

A Report on CpG Methylation in Early Mouse Embryonic Development

WGBS samples - (sequencing of the whole bisulfite genome)

Chromosome Used: Chromosome 19 of Genome mm10 has been used.

FastQC Reports for the three samples

Replicate	Read S	Sequences		GC%		Sequence Length	
		Before Trimming	After Trimming	Before Trimming	After Trimmin	Before Trimmin	After Trimming
Replicate1							
	Epiblas t D1	163183367	15925094 9	22	22	19-140	20-140
	Epiblas t	163183367	15925094 9	22	22	19-140	20-140
	ICM R1	108333570	10555753	26	25	19-140	20-140
	ICM R2	108333570	10555753	27	25	19-140	20-140
Replicate 2							
	Epiblas t D1	210268052	20732474 2	22	22	19-140	20-140
	Epiblas t	210268052	20732474 2	22	22	19-140	20-140
	ICM R1	98895869	98895869	26	26	19-140	20-140
	ICM R2	98976468	98895869	26	27	19-140	20-140

Bismark:

ICM Methylation rep_1:

Reads uniquely mapped to genome:

No. of reads: 2843419

% of reads: 2.69%

Before extraction:

Methylated C's in CpG context: 2274739

Unmethylated C's in CpG context: 5523775

Percentage Methylation CpG context: 29.2%

After extraction:

Methylated C's in CpG context: 1398563

Unmethylated C's in CpG context: 3835210

Percentage Methylation CpG context: 26.7%

Epiblast Methylation rep_1:

Reads uniquely mapped to genome:

No. of reads: 4626992

% of reads: 2.91%

Before extraction:

Methylated C's in CpG context: 9294581

Unmethylated C's in CpG context: 2770030

Percentage Methylation CpG context: 77.0%

After extraction:

Methylated C's in CpG context: 7311455

Unmethylated C's in CpG context: 2203413

Percentage Methylation CpG context: 76.8%

ICM Methylation rep_2:

Reads uniquely mapped to genome:

No. of reads: 2602002

% of reads: 2.63%

Before extraction:

Methylated C's in CpG context: 2137859

Unmethylated C's in CpG context: 5363783

Percentage Methylation CpG context: 28.5%

After extraction:

Methylated C's in CpG context: 1208182

Unmethylated C's in CpG context: 3509345

Percentage Methylation CpG context: 25.6%

Epiblast Methylation rep_2:

Reads uniquely mapped to genome:

No. of reads: 5610925

% of reads: 2.67%

Before extraction:

Methylated C's in CpG context: 11919864

Unmethylated C's in CpG context: 3525298

Percentage Methylation CpG context: 77.2%

After extraction:

Methylated C's in CpG context: 8994674

Unmethylated C's in CpG context: 2695876

Percentage Methylation CpG context: 76.9%

Bismark reports:

