Roche

Authors

Adriana Geldart
Support & Applications Manager

Rachel W. Kasinskas
Director of Scientific Support & Applications

Roche Sequencing & Life Science
Wilmington, MA, USA

Benjamin G. Barwick
Graduate student

Christopher D. Scharer
Instructor

Jeremy M. Boss
Professor and Chair of the Department of Microbiology and Immunology

Emory University
Atlanta, GA, USA

Date of first publication: 2016

KAPA HyperPrep and HiFi Uracil+ Kits allow for reduced representation bisulfite sequencing of ultra-low amounts of DNA in B Cells

Methyl-Seq is a powerful tool for understanding genome-wide methylation with base-pair resolution. Several strategies have been developed to understand the methylation status of a genome, including reduced representation bisulfite sequencing (RRBS). Bisulfite treatment results in a significant decrease in DNA input and quality; therefore, the construction of high-quality libraries and efficient downstream amplification is critical. The experimental design and results here provide a robust methodology for RRBS analysis using the streamlined chemistry of the KAPA HyperPrep Kit and HiFi HotStart Uracil+ DNA Polymerase and demonstrates its utility in accurately quantifying the DNA methylation status from sample-limiting conditions.

Introduction

DNA methylation is essential for mammalian development and understanding its regulation can provide insight into disease etiology. Base-pair resolution analysis of DNA methylation is facilitated by sequencing of bisulfite-treated DNA, which is employed to convert cytosine residues into uracils, subsequently read as thymines during sequencing. Methylated residues, on the other hand, are left unmodified by the treatment. A comparison of bisulfite-converted to unconverted DNA enables the identification of methylated loci.

Because of the high coverage depth necessary and associated cost to analyze the methylation state of the whole genome, RRBS was developed to reduce the effective sequencing to 1%.[1] In this technique, genomic DNA is first digested with a restriction enzyme that is not affected by the presence of a methylation site, as this will allow for both methylated and unmethylated sites to be digested. *MspI* is a commonly used enzyme for this application.

The construction of high-quality libraries and subsequent downstream amplification is essential, due to the decrease in both input and quality of bisulfite-treated libraries. The use of the KAPA HyperPrep Kit in such a workflow allows for increased yields of adapter-ligated library molecules and reduced amplification bias, translating to higher library diversity, lower duplication rates and more uniform coverage. KAPA HiFI HotStart Uracil+ DNA Polymerase is particularly well suited for the amplification of bisulfite-converted libraries, as it has been engineered for tolerance to uracil residues incorporated during bisulfite treatment.

Experimental design and methods

The experimental design and results described here provide a method to investigate the DNA methylation status of B cells and identifies specific conditions that optimize RRBS library construction.

In this study, B cell genomic DNA was isolated from an eightweek-old C57BL/6 female mouse and used to prepare libraries following the workflow depicted in Figure 1. To summarize, DNA was digested with *MspI* (New England Biolabs), which targets 5'-CCGG-3' sequences. Following AMPure® cleanup, libraries were generated using the KAPA HyperPrep Kit using both 500 ng and 10 ng inputs. To limit adapter cost and increase adapter ligation, a short-adapter strategy was implemented,[2] and adapters were universally methylated on all cytosine residues (Integrated DNA Technologies). The final adapter concentration was titrated for both DNA input amounts, at 500:1, 100:1 and 50:1 adapter:insert molar ratios. Bisulfite conversion was performed after the post-ligation cleanup using the Qiagen Epitect® Bisulfite Kit with minor adjustments (QIAGEN). Bisulfite-converted libraries were amplified for 10 and 15 cycles for the 500 ng and 10 ng input libraries, respectively, using indexed primers and KAPA HiFi HotStart Uracil+ ReadyMix. Dual-SPRI® size selection was performed postamplification to obtain fragments between 200 bp and 600 bp using Agencourt AMPure XP Beads (Beckman Coulter).

QC samples were taken after 5 cycles of amplification to generate full-length adapters. Library size distributions were assessed using the Agilent Bioanalyzer, and library concentrations at the two different stages during library amplification were determined via qPCR using the KAPA Library Quantification Kit on the BioRad CFX96.

