

Comparing HipSTR with Strawman benchmark on chr20

Helia Ziaei

Preparing region file for HipSTR

- Decompose regions with more than 1 repeats into multiple repeats
 - Alternate strategies: remove them, keep the longest
- Only keep the repeats with motif length < 7
- 42,483 repeats remained in chr20

Running HipSTR

```
HipSTR --bams $bam_file \  
  --fasta $ref \  
  --regions $regions \  
  --str-vcf $output \  
  --log $log \  
  --bam-samps HG002 --bam-libs HG002 \  
  --def-stutter-model \  
  --output-filters \  
  --min-reads 15
```

- Genotyping succeeded for 37838/41144 loci
- Skipped loci:
 - 843 with lengths above the maximum threshold
 - 496 with too few reads
- Failed genotyping:
 - Repetitive flanking regions

Truvari benchmark

After refine with --regions hipstr.bed

--use-includebed

- #FP = 14
- #FN = 147

	bench	refine
Precision	0.016	0.99
Recall	0.009	0.87
F1	0.012	0.92

Some examples from FP file, problem in coordinates?

This got fixed after fixing END position of HipSTR calls

Li
chr

Some examples from FP file, problem in coordinates?

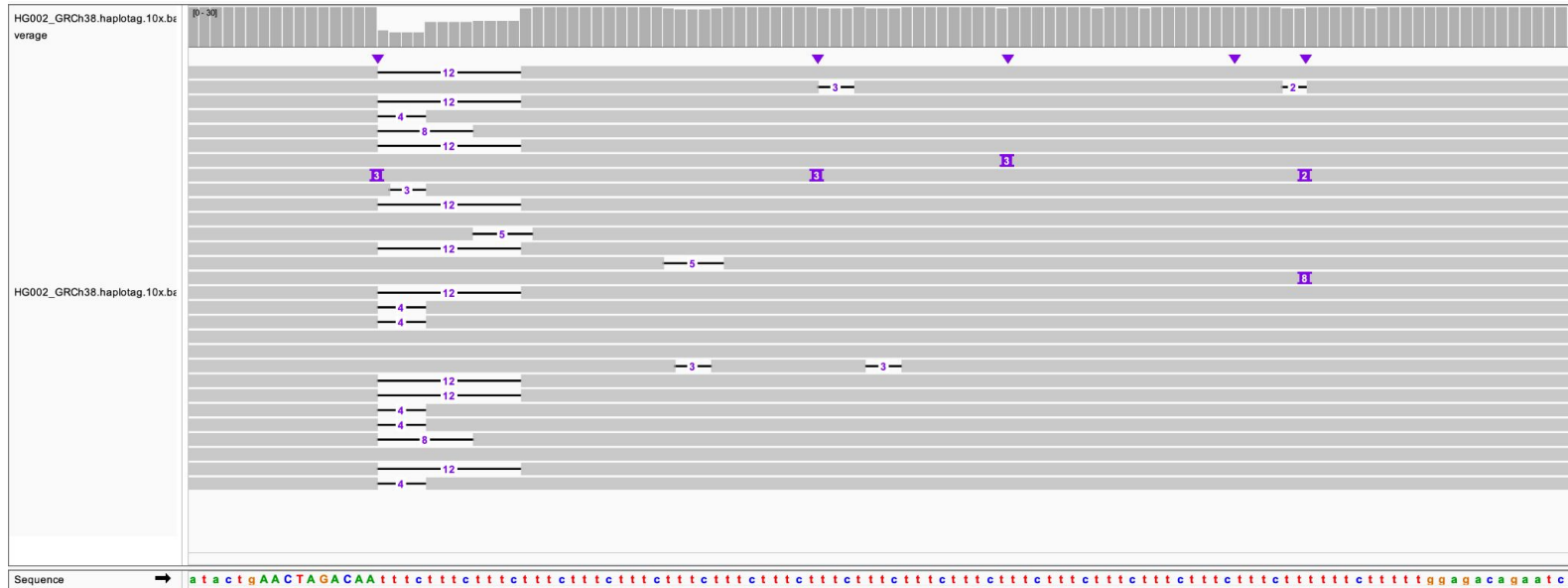
This got fixed after fixing END position of HipSTR calls

