illumina®

TECHNICAL NOTE

High-accuracy methylation and variant detection with Illumina 5-Base DNA Prep

DNA methylation detection comparable to standard methylation assays

Precise germline variant calling, comparable to whole-genome sequencing

Dual insights from a single assay with exceptional performance and coverage

Introduction

Multiomic approaches that combine DNA methylation and genomic information can illuminate complex pathways of development and disease. However, profiling both genomic variation and DNA methylation usually requires separate assays. Illumina 5-Base DNA Prep uses unique chemistry and optimized analysis to detect five bases in the same assay: adenine (A), thymine (T), guanine (G), cytosine (C), and 5-methylcytosine (5mC).[1] Dual insights from Illumina 5-Base DNA Prep enable comprehensive discovery across the genome and epigenome.

In this technical note, we compare Illumina 5-Base DNA Prep, enzymatic methylation sequencing (EM-Seq v2), bisulfite sequencing, methylation arrays, and whole-genome sequencing (WGS). We show that Illumina 5-Base DNA Prep has high concordance of methylation measurements compared to other methylation methods, and germline variant calling comparable to WGS.

Methods

Methylation sequencing libraries were prepared from 50 ng genomic DNA (gDNA) from human reference genome sample NA12878 (Coriell Institute for Medical Research) using three approaches for comparison:

- Illumina 5-Base DNA Prep (Illumina, Catalog no. 20140364)
- NEBNext Enzymatic Methyl-seq v2 Kit (New England Biolabs, Catalog no. E8015S)
- xGen Methyl-Seq DNA Library Prep Kit (IDT, Catalog no. 10009860) with bisulfite conversion using EZ DNA Methylation-Gold (Zymo Research, Catalog no. D5005)

First, gDNA was mechanically sheared to ~450 bp. Each library prep was performed according to manufacturer's instructions and reagents. Libraries were sequenced on the NovaSeq™ 6000 System with 2 × 151 bp reads.

For methylation analysis, data was downsampled to 450M clusters/900M paired-end reads for comparison. Methylation data was analyzed using DRAGEN™ Germline v4.4.6 for Illumina 5-Base DNA Prep or DRAGEN Methylation v4.4.4 for EM-Seq v2 and bisulfite sequencing. Array data was generated with Infinium™ MethylationEPIC v2.0 microarrays (Illumina, Catalog no. 20087706) and the iScan™ System. Data were analyzed using the SeSAMe package v1.26.0 in R v4.5.1.

For germline variant calling, data were downsampled to 500M clusters/1B paired-end reads for comparison to WGS. Illumina DNA PCR-Free Prep data were obtained from BaseSpace™ Sequence Hub demo data (basespace.illumina.com/s/2teNclBsUsmH) and subsampled to 400M clusters/800M paired-end reads. For Illumina 5-Base DNA Prep and Illumina DNA PCR-Free Prep, variant calling was performed using DRAGEN Germline v4.4.6. For EM-Seq v2 and bisulfite sequencing, variant calling was performed using Bis-SNP v0.7.1.

Results

Methylation detection comparable to standard methods

Illumina 5-Base DNA Prep demonstrates high methylation detection specificity and sensitivity, comparable with two standard methylation assays, EM-Seq v2 and bisulfite sequencing (Figure 1, Figure 2). Weighted methylation levels for small genome control spike-ins show true positive and true negative rates are consistent across the three methods (Figure 1). Illumina 5-Base DNA Prep shows methylation measurement at CpG islands that is highly concordant with EM-Seq v2, bisulfite sequencing, as well as Infinium MethylationEPIC v2.0 methylation microarrays (Figure 2).

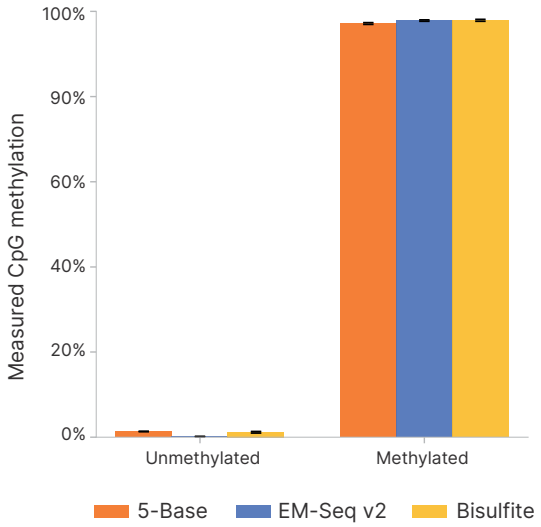

Figure 1: Comparable true positive and true negative rates of methylation detection
Illumina 5-Base DNA Prep (orange) offers true positive and true negative methylation rates comparable to EM-Seq v2 (blue) and bisulfite sequencing (yellow). Small genome spike-ins were unmethylated lambda bacteriophage DNA as a negative control and methylated pUC19 DNA as a positive control.