For both inputs, the 500:1 adapter:insert molar ratio samples were sequenced on an Illumina HiSeq® System with single-end 50 bp reads. Sequencing was performed by the Genome Technology Center at New York University. Bismark was used for bisulfite mapping and methylation calling.

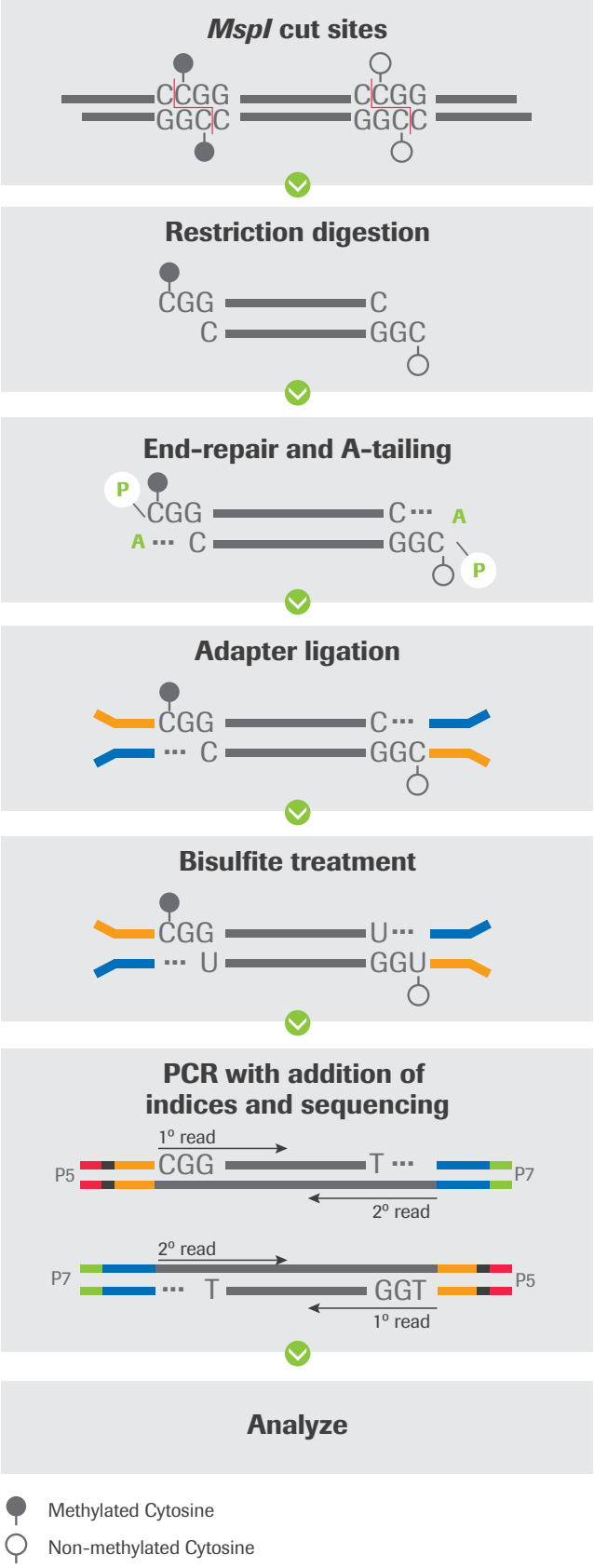

Figure 1. Schematic of reduced representative bisulfite sequencing. 500 ng and 10 ng inputs of gDNA were digested with MspI enzyme (New England Biolabs), followed by library construction using the KAPA HyperPrep Kit. Bisulfite conversion was performed after the post-ligation cleanup using the Qiagen Epitect® Bisulfite Kit (QIAGEN) and libraries amplified using indexed primers and KAPA HiFi HotStart Uracil+ ReadyMix.