Li

chr

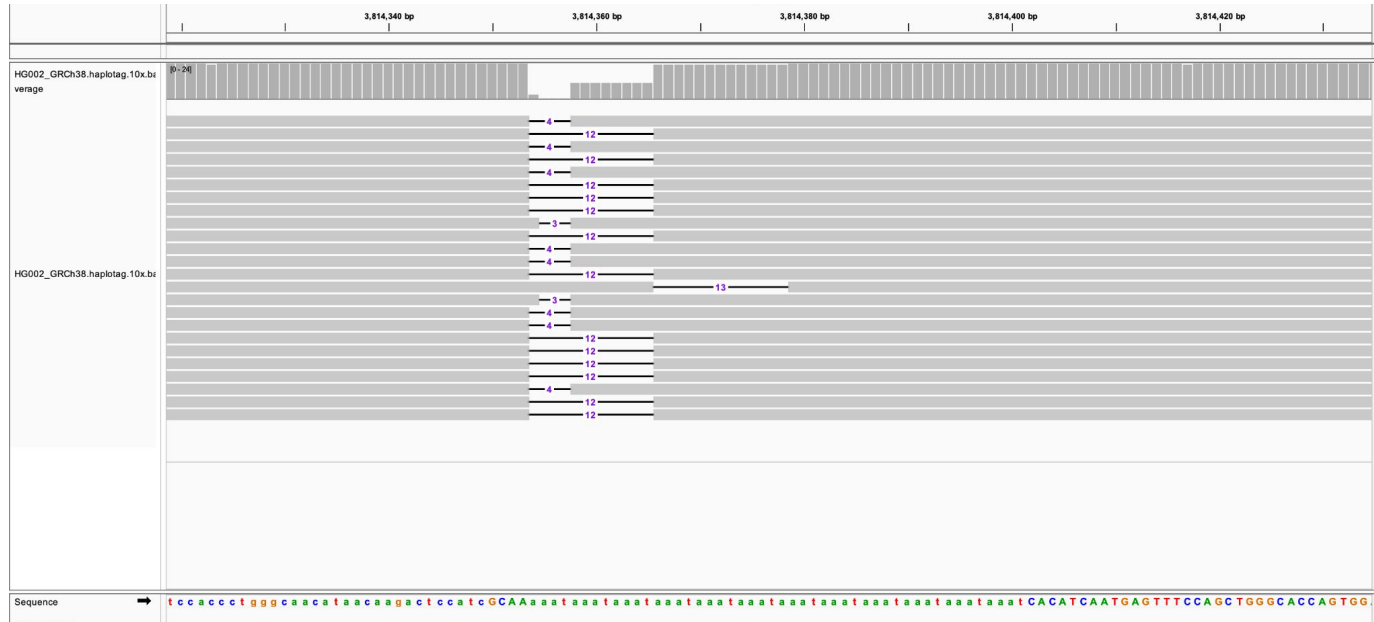
FPs: CTTT repeat in chr20:2,281,317-2,281,402

- HipSTR reports a deletion of -12/-4bp
- Strawman reports a -5bp homozygous deletion at 2,281,389



