What does secondary (and some other fields) mean in flagstat output?

6

This is my samtools flagstat output for the alignment of a paired-end sample using HISAT (filtered with MQ<1):

```
45084184 + 0 in total (QC-passed reads + QC-failed reads)
4717987 + 0 secondary
0 + 0 supplementary
0 + 0 duplicates
45084184 + 0 mapped (100.00%:-nan%)
40366197 + 0 paired in sequencing
20273012 + 0 read1
20093185 + 0 read2
39146254 + 0 properly paired (96.98%:-nan%)
39644363 + 0 with itself and mate mapped
721834 + 0 singletons (1.79%:-nan%)
260722 + 0 with mate mapped to a different chr
235361 + 0 with mate mapped to a different chr (mapQ>=5)
(END)
```

There's quite a few questions about flagstat on Biostars, but I haven't seen the secondary, supplementary or duplicates flag before. Assuming that this is HISAT specific, are is secondary the number of mapped reads that get mapped using the second (non-global) HISAT index? What would supplementary mean? And are duplicates the amount of reads that map non-uniquely?

flagstat samtools • 16k views

ADD COMMENT • link

9.9 years ago
Niek De Klein ★ 2.6k

updated 2.5 years ago by Ram 45k • written 9.9 years ago by Niek De Klein ★ 2.6k

28

These are defined more formally in the SAM specification, though perhaps the wording there isn't great.

Secondary: One of the many places a multimapper can align. Note that multimappers will have one primary and 0 or more such secondary alignments.

Supplementary: For chimeric/fusion/non-linear alignments, this is the location of one part of the alignment. For reads that align in a chimeric fashion, one segment will be designated as primary and the remainder supplementary.

Duplicate: If you've marked possible PCR duplicates, then this will be set. The definition of a duplicate is somewhat dependent on the tool used (N.B., aligners don't typically set this flag, it's down by picard's markDuplicates command or similar).

ADD COMMENT • link

9.9 years ago
Devon Ryan 105k

updated 2.5 years ago by Ram 45k • written 9.9 years ago by Devon Ryan 105k

Login before adding your answer.

Similar Posts

Problem mapping with STAR • updated 6.7 years ago by WouterDeCoster 48k • written 6.7 years ago by alvarocentron91 10

Hello, I'm mapping with three different programs: TopHat, HISAT2, and STAR. I'm using default values for all but STAR (I have short read…

"Secondary", "Supplementary", "Duplicates" and "Paired in sequencing" in samtools flagstat • updated 2.5 years ago by Ram 45k • written 3.2 years ago by pasha64t • 0

I wonder if someone please explain what secondary, supplementary, duplicates and paired in sequencing mean in samtools flagstat Example: …

How to extract uniquely aligned paired end reads obtained from freebayes using diffrent parmeter combinations of samtools? • 4.9 years ago by nadiabeg.comsats 10

I have 10x reads and I need uniquely aligned reads for SNP calling. I am using samtools to extract uniquely aligned reads but I am not sat…

Extracting supplementary reads from bam file • updated 9 months ago by cmdcolin ★ 4.2k • written 9 months ago by g1ang 20

Hi everyone, bioinformatics noob here :) I'm trying to extract supplementary reads (Illumina paired-end 150bp) from a series of bam files…

Separating unmapped reads from bam files is not working using samtools -f 4 • updated 5.0 years ago by GenoMax 151k • written 5.0

years ago by sorrymouse ▲ 120

I mapped reads using bwa mem and created a sorted bam file. An example flagstat is as follows: 12446425 + 0 in total (QC-passed rea…

BWA MEM Definition of a Primary Alignment •
updated 4.0 years ago by d-cameron ★ 2.9k • written 6.0 years ago by QVINTVS_FABIVS_MAXIMVS ★ 2.6k

Quick question: When bwa mem splits a read into a primary and a secondary/supplementary alignment, is the primary always the "left-most…

Interpreting mapping contaminants •
updated 6.5 years ago by h.mon ϟ 35k • written 6.5 years ago by cecilio11 ▲ 120

Dear Biostars, I am posting this question in this forum because I found this wonderful site is populated mostly by very kind and helpful…

Different mapping results between HISAT and TopHat •
5.0 years ago by concetta ▲ 10

Hi all! I am performing a genome-guided transcriptome assembly with Stringtie. Before the transcriptome assembly, I have mapped my pa…

Difference between chimeric alignments and multiple mapping •
updated 8.0 years ago by Devon Ryan ϟ 105k • written 8.0 years ago by Vanilla ▲ 110

Hi all: I recently got quite confused with two SAM flags got from BWA alignment, which is "**supplementary alignment**" from chimeric al…

Total number of trimmed reads less than number of mapped reads •
4.3 years ago by kspata ▲ 90

Hi All, I have forward and reverse trimmed reads. The total number of forward and reverse trimmed reads calculated using fastqc was 9,83…

