

Bed file

Bed : Browser Extensible Data

Specifications

Officially none,
but UCSC browser propose one that can be used :
<http://www.genome.ucsc.edu/FAQ/FAQformat.html#format1>
Most prog (I've used) consider only the 3 first column
(seq_id, start, end).

Warning :

UCSC spec. use 0-based coordinates, half-open region, not all
prog consider this.

Samples

S1

```
chr7 127471196 127472363  
chr7 127472363 127473530  
chr7 127473530 127474697
```

S2

```
chr7 0 1 base_une
```

Recommended library

?

Recommended tools

bedops : <https://bedops.readthedocs.io>

bedtools : bedtools.readthedocs.io

Comments

- Spec seems to imply that `start < end`, but as it is not strictly forbidden, I use “`seq end start`” to a feature on the opposite strand.
- To load data I used a personal class `RegionLink` that aims to manage file like “`chrom,start,end <=> chrom',start',end'`”
It can be found in https://gitlab.inria.fr/sletort/Brassica_scripts/blob/master/lib/RegionLink.py
- Bedops propose the format `starch` which is a `gzip` bed with an integrated index, which is great to work and save space.