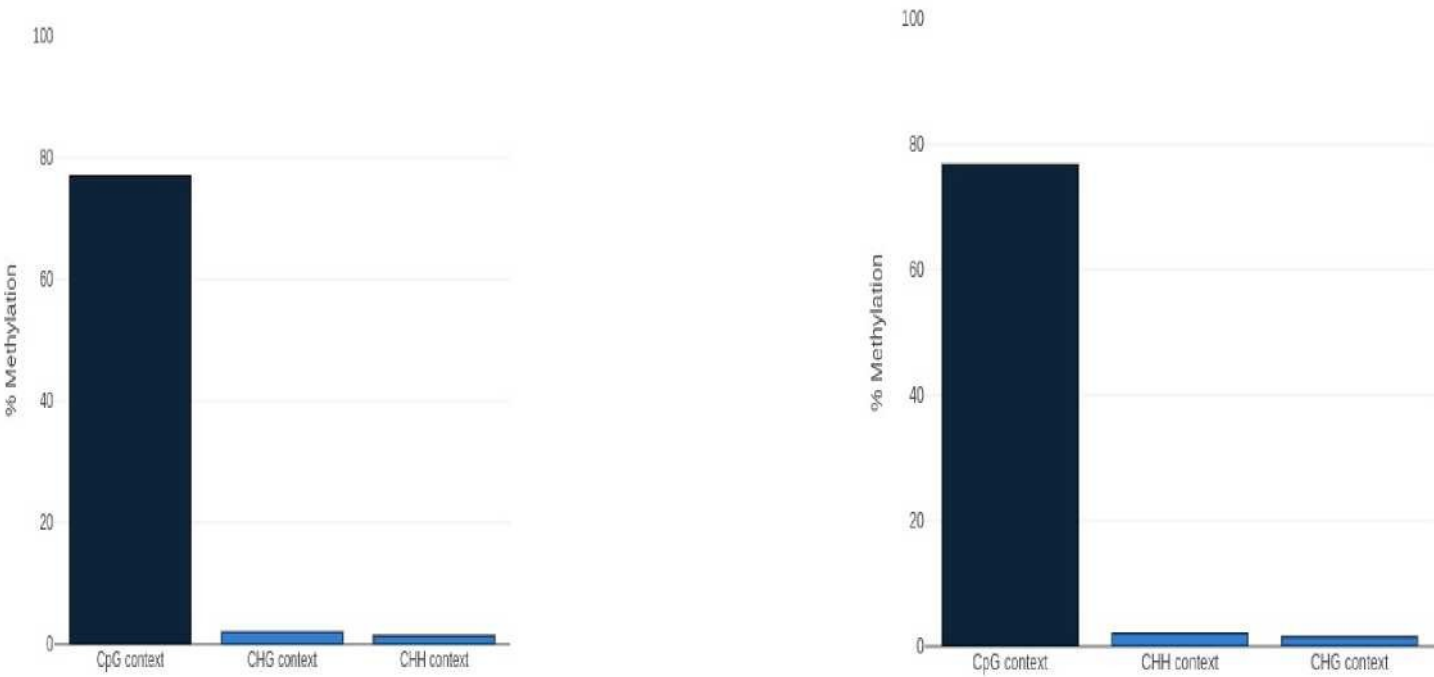


Fig.1: Bismark reports: Before Bismark Methylation extraction (left) and after Bismark Methylation Extraction (right) in Epiblast Rep1.

