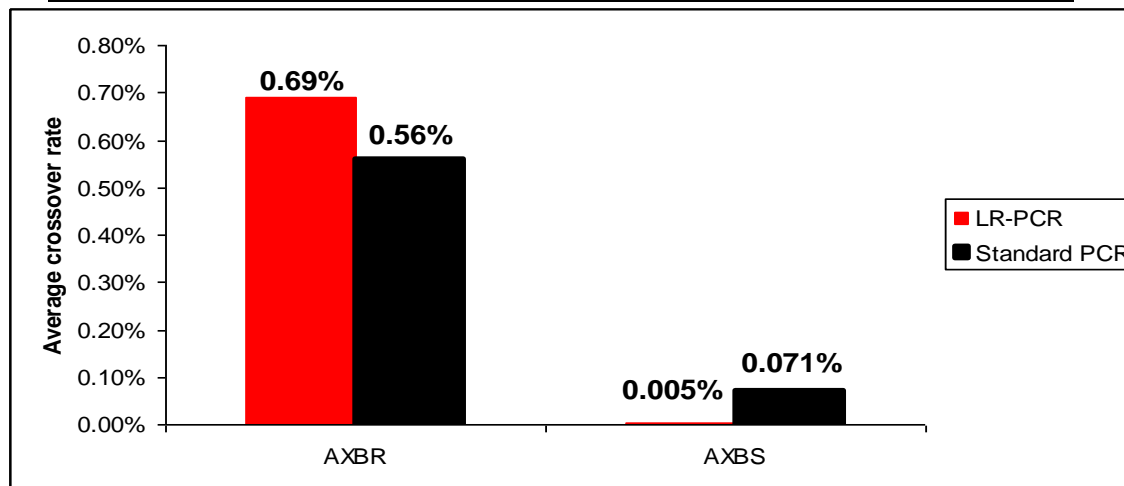
Question 2: Point/Indel sequencing errors

- 1. Percent of sequencing errors**
- 2. Distribution of sequencing errors**
- 3. Sources of sequencing errors**

Recombinant Detection: Low recombination PCR (LR-PCR) reduces recombinant rate by 27 fold

	LR-PCR	Standard PCR	Reduction	p value
Recombinants	0.43%	11.65%	27 fold	
Point errors	0.084+/-0.125%	0.123+/-0.150%		<0.01
Indel errors	0.299+/-1.717%	0.325+/-1.615%		<0.05

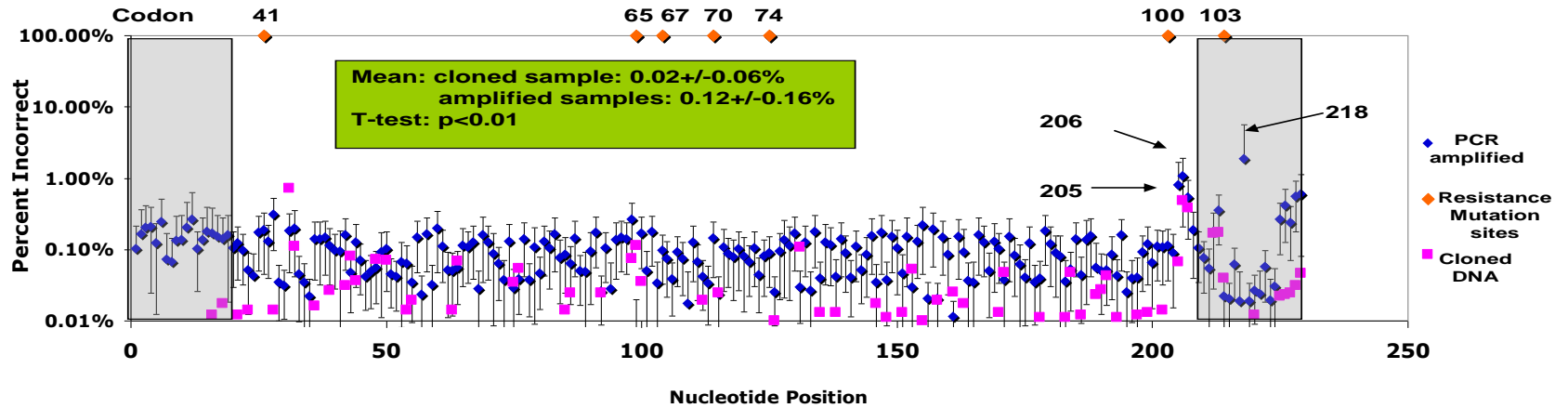
Fewer parental molecules are involved in LR-PCR mediated crossovers



