Characterization of Rectal Transmission Bottleneck Using SIV-Macaque Model

Zhe Yuan[a], Fangrui Ma[a], Andrew J. Demers[a], Dong Wang[b], Jianqing Xu[c], Mark G. Lewis[d], Qingsheng Li[a]

a, Nebraska Center for Virology, School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE, USA. b, Dow AgroSciences, LLC, Indianapolis, IN, USA.
c, Shanghai Public Health Clinical Center and Institutes of Biomedical Sciences, Fudan University, Shanghai, China. d, BIOQUAL, Inc., Rockville, MD, USA

Abstract Presentation Number: 199

ABSTRACT

A better understanding of HIV-1 transmission is critical for developing preventative strategies. To that end, we analyzed 524 full-length env sequences of SIVmac251 at 6 and 10 days post intrarectal infection of rhesus macaques. We found that there was no tissue compartmentalization of founder viruses across plasma, rectal and distal lymphatic tissues for most animals; however one animal had evidence of viral tissue compartmentalization. Despite identical viral inoculums, founder viruses were animal-specific, primarily derived from rare variants in the inoculum, and have a founder virus signature that can distinguish dominant founder variants from minor founder or untransmitted variants in the inoculum. Importantly, the sequences of post-transmission defective viruses were phylogenetically associated with competent viral variants in the inoculum and were mainly converted from competent viral variants by frameshift rather than APOBEC mediated mutations, suggesting the converting the transmitted viruses into defective viruses through frameshift mutation is an important component of rectal transmission bottleneck.

EXPERIMENTAL DESIGN

SIV macaque model

Male adult rhesus macaques of Indian origin were intrarectally inoculated with cell-free SIVmac251 at a dose of 3.1 x10^4 $TCID_{50}$ and euthanized at 6 (n=3) and 10 days (n=3) post infection (dpi).

Single genome amplification of full-length env

The 2.7 kb full length SIVmac251 env from rectal tissues, axillary lymph node tissues, and blood was amplified using SGA as previously reported[5] and sequenced using Sanger sequencing method.

Sequence analysis and statistical analysis

The env sequences passed quality control were assembled to obtain the full-length env individually using Sequencher 5.0 (Gene Codes Inc., Ann Arbor Michigan). Phylogenetic analysis on both nucleotide and amino acid sequences of founder viruses and in inoculum were performed using maximum likelihood method in PHYML.

Pearson's chi-squared test and probability theory were used for statistical analysis. Fisher's method was used to combine the results from the six monkeys to form a single p-value for the test.

RESULTS

Figure 1. Maximum likelihood amino acid tree and the dominant cluster of founder viruses across multiple tissue compartments at 10 dpi (Rh4976).

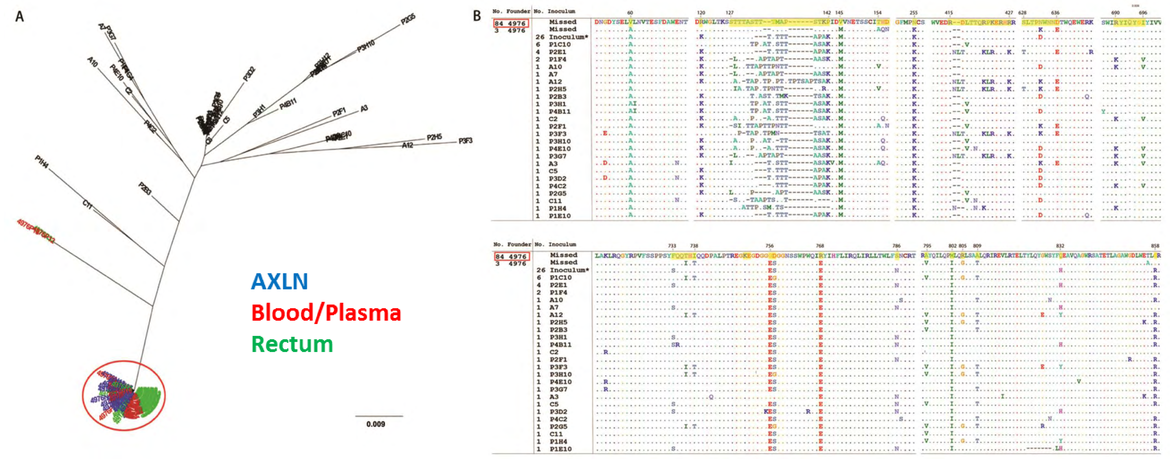


Table 1. Test of tissue compartmentalization.

