illumina®

TruSeq™ Methyl Capture EPIC Library Prep Kit

Combining the power of next-generation sequencing with epigenetic insights to accelerate biomarker discovery.

Highlights

- **Comprehensive Methylation Site Content**
  Access content carefully designed with a proprietary algorithm, driven by customer feedback and Infinium™ MethylationEPIC BeadChip content
- **Maximize Discovery Power, Minimize Cost**
  Query more differentially methylated bases (DMBs) per research dollar compared to TruSeq whole-genome bisulfite sequencing (WGBS)[1]
- **Convenient, Hassle-Free Kit Configuration**
  Order bisulfite conversion, library prep, target enrichment, and purification reagents, all in a single kit
- **User-Friendly Data Analysis**
  Analyze sequence data in the BaseSpace™ Sequence Hub with methylation apps designed for biologists

Introduction

Diseases such as cancer and autism are complex, multigenic disorders. Adding another layer to the challenge of understanding complex diseases, is the fact that environmental factors interact with the genome to mediate disease progression.[1] While researchers have traditionally focused on genetic mutations as a primary cause of disease, there is an increasing body of evidence that aberrant DNA methylation and its impact on genome structure and gene expression play a critical role in the pathology of diseases such as cancer, obesity, and diabetes.[2–4] Methylation studies have also enabled investigators to shift from correlative to causative research results.[5,6] By delving into the molecular interplay between methylation status and transcriptional activity, studies have shown that the promoters of several tumor suppressor genes in cancer are hypermethylated: a methylation state associated with gene silencing.[5,6] Furthermore DNA hypomethylation throughout the majority of the cancer genome seems to contribute to genome instability and dramatic changes in genome/chromatin structure.[5,6]

While methylation arrays continue to be used successfully, researchers are now leveraging the power of next-generation sequencing (NGS) to perform whole-genome bisulfite sequencing (WGBS) and targeted bisulfite sequencing (Methyl-Seq) for a deeper understanding of the complexity and dynamic state of the methylome. To support these efforts, Illumina offers the TruSeq Methyl Capture EPIC Library Prep Kit, a library preparation kit that includes nearly all the reagents needed to prepare targeted Methyl-Seq libraries for NGS.

With a content design driven by customer-based requests and emerging regions of interest identified by ENCODE,[9] FANTOM5,[10] and the Epigenomics RoadMap Consortium,[11] the TruSeq Methyl Capture EPIC Library Prep Kit spans the full human methylome (Table 1). For unbiased target enrichment during library preparation, the target capture step is performed before bisulfite conversion (Figure 1).

Table 1: Epigenetic Regions Covered in Targeted Methyl-Seq Kits

Epigenetic Regions[a]	Genome	TruSeq Methyl Capture EPIC Kit	Competitor X Methyl-Seq Kit[12]	Competitor Y Methyl-Seq Kit[13]
Total Size	3000 Mb	107 Mb	84 Mb	81 Mb
CpG Sites	28,217,009	3,340,894	3,151,840	2,804,771
CpG Islands	27,718	26,981	26,252	26,002
WGBS DMRs	1,853,606	345,327	240,042	227,252
GenCode Promoters	149,463	141,527	129,958	105,924
Open Chromatin (synthesis)	303,304	199,938	79,951	76,129
TFBS	229,067	180,047	62,123	61,642
MEDs	264,834	253,022	77,967	76,537
FANTOM5 Enhancers	28,563	28,172	7326	6998
CTCF ChIP Peaks	433,466	93,242	56,084	53,716

a. Cytosine phosphate guanine (CpG), whole-genome bisulfite sequencing (WGBS), differentially methylated regions (DMRs), transcription factor binding sites (TFBS), multifunctional epigenetic domains (MEDs), the *CTCF* gene encodes a transcription factor

Table 2: Samples Sequenced Per Run With the TruSeq Methyl Capture EPIC Library Prep Kit

Sequencing System	Samples/Run[a]
NextSeq™ Series	
High-Output Flow Cell	8
HiSeq™ 2500 System	
Rapid Run Mode, Dual Flow Cell	8-12
High-Output Mode, Dual Flow Cell	64
HiSeq 3000 System	
Single Flow Cell	48
HiSeq 4000 System	
Dual Flow Cell	96

a. Samples at > 40× mean coverage and > 90% of target bases are covered at ≥ 10×.