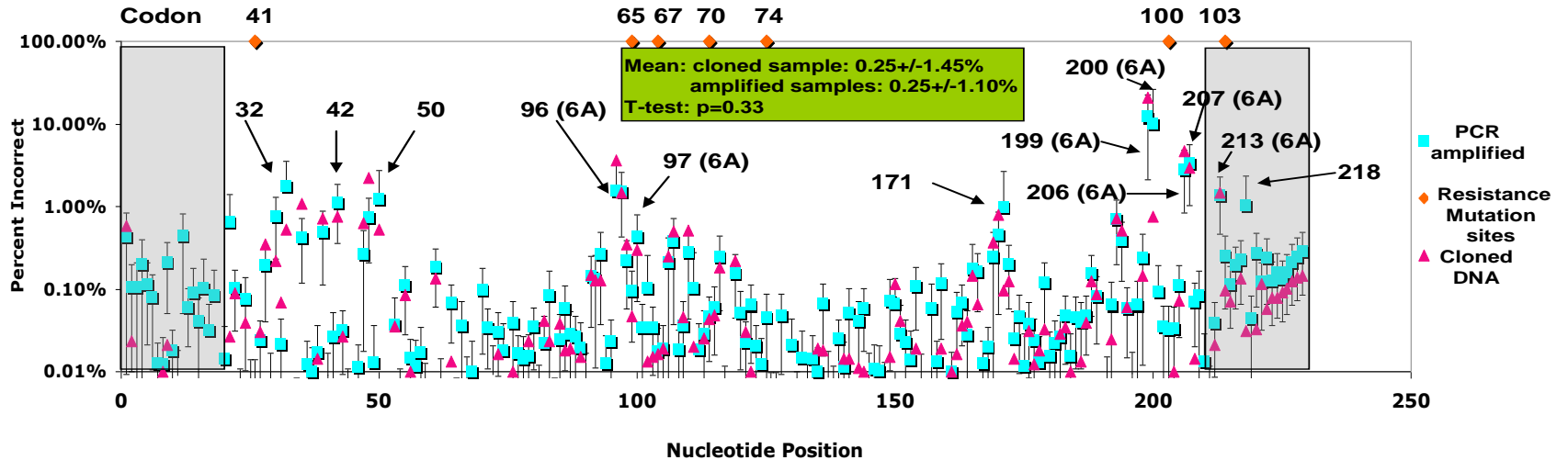
AXBR: Average Crossover/Base/Recombinant

AXBS: Average Crossover/Base/Sequence

PCR amplified samples vs. cloned sample: a large fraction of point errors occurred during the PCR step



PCR amplified samples vs. cloned sample: a large fraction of indel errors occurred during the sequencing step



Transition error rate > Transversion error rate

Correct Base	Read As:				Total
	A	C	G	T	
Run1					
A		0.016%	0.096%	0.018%	0.130%
C	0.006%		0.011%	0.036%	0.053%
G	0.087%	0.010%		0.006%	0.103%
T	0.027%	0.083%	0.020%		0.130%
Run2					
A		0.005%	0.069%	0.016%	0.090%
C	0.002%		0.002%	0.026%	0.031%
G	0.048%	0.002%		0.006%	0.055%
T	0.010%	0.051%	0.014%		0.075%

Errors at drug resistance sites

Position	Mutation	Run1	Run2
21	M41L(A->C)	0.0000%	0.0103%
94	K65R(A->G)	0.3057%	0.0861%
99	D67W(G->A)	0.0706%	0.0185%
109	K70R(A->G)	0.0941%	0.0595%
120	L74V(T->G)	0.0000%	0.0021%
198	L90I(T->A)	0.0000%	0.0000%
209	K103N(A->C)	0.0000%	0.0082%
210	K103N(A->T)	0.0000%	0.0185%

Conclusions

- 1) 454 sequencing of PCR amplified DNA is error prone. The errors are produced both in the sequencing step and the PCR step.**
- 2) Our analyses demonstrate that 454 sequencing can reliably detect HIV-1 resistance mutations at frequencies down to 1% (or much less in specific cases).**
- 3) 454 sequencing coupled with standard PCR can produce artifactual recombinant sequences between mutant and wt as high as 14%.**
- 4) With low recombination PCR, the percentage of artifactual recombinant sequences can be reduced to 0.43%. This significant reduction makes 454 sequencing technology a potential tool to study mutation linkage and construct haplotypes.**

Acknowledgements

NCI HIV Drug Resistant Program

John Coffin
Valerie Fliakas Boltz
Jon Spindler
Mary Kearney
Frank Maldarelli
Ann Wiegand
Helene Mens

Division of Infectious Disease University of Pittsburgh

John Mellors

LMT, SAIC Frederick

Claudia Stewart

ABCC, SAIC Frederick

Alexander Levitsky
Natalia Volfovsky
Robert Stephens