How to interpret flagstat output •
updated 3.0 years ago by Istvan

Albert 102k • written 3.0 years ago by deniselavezzari • 0

Hi, I have an issue similar to some published. FastQC gives me 2,941,170 Total sequence, whereas the flagstat outputs these: ``` 351046…

High number of secondary alignments with HISAT2 • updated 7.2 years ago by Biostar 20 • written 7.3 years ago by JJ 760

Hi all, So I am working with a public dataset and I am a bit worried about the high number of secondary alignment I get. Here is the sam…

The meaning of uniquely mapped •

MISSING

Can anyone suggest resources for troubleshooting why my alignment using Bowtie1.1.1 is incorrect? • 4.1 years ago by nessj • 0

My bowtie command : bowtie GCF_000001215.4_Release_6_plus_ISO1_MT_genomic -q SRR8191524.fastq -v 2 -m 1 -31 -S 2> ./SRR8191524.ou…

Different alignments rates with bwa mem (0%) and bowtie2 (82%) • 8.3 years ago by James Ashmore ★ 3.5k

I have ChIP-seq data which has been trimmed using trim galore (all default settings). When I align these trimmed reads using bwa mem (all d…

Samtools rmdup and Piccard Markduplicates • updated 7.8 years ago by lakhujanivijay 5.9k • written 7.8 years ago by Prakash ★ 2.2k

Hello Bio Stars, I have doubt regarding duplicate removal from BAM file. I used two tools "**samtools rmdup**" and **Piccard MarkDuplic…

Can't convert paired end BAM to bed using bedtools • updated 2.3 years ago by GenoMax 151k • written 2.3 years ago by oksana03fel • 0

Hello, I got .bam files from my pipeline and i merged them with

samtools merge ALL.bam *.bam and I got this % of mapping 3825773 + …

PCR duplicates in RNASeq • updated 2.1 years ago by Ram ϟ 45k • written 7.8 years ago by Prakash ★ 2.2k

Hello Bio stars, I have small query regarding identification and removal of PCR duplicates from RNASeq data. The Tophat2 alignment stats c…

Variant calling and alignment stats • 5.0 years ago by nadiabeg.comsats ▲ 10

Hi. I am using samtools flagstats to see the statistics of my alignment file. It looks something like this: I have 10x genomics reads. …

different result using minimap2 and pbmm2 • updated 2.6 years ago by gconcepcion ▲ 410 • written 3.2 years ago by pingu77 ▲ 40

Hi all! I am analysing CSS Pacbio data and each sample came from different run, in particular I have three files for each sample. I teste…

Variation in mapping percentage with genome • updated 24 months ago by Ram ϟ 45k • written 3.4 years ago by onkar ▲ 10

I have a few resequencing data (Illumina DNA Seq) for various cultivars/varieties of same plant. I had mapped the reads with a published c…

Bowtie2 and BWA-MEM giving very different results in metagenomic data • 5.5 years ago by Antonio Camargo ▲ 160

I've assembled a metagenome using MEGAHIT and begun testing different mapping options to perform the binning of the contigs. However, I've …

Cannot align reads to plasmid • updated 7.3 years ago by h.mon ϟ 35k • written 7.3 years ago by David ▲ 240

Hi, I have sequenced a bacterial genome for which i have a reference genome (98% similarity). I have used bwa to

map reads to the refe…

Samtools: How can I extract properly-paired QC-passed reads instead of extracting only properly-paired? •
8.0 years ago by bioinfo8 ▲ 230

Hi, Here is the `flagstat` output of my bam file: 37750740 + 352032 in total (QC-passed reads + QC-failed reads) 0 + 0 secon…

Samtools flagstat confusing result of a merged bam file •
MISSING

How does Base Call casing affect BWA •
4.8 years ago by oconnwald ▲ 20

I was wondering how the casing of base calls affected BWA MEM performance. I ran three sets of data to see if there was any difference a…

tophat --max-multihits impacts 'mapped in a proper pair' •
updated 8.6 years ago by John ⚡ 13k • written 8.6 years ago by Carlo Yague ⚡ 9.0k

Hi everyone, I have a RNA-seq library (paired-end) with a large amount of multimappers because it comes from a total RNA extract with pa…

samtools flagstat interpretation •
updated 5.7 years ago by ATpoint ⚡ 88k • written 5.7 years ago by asmaaaljuhani • 0

How can I interpret the following results? and is there a specific percentage that we accept the allignment? samtools flagstat aln-p…

Different flagstat number after using MergeBamAlignment •
13 months ago by ThomasLam • 0

Hi everyone, I'm utilizing the GATK Best Practices for analyzing mitochondrial NGS data. I utilized FastqToSam to generate a ualign BAM fi…

Samtools Flagstat Comments •
updated 3.9 years ago by GenoMax ⚡ 151k • written 3.9 years ago by santos48 ▲ 40

Hi, I am trying to solve [total alignments :0 results of

featureCounts[1] this problem.
Now I checked my bam files
with that command `sam…

Content	Help	Access
Search	About	RSS
Users	FAQ	API
Tags		Stats
Badges		