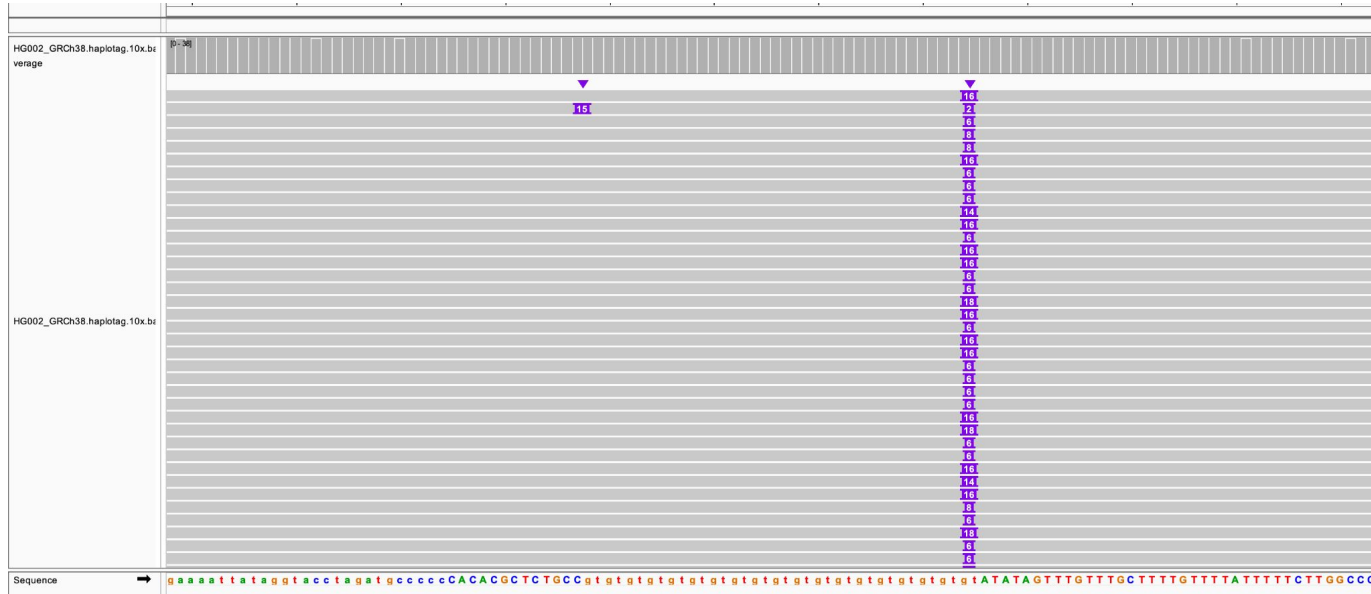
FPs: AAAT repeat in chr20:3,814,353-3,814,400

- HipSTR reports a deletion of -12/-4bp
- Strawman reports a -5bp homozygous deletion at 3,814,352



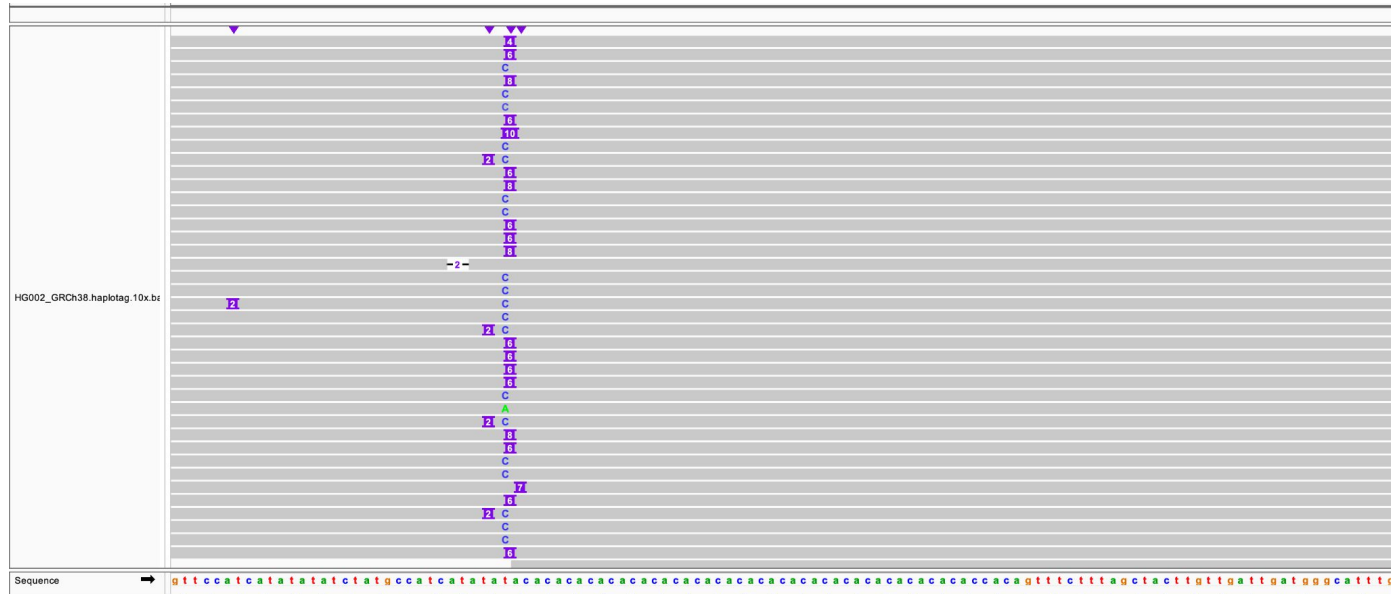
FPS: GT repeat in chr20:11,753,707-11,753,744

