Bismark analysis

#bismark genome prepration bismark_genome_preparation /path/to/genome/

```
ibab@IBAB-MSc14-Comp003:~/data_analysis/bismark$ bismark_genome_preparation referance/
Writing bisulfite genomes out into a single MFA (multi FastA) file

Bisulfite Genome Indexer version v0.25.1 (last modified: 19 May 2022)

Step I - Prepare genome folders - completed

Total number of conversions performed:
C->T: 209834
G->A: 209332

Step II - Genome bisulfite conversions - completed

Bismark Genome Preparation - Step III: Launching the Bowtie 2 indexer Please be aware that this process can - depending on genome size - take several hours! Settings:
```

Alignment

bismark -bowtie2 referance/ R1.fastq.gz

```
Number of alignments to (merely theoretical) complementary strands being rejected in total:
Final Cytosine Methylation Report
Total number of C's analysed: 875967
Total methylated C's in CpG context:
                                     29030
Total methylated C's in CHG context:
                                     851
Total methylated C's in CHH context:
                                     2211
Total methylated C's in Unknown context:
Total unmethylated C's in CpG context: 17594
Total unmethylated C's in CHG context:
Total unmethylated C's in CHH context: 604208
Total unmethylated C's in Unknown context:
                                             532
C methylated in CpG context:
                             62.3%
C methylated in CHG context:
                           0.4%
C methylated in CHH context:
                             0.4%
C methylated in Unknown context (CN or CHN):
                                            0.6%
Bismark completed in 0d 0h 0m 12s
===========
Bismark run complete
============
```

bismark -bowtie2 referance/ R2.fastq.gz

```
Number of alignments to (merely theoretical) complementary strands being rejected in total: 0
Final Cytosine Methylation Report
Total number of C's analysed:
Total methylated C's in CpG context:
                                      10
Total methylated C's in CHG context:
                                      8
Total methylated C's in CHH context:
                                      282
Total methylated C's in Unknown context:
Total unmethylated C's in CpG context:
Total unmethylated C's in CHG context:
Total unmethylated C's in CHH context:
                                      143
Total unmethylated C's in Unknown context:
C methylated in CpG context:
                              83.3%
C methylated in CHG context: 16.7%
C methylated in CHH context:
                              66.4%
C methylated in Unknown context (CN or CHN):
                                              100.0%
Bismark completed in Od Oh Om 6s
Bismark run complete
```

#methylation site extraction

bismark_methylation_extractor --gzip --bedGraph R1_bismark_bt2_pe.bam

```
Methylation information will now be written into a bedGraph and coverage file
_______
Using the following files as Input:
/home/ibab/data_analysis/bismark/CpG_OT_R1_bismark_bt2.txt.gz /home/ibab/data_analysis/bisma
rk/CpG_OB_R1_bismark_bt2.txt.gz
Writing bedGraph to file: R1 bismark bt2.bedGraph.gz
Also writing out a coverage file including counts methylated and unmethylated residues to file
: R1_bismark_bt2.bismark.cov.gz
Now writing methylation information for file >>CpG_OT_R1_bismark_bt2.txt.gz<< to individual fi
les for each chromosome
Finished writing out individual chromosome files for CpG_OT_R1_bismark_bt2.txt.gz
Now writing methylation information for file >>CpG_OB_R1_bismark_bt2.txt.gz<< to individual fi
les for each chromosome
Finished writing out individual chromosome files for CpG OB R1 bismark bt2.txt.gz
Collecting temporary chromosome file information... Processing the following input file(s):
CpG_OT_R1_bismark_bt2.txt.gz.chrchrREF.methXtractor.temp
Sorting input file CpG_OT_R1_bismark_bt2.txt.gz.chrchrREF.methXtractor.temp by positions (usin
q -S of 2G)
Successfully deleted the temporary input file CpG_OT_R1_bismark_bt2.txt.gz.chrchrREF.methXtrac
tor.temp
```

```
______
Methylation information will now be written into a bedGraph and coverage file
______
Using the following files as Input:
/home/ibab/data_analysis/bismark/CpG_OT_R2_bismark_bt2.txt.gz /home/ibab/data_analysis/bisma
rk/CpG_OB_R2_bismark_bt2.txt.gz
Writing bedGraph to file: R2 bismark bt2.bedGraph.gz
Also writing out a coverage file including counts methylated and unmethylated residues to file
: R2_bismark_bt2.bismark.cov.gz
Now writing methylation information for file >>CpG_OT_R2_bismark_bt2.txt.gz<< to individual fi
les for each chromosome
Finished writing out individual chromosome files for CpG_OT_R2_bismark_bt2.txt.gz
Now writing methylation information for file >>CpG_OB_R2_bismark_bt2.txt.gz<< to individual fi
les for each chromosome
Finished writing out individual chromosome files for CpG_OB_R2_bismark_bt2.txt.gz
Collecting temporary chromosome file information... Processing the following input file(s):
CpG_OT_R2_bismark_bt2.txt.gz.chrchrREF.methXtractor.temp
Sorting input file CpG_OT_R2_bismark_bt2.txt.gz.chrchrREF.methXtractor.temp by positions (usin
g -S of 2G)
Successfully deleted the temporary input file CpG OT R2 bismark bt2.txt.gz.chrchrREF.methXtrac
tor.temp
```

