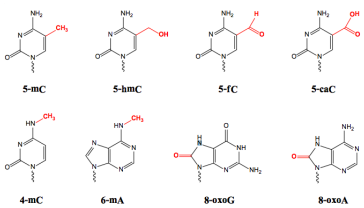
Detecting epigenetic modifications by Single Molecule, Real-Time (SMRT) Sequencing

• Introduction

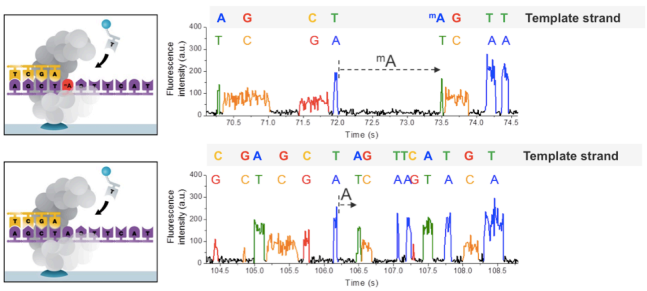
DNA modifications are actively involved in many biological functions such as gene expression, host-pathogen interactions DNA damage, and DNA repair. Currently most high-throughput techniques only focus on cytosine methylation. Furthermore, these techniques usually need to convert unmethylated cytosine to uracil nucleotides by bisulfite treatment and then compare sequence reads from bisulfate-treated and untreated samples*. In contrast, SMRT sequencing can directly detect base modifications by measuring the kinetics of base addition during the normal course of sequencing. These kinetic measurements present characteristic patterns for a wide variety of base modifications. SMRT sequencing has the potential to revolutionize the study of DNA modifications.

• Types of Base Modification



- Markers that influence gene expression: 5-mC, 5-hmC, 5-fC, and 5-caC
- Bacterial markers that affect host-pathogen interactions: 6-mA, 4-mC, and 5-mC
- Bacterial markers for regulating DNA replication and repair and transcription regulation: 6-mA
- DNA damage products : 8-oxoG and 8-oxoA

• Principle



- Interpulse duration (IPD): the time duration between two successive base additions, which can be altered by a modified base in the DNA template.
- IPD ratio: the ratio of the mean IPD at a site in the native sample to the mean IPD at the same site in the amplified control.

The presence of a modified base (6-mA) in the template (top) causes a delayed incorporation of the corresponding T nucleotide (longer IPD) compared to a control DNA template lacking modification (bottom).

• Experiment

1. Obtain the DNA template of interest and prepare a SMRT™ library for SMRT sequencing
2. Take a small aliquot of DNA sample and perform a whole-genome amplification reaction to get the control DNA sample lacking any base modifications. Then prepare a SMRT™ library for the control sample
3. Perform adequate SMRT sequencing for both samples to get enough coverage needed for the modifications under study, and also sufficient overall coverage to characterize the genome of interest
4. Perform kinetic analysis for base modification with bioinformatics tools.

• Applications

- Find the modification (5-mC, 4-mC, and 6-mA) sites in *E. coli* genome when studying virulence, gene expression, or pathogen-host interactions
- Detecting modification sites in some small genomes such as mitochondria, chloroplast
- Enriching portions of a larger genome containing modification sites and detect the modified bases

• References

Detecting DNA Base Modifications Using Single Molecule, Real-Time Sequencing, Pacific biosciences (http://www.pacb.com/wp-content/uploads/2015/09/WP_Detecting_DNA_Base_Modifications_Using_SMRT_Sequencing.pdf)

* Clark S.J., Statham A., Stirzaker C., Molloy P.L. & Frommer, M. DNA methylation: bisulphite modification and analysis. Nat. Protocols 1, 2353–2364 (2006).