- HipSTR reports an insertion of 16/6 bp
- Strawman reports a 6bp homozygous insertion at 11,753,744



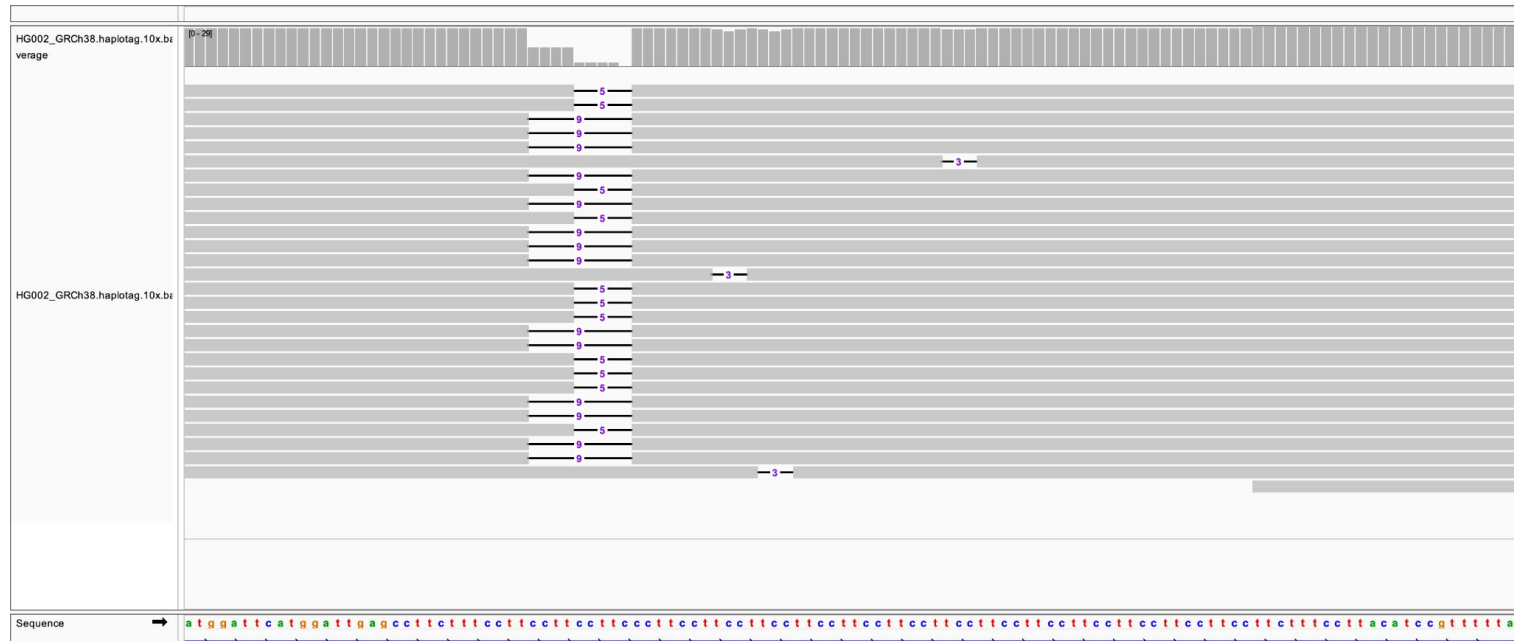
FPs: AC repeat in chr20:14,215,337-14,215,384

