Hydroxymethylation technologies

Mikhail Dozmorov

Spring 2018

Methylation

- The two major epigenetic modifications of cytosines, 5-methylcytosine (5-mC) and 5-hydroxymethylcytosine (5-hmC), coexist with each other in a range of mammalian cell populations.
- 5-hmC was discovered in 2009, abundant in embtyonic stem cells and neurons.
- Sequence- and strand-specific mark
- Near but not on transcription factor binding sites
- 5-hmC is reciprocal to 5-mC - high level of one suggests low level of the other

Hydroxymethylation

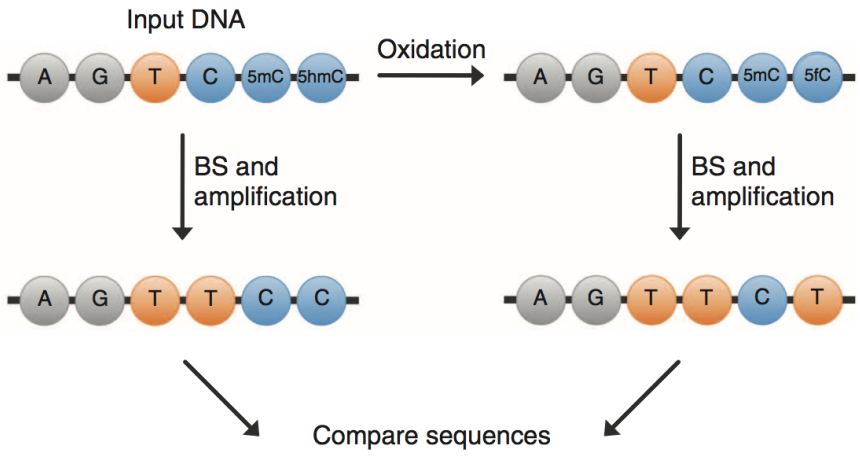
- 5-hydroxymethylcytosine (5-hmC) is an oxidation product of the extensively studied 5-methylcytosine (5-mC) modification
- It has been observed at substantial levels in both somatic and embryonic stem cells (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009).
- 5-hmC is involved in epigenetic regulation

Technology

- **Bisulfite sequencing (BS-seq)** - treatment with sodium bisulfite converts unmethylated cytosines to uracils, but does not distinguish between 5-mC and 5-hmC (Huang et al., 2010), and consequently the yield of methylation from BS-seq is the sum of 5-mC and 5-hmC levels
- Two recently developed techniques, **oxidative bisulfite sequencing (oxBS-seq)** (Booth et al., 2012) and **Tet-Assisted Bisulfite sequencing (TAB-seq)** (Yu et al., 2012), provide high-throughput single-base resolution measurements of 5-mC and 5-hmC, respectively.
- Combining two technologies can be used to estimate joint levels of 5-mc and 5-hmc

oxBS-seq

- Oxidative bisulfite sequencing
- Chemical oxidation of 5hmC to 5-formylcytosine (5fC) enables bisulfite conversion of 5fC to uracil

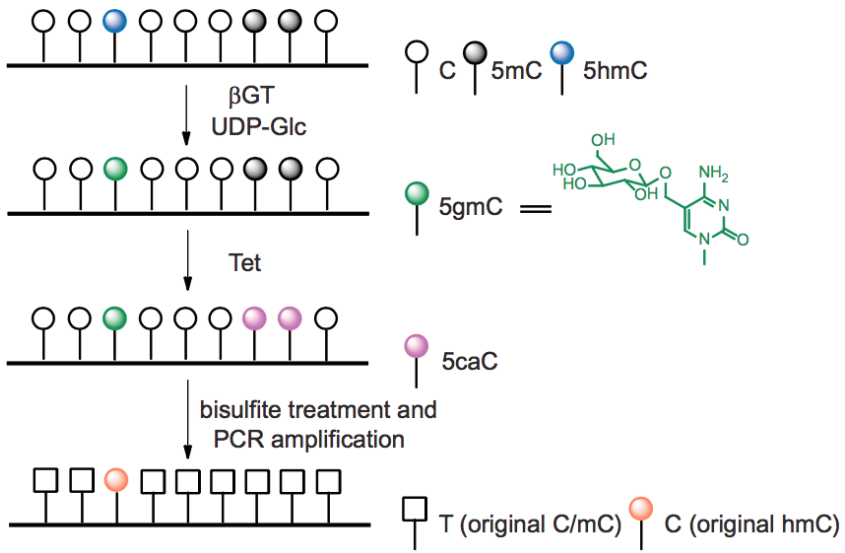


Base	Sequence	BS Sequence	oxBS Sequence
C	C	T	T
5mC	C	C	C
5hmC	C	C	T

Booth, Michael J., Miguel R. Branco, Gabriella Ficz, David Oxley, Felix Krueger, Wolf Reik, and Shankar Balasubramanian. "Quantitative Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine at Single-Base Resolution." Science (New York, N.Y.) 336, no. 6083 (May 18, 2012): 934–37. https://doi.org/10.1126/science.1220671.

TAB-seq

- Tet-assisted bisulfite sequencing
- TET proteins oxidize 5mc fo 5hmc to 5caC - reads as uracil after bisulfite treatment
- 5hmC glycosylation protects 5hmC from TET oxidation - reads as C after bisulfite treatment



Yu, Miao, Gary C. Hon, Keith E. Szulwach, Chun-Xiao Song, Liang Zhang, Audrey Kim, Xuekun Li, et al. "Base-Resolution Analysis of 5-Hydroxymethylcytosine in the Mammalian Genome." Cell 149, no. 6 (June 2012): 1368–80. https://doi.org/10.1016/j.cell.2012.04.027.

Analysis

- When combining BS-seq with TAB-seq or oxBS-seq, the 5-mC level at a given CpG site can be estimated by subtracting the 5-hmC level (TAB/oxBS-seq) from the combined 5-mC:5-hmC level (BS-seq).
- Maximum likelihood methylation levels (MLML) for simultaneous estimation of 5-mC and 5-hmC, combining data from any two of BS-seq, TAB-seq or oxBS-seq, or all three when available.

http://smithlabresearch.org/software/mlml/

Qu, Jianghan, Meng Zhou, Qiang Song, Elizabeth E. Hong, and Andrew D. Smith. "MLML: Consistent Simultaneous Estimates of DNA Methylation and Hydroxymethylation." Bioinformatics (Oxford, England) 29, no. 20 (October 15, 2013): 2645–46. https://doi.org/10.1093/bioinformatics/btt459.