Subjects	Days post inoculation (dpi)	Test of Compartmentalization				Compartmentalization
		F_{st}			S_{nn}	
		HSM	S	HBK		
4809 N vs P	6	0.482	0.482	0.482	0.354	No
4809 N vs R	6	0.993	0.993	0.993	0.02	No
4809 P vs R	6	0.117	0.117	0.117	0.113	No
4976 N vs P	10	0.729	0.729	0.729	0.162	No
4976 N vs R	10	0.396	0.396	0.396	0.039	No
4976 P vs R	10	0.168	0.168	0.168	0.762	No
4978 N vs P	10	0.385	0.385	0.385	0.002	No
4978 N vs R	10	0.458	0.448	0.422	0.299	No
4978 P vs R	10	0.642	0.642	0.641	0.104	No
5053 N vs P	6	0.086	0.086	0.086	0.131	No
5053 N vs R	6	0	0	0	0	Yes
5053 P vs R	6	0	0	0	0	Yes

N – lymph node, P – plasma, and R – rectal; HSM – Hudson, Slatkin and Maddison, S – Slatkin, and HBK – Hudson, Boos and Kaplan; 5052 or 4979 sequences from some tissue type cannot reach statistical power.

Figure 2. The animal-specific founder virus signatures (ASFVS).

A

Animal ID	ASFVS	Glycoprotein	Region	Note	
Rh4976	^{68}V	GP120	C1		3D Structure available
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{126}K/R		V1	O-linked glycan attachment site	
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{127}STTTASTTTTAPAK^{141} ★				
Rh4976	^{145}V				
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{155}THD^{155} ★				
Rh4976	^{214}I		C2		
Rh4976	^{255}N				
Rh4809	^{337}D		V3		
Rh4978	^{342}S				
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{414}RDLLTQPPKERHR^{426} ★		V4		
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{627} (S/T) LTP (K/D/N) WNN (D/E) ^{635}	GP41	Loop linking GP120	one of 23 N-glycosylation positions	
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{680}RYIQYGI^{695} ★		MPER	GalCer binding	3D Structure unavailable
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{732}(F/S) QQ (I/T) H (T/I) ^{737}		cytosolic tail		
Rh4976 Rh4979	^{756}GD^{757}			homologous position with Kennedy epitope	
Rh4976	^{768}R				
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{785}S/N				
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{794}V/A				
Rh4976	^{802}M			Lentivirus Lytic Peptide (LLP)	
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{804}G/R				
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{808}T/A				
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{831}Q/H				
Rh4809 Rh5052 Rh5053 Rh4976 Rh4978 Rh4979	^{855}AL(G/R)^{857}				

B

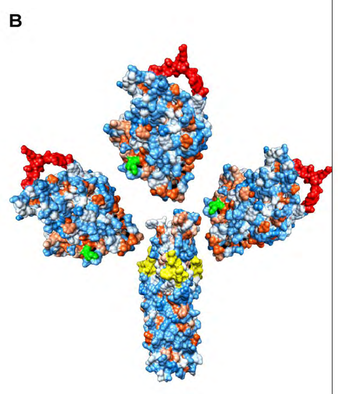
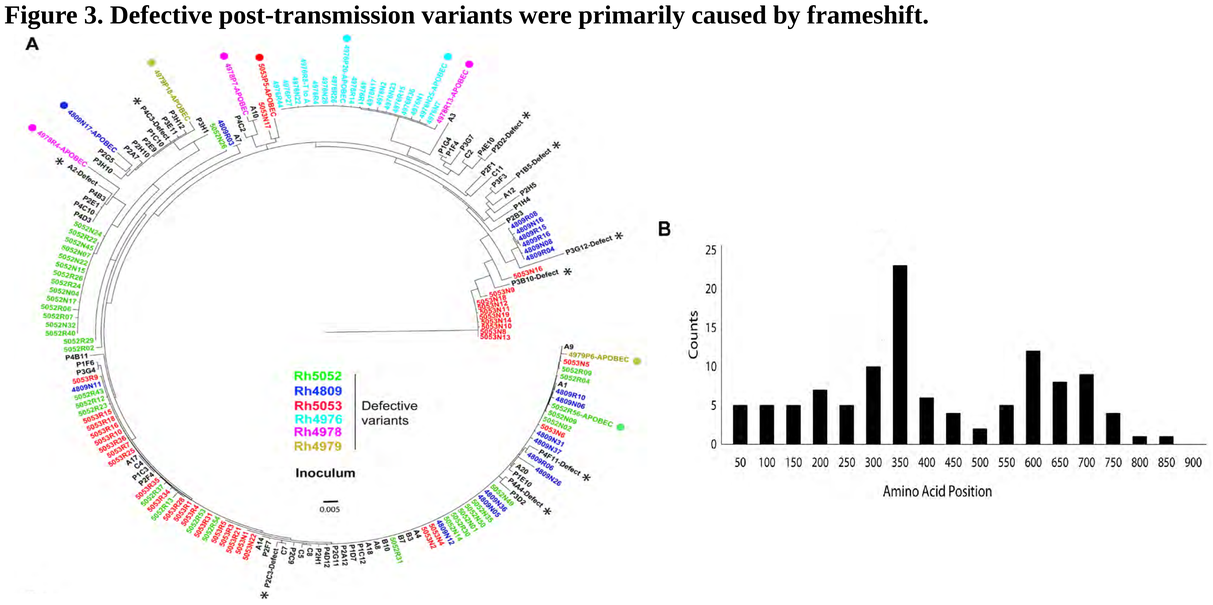


