



Samtools flagstat results



0



```
221372 + 0 in total (QC-passed reads + QC-failed reads)
0 + 0 secondary
0 + 0 supplementary
20419 + 0 duplicates
218469 + 0 mapped (98.69% : N/A)
155851 + 0 paired in sequencing
77895 + 0 read1
77956 + 0 read2
142663 + 0 properly paired (91.54% : N/A)
150045 + 0 with itself and mate mapped
2903 + 0 singletons (1.86% : N/A)
4938 + 0 with mate mapped to a different chr
2120 + 0 with mate mapped to a different chr (mapQ>=5)
```

I have the following questions:

1. From previous posts I understood that read1 may not be equal to read2, as there may be reads whose mates didn't align. These are singletons. So doesn't that mean that read2-read1 must be equal to singletons? What am i missing here?
2. What does the field "with itself and mate mapped mean?"

[next-gen](#) • 10k views

[ADD COMMENT](#) • [link](#)

 updated 9 months ago by [virog](#) • 0 • written 7.8 years ago by [Gene_MMP8](#) ▾ 240

Traffic: 2389 users visited in the last hour



What preprocessing steps have been applied to get this file ?

0

[ADD REPLY](#) • [link](#)

 7.8 years ago by [GouthamAtla](#) ⚡ 12k


I wrote this command: samtools flagstat example.bam. Does this help?

0

[ADD REPLY](#) • [link](#)

 7.8 years ago by [Gene_MMP8](#) ▾ 240


7

I realized there aren't accessible resources that clearly explain what the different lines of the flagstat output mean (This [answer](#) was useful). Here is what I found based on some other answers on forums and playing around with my own files (paired-end reads mapped using bwa mem). Some of these lines may apply regardless of aligner but others may not.



total (QC-passed reads + QC-failed reads): Total number of reads including other such as supplementary

primary: Total number of reads that were provided as input for mapping

secondary: see [here](#)


 7.8 years ago
[Gene_MMP8](#) ▾ 240

Similar Posts

[How to extract uniquely aligned paired end reads obtained from freebayes using different parameter 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...

[BWA mem paired end vs single end shows unusual flagstat summary](#) • 8.4 years ago by [rrdavis](#) ▾ 60

I used BWA mem to align some Miseq (2X300bp) reads to a human reference expecting very low mapping since human reads would be considered co...

[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...

[samtool flagstat question](#) • updated 2.3 years ago by [Ram](#) ⚡ 45k • written 10.1 years ago by [alex](#) ▾ 250

Hi, So I have created a bam from a NIST sample. I did realignment and recalibration but the flagstat numbers don't change. Should I be wor...

[Thorough Documentation Of Samtools Flagstat](#) • updated 11.6 years ago by [GouthamAtla](#) ⚡ 12k • written 11.6 years ago by [Biomonika \(Noolean\)](#) ⚡ 3.2k

I am trying to find thoroughfull documentation of [samtools]


 2.8 years ago
[aiswarya.prasad1998](#) ▾ 70

supplementary: see [here](#)

duplicates: see [here](#)

primary duplicates: primary reads that were marked as duplicates

mapped: number of mapped reads including supplementary

mapped %: percentage of mapped reads including supplementary (denominator is the number of **total** reads)

primary mapped: number of mapped reads that are labelled primary (just the number of mapped reads out of the input reads). i.e. excludes the number of reads that are supplementary.

primary mapped %: percentage of mapped primary reads (denominator is number of **primary** reads)

paired in sequencing: This is the number of paired reads. If you used only paired reads after trimming, this will be the same as the number in the **primary** field

read1: This is the number of forward (R1) reads. If you used only paired reads after trimming, this will be half of the number in the **primary** field

read2: This is the number of reverse (R2) reads. If you used only paired reads after trimming, this will be the same as the **read1** field.

properly paired: This is the number of reads that map in a way that makes sense (Not too far apart, on different chromosomes, R1 read maps to the forward strand and R2 to the reverse strand etc. depending on the aligner). See [this](#) for some more information. This is suitable if you want to be very conservative with the number of reads that you consider mapped.

properly paired %: percentage of properly paired reads (denominator is number of **primary** reads)

with itself and mate mapped: Number of reads with its corresponding reverse / forward read also mapped. This is less strict than **properly paired** but more than **primary mapped**.

singletons: This is the number of reads that are mapped but their corresponding reverse / forward read did not map. (primary mapped - with itself and mate mapped = singletons)

singletons %: percentage of singletons (denominator is number of **primary** reads)

with mate mapped to a different chr: This is the number of reads that are mapped but their corresponding reverse / forward read mapped to a different chromosome. Remove these from **properly paired** to get an even more conservative estimate of number of mapped reads.

with mate mapped to a different chr (mapQ>=5): This is the number of reads that are mapped but their corresponding reverse / forward read mapped to a different chromosome with good quality for the alignment. Remove these from **properly paired** to get a more conservative estimate of number of mapped reads.