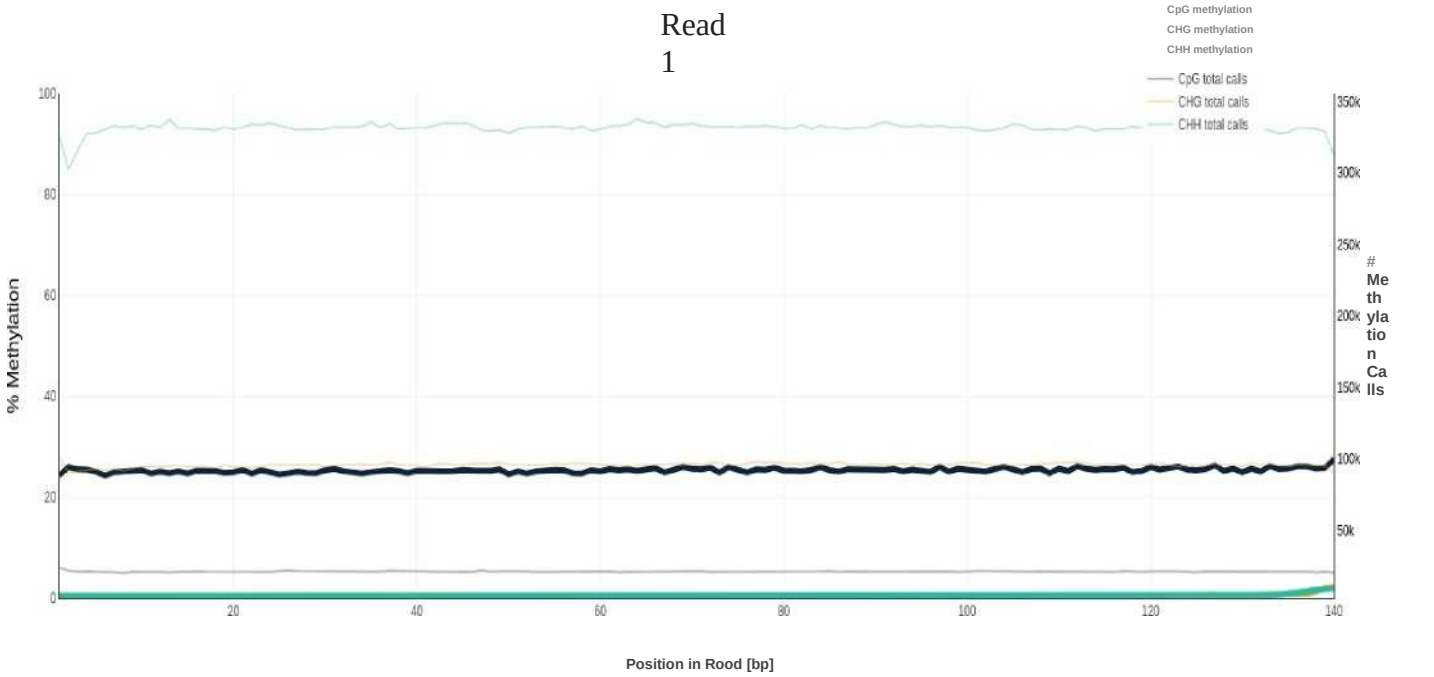
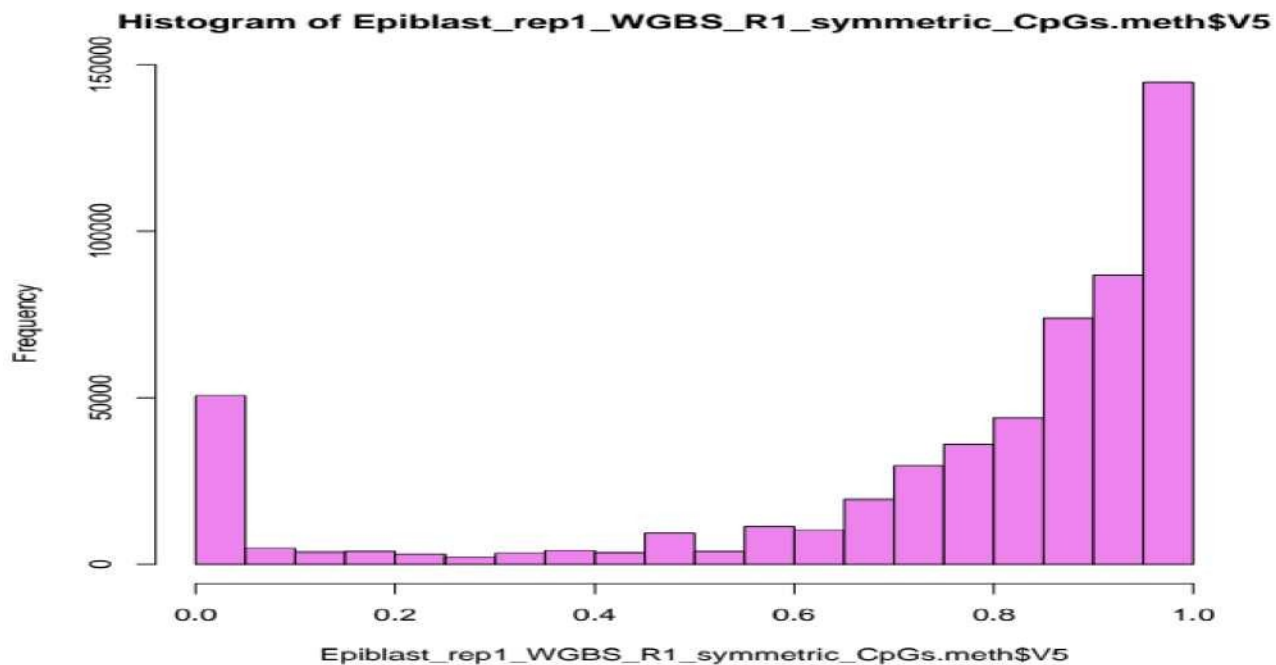
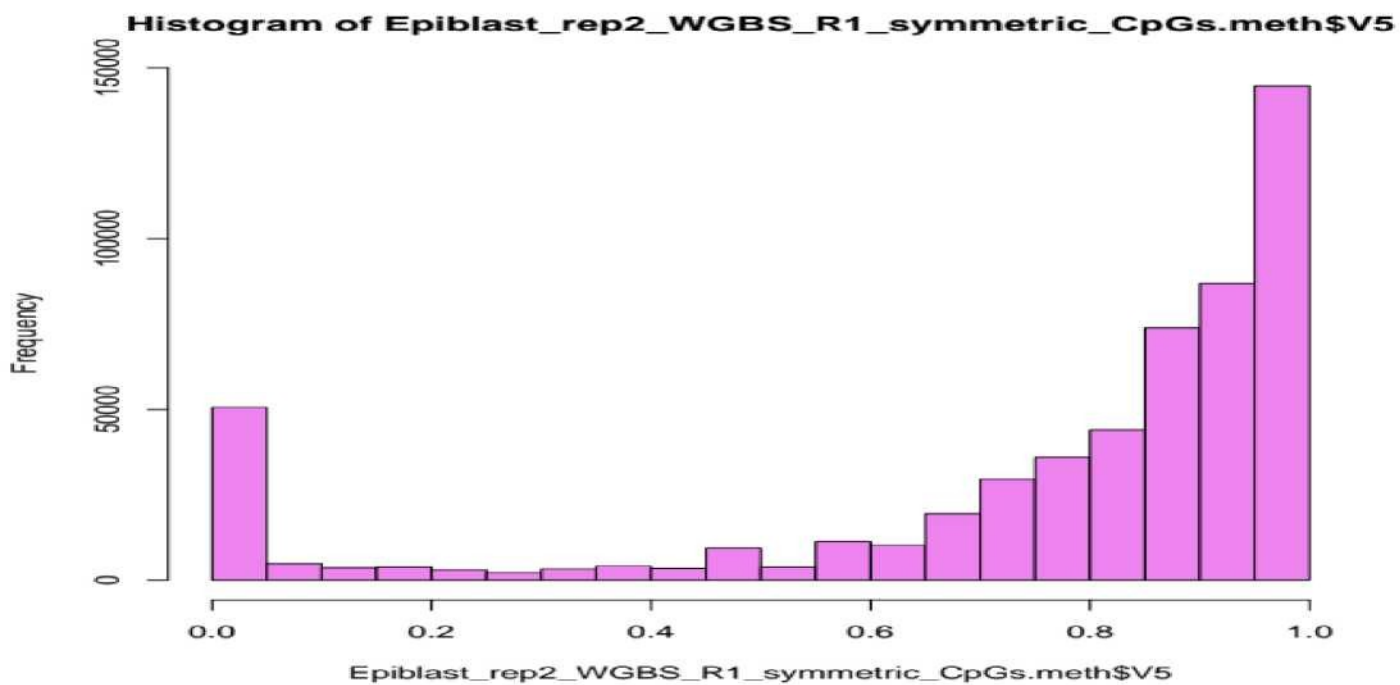