Table 2. The distribution of functional and defective variants and the causes of defects.

	Animal ID	Total sequences	Functional variants (%)	Defective variants (%)	Defective caused by frameshift (%)	Defective caused by APOBEC (%)
6 dpi	4809	74	75.68	24.32	22.97	1.35
	5052	118	69.49	30.51	29.66	0.85
	5053	70	50.00	50.00	48.57	1.43
10 dpi	4976	105	82.86	17.14	14.29	2.86
	4978	53	94.34	5.66	0.00	5.66
	4979	36	94.44	5.56	0.00	5.56

Figure 3. Defective post-transmission variants were primarily caused by frameshift.



CONCLUSIONS

- **There is no tissue compartmentalization of founder viruses in most of studied animals at very early infection.** Statistical analyses using three independent methods all indicate there is no virus tissue compartmentalization for most of studied animals except an animal of 6 dpi.
- **Founder viruses are mainly derived from rare variants in the inoculum.** For all six macaques, nine dominant founder virus groups were identified. Six dominant founder virus groups came from rare variants in the inoculum (66.6% derived from 35.6%), whereas only three dominant founder viruses came from the high-abundant group in the inoculum (33.3% derived from 64.4%). Our statistical analyses show that the abundance of founder viruses for each animal is highly independent of their ancestor variant proportion in the inoculum (p=3.3E-16).
- **Founder viruses are animal-specific with founder virus signatures (FVS).** The locations and composition of founder virus sequences among macaques were not phylogenetically shared. There are significant differences in the composition of founder variants among animals at 6 dpi (p=1.4e-8), among animals at 10 dpi ($p < 2.2e-16$), and between the two groups of macaques at 6 and 10 dpi ($p < 2.2e-16$). FVS can distinguish dominant founder variants from minor founder variants and untransmitted variants in the inoculum for each animal.
- **Sequences of post-transmission defective virus variants were mainly caused by frameshift rather than APOBEC mediated mutations.** the ratio of defective to functional founder variants significantly decreased from 6 dpi (34.9 ± 13.4%) to 10 dpi (9.5 ± 6.6%) (p=1.1e-7). Frameshift mutations were detected in the majority of defective variants at 6 dpi, but were significantly decreased at 10 dpi (p=6.0e-10) corresponding with the overall decrease of defective variants.

REFERENCES

1. Edwards C, et al. (2006) BMC Evolutionary Biology 6: 28.
2. Abrahams M-R, et al. (2009) Journal of Virology 83: 3556-3567.
3. Boeras DI, et al. (2011) Proceedings of the National Academy of Sciences 108: E1156–E1163.
4. Keele BF, Estes JD (2011) Blood 118: 839-846.
5. Stone M, et al. (2010) Journal of Virology 84: 7083-7095.

ACKNOWLEDGMENTS

We thank Dr. Brandon Keele for SGA technical help, Drs. James Mullins, Lance Daharsh, and Dr. John West for discussion and critical reading of the manuscript. We thank Dr. Nancy Miller for providing the SIVmac251 stock. NIH supports via NIDDK R01 DK087625-01 (Q. Li), NIH NIAID R01 AI111862-01 (Q. Li and J. Guo).