Maximized mapping and coverage

Illumina 5-Base DNA Prep shows exceptional sequencing data quality with higher percent mapped reads and median insert length than EM-Seq v2 or bisulfite sequencing (Figure 3). Because Illumina 5-base conversion is nondamaging to DNA, it enables rentention of longer fragments. The Illumina 5-base workflow is also optimized to select for longer insert sizes (~350 bp) to maximize variant calling performance.

Illumina 5-Base DNA Prep provides more comprehensive coverage of the methylome for a given read length, compared to EM-Seq v2 and bisulfite sequencing (Figure 4). Illumina 5-Base DNA Prep also offers uniform GC coverage across the human genome with higher normalized coverage than EM-Seq v2 at low-GC regions and higher normalized coverage than bisulfite sequencing at high-GC regions (Figure 5).

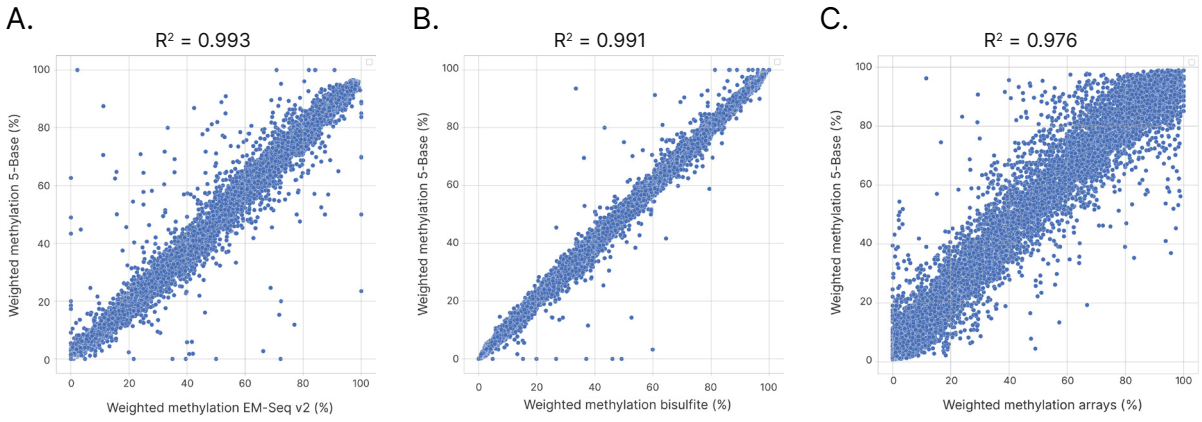

Figure 2: Highly accurate methylation measurement at CpG islands
Weighted methylation levels of individual CpG islands using Illumina 5-Base DNA Prep are highly concordant with (A) EM-Seq v2, (B) bisulfite sequencing, and (C) Infinium MethylationEPIC v2.0 arrays.

