

# LTR Parser Results

Kevin McCormick

22 June, 2020

## Overview

This report contains the results of attempted primer & LTR identification in the following FASTQ file:

/home/kevin/dev/ltrparser/testdata/demo/fastq.gz/hivTest8\_u3.R2.fastq.gz

ODoherty project, sample: GTSP2314b (U3)

## Analysis of common primer-LTR pairs

A total of 5 potential primer-LTR pairs were identified.

This table shows how the reads captured by each primer-LTR pair compare to each other and to the raw FASTQ file.

id	primer	LTR	expectedLTR	LANL
1	TGTGTGGT	AGACCCACAGGTCAAGGATCTCTTGCTCTTTTGGGAGTAAATTAGCCCTTCCA	FALSE	TRUE
2	TGTGTGGT	AGACCCACAGGTCAAGGATCTCTTGCTCTTTTGGGAGTAAATCAGCCCTTCCA	FALSE	TRUE
3	TGTGTGGT	AGACCCACAGGTCAAGGACCTCTTGCTCTTTTGGGAGTAAATTAGCCCTTCCA	FALSE	TRUE
4	TGTGTGGT	AGACCCAAAGGTCAAGGATTCTTGCTCTTTTGGGAGTAAATTAGCCCTTCCA	FALSE	TRUE
5	TGTGTGGT	AGACCCACAGGTCAAGGATCTCTTGCTCTTTTGGGAGTAAATTAGCCCTTCCACTTGTC	FALSE	TRUE

Variable	Total	1	2	3	4	5
Distinct	9181	6870 (74.8%)	6371 (69.4%)	5292 (57.6%)	2437 (26.5%)	937 (10.2%)
DistinctHuman	1393	1192 (85.6%)	1149 (82.5%)	859 (61.7%)	316 (22.7%)	133 (9.5%)
DistinctLinker	1944	1243 (63.9%)	1155 (59.4%)	879 (45.2%)	282 (14.5%)	0 (0%)
Max	914	914 (100%)	914 (100%)	914 (100%)	914 (100%)	164 (17.9%)
MaxHuman	576	576 (100%)	576 (100%)	576 (100%)	144 (25%)	14 (2.4%)
MaxLinker	576	576 (100%)	576 (100%)	576 (100%)	203 (35.2%)	0 (0%)
Total	23172	17197 (74.2%)	16162 (69.7%)	14834 (64%)	7421 (32%)	2235 (9.6%)
TotalHuman	3275	3055 (93.3%)	2853 (87.1%)	2475 (75.6%)	799 (24.4%)	183 (5.6%)
TotalLinker	4593	3160 (68.8%)	2931 (63.8%)	2586 (56.3%)	709 (15.4%)	0 (0%)

## Visualization of reads

Reads are sectioned off into sections separated by white (nucleotide color legend is below the plot).

Primer /// LTR /// CA /// (unmatched) /// human-matching /// linker-matching /// (unmatched)