[1]Calculations on file.

Although WGBS is the most comprehensive approach for methylation studies, it is also the most cost-prohibitive, particularly for studies with large sample numbers.[7] In contrast, methylation arrays are a highly cost-effective method for large-scale screening studies, but have a higher limitation on the number of sites interrogated and provide little detail about methylation heterogeneity in a given region compared to WGBS and targeted Methyl-Seq.[7] Targeted Methyl-Seq offers a balanced, cost-effective choice between WGBS and methylation arrays that can support both screening and biomarker discovery study objectives.[8] By analyzing preselected epigenetic areas of interest, the TruSeq Methyl Capture EPIC kit offers several distinct advantages over WGBS:

- Cost-effective discovery of DMBs
- Minimizes data analysis burden due to a smaller, more manageable data set
- Convenient kit configuration for hassle-free ordering

Comprehensive Methylation Site Content

The TruSeq Methyl Capture EPIC Library Prep Kit builds upon content included in the Infinium™ MethylationEPIC BeadChip with additional regions of importance identified by ENCODE, FANTOM5, the Epigenomics RoadMap Consortium, and customer-based requests. This carefully curated content contains traditional as well as emerging epigenetic regions of interest—all at a fraction of the price of WGBS.

To compare how the TruSeq Methyl Capture EPIC Library Prep Kit compares to other Methyl-Seq kits currently on the market, the coverage of various epigenetic regions of interest are shown (Figure 2). While coverage of traditional epigenetic regions, such as CpG islands and CpG shores appear comparable between the 3 kits, the TruSeq Methyl Capture EPIC Library Prep Kit offers significantly higher coverage of emerging epigenetic regions of interest. Beyond CpG sites, the TruSeq Methyl Capture EPIC Library Prep Kit also contains significant coverage of CHG (H stands for A, T, or C nucleotides) and CHH sites (Table 3).

Table 3: Non-CpG Methylation Sites Covered by the TruSeq Methyl Capture EPIC Library Prep Kit

	CHG	CHH
HeLa Cells	8,636,751	23,363,530
Jurkat Cells	8,752,284	23,534,284
NA12878 Cells	8,668,229	23,859,152
Normal HCC1187	9,126,118	25,084,617
Tumor HCC1187	8,917,380	24,527,767

Five cell lines were sequenced with the TruSeq Methyl Capture EPIC Library Prep Kit at a sequencing depth of 11 Gb (55 million, 2× 100, paired-end reads). Each site covered at ≥ 10×. Data analyzed with the BaseSpace MethylSeq[14] and BaseSpace MethylKit[15] Apps.

For investigators who have used or who are planning to use the Infinium MethylationEPIC BeadChip, using an enrichment panel with content that overlaps with the MethylationEPIC BeadChip means that comparative bioinformatics analysis will be greatly simplified and research continuity can be maintained. The TruSeq Methyl Capture EPIC target panel was designed to have > 90% overlap with the Infinium MethylationEPIC BeadChip target regions. To demonstrate backward compatibility between the TruSeq Methyl Capture EPIC Library Prep Kit and the Infinium MethylationEPIC BeadChip, differentially methylated regions (DMRs) between tumor-normal pair HCC1187 were characterized (Figure 3). Results demonstrated that 97% of DMRs identified with the MethylationEPIC BeadChip were also covered by the TruSeq Methyl Capture EPIC Library Prep Kit. In addition, the TruSeq Methyl Capture EPIC Library Prep Kit identified over 7000 unique DMRs from ENCODE, FANTOM5, the Epigenomics RoadMap Consortium, and customer-based requests.

