# 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/hivTest3_u5.R2.fastq.gz

ODoherty project, sample: GTSP3456-1 (U5)

## Analysis of common primer-LTR pairs

A total of 4 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.
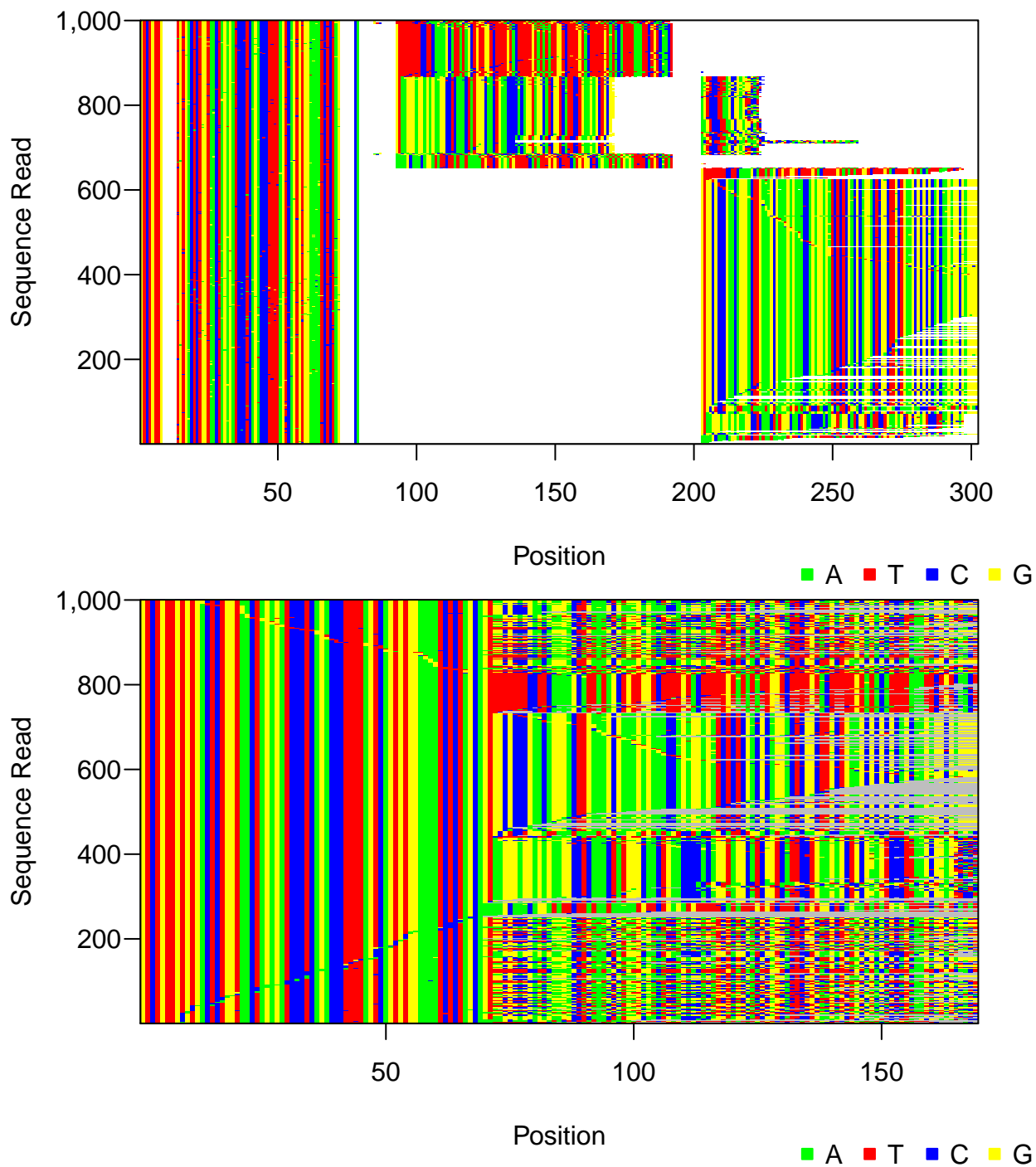
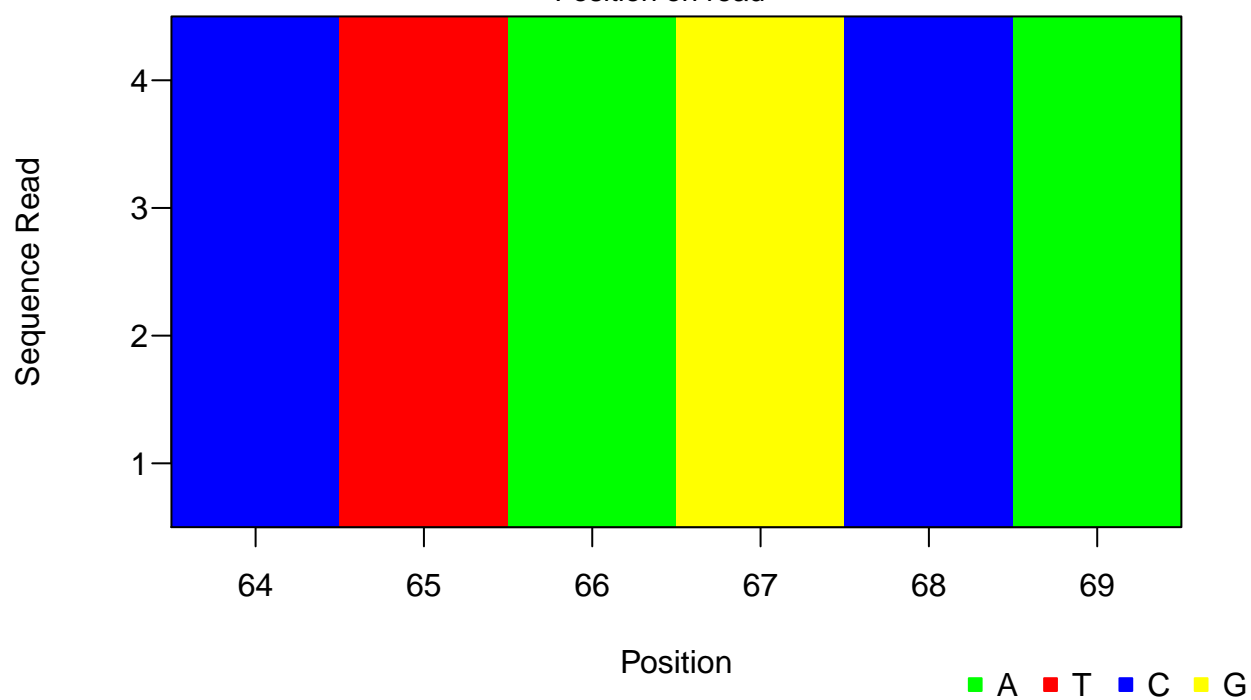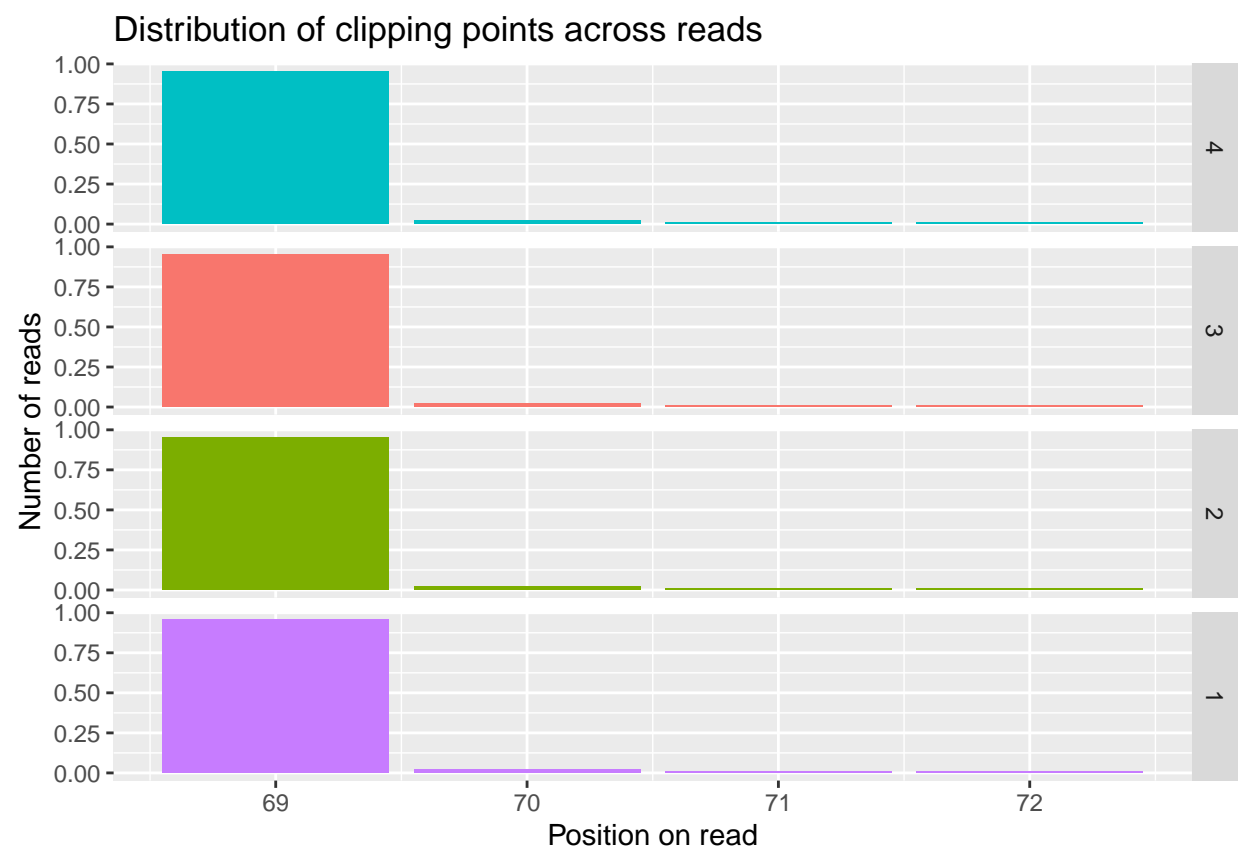| id | primer | LTR | expectedLTR | LANL |
|----|--------|-----|-------------|------|
| 1 | GTCTGTTG | TGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA | TRUE | TRUE |
| 2 | GTCTGTTG | TGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGCCAGTGTGGAAAATCTCTAGCA | FALSE | TRUE |
| 3 | GTCTGTTG | CGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA | FALSE | TRUE |
| 4 | GTCTGTTG | TGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCCCTAGCA | FALSE | TRUE |

