capC-MAP File Conversion

Chris Brackley

19 November 2018

capC-MAP generates chromosome interaction profiles for each target in the standard bedGraph file format. While this is a very common format which can be read by many genome browsers and downstream analysis software, it can also be useful to convert this to alternative formats.

BigWig format

bedGraph files can be large. A common compressed file format is BigWig, and a conversion tool is available from UCSC at http://hgdownload.soe.ucsc.edu/downloads.html#utilities_downloads. bedGraphs can be converted to BigWig using the following command:

```
bedGraphToBigWig myfile.bdg chrom.sizes myfile.bw
```

where chrom.sizes is a file containing a list of chromosome and their size in bp. Such a file is generated automatically by capC-MAP during an analysis run, or can be generated from the restriction enzyme file, for example as follows

```
capC-MAP getchromsizes -f dpnII_map_mm9.bed -o chrom.sizes
```

Wig format

A common alternative to bedGraph format is the 'wiggle' or 'wig' format. There are several types of wig file, but bedGraph can most easily be converted to a 'variableStep' wig file using the following Unix command

```
awk '{
if (NR==1) {
   if ($1=="track") {$2="type=wiggle_0"; print}
   else {print "track type=wiggle_0"}
}
span=$3-$2;
chr=$1;
if ($1!="track") {
   if (chr!=lastchr || span!=lastspan) {
      print "variableStep chrom=" chr " span=" span
   }
   print $2,$4
}
lastspan=span;
lastchr=chr;
}' MybedGraphFile.bdg > MyWigFile.wig
```

where the bedGraph file must be sorted.

An example recipe for converting a set of capC-MAP outputs to wig files using Unix commands is:

```
dir="MyOutput"
step=200
window=2000
while read -r line
```

```
do
    line=($line)
    name=${line[3]}

    sort -k1,1 -k2,2n ${dir}/captured_bin_${step}_${window}_RPM_${name}.bdg \
    | awk '{
      if (NR==1) {
        if ($1=="track") {$2="type=wiggle_0"; print}
        else {print "track type=wiggle_0"}
      }
      span=$3-$2;
      chr=$1;
      if ($1!="track") {
        if (chr!=lastchr || span!=lastspan) {
            print "variableStep chrom=" chr " span=" span
        }
        print $2,$4
      }
      lastspan=span;
      lastchr=chr;
    }' > ${name}.wig

done < "targets.bed"
```