Exploration of 5hmC enrichment in 5dpp mouse liver

Jordan Ellis-Pugh and Tamara L. Davis
Biology Department, Bryn Mawr College, Bryn Mawr PA 19010-2899

Introduction to Genomic Imprinting

Genomic imprinting is an epigenetic phenomenon that causes one of the two alleles offspring receive from their parents to be silenced. This silencing occurs when methyl groups are added to DNA, specifically to CpG dinucleotides. Methylation of one allele creates a difference between the two copies of a single gene, promoting monoallelic expression. The silencing of one allele of imprinted genes is necessary for proper development. When the expression patterns of imprinted genes are disrupted, this can lead to developmental disorders, such as Prader-Willi and Angelman syndromes. The potential for these disorders to occur makes the search for the regulatory factors of genomic imprinting essential.

The regulatory elements associated with imprinted genes are classified as primary or secondary based on when the methylation appears on the locus. Primary differentially methylated regions (DMRs) acquire methylation during gametogenesis and function to control imprinting across a cluster of imprinted genes. Secondary DMRs acquire their methylation during embryonic development. These DMRs influence individual genes and are not as well maintained as their primary counterparts. DNA methyltransferases (DNMTs) are responsible for the establishment and maintenance of the methylation at the DMRs. We are investigating the reasons for the difference in the stability of methylation between the two types of DMRs.

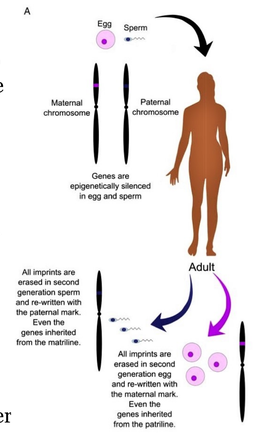


Figure 1. Representation of genomic imprinting, as it is passed down through generations.
https://bjsm.bmj.com/content/49/24/1567

oxBS Sequencing and Analysis

The DNA base 5-methylcytosine (5mC) is a well-known epigenetic modification, playing an important role in the silencing of methylated alleles. While the mechanisms of methylation have been investigated at length, the processes of demethylation are still being studied. The TET enzyme-mediated oxidation of 5mC into 5-hydroxymethylcytosine (5hmC) is one proposed mechanism of demethylation. 5hmC can be further oxidized into 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). These less stable forms of cytosine can be removed either actively or passively. ***We hypothesize that secondary DMRs display less stable methylation patterns due to enrichment of 5hmC*** which signifies a loss of stable methylation. This could explain why primary and secondary DMRs show different levels of methylation maintenance.

In order to detect and quantify methylation in the forms of 5mC and 5hmC, a method known as the oxidative bisulfite (oxBS) protocol was implemented. Samples were processed in parallel, with one receiving an oxidative treatment. This treatment converted any 5hmC to 5fC. Both samples then received a bisulfite treatment, which converted unmodified cytosines and cytosines with unstable modifications (i.e. 5fC) to uracil. The sample that did not receive the oxidative treatment conveys the amount of 5mC and 5hmC present at the target locus (Figure 2, left). The sample that was oxidized indicates the amount of 5mC present (Figure 2, right). Subtracting the amount of cytosine in the oxidized sample from the cytosine in the bisulfite only treated sample reveals the amount of 5hmC present at the locus, illustrating the level of 5hmC enrichment.