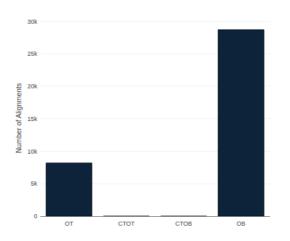
coverted in sam

```
ibab@IBAB-MSc14-Comp003:~/data_analysis/bismark$ head R1_aligned.sam
HISEQ:105:C2UE1ACXX:3:1101:11160:2245_1:N:0:CAGATC 16
                                                                                                       chrREF 87709 8
                          IFBFFIBIIIIBIIIIIFFFFFBFFFFBBB NM:i:22 MD:Z:7G6G1G5C2G5G6G1G4G1G0G1G3G0G1G1G1G13G2G5G3G7G4 X
M:Z:\ldots\ldots h\ldots x.h\ldots x.h\ldots x.h\ldots x.hh\ldots h.h.h.h.h.h.h.h.h.
                         XR:Z:CT XG:Z:GA
HISEQ:105:C2UE1ACXX:3:1101:19338:2197_1:N:0:CAGATC
                                                                                           0
                                                                                                        chrREF 32493 8
                                                                                                                                                101M *
                          0
TATAAGTAGGATTTAGGTCTATT BBBFFFFFFB0<BF7BBBF7BFFIIIII7BF'0<0BBFFF'<BB'<B7<B07<B7<BFBBF0<BBBBB0<
<B0BB<<000<BF00<<'0<BBB0'00BBF# NM:i:29 MD:Z:1C0C0C3C8C0C3C5C1C6C1C6C2C7C2C3C0C0C0C4C0A0C15C0C</pre>
0C3C1C1C0G0
                         XM:Z:.hhh...h...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x..h.x...h.x...h.x...h.x..h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.x...h.
                                                  XR:Z:CT XG:Z:CT
x....hHh.z.
HISEQ:105:C2UE1ACXX:3:1101:19467:2281 1:N:0:CATATC
                                                                                           16
                                                                                                        chrREF 712646 42
                          ACTATACATCTCACCCTAACCAA BB700007<7<77FBB7<7<8<<<BBBBBB777FFBFBB77FFFBFFBB7FBBIIFFBIIIIFFBBBII
FFFB<<IFFBIFIFFFFFFFBBFFF<<B< NM:i:19 MD:Z:0G0G6G7G10G8G23G0G0GG5G0G1G4G0G2G1G12G3G0 XM:Z:x
h.....x.....h....xh.h....xh.h....xh.h....xh.h....xh.h....xh.h....xh.
.h...x XR:Z:CT XG:Z:GA
HISEQ:105:C2UE1ACXX:3:1101:7009:2740_1:N:0:CAGATC
                                                                                                        chrREF 32598 42
                                                                                           16
                                                                                                                                                101M *
                         0
IFBBIFFFFFF<BIIIIIFFFFFFFFBBB NM:i:24 MD:Z:2G6G3G7G5G2G3G0G2G0G5G6G0G1G1G4G3G0G9G9G9GGGGGGGG
...h..xh.... XR:Z:CT XG:Z:GA
HISEO:105:C2UE1ACXX:3:1101:4999:3182 1:N:0:CAGATC
                                                                                                        chrREF 42744 42
                                                                                                                                                101M *
```

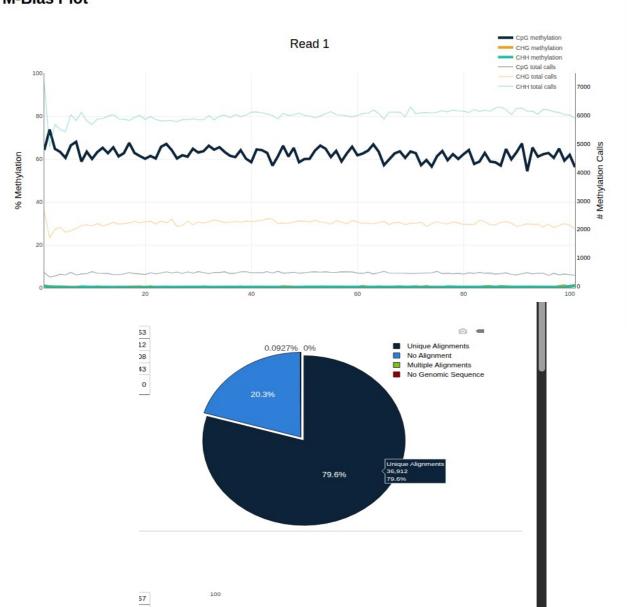
ope context one context of in conte

Alignment to Individual Bisulfite Strands

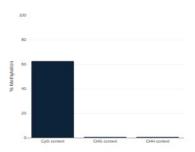
OT	8195	original top strand	
стот	0	complementary to original top strand	
стов	0	complementary to original bottom strand	
ОВ	28717	original bottom strand	



M-Bias Plot

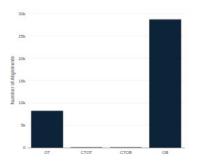


Total C's analysed	875967
Methylated C's in CpG context	29030
Methylated C's in CHG context	851
Methylated C's in CHH context	2211
Methylated C's in Unknown context	3
Unmethylated C's In CpG context	17594
Unmethylated C's In CHG context	222073
Unmethylated C's In CHH context	604208
Unmethylated C's in Unknown context	532
Percentage methylation (CpG context)	62.3%
Percentage methylation (CHG context)	0.4%
Percentage methylation (CHH context)	0.4%
Methylated C's in Unknown context	0.6%



Alignment to Individual Bisulfite Strands

original top strand	8195	ОТ
complementary to original top strand	0	стот
complementary to original bottom strand	0	стов
original bottom strand	28717	ОВ



Cytosine Methylation after Extraction

Total C's analysed	875967
Methylated C's in CpG context	29030
Methylated C's in CHG context	851
Methylated C's in CHH context	2211
Unmethylated C's in CpG context	17594
Unmethylated C's in CHG context	222073
Unmethylated C's in CHH context	604208
Percentage methylation (CpG context)	62.3%
Percentage methylation (CHG context)	0.4%
Percentage methylation (CHH context)	0.4%