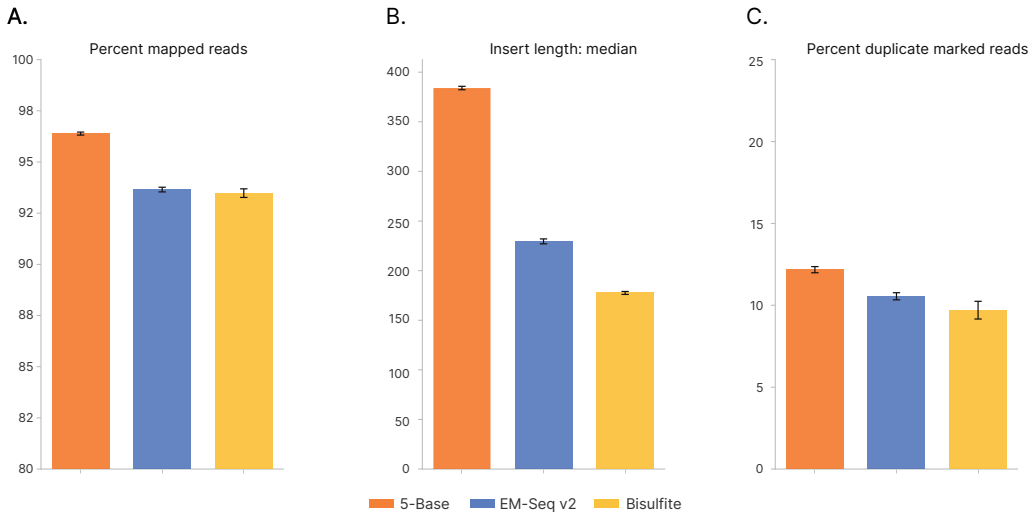

Figure 3: Exceptional sequencing data quality with Illumina 5-Base DNA Prep
Illumina 5-Base DNA Prep (orange) offers (A) higher percentage mapped reads, (B) superior median insert lengths, and (C) comparable PCR duplicate rate to EM-Seq v2 (blue) and bisulfite sequencing (yellow).

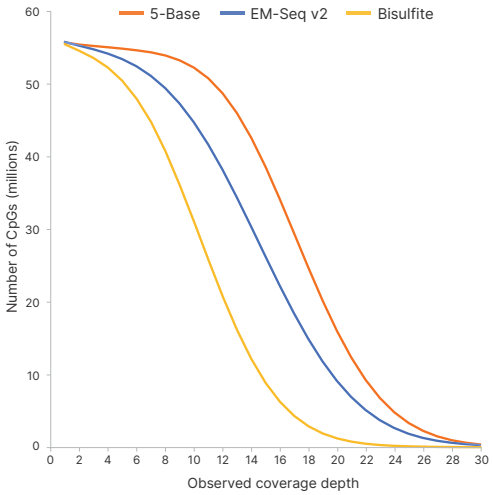

Figure 4: Higher coverage of CpG sites with Illumina 5-Base DNA Prep
The number of CpGs at each observed coverage depth was calculated by the total number of cytosines in a CpG context at or above the threshold. Illumina 5-Base DNA Prep (orange) shows ~20% more CpGs with at least 10× coverage compared to EM-Seq v2 (blue) and 68% more CpGs with at least 10× coverage compared to bisulfite sequencing (yellow) for 450M clusters/900M paired-end reads. Replicates for each method were averaged.

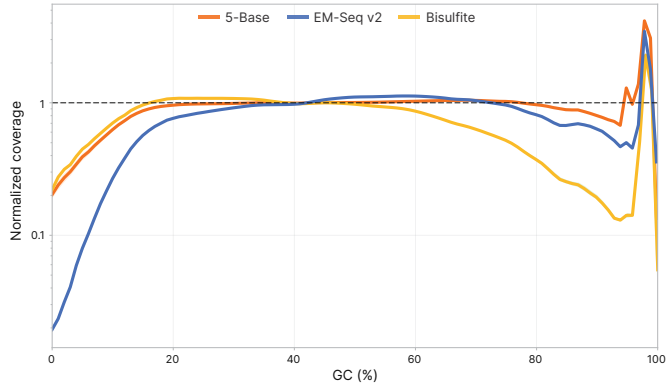

Figure 5: Uniform coverage across a range of GC content in the human genome with Illumina 5-Base DNA Prep
Illumina 5-Base DNA Prep (orange) shows reduced GC bias with normalized coverage that is consistently higher at low-GC percentages (1%–20%) compared with EM-Seq v2 (blue) and at high-GC percentages (70%–100%) compared with bisulfite sequencing (yellow).

Included germline variant calling

In addition to methylation detection, Illumina 5-Base DNA Prep offers high-accuracy germline variant calling in concordance with gold-standard WGS assays such as Illumina DNA PCR-Free Prep (Figure 6, Table 1). Standard methylation detection assays, EM-Seq v2 and bisulfite sequencing, have reduced library complexity and data yield and low compatibility with 4-base variant calling algorithms. These limitations affect the accuracy and usability of DNA variant results from EM-Seq v2 or bisulfite sequencing data.