Results and discussion

To evaluate the efficiency of library construction, the KAPA Library Quantification Kit was used to assess both adapter-ligated molecules and final library concentrations. Further, final library size distributions were determined by the Agilent Bioanalyzer.

QC samples analyzed after 5 cycles are shown in Figure 2A for 500 ng and 10 ng inputs. Five cycles of PCR were performed to examine post-ligation material once full-length adapter sequences were generated during initial cycling of PCR. As shown, increased adapter concentration in DNA-limiting conditions (e.g., 10 ng) improved the efficiency of the reaction by increasing the amount of adapter-ligated material and indicated an optimal adapter:insert molar ratio of minimally 500:1. In contrast, no improved reaction efficiency was observed above 50:1 for non-limiting input concentrations (e.g., 500 ng). Final library concentrations are shown after an additional 5 and 10 cycles for the 500 ng and 10 ng inputs, respectively (Figure 2B).

Electropherograms of final libraries are shown in Figure 3. Mean peak sizes for 500:1, 100:1 and 50:1 adapter:insert ratios for 500 ng inputs were 379 bp, 384 bp and 366 bp, respectively. Mean peak sizes for 500:1, 100:1 and 50:1 adapter:insert molar ratios for 10 ng inputs were 374 bp, 357 bp and 310 bp, respectively. Consistent with the dual-SPRI size selection performed, the fragment sizes ranged from 200 – 600 bp in length.

To assess sequencing quality, both the 500 ng and 10 ng input samples were compared with respect to general sequencing metrics as shown in Table 1. Results were comparable between both inputs assayed. In total, 12.96 million and 13.16 million reads were obtained per sample, and 62.1% – 65.5% of those reads were uniquely mappable. Further, there is a high correlation between DNA methylation calls between the 10 ng and 500 ng libraries and a 500 ng library sequenced in duplicate. (Figure 5A). As discussed, *MspI* targets 5'-CCGG-3' and cleaves the phosphodiester bonds upstream of the CpG dinucleotide, and therefore 98% of all reads start with an *MspI* cut site (Figure 5B).

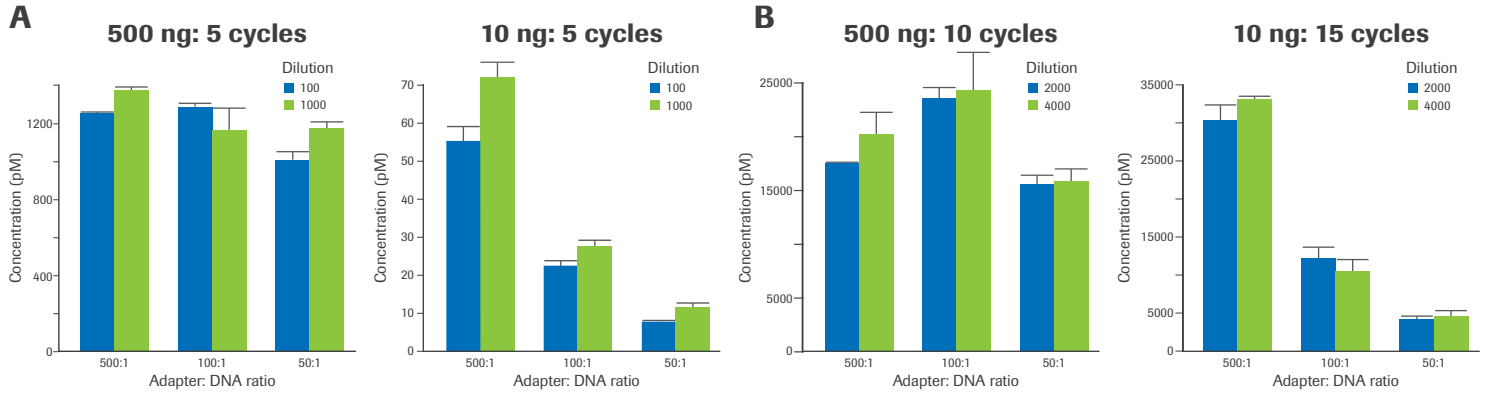

Figure 2. Quantification of adapter-ligated libraries. Library concentrations were determined at two different stages during library amplification during qPCR using the KAPA Library Quantification Kit on the BioRad CFX96. **(A)** Quantification of libraries after 5 cycles of PCR at 1:100 and 1:1000 dilutions for 500 ng (left) and 10 ng (right) inputs. **(B)** Quantification of libraries after 10 cycles for 500 ng (left) and 15 cycles for 10 ng (right) at 1:2000 and 1:4000 dilutions. For both inputs, adapter ligation was performed with 500:1, 100:1, or 50:1 adapter:insert molar ratios.

