

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/hivTest1_u5.R2.fastq.gz

Siliciano project, sample: GTSP3034-4 (U5)

Analysis of common primer-LTR pairs

A total of 2 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.

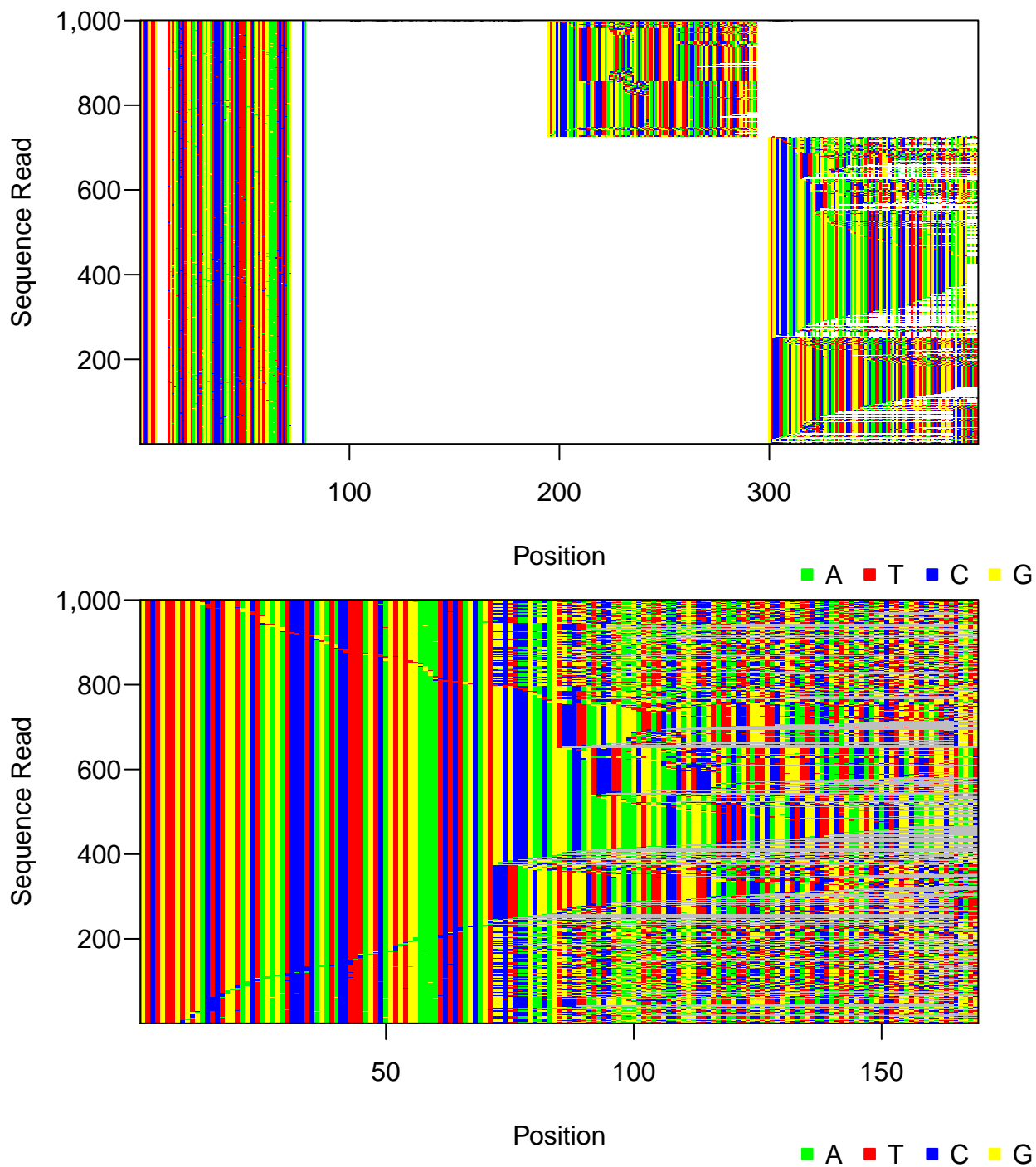
id	primer	LTR	expectedLTR	LANL
1	GTCTGTTG	TGTGACTCTGGTAGCTAGAGATCCCTCAGATACCTTTAGTCAGTGTGGAAAATCTCTAGCA	FALSE	TRUE
2	GTCTGTTG	TGTGACTCTGGTAGCTAGAGATCCCTCAGATACCTTTAGTCAGCGTGGAAAATCTCTAGCA	FALSE	TRUE