- HipSTR reports an insertion of 0/6 bp
- Strawman reports an insertion of 0/4 bp



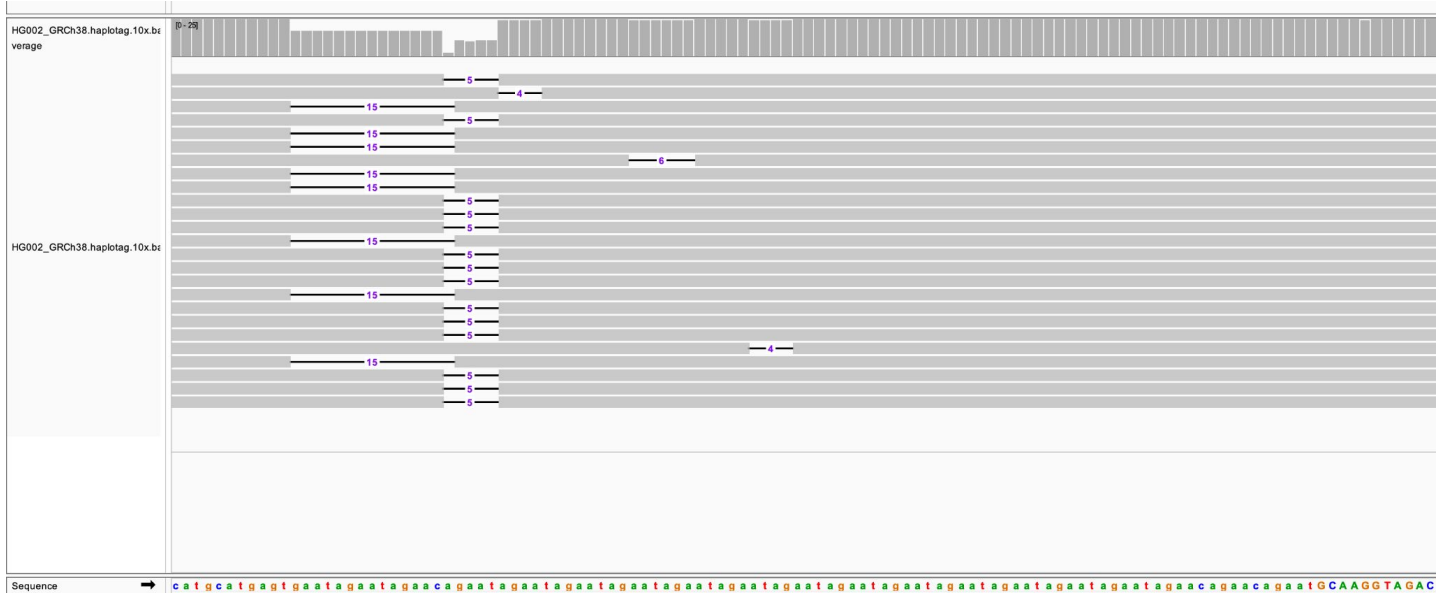
FPs: CCTT repeat in chr20:19,322,344-19,322,428

- HipSTR reports a deletion of -5/-9 bp
- Strawman reports a homozygous deletion of 6 bp at 19,322,351



FPs: AGAAT repeat in chr20:33,544,931-33,545,024

- HipSTR reports a deletion of -5/-15 bp
- Strawman reports a homozygous deletion of 5bp at 33,544,945



FPs: TTTTG repeat in chr20:40,765,189-40,765,246

- HipSTR reports an insertion of 10/15 bp
- Strawman reports a homozygous insertion of 10bp at 40,765,204