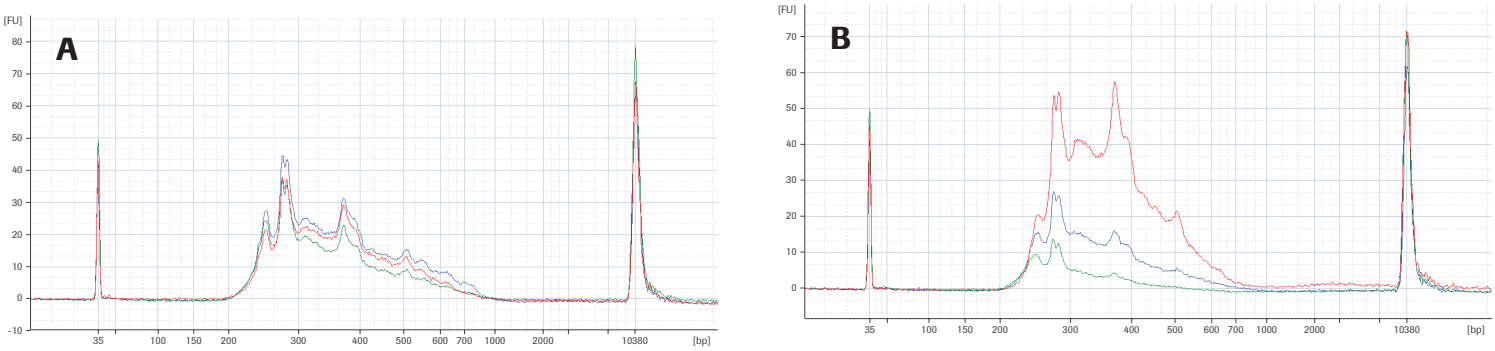

Figure 3. Size distribution of final libraries. Libraries prepared from 500 ng **(A)** and 10 ng **(B)** from B cell genomic DNA with 500:1 (red), 100:1 (blue), and 50:1 (green) adapter:insert molar ratios were analyzed using a 2100 Bioanalyzer Instrument and High Sensitivity DNA Kit (Agilent Technologies).

Table 1. General sequencing metrics

Sample	Total reads	Uniquely mappable		Non-uniquely mappable		Non-mappable	
10 ng input	12,958,987	8,052,046	62.1%	3,278,991	25.3%	1,627,950	12.6%
500 ng input	13,164,796	8,621,114	65.5%	3,271,167	24.8%	1,272,515	9.7%

Data on file.