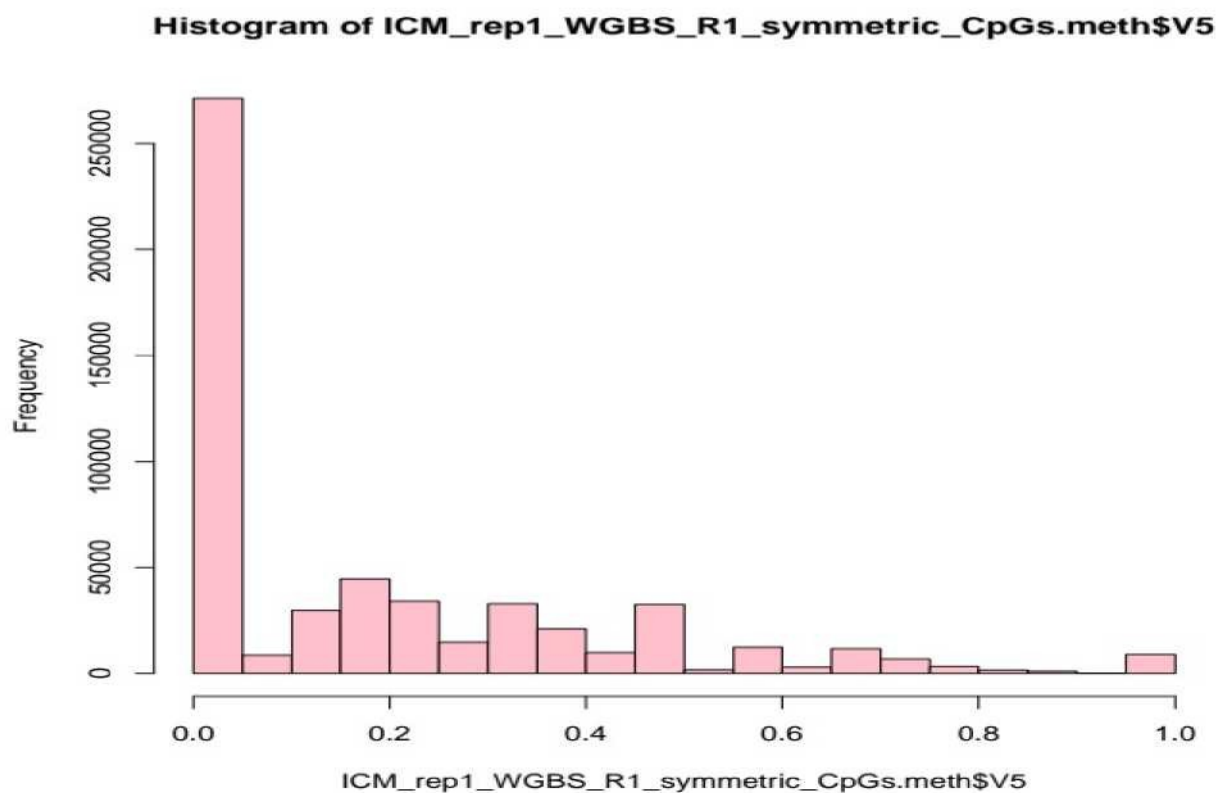
Fig.2: ICM Rep 2 M-Bias plot



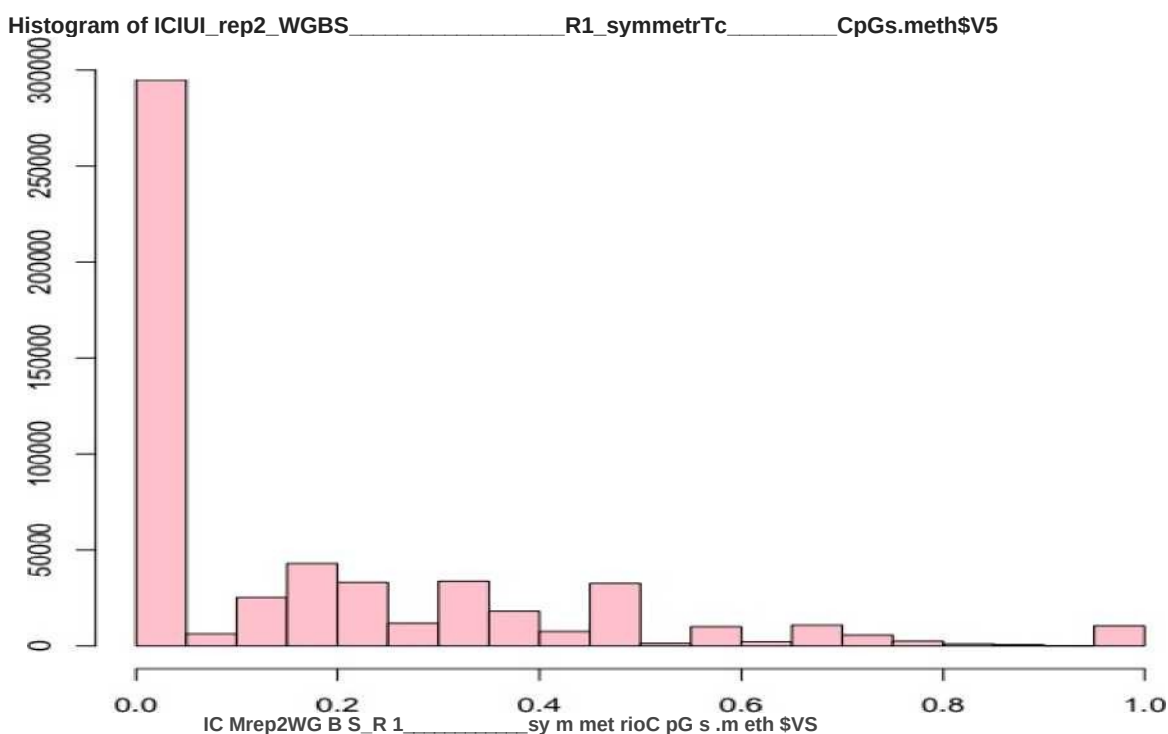
Average methylation level for Epiblast_rep1 =
0.7588857