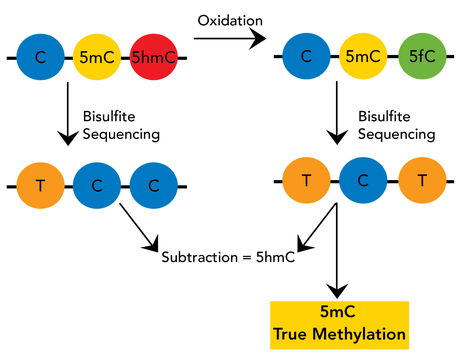
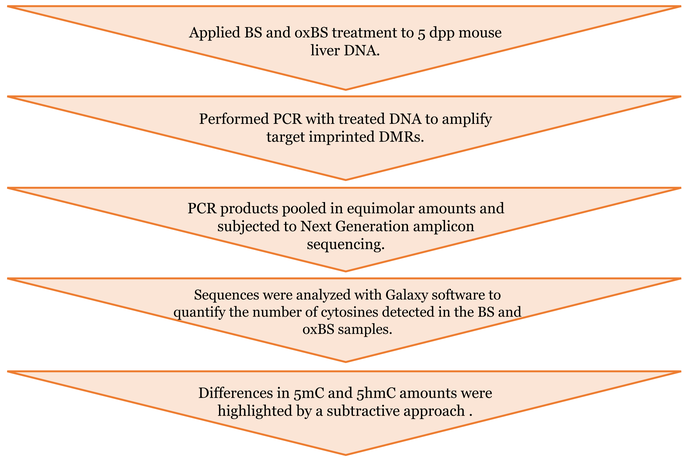


Figure 2. Model of the oxBS protocol. Depicts example starter sequences, before and after oxidation. The 5hmC seen in the first sequence is converted, via oxidation, into 5fC. The bisulfite treatment followed by PCR and sequencing converts unmodified cytosine to thymine in both cases. 5fC is also read as thymine. 5mC and 5hmC are read as cytosine. The resulting cytosine from the oxBS sample demonstrated the 5mC present at the gene's DMR. The subtraction of the cytosine in the oxBS sample from the cytosine in the BS sample represents the amount of 5hmC at the DMR. https://www.nugen.com/products/truemethyl-oxbs-seq-module

Objective

We characterized the methylation patterns of a group of 6 primary and 9 secondary imprinted DMRs. The oxBS protocol was utilized to detect enrichment of 5hmC. Focusing on 5-day post-partum (5 dpp) mouse liver, I analyzed and compared the amounts of 5mC and 5hmC present at targeted DMRs. Levels of 5mC and 5hmC will illuminate how methylation patterns vary across different DMRs, tissues and developmental stages.

Methods



Troubleshooting Data Analysis

Issue: Mapping NGS reads against all target sequences yielded high % of cytosine detected in non-CpG context which could indicate inefficient bisulfite mutagenesis.

Solution: Mapped NGS reads against single DMR sequences yielded non-CpG methylation percentages <1%.

Original Analysis

	BS	oxBS
CHG	1.2%	16.7%
CHH	1.7%	24.9%

Sample Single DMR Analysis

	Snrpn BS	Snrpn oxBS
CHG	0.5%	0.2%
CHH	0.4%	0.2%

Figure 3. Percentages of methylated cytosines analyzed in non-CpG contexts. We are specifically interested in methylation in CpG contexts. The Original Analysis refers to the alignment of the data against the all the target sequences. The Sample Single DMR Analysis refers to the alignment of the data against a representative individual target sequence. The Sample Single DMR Analysis illustrated better quality control, with lower amounts of non-CpG methylation.

Issue: Analysis of the same data set with different order of input reads yielded different results.

Solution: Ran paired-read analysis putting Read 2 before Read 1, yielding higher cytosine counts than previous runs, resulting in more data to analyze. Offered a consistent way to analyze data in the future.

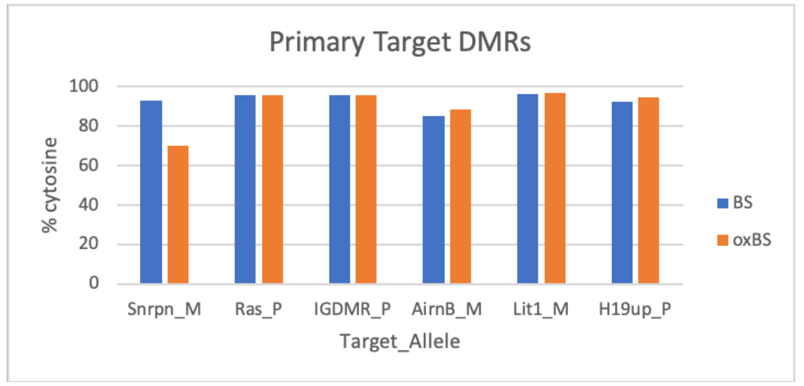
Target DMR	Old BS Count	New BS Count	Old oxBS Count	New oxBS Count
Snrpn	4096	12317	12383	28901
IGDMR	6011	23418	25750	106189
Lit1	739	25679	3513	117201
H19 ICR	196	187	856	1151
Magel2	23743	97476	68400	243258
Dlk1	14728	72580	86782	304845
Ndn	2432	36092	545	742
Gtl2	147	14	6197	118
Igf2r DMR1	2805	26993	140064	221204
p57	112	279	100	119