[ADD COMMENT](#) • [link](#)

2.8 years ago by [aiswarya.prasad1998](#) ▲ 70

1 Thank you for your simple and understandable explanation. I've read several definitions but still didn't quite understand how to
0 read Samtools flagstat results

[ADD REPLY](#) • [link](#)

18 months ago by [LuciaNhu](#) ▲ 10

1 Thank you very much for this very clear simple explanation. This is really helpful.
0 [ADD REPLY](#) • [link](#)

9 months ago by [virog](#) ▲ 0

(<http://samtools.sourceforge.net/>) flagstat for every row for file [Un.bam](ht...)

RNA seq samtools flagstat • updated 8.9 years ago by [GouthamAtla](#) ↗ 12k • written 8.9 years ago by [pshubhamoy](#) ▲ 20
Hi all I am getting flagstat outputs (alining with 'STAR') 0 + 69618651 in total (QC-passed reads + QC-failed reads) 0 + 0
...

Remove singletons from bam • updated 7.2 years ago by [Ram](#) ↗ 45k • written 7.2 years ago by [Gene_MMP8](#) ▲ 240

I have a tumor bam file and running samtools flagstat gave the following output
1020505173 + 0 in total (QC-passed reads + QC-fa...

Interpreting mapping contaminants • updated 6.5 years ago by [h.mon](#) ↗ 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...

samtools filter proper paired alignment • 6.3 years ago by [mynameluxiang](#) ▲ 0

I want get proper paired alignment here is my scripts \$ samtools view test1.v2.4.reordered.bam -f 2 -bS -1> test1.uniq.prop....

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...

Tophat mapping less reads (and less properly paired) when more real-r is specified? • updated 2.6 years ago by [Ram](#) ↗ 45k • written 9.9 years ago by [manekineko](#) ▲ 150

First, I have run tophat with -r 50, however I have calculated than the mean inner distance is around 0 mean insert / 1051



1. Singletons occur when only one mate in a pair aligns. You can also have situations where one mate aligns



4

multiple times (e.g., to a simple repeat) and the other only once. Then one will have a single entry and the other may have multiple. Also, if you did any filtering then that'd affect this as well.

2. It means exactly what it says, both mates mapped. They may be "properly paired" or they may not be.

Regardless, if they both align somewhere at least once then they count toward this.



[ADD COMMENT](#) • [link](#)

7.8 years ago by Devon Ryan 105k

1 When I subtract number of reads mapped from the total number of reads ($221372 - 218469 = 2903$), which is the number of singltons. So can I say that singltons are the reads which didn't map to any reference?

[ADD REPLY](#) • [link](#)

7.8 years ago by Gene_MMP8 ▲ 240

1 By definition a singleton cannot be unmapped. If it is, it's not a singleton.

0 [ADD REPLY](#) • [link](#)

7.8 years ago by Devon Ryan 105k

1 Thanks for your reply. Can you explain what you meant by "You can also have situations where one mate aligns multiple times (e.g., to a simple repeat) and the other only once.". I know it is a trivial question, but I am entirely new to this field. Hence asking.

[ADD REPLY](#) • [link](#)

7.8 years ago by Gene_MMP8 ▲ 240

1 Suppose you have the sequence **ATATATATATATATAAGCGCTAGCTAGTCGATCTAGCTAGCTGTAGCTGGTCAGAC**. You might 5 have reads **ATATATAT** and **GCGCTAGC**. The latter read can only align to one place in that sequence. The former read can align 5 equally well to multiple places. Consequently, some aligners will produce multiple entries for **ATATATAT** and a single one for **GCGCTAGC**.