Average methylation level for Epiblast_rep2 =
0.7588857



Average methylation level for ICM_rep1 = 0.1850298



Average methylation level for ICM_rep2 = 0.1711466

Fig.3: CpG Cytosine Methylation Levels Histograms: These graphs were generated using R

Genome Browser Tracks:

Methylation and CpG analysis:

track type=bigWig visibility=full name="Epiblast_rep1_WGBS_1.methylation.bigWig"
color=197,172,24 maxHeightPixels=24 viewLimits=0:10 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
Epiblast_rep1_WGBS_1.methylation.bigWig?raw=true/
Epiblast_rep1_WGBS_1.methylation.bigWig](https://github.com/j08160457248/methResult/blob/master/Epiblast_rep1_WGBS_1.methylation.bigWig?raw=true/Epiblast_rep1_WGBS_1.methylation.bigWig)

track type=bigWig visibility=full name="Epiblast_rep1_WGBS_1.coverage.bigWig"
color=0,0,0 maxHeightPixels=24 viewLimits=0:20 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
Epiblast_rep1_WGBS_1.coverage.bigWig?raw=true/
Epiblast_rep1_WGBS_1.coverage.bigWig](https://github.com/j08160457248/methResult/blob/master/Epiblast_rep1_WGBS_1.coverage.bigWig?raw=true/Epiblast_rep1_WGBS_1.coverage.bigWig)

track type=bigWig visibility=full name="Epiblast_rep2_WGBS_1.methylation.bigWig"
color=197,172,24 maxHeightPixels=24 viewLimits=0:10 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
Epiblast_rep2_WGBS_1.methylation.bigWig?raw=true/
Epiblast_rep2_WGBS_1.methylation.bigWig](https://github.com/j08160457248/methResult/blob/master/Epiblast_rep2_WGBS_1.methylation.bigWig?raw=true/Epiblast_rep2_WGBS_1.methylation.bigWig)

track type=bigWig visibility=full name="Epiblast_rep2_WGBS_1.coverage.bigWig"
color=0,0,0 maxHeightPixels=24 viewLimits=0:20 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
Epiblast_rep2_WGBS_1.coverage.bigWig?raw=true/
Epiblast_rep2_WGBS_1.coverage.bigWig](https://github.com/j08160457248/methResult/blob/master/Epiblast_rep2_WGBS_1.coverage.bigWig?raw=true/Epiblast_rep2_WGBS_1.coverage.bigWig)

track type=bigWig visibility=full name="ICM_rep1_WGBS_1.methylation.bigWig"
color=197,172,24 maxHeightPixels=24 viewLimits=0:10 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
ICM_rep1_WGBS_1.methylation.bigWig?raw=true/ICM_rep1_WGBS_1.methylation.bigWig](https://github.com/j08160457248/methResult/blob/master/ICM_rep1_WGBS_1.methylation.bigWig?raw=true/ICM_rep1_WGBS_1.methylation.bigWig)

track type=bigWig visibility=full name="ICM_rep1_WGBS_1.coverage.bigWig" color=0,0,0
maxHeightPixels=24 viewLimits=0:20 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
ICM_rep1_WGBS_1.coverage.bigWig?raw=true/ICM_rep1_WGBS_1.coverage.bigWig](https://github.com/j08160457248/methResult/blob/master/ICM_rep1_WGBS_1.coverage.bigWig?raw=true/ICM_rep1_WGBS_1.coverage.bigWig)

track type=bigWig visibility=full name="ICM_rep2_WGBS_1.methylation.bigWig"
color=197,172,24 maxHeightPixels=24 viewLimits=0:10 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
ICM_rep2_WGBS_1.methylation.bigWig?raw=true/ICM_rep2_WGBS_1.methylation.bigWig](https://github.com/j08160457248/methResult/blob/master/ICM_rep2_WGBS_1.methylation.bigWig?raw=true/ICM_rep2_WGBS_1.methylation.bigWig)

track type=bigWig visibility=full name="ICM_rep2_WGBS_1.coverage.bigWig" color=0,0,0
maxHeightPixels=24 viewLimits=0:20 autoScale=off alwaysZero=on
bigDataUrl=[https://github.com/j08160457248/methResult/blob/master/
ICM_rep2_WGBS_1.coverage.bigWig?raw=true/ICM_rep2_WGBS_1.coverage.bigWig](https://github.com/j08160457248/methResult/blob/master/ICM_rep2_WGBS_1.coverage.bigWig?raw=true/ICM_rep2_WGBS_1.coverage.bigWig)