| Variable | Total | 1 | 2 | 3 | 4 |
|----------|-------|---|---|---|---|
| Distinct | 8378 | 1203 (14.4%) | 1142 (13.6%) | 1141 (13.6%) | 1139 (13.6%) |
| DistinctHuman | 527 | 429 (81.4%) | 414 (78.6%) | 414 (78.6%) | 412 (78.2%) |
| DistinctLinker | 4213 | 2 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| Max | 1726 | 1325 (76.8%) | 1325 (76.8%) | 1325 (76.8%) | 1325 (76.8%) |
| MaxHuman | 312 | 312 (100%) | 312 (100%) | 312 (100%) | 312 (100%) |
| MaxLinker | 1726 | 1 (0.1%) | 0 (0%) | 0 (0%) | 0 (0%) |
| Total | 24128 | 4037 (16.7%) | 3976 (16.5%) | 3975 (16.5%) | 3973 (16.5%) |
| TotalHuman | 1220 | 1114 (91.3%) | 1099 (90.1%) | 1099 (90.1%) | 1097 (89.9%) |
| TotalLinker | 13370 | 2 (0%) | 0 (0%) | 0 (0%) | 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)

Distribution of clipping points across reads

```
logo          [sequence logo]

4        TGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCCCTAGCA     61
3        CGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA     61
2        TGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGCCAGTGTGGAAAATCTCTAGCA     61
1 (HXB2) TGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA     61
consensus *!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!*!!!!!!!!!!!!!!!*!!!!!
```

☒  non conserved
☒  ≥ 50% conserved