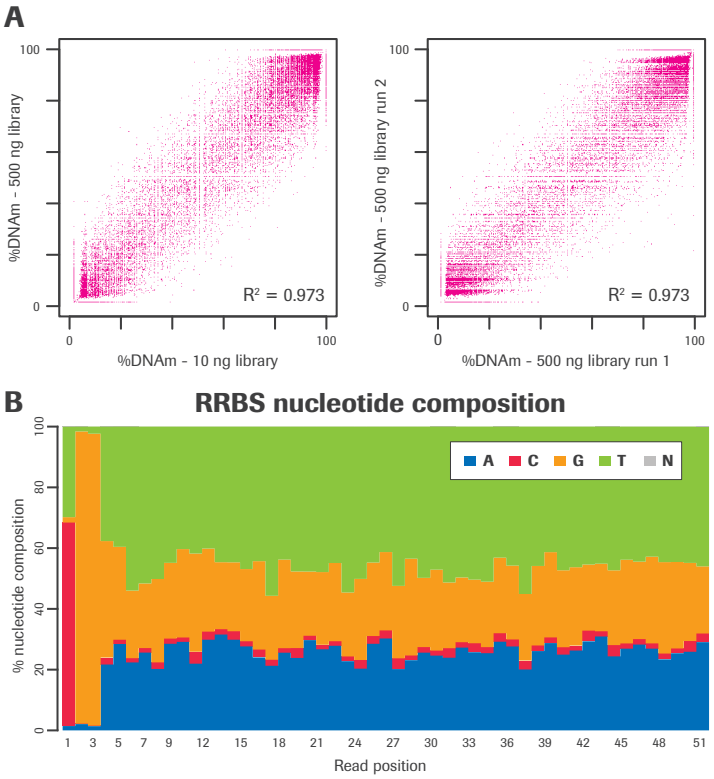

Figure 5. Quality control and reproducibility of libraries. (A) Scatter plot of CpG methylation values for a 10 ng and 500 ng library made from the sample (left) and the 500 ng library sequenced in duplicate (right) indicates strong correlation. **(B)** Plot of nucleotide composition by read position indicates that >95% of total reads start with an *MspI* digestion cut site (CCGG).

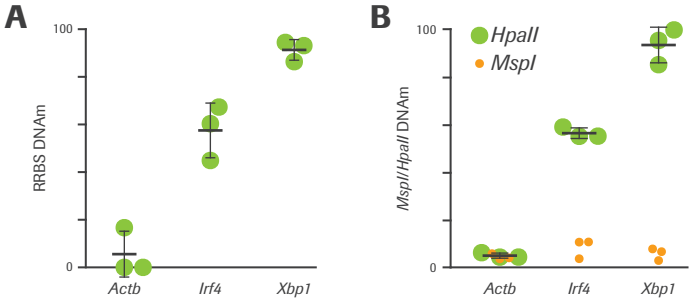

Figure 6. Validation of reduced representation bisulfite sequencing (RRBS.) (A) RRBS data at CpGs near the genes *ActB* (chr5:143668060), *Irf4* (chr13:30839747), and *Xbp1* (chr11:5417784) for three biological replicates of murine B cells. **(B)** DNA methylation for the CpGs shown in Figure 6A measured using qPCR of mock digested DNA compared to DNA digested with the methyl-sensitive restriction enzyme *HpaII*. As a negative control, DNA was digested with the methyl-insensitive isoschizomer *MspI*. Coordinates are from the mouse genome build mm9.

RRBS data at CpG dinucleotides near the genes *ActB*, *Irf4*, and *Xbp1* are shown in Figures 6A and 6B. To validate the sequencing results, qPCR was performed across CCGG cut sites for mock digested DNA, compared to DNA digested with the methyl-sensitive restriction enzyme *HpaII*. DNA digested with the methyl-insensitive isochizomer *MspI* was used as a negative control. As shown, the RRBS sequencing results were replicated when digesting the same regions with *HpaII*.

Conclusions

The KAPA HyperPrep Kit is the ideal solution for RRBS library generation, as it is well suited for low-input samples due to the increased conversion of adapter-ligated molecules. Further, KAPA HiFi HotStart Uracil+ DNA Polymerase is particularly well suited for the amplification of bisulfite-converted libraries, as it has been engineered for tolerance to uracil residues incorporated during bisulfite treatment. KAPA Library Quantification Kits offer accurate qPCR-based library quantification prior to normalization and pooling for multiplexed sequencing. The combination of these products provides a method to interrogate the DNA methylation status of B cells and demonstrates their utility to accurately quantitate DNA methylation from a limited number of cells.

References

1. Meissner, et al., *Nucleic Acids Research*, 33, 18 (2005).
2. Bowman, et al., *BMC Genomics*, 14, 466 (2013).

Published by:

Roche Sequencing and Life Science
9115 Hague Road
Indianapolis, IN 46256

sequencing.roche.com

MC-US-10905 APP108001 A520 3/23