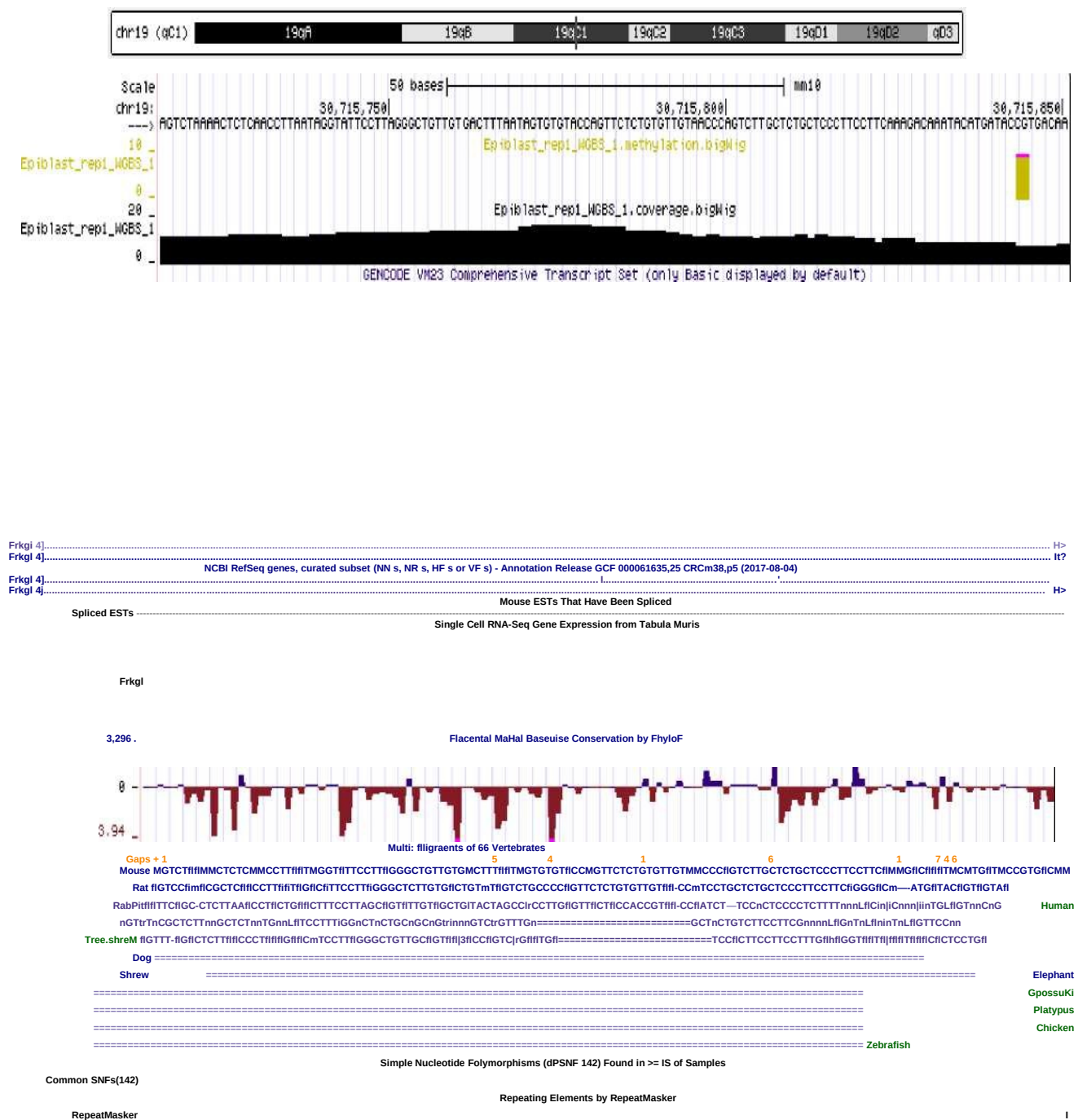
Session Links: <https://genome.ucsc.edu/cgi-bin/hgSession>

