

Laurie Kurihara¹, Chongyuan Luo^{2,3}, Eran A. Mukamel⁴, Rosa Castanon³, Jacinta Lucero⁵, Joseph R. Nery³, Christopher L. Keown⁴, Cassie Schumacher¹, Tim Harkins¹, M. Margarita Behrens⁵ and Joseph R. Ecker^{2,3}

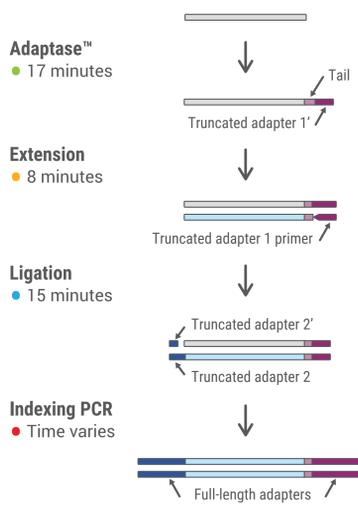
1. Swift Biosciences Inc., Ann Arbor, MI; 2. Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, CA; 3. Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, CA; 4. Department of Cognitive Science, University of California, San Diego, La Jolla, CA; 5. Computational Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA



Abstract

When performing whole genome bisulfite sequencing (WGBS), highly efficient conversion of DNA fragments into library molecules is necessary when input quantity is limited. To meet this need, we developed an efficient library preparation using Adaptase™ for NGS adapter ligation to single-stranded, bisulfite-converted DNA fragments. This method significantly improves library complexity compared to existing commercially available methods. Since comprehensive methylome coverage was achievable from low DNA inputs, this method was modified and applied to single cells for classification of mammalian brain cell types based on methylation pattern. Approximately 20% of the mouse genome contains differential methylation that allows neuronal cell types to be distinguished by low pass WGBS. For example, >200,000 differentially methylated regions were identified among three cortical excitatory and inhibitory neuron types. Starting from single neuronal nuclei isolated from mouse frontal cortex that had undergone lysis and bisulfite conversion, NGS adapters were incorporated onto DNA fragments using a single cycle of random priming followed by Adaptase. The resulting low pass WGBS sequence demonstrated robust neuron type classification, readily separating excitatory and inhibitory cells, and further identified distinct populations of inhibitory cells. This method enabled unbiased characterization of brain epigenomic diversities without the need for the isolation of specific cell populations. Compared to other single cell bisulfite sequencing methods, the Adaptase workflow was faster and produced a higher percentage of aligned reads that increased the per cell data output. This method will meet the need of large-scale single cell methylome profiling of thousands of cells to enable discovery of single cell epigenomic variation.

Accel-NGS® Methyl-Seq



The first adapter is efficiently attached to sheared, bisulfite converted ssDNA fragments using Adaptase.

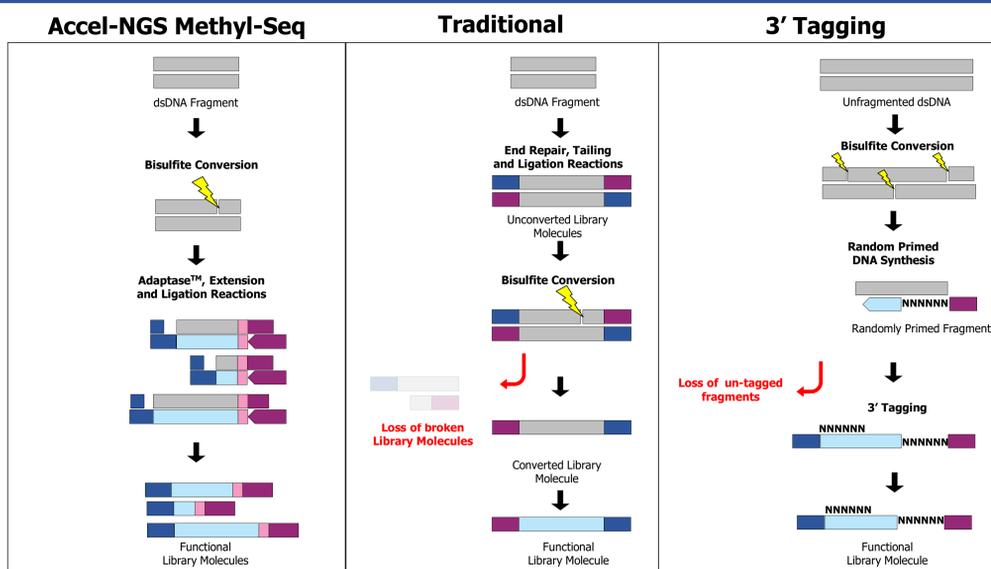
Primer extension produces a non-uracil containing complementary strand.