Figure 4. Cytosine counts from a selection of target DMRs. Old counts refer to cytosine counts collected putting Read 1 before Read 2 during the mapping of the sequence data against the known target sequences. New counts refer to cytosine counts collected when Read 2 was put before Read 1 during mapping. Comparisons are set up between each treatment, BS and oxBS. In many cases, there was a large increase in analyzed cytosines in the new runs compared to the old runs.

Results and Interpretations

Secondary DMRs display less stable methylation than primary DMRs. We propose that this instability is due to 5hmC enrichment at secondary DMRs. Primary DMRs would be expected to show low levels of 5hmC, which was the case for 5 of our 6 targets. *Snrpn* was the exception, displaying dramatically different levels of cytosine in the BS vs. the oxBS samples. We further predicted that 5hmC would be enriched at secondary DMRs, resulting in less cytosine in oxBS samples vs. the BS samples. Three of our 9 secondary targets, *Dlk1, Ndn* and *Gtl2*, displayed 5hmC enrichment.

For some loci (such as *Gtl2* and *p57*), the sample sizes were very small and may not accurately represent the methylation status at those loci. Experimental replicates are necessary to capture an accurate image of methylation at our target DMRs.



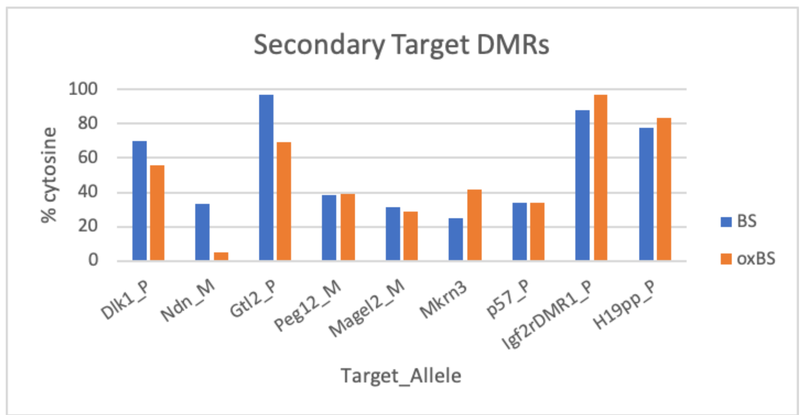


Figure 5. Charts of the % cytosine at the 15 target loci. The top chart shows the primary DMRs, with the secondary DMRs depicted in the bottom chart. The blue bars indicates samples with only the bisulfite treatment (BS). The orange bars indicate samples with both the oxidative and bisulfite treatments (oxBS). The x-axis indicates the target locus followed by the methylated parental allele (M = maternal, P = paternal). The y-axis indicates the % cytosine found in each sample. A decrease in % cytosine in the oxBS samples vs. the BS samples indicates 5hmC enrichment, based on the subtractive approach mentioned earlier. The *Mkrn3* locus does not have a designated allele due to a lack of SNP to differentiate between maternal and paternal inheritance.

Future Directions

- Perform BS and oxBS experiments on different tissues and developmental stages as a means of comparing 5hmC enrichment.
- Run paired-read analysis with data collected by previous research students using new conventions (i.e. putting Read 2 before Read 1).

Acknowledgements

I'd like to thank the Bryn Mawr College Summer Science Research program for support. Additional support for this research was provided by awards to TLD from the National Science Foundation (NSF grant #1514600) and the Bryn Mawr College Data Science initiative.