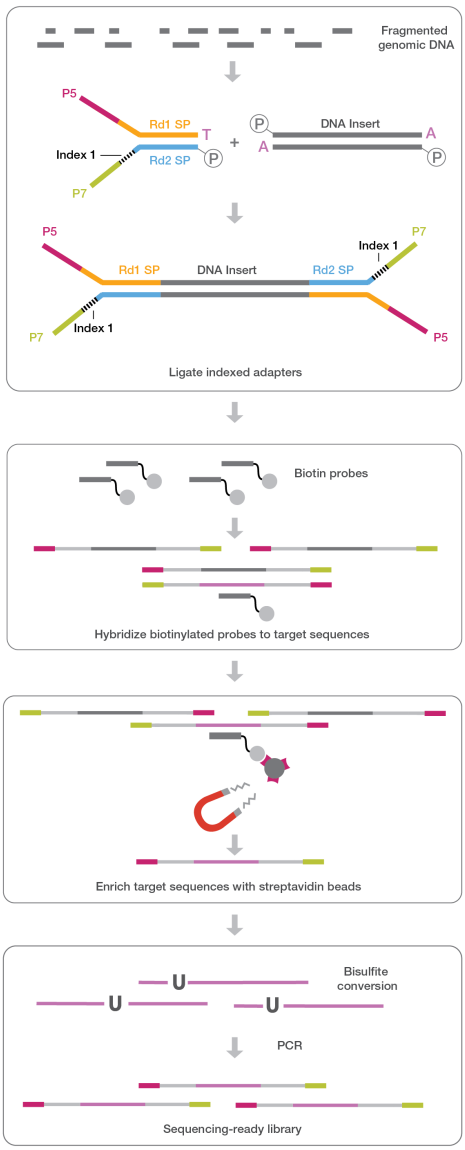

Figure 1: TruSeq Methyl Capture EPIC Library Preparation Workflow — The TruSeq Methyl Capture EPIC Library Prep Kit contains the bisulfite conversion, library prep, target enrichment, and purification reagents all in a single kit for convenient, hassle-free targeted Methyl-Seq library preparation.