Conventional ligation introduces the second adapter to the bottom strand.

Library amplification introduces single or dual indexing for multiplexed sequencing.

This enables a high conversion rate of bisulfite converted fragments compared to other methods that produce broken library molecules when preparing the library prior to bisulfite conversion (see Traditional method below), or methods with inefficient adapter attachment (see 3' Tagging method below).

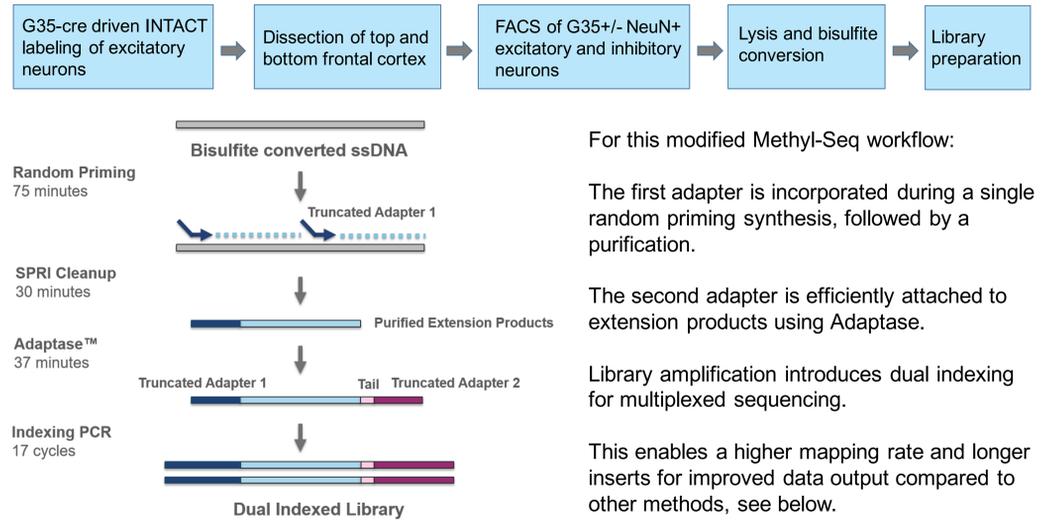
Superior Performance Compared to Other Kits



	METHOD	% READS ALIGNED	GENOME COVERAGE	% DUPLICATE READS	EST. LIBRARY SIZE (MILLIONS)	% CpX MISSING	% CpX COVERED ≥ 10X
100 ng <i>Arabidopsis</i>	Methyl-Seq	89.6	22X	1.9	714	0.56	92.2
	Traditional	80.2	21X	2.7	604	0.57	88.1
	3' Tagging	71.4	16X	22.1	48	7.70	39.4
10 ng <i>Arabidopsis</i>	Methyl-Seq	87.8	22X	2.7	406	0.58	90.4
	Traditional	76.7	19X	11.9	70	0.57	83.9
	3' Tagging	71.9	16X	22.2	45	5.20	45.2
1 ng <i>Arabidopsis</i>	Methyl-Seq	83.3	18X	18.2	38	0.59	77.1
	Traditional	80.7	10X	62.3	6	2.00	17.0
	3' Tagging	73.4	12X	46.1	12	6.60	31.3

Using *Arabidopsis thaliana*, a small genome model organism for methylation analysis, the Swift Accel-NGS Methyl-Seq kit constructed higher complexity libraries and provided comprehensive coverage of CpX (CpG + CpH) sites, making it an ideal choice for developing a single cell method.