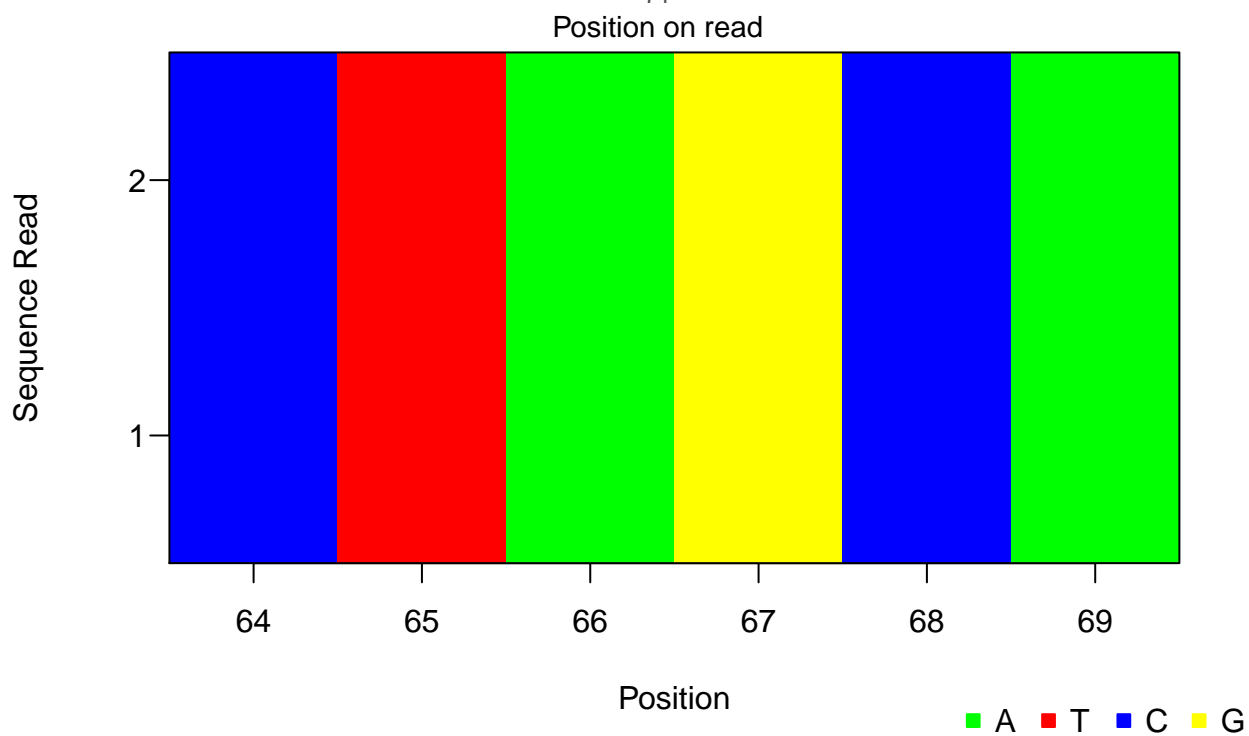
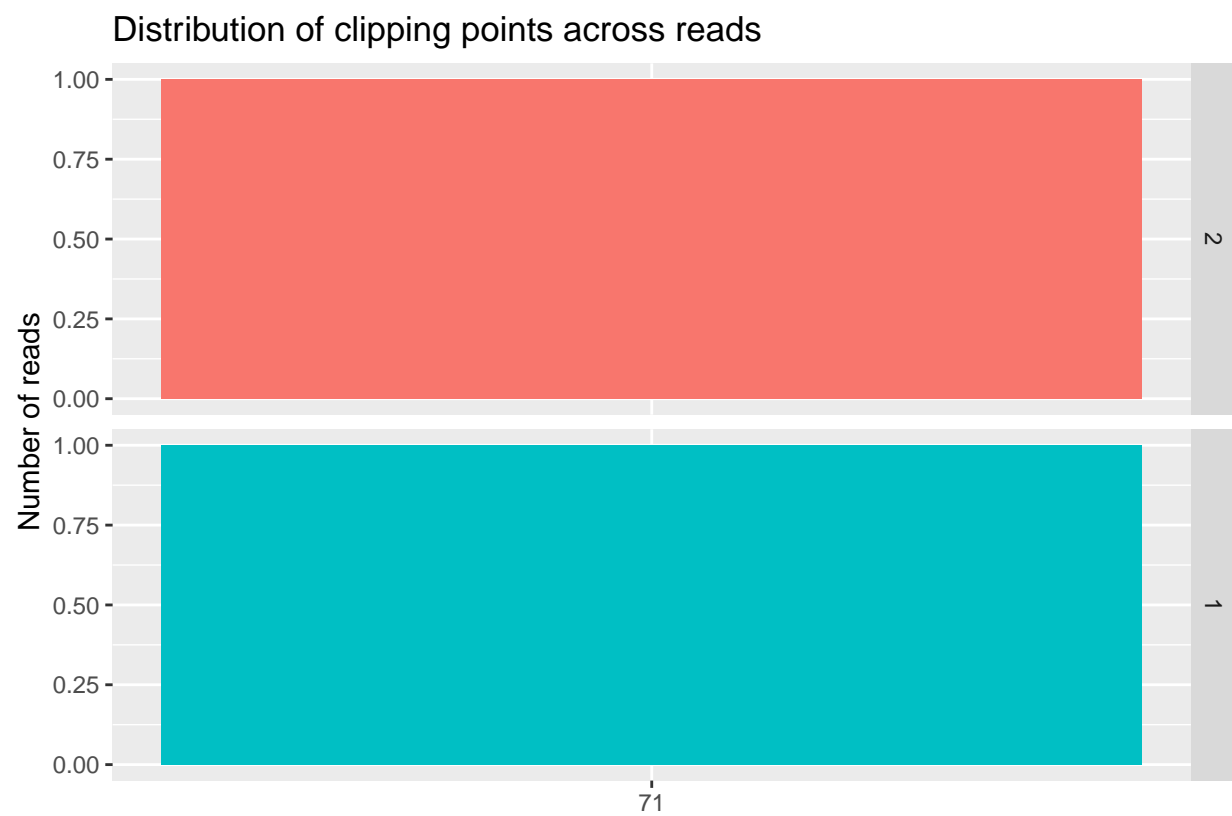
Variable	Total	1	2
Distinct	6669	2413 (36.2%)	2265 (34%)
DistinctHuman	79	3 (3.8%)	3 (3.8%)
DistinctLinker	2421	642 (26.5%)	590 (24.4%)
Max	1349	1349 (100%)	1349 (100%)
MaxHuman	4	1 (25%)	1 (25%)
MaxLinker	855	655 (76.6%)	655 (76.6%)
Total	22660	8541 (37.7%)	8386 (37%)
TotalHuman	89	3 (3.4%)	3 (3.4%)
TotalLinker	10438	2499 (23.9%)	2440 (23.4%)

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)







 non conserved
 $\geq 50\%$ conserved