[ADD REPLY](#) • [link](#)

7.8 years ago by Devon Ryan 105k

1 Excellent explanation. Thanks a lot!

0 [ADD REPLY](#) • [link](#)

7.8 years ago by Gene_MMP8 ▲ 240

[Login](#) before adding your answer.

GROUP & MEETINGS (100)
2xread_siz...

How to remove singletons in a

BAM file using samtools •

updated 6.5 years ago by

finswimmer 16k • written 6.5

years ago by neranjan 70

Dear All, I want to remove the singletons from the aligned bam file. Downloaded fastq files using fastq-dump --split-files SR...

Why number of #read1 and #read2 is different in samtools flagstat output? •

updated 7.3 years ago by Biostar 20 • written 8.4 years ago by cristina.osuna.cruz ▲ 10

Dear All, I am using BWA mem algorithm to map my Illumina reads (paired-ends) data against to my assembled genome (diatom species). Aft...

Getting uniquely aligned

properly paired reads •

5.6 years ago by

nadiabeg.comsats ▲ 10

Hi all, I have bam files obtained from mapping 10x genomics data. My goal is to get the VCF file. To do the variant calling I want to ge...

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 -3 1 -S >; ./SRR8191524.ou...

Convert processed bam file to fastq •

updated 4.3 years ago by

GenoMax 151k • written 4.3

years ago by coltonrobbins73 0

I found some 10x single-cell RNA seq data on GEO that I want to use for a study ([SRR7754181][1]). Unfortunately, it looks like the author ...

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...)

[Getting Confused With The Flagstat After Pcr Duplicates Removed](#) •

updated 13.1 years ago by [swbarnes2](#) ↗ 15k • written 13.1 years ago by [KJ Lim](#) ▲ 140

Good day. I encountered a situation like below: The flagstat before PCR duplicates removed from paired end mapped reads.
.....

[How to extract properly paired alignments from a bam file using samtools?](#) •

updated 3.1 years ago by [Ram](#) ↗ 45k • written 10.3 years ago by [pedrodcb](#) ▲ 110

This is what samtools flagstat returns : ```` 87410692 + 0 in total (QC-passed reads + QC-failed reads) 17796957 + 0 secondary 0 + 0 suppl...

[Converting BAM to FASTQ...to new reference genome](#) •

updated 23 months ago by [Ram](#) ↗ 45k • written 2.8 years ago by [kcarey](#) • 0

Hey, I was wondering if someone can help me to see what I am getting wrong here? I have paired-end RNA seq data bam files but do not have ...

[get identical number of read1 and read2 aligned](#) •

7.0 years ago by [oghzzang](#) ▲ 50

Dear Biostar users. This is my samtools flagstat output.
39219750 + 0 in total (QC-passed reads + QC-failed reads) 0 + 0 sec...

[PCR duplicates in RNASeq](#) •

updated 2.1 years ago by [Ram](#) ↗ 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. ...

[Samtools rmdup and Picard 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 **Picard MarkDuplic...

[Very low mapping rate of a bam file, reasons?](#) • updated 2.2 years ago by [Ram](#) ↗ 45k • written 9.1 years ago by [nidoo.beg](#) ▲ 20

I have got following results for my bam file. 186895644 + 0 in total (QC-passed reads + QC-failed reads) 0 + 0 duplicates 1627...

[Samtools view only showing read2 for PE mapped reads](#) • updated 3 months ago by [Pierre Lindenbaum](#) ↗ 166k • written 3 months ago by [knickknack](#) • 0

Hello! I've mapped illumina paired-end reads to a reference genome using bwa mem2. When I Samtools flagstat the output.sam, it looks like t...

[Extract all paired reads from sam file](#) • updated 7.7 years ago by [ATpoint](#) ↗ 88k • written 7.7 years ago by [Picasa](#) ▲ 680

Hi, I had to filter manually a sam file by keeping only reads that fall into a specific regions. Using flagstat, my sam file is looki...

[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 + ...

[Cannot align reads to plasmid](#) • updated 7.3 years ago by [h.mon](#) ↗ 35k • written 7.3 years ago by [David](#) ▲ 240

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...

Content
[Search](#)
[Users](#)
[Tags](#)
[Badges](#)

Help
[About](#)
[FAQ](#)

Access
[RSS](#)
[API](#)
[Stats](#)

Use of this site constitutes acceptance of our [User Agreement and Privacy Policy](#).

Powered by the [biostar](#) [forum](#) version 2.3.6