Single Cell Methyl-Seq



For this modified Methyl-Seq workflow:

The first adapter is incorporated during a single random priming synthesis, followed by a purification.

The second adapter is efficiently attached to extension products using Adaptase.

Library amplification introduces dual indexing for multiplexed sequencing.

This enables a higher mapping rate and longer inserts for improved data output compared to other methods, see below.

Achieve Higher Mapping Rate

Single cell	Mapping rate to mm10 (mouse)	Mapping rate to hg19 (human)
mouse	40.00%	0.40%
mouse	40.40%	0.30%
human	0.10%	44.00%
human	0.10%	44.20%
mouse	34.50%	0.50%
mouse	36.30%	0.60%
human	0.10%	36.90%
human	0.10%	38.50%
mouse	38.00%	0.30%
mouse	37.30%	0.40%
human	0.10%	40.80%
human	0.00%	40.40%
mouse	37.80%	0.40%
mouse	37.30%	0.80%
human	0.50%	2.20%
human	0.00%	42.60%

To assess mapping efficiency and workflow quality (avoidance of cross-contamination), FACS mouse and human single neurons were processed simultaneously.

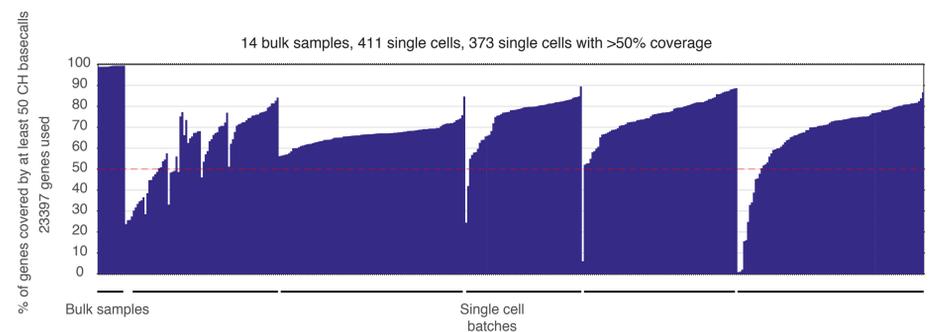
WGBS was performed on a HiSeq® 4000 to an average 4M read depth and aligned to both reference genomes.

Alignment to the corresponding reference achieved an average of ~39% mapped, compared to alignment to the non-reference at ~0.30% mapped.

This is a two-fold increase in mapping rate compared to a single cell workflow that demonstrated an average of 20.1% mapping rate (Smallwood *et al.*, Nature Methods 2014).

The larger insert size of these libraries (>400 bp) also improved per cell data recovery compared to a 3' tagging single cell method (data not shown).

Robust Neuron Type Classification



As shown above, coverage metrics for 373/411 single cells indicate >50% of genes had at least 50 CH calls.

This data was profiled for total non-CG methylation (mCH) in the gene bodies of 1,183 select genes, and the bulk methylomes were downsampled to equalize coverage with single cells.

The data was analyzed using dimensional reduction, clustering and beta-binomial likelihood ratios to classify cell type using >200,000 previously identified DMRs among three cortical excitatory and inhibitory neuron types.

The cell clusters generated were in agreement with known cell type marker genes, as well as known layer-specific marker genes, indicating robust and unbiased neuron classification, without the need for cell-type specific isolation methods. We have modified this workflow and applied to a large scale study (in progress).

Attend Our Exhibitor Event to Learn More

Advancing Epigenetics NGS Sequencing and Analysis to the Single-Cell Level
Thursday October 20 at 1:00 to 2:30 PM
Room 8/15, Convention Centre East Building

Dr. Joseph Ecker, Salk Institute for Biological Studies

“Single Cell Methylomes Distinguish Brain Cell Types”

Dr. Ecker will present results of a large scale study in progress using the Single Cell Methyl workflow.

Dr. Adam Blattler, Active Motif

“New Tools for Studying the Epigenomes of Clinical Samples”

Dr. Blattler will present a novel workflow for low input ChIP-Seq using molecular identifiers (MIDs).