https://genome.ucsc.edu/s/Josephluper/Epiblast_rep1_WGBS_1

https://genome.ucsc.edu/s/Josephluper/Epiblast_rep2_WGBS_1

https://genome.ucsc.edu/s/Josephluper/ICM_rep1_WGBS_1

https://genome.ucsc.edu/s/Josephluper/ICM_rep2_WGBS_1



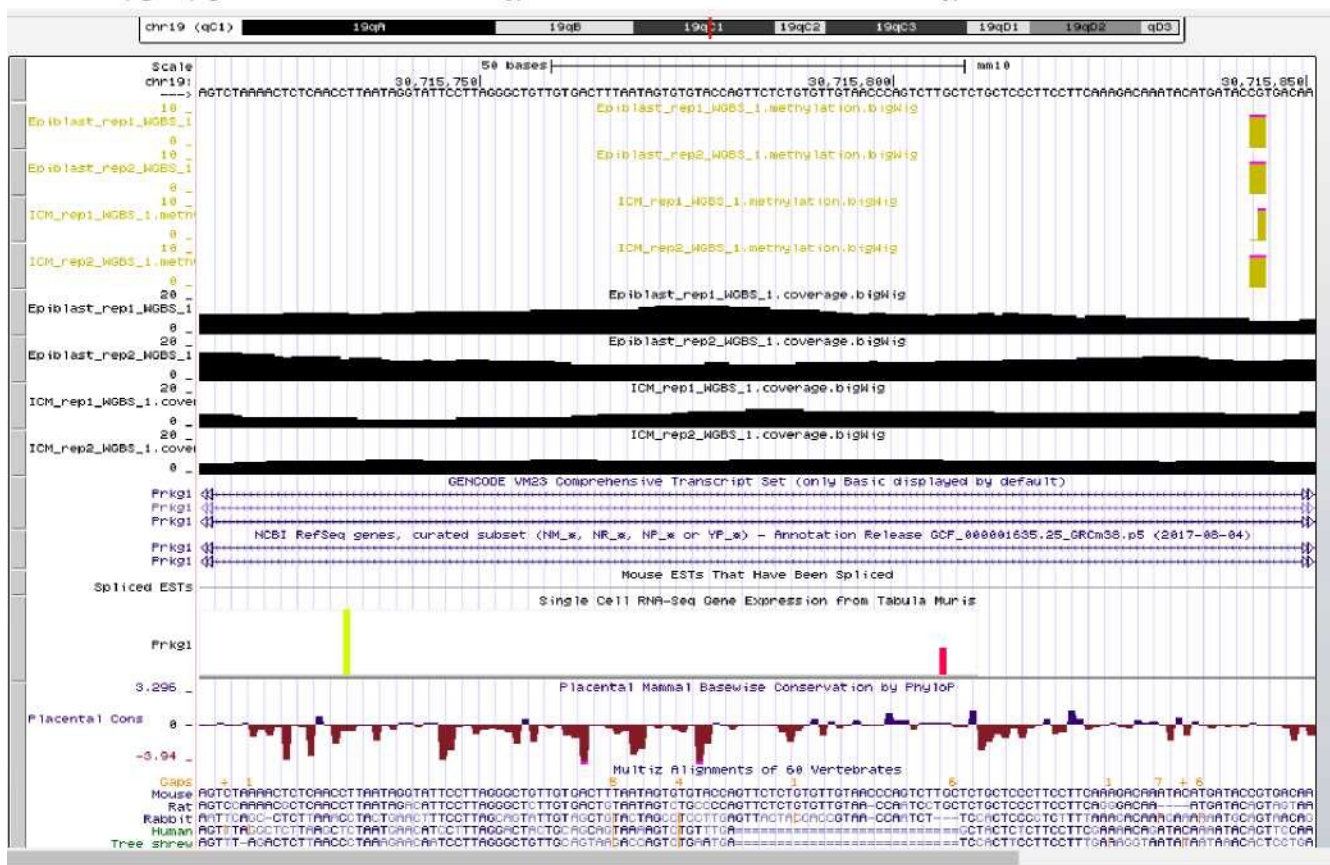


Fig.4:UCSC Genome Browser, cytosine Methylation and coverage overview at position chromosome 19

DISCUSSION/CONCLUSION

From the analysis conducted on ICM and Epiblast, while ICM cells were hypomethylated more, epiblast were hypermethylated more. From my bismark extraction result, it can be seen that the percentage of methylation CpGs in epiblast far exceeds that of ICM in both replicates. This means that epiblast has been methylated more than ICM. On the otherhand, the number of unmethylated C's in CpG context of ICM is in abundance compared to epiblast unmethylated C's in CpG context. This means that ICM is hypomethylated while epiblast is hypermethylated. This confirms the believe that during the development of the embryo, CpG methylation decreases (about 25%) in the early stages as can be seen here in the case of ICM while during tissue differentiation, the CpG content increases significantly (about 90%) and remains throughout the life of the body as we can see here in the case of epiblast. This means that CpG methylation plays a crucial role in ICM cells pluripotency and cell differentiation. This therefore accounts for the global reduction levels of CpGs in the inner cell mass (ICM) which is accompanied by undetectable levels of DNA methyltransferase of each class in the nuclei of ICM while the levels of CpGs in epiblast increases as cell differentiation occurs.

