Samtools fixmate: Fill in mate coordinates, ISIZE and mate related flags from a name-sorted or name-collated alignment.

-r Removes secondary and unmapped reads.

-p Disable FR proper pair check.

-c Add template cigar ct tag.

-m Add ms (mate score) tags. These are used by markdup to select the best reads to keep.

-o format write the final output as sam, bam or cram.

-@ : Number of input/output compression threads to use in addition to main thread.

Because BWA can sometimes leave unusual FLAG information on SAM records, it is helpful when working with many tools to first clean up read pairing information and flags.

samtools view tmp.bam | cut -f 2 | sort | uniq -c

2534561 147 {1+2+16+128 (first read), 1+2+32+64 (second read)}

2543330 163 {1+2+32+128 (first read), 1+2+16+64 (second read)}

2543820 83 {1+2+16+64 (first read), 1+2+32+128 (second read)}

2534983 99 {1+2+32+64 (first read), 1+2+16+128 (second read)}

samtools fixmate tmp.bam - | samtools view | cut -f 2 | sort | uniq -c

1348 128 {128}

1331 144 {16 + 128}

2533230 147 {1+2+16+128 (first read), 1+2+32+64 (second read)}

2541982 163 {1+2+32+128 (first read), 1+2+16+64 (second read)}

1753 64 {64}

1838 80 {16+64}

2541982 83 {1+2+16+64 (first read), 1+2+32+128 (second read)}

2533230 99 {1+2+32+64 (first read), 1+2+16+128 (second read)}

**-r**

| **#** | | **Decimal** | | **Description of first read** | |
| --- | --- | --- | --- | --- | --- |
| 1 | | 1 | | Read paired | |
| 2 | | 2 | | Read mapped in proper pair | |
| 3 | | 4 | | Read unmapped | |
| 4 | | 8 | | Mate unmapped | |
| 5 | | 16 | | Read reverse strand | |
| 6 | | 32 | | Mate reverse strand | |
| 7 | | 64 | | First in pair | |
| 8 | | 128 | | Second in pair | |
| 9 | | 256 | | Not primary alignment | |
| 10 | | 512 | | Read fails platform/vendor quality checks | |
| 11 | | 1024 | | Read is PCR or optical duplicate | |
| 12 | | 2048 | | Supplementary alignment | |
| **Decimal** | | **Description of second read** | |
| 1 | | Read paired | |
| 2 | | Read mapped in proper pair | |
| 4 | | Read unmapped | |
| 8 | | Mate unmapped | |
| 16 | | Read reverse strand | |
| 32 | | Mate reverse strand | |
| 64 | | First in pair | |
| 128 | | Second in pair | |
| 256 | | Not primary alignment | |
| 512 | | Read fails platform/vendor quality checks | |
| 1024 | | Read is PCR or optical duplicate | |
| 2048 | | Supplementary alignment | |
|  | |  | |

67 {1+2+64 (read paired, read mapped in proper pair, first in pair) }, 131 {Read paired, Read mapped in proper pair, second in pair}

65 { 1+64 (read paired, first in pair) }, 129 {read paired, second in pair) }

**#!/bin/bash**

 **raw = “/home/genome1/mitochondria\_assembly/raw”**

**while read p**

**do**

**nam = $ (echo $p | xargs -n 1 basename | cut -f 1-2 -d “ \_”);**

**$ time python norgal.py -i $p $raw/${nam}\_2.fastq.gz – o “$raw/Norgal” > output.txt 2>&1**

 **done < “$raw/alignments/fastqlist”**