

Table S1. Performance of different BLASTN protocols when searching simulated HIV reads.

	Default megablast	Default blastn	Blastn sensitive	Blastn w7
Runtime				
Real time (s)	4.28	12.38	49.26	569.69
Sensitivity (%)				
HIV starting genome reads	100	100	100	100
<i>10% divergent reads</i>				
replicate 1	46	100	100	100
replicate 2	54	100	100	100
replicate 3	54	100	100	100
<i>20% divergent reads</i>				
replicate 1	6	96	100	100
replicate 2	7	99	100	100
replicate 3	9	98	100	100
<i>30% divergent reads</i>				
replicate 1	3	73	95	100
replicate 2	0	77	96	97
replicate 3	0	76	94	100
<i>40% divergent reads</i>				
replicate 1	0	34	69	98
replicate 2	0	36	67	90
replicate 3	0	34	69	90
<i>50% divergent reads</i>				
replicate 1	0	15	48	76
replicate 2	0	12	46	77
replicate 3	0	15	40	67

BLASTN protocols are named following the text in online supplementary File S1. Run-time was the execution time recorder for a BLASTN search (for each protocol) of 100 reads simulated from the starting HIV genome in the viral genomes database. Sensitivity (%) is defined as the percentage of reads from each simulation that effectively aligned to an HIV genome, using the specific BLASTN protocol, when searched in the viral genomes database.