Commands Used

Running fastq

```
```bash
fastq 8cell_rep1_trimmed_WGBS_R1.fastq.gz 8cell_rep1_trimmed_WGBS_R2.fastq.gz
8cell_rep2_trimmed_WGBS_R1.fastq.gz 8cell_rep2_trimmed_WGBS_R2.fastq.gz
Epiblast_rep1_trimmed_WGBS_R1.fastq.gz Epiblast_rep1_trimmed_WGBS_R2.fastq.gz
Epiblast_rep2_trimmed_WGBS_R1.fastq.gz
Epiblast_rep2_trimmed_WGBS_R2.fastq.gz ICM_rep1_trimmed_WGBS_R1.fastq.gz
ICM_rep1_trimmed_WGBS_R2.fastq.gz ICM_rep2_trimmed_WGBS_R1.fastq.gz
ICM_rep2_trimmed_WGBS_R2.fastq.gz
```
```

Bismark Genome Preparation

```
```bash
to-mr -v -m bismark -o Epiblast_rep1_WGBS_1.deduplicated.mr
Epiblast_rep1_WGBS_1.deduplicated.bam
```
```

Sorting

```
```bash
sort -k 1,1 -k 2,2n -k 3,3n -k 6,6 -o
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.mr
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.mr
```
```

Alignment to chr19

```
```bash
methpipe/methpipe-3.4.3/bin/analysis/methcounts -cpg-only -c mouse_chr/chr19.fa -o
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.all_CpGs.meth
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated
```
```

Getting mr

```
```bash
methpipe/methpipe-3.4.3/src/analysis/methcounts -cpg-only -c mouse_chr/chr19.fa -o
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.all_CpGs.meth
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.mr
```
```

Getting symmetrics

```
```bash
methpipe/methpipe-3.4.3/src/utils/symmetric-cpgs -v -o
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.symmetric_CpGs.meth
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.all_CpGs.meth
```
```

```
```
```

Generating BigWig files

```
```bash
```

```
bedGraphToBigWig
```

```
Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.bedGraph
```

```
mm10.chromosomeSizes
```

```
Epi1.bw
```

```
```
```

Getting bg

```
```bash
```

```
bedtools
```

```
genomecov
```

```
-ibam
```

```
ICM_rep2_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.bam -
```

```
bg | head
```

```
```
```

```
```bash
```

```
bedtools genomecov -ibam
```

```
../bedGraphs/Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.bam -bg
```

```
```
```

Getting epicov.bg

```
```bash
```

```
bedtools genomecov -ibam
```

```
../Epiblast_rep1_trimmed_WGBS_R1_bismark_bt2_pe.deduplicated.bam -bg | cat >
```

```
epicov.bg
```

```
```
```

Getting epi2cov.bw

```
```bash
```

```
bedGraphToBigWig epi2cov.bg mm10.chromosomeSizes epi2cov.bw
```

```
```
```

Getting methdiff

```
```bash
```

```
methpipe/methpipe-3.4.3/bin/methdiff -o Epiblast_Rep1_Rep2.methdiff
```

```
Epiblast_rep1_WGBS_R1_symmetric_CpGs.meth
```

```
Epiblast_rep2_WGBS_R1_symmetric_CpGs.meth
```

```
```
```

Generating hmr file

```
```bash
```

```
methpipe/methpipe-3.4.3/bin/hmr -o Epiblast_rep1.hmr
```

```
Epiblast_rep1_WGBS_R1_symmetric_CpGs.meth
```

```
```
```

```
```bash
```

Desktop/methpipe/methpipe-3.4.3/bin/dmr

Epiblast_rep1_rep2.methdiff

Epiblast_rep1.hmr

Epiblast_rep2.hmr \ DMR_rep1_lt_rep2 DMR_rep2_lt_rep1

...

Each step above was carried out on all the four samples

R Commands Used in generating histogram

```
```bash
```

```
hist(ICM_rep2_WGBS_R1_symmetric_CpGs.meth$V5, col="pink")
```

```
```
```

```
```bash
```

```
pdf(file="saving_ICM_rep2.pdf")
```

```
```
```

```
```bash
```

```
hist(ICM_rep2_WGBS_R1_symmetric_CpGs.meth$V5, col="pink")
```

```
dev.off()
```

```
```
```

References

1. Smith ZD, Shi J, Gu H, Donaghey J, Clement K, Cacchiarelli D, Gnirke A, Michor F, Meissner A. Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature*. 2017 Sep 28;549(7673):543-547. doi: 10.1038/nature23891. Epub 2017 Sep 20. PMID: 28959968; PMCID: PMC5789792.
2. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011 Jun 1;27(11):1571-2. doi: 10.1093/bioinformatics/btr167. Epub 2011 Apr 14. PMID: 21493656; PMCID: PMC3102221.
3. Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. BigWig and BigBed: enabling browsing of large distributed data sets. *Bioinformatics*. 2010 Sep 1;26(17):2204-7.
4. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res*. 2002 Jun;12(6):996-1006.
5. Song Q, Decato B, Hong E, Zhou M, Fang F, Qu J, Garvin T, Kessler M, Zhou J, Smith AD (2013) A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. *PLOS ONE* 8(12): e81148 [PDF] [Publisher Site]

6. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.