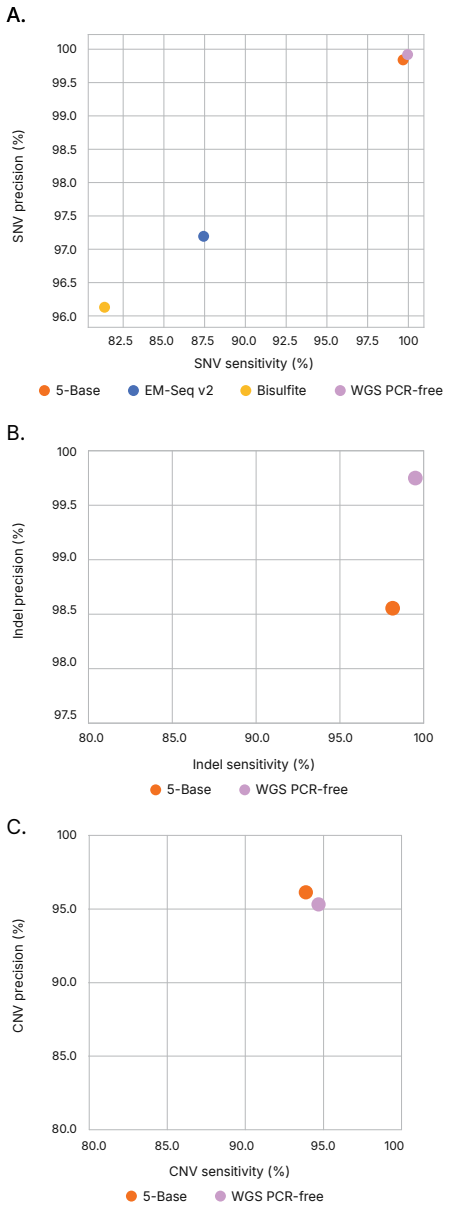

Figure 6: High-accuracy germline variant calling with Illumina 5-Base DNA Prep

Sensitivity and precision plot for (A) germline single nucleotide variant (SNV) calling shows Illumina 5-Base DNA Prep (orange) is very close to standard WGS with Illumina DNA PCR-Free Prep (purple), and higher than EM-Seq v2 (blue) and bisulfite sequencing (yellow). SNV calling performance was assessed against the NIST 4.2.1 truth set for NA12878 in confident regions. Results from technical replicates were averaged. Illumina 5-Base DNA Prep also shows (B) insertions-deletions (indels) variant calling with high accuracy and (C) copy number variants (CNV) calling comparable to Illumina DNA PCR-Free Prep. Differences in indel calling are attributed to use of PCR vs PCR-free.

Table 1: Illumina 5-Base DNA Prep offers germline variant calling performance comparable to gold-standard WGS

Assay	False positive	True positive	False negative	Precision	Sensitivity	F1 Score
SNV						
Illumina 5-Base DNA Prep	5777	3,246,292	10,731	0.9982	0.9967	0.9975
Illumina DNA PCR-Free Prep	2348	3,252,713	1345	0.9993	0.9996	0.9996
EM-Seq v2	90,820	3,144,462	451,770	0.9720	0.8744	0.9206
Bisulfite sequencing	117,749	2,925,648	670,584	0.9613	0.8136	0.8813
Indel[a]						
Illumina 5-Base DNA Prep	6712	457,155	8677	0.9855	0.9814	0.9835
Illumina DNA PCR-Free Prep	1172	465,589	2349	0.9975	0.9950	0.9963
CNV[a]						
Illumina 5-Base DNA Prep	82,271	2,345,550	134,484	0.9661	0.9458	0.9558
Illumina DNA PCR-Free Prep	101,835	2,364,830	116,446	0.9587	0.9531	0.9559

a. Bis-SNP software only provides SNVs, thus indel and CNV benchmarks are not available for EM-Seq v2 and bisulfite sequencing datasets.

Summary

Illumina 5-Base DNA Prep has the methylation detection accuracy of EM-Seq and bisulfite methylation assays, with more CpGs covered for equivalent read depth. The Illumina 5-base solution also provides high-accuracy variant calling, enabling researchers to obtain multiomic insights from the same DNA sample in a single assay, enhancing potential discovery power.

Learn more

Illumina 5-Base DNA Prep

References

1. Illumina. Illumina 5-Base DNA Prep data sheet. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-5-base-dna-prep-data-sheet-m-gl-03689/illumina-5-base-dna-prep-data-sheet-m-gl-03689.pdf. Published 2025. Accessed November 11, 2025.

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-03816 v1.0