Cheatography

# Bedtools Cheat Sheet

by ilevantis via cheatography.com/59886/cs/15673/

## BED file format

| Column | *e.g.* | Definition |
|---|---|---|
| **chrom** | *Sc112.1* | <STR> name of chromosome/scaffold |
| **start** | *2134* | <INT> start position of feature |
| **end** | *2565* | <INT> end position of feature |
| **name** | *gene123* | <STR> name of feature |
| **score** | *544* | <NUM> score for the feature e.g. bit score |
| **strand** | *+* | <+/-/.> strand on which feature is located |
| **thickStart** | *2235* | |
| **thickEnd** | *2489* | |
| **itemRgb** | *255,0,0* | |
| **blockCount** | *2* | |
| **blockSizes** | *150,80* | |
| **blockStarts** | *0,2333* | |

## GFF vs BED indexing

```
GFF    ┌─1   2   3─┐ 4   ...
         G---A---T   C   ...
BED    └─0   1   2 └─3   ...
```

**gff > bed:**

bed_start = gff_start - 1,

bed_end = gff_end

**bed > gff:**

gff_start = bed_start + 1,

gff_end = bed_end

## getfasta

```
$ bedtools getfasta [OPTIONS] -fi <input FASTA> -bed <BED/GFF/VCF>
```

**options**

| | |
|---|---|
| **-fo** | Specify an output file name. By default, output goes to stdout. |
| **-name** | Use the “name” column in the BED file for the FASTA headers in the output FASTA file. |
| **-tab** | Report extract sequences in a tab-delimited format instead of in FASTA format. |
| **-bedOut** | Report extract sequences in a tab-delimited BED format instead of in FASTA format. |

## getfasta (cont)

| | |
|---|---|
| **-s** | Force strandedness. If the feature occupies the antisense strand, the sequence will be reverse complemented. Default: strand information is ignored. |
| **-split** | Given BED12 input, extract and concatenate the sequences from the BED “blocks” (e.g., exons) |

## maskfasta

```
$ bedtools maskfasta [OPTIONS] -fi <input FASTA> -bed <BED/GFF/VCF> -fo <output FASTA>
```

OPTIONS

| | |
|---|---|
| **-soft** | Soft-mask (that is, convert to lower-case bases) the FASTA sequence. By default, hard-masking (that is, conversion to Ns) is performed. |
| **-mc** | Replace masking character. That is, instead of masking with Ns, use another character. |

```
FASTA   ACTGATCATGATACATGATACCATTAGGATACAATA
BED           ████       █████      ████
FASTA'  ACTGATNNNNATACATGNNNNNATTAGGNNNNAATA
```

By **ilevantis**
cheatography.com/ilevantis/