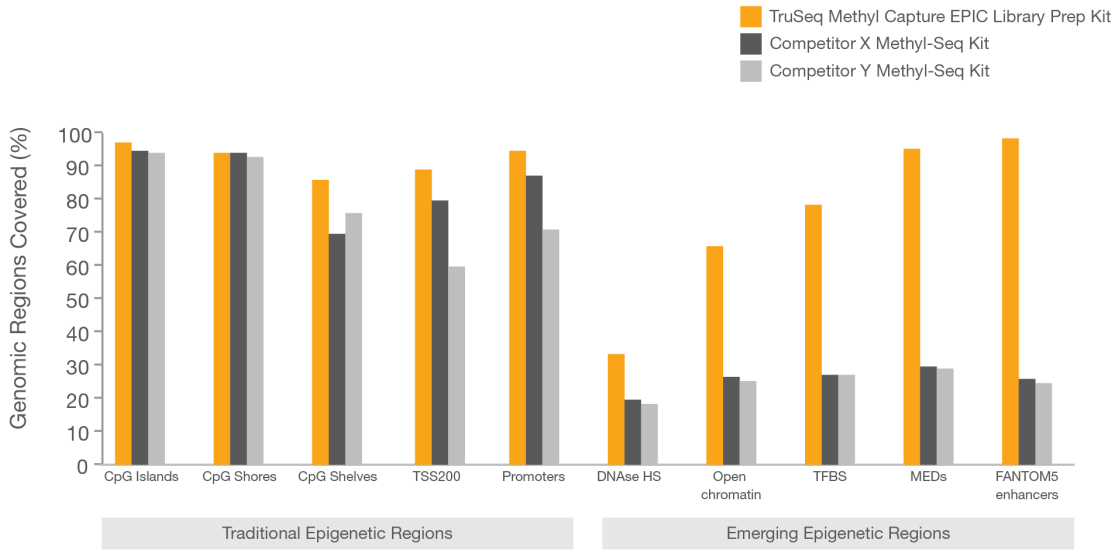

Figure 2: Epigenetic Regions Covered— The TruSeq Methyl Capture EPIC Library Prep Kit uniquely targets emerging epigenetic regions of interest such as open chromatin, ENCODE, and FANTOM regions. The percentage of genomic regions covered using the TruSeq Methyl Capture EPIC Library Prep Kit, the Competitor X Methyl-Seq Kit,[12] and the Competitor Y Methyl-Seq Kit[13] are shown. The first 5 regions (CpG islands, CpG shores, CpG shelves, TSS200, and promotor regions) are traditional epigenetic regions of interest, while the last 5 regions (DNase Hypersensitive Regions, Open chromatin, transcription factor binding sites (TFBS), multifunctional epigenetic domains (MEDs), and FANTOM5 enhancers) represent emerging epigenetic regions of interest identified by ENCODE, FANTOM5, the Epigenomics RoadMap Consortium, and customer-based requests.

Maximize Discovery Power, Minimize Cost

The growing Illumina methylation portfolio has solutions for a variety of research needs and budgets. WGBS with the TruSeq DNA Methylation Library Prep Kit provides a comprehensive view of methylation patterns across the whole genome at single-base resolution. While WGBS delivers the most comprehensive approach for discovery-related applications, methylation arrays offer a cost-effective solution for studies with high sample numbers and are the preferred choice among researchers for epigenome-wide array studies (EWAS).[7] The TruSeq Methyl Capture EPIC Library Prep Kit offers a balanced choice between these options that can support both screening and discovery-related study designs.[8]

To compare cost effectiveness of the TruSeq DNA Methylation Library Prep, TruSeq Methyl Capture EPIC Library Prep, and Infinium MethylationEPIC BeadChip Kits, tumor-normal HCC1187 samples were analyzed by each of the 3 methods. The total number of DMBs were identified and the cost per sample calculated for each method. The data show that for less than 15% of the per sample cost of WGBS, the EPIC Library Prep Kit provides over 70% of the same DMBs (Figure 4A). Another way to compare cost-effectiveness is to calculate the number of DMBs identified per experimental dollar. For the tumor-normal HCC1187 sample set, the TruSeq Methyl Capture Library Prep Kit identified nearly 5 times more DMBs per research dollar compared to TruSeq WGBS and the Infinium MethylationEPIC BeadChip (Figure 4B).

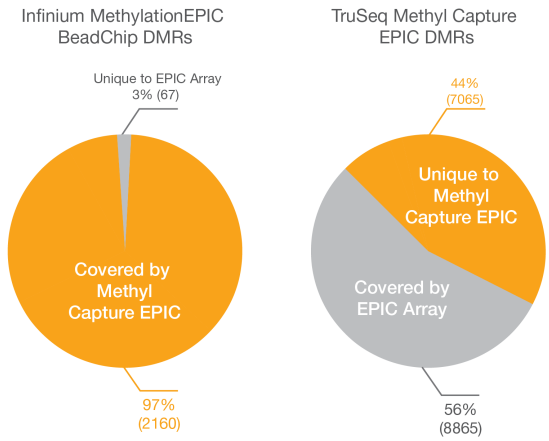

Figure 3: Differentially Methylated Regions in the Infinium MethylationEPIC BeadChip vs the TruSeq Methyl Capture EPIC Library Prep Kit— DMRs from tumor-normal HCC1187 sample set were compared. 97% of DMRs identified by MethylationEPIC BeadChip (left) were also covered by the TruSeq Methyl Capture EPIC Library Prep Kit. Of the DMRs identified with TruSeq Methyl Capture EPIC Library Prep kit (right), over 7000 DMRs were unique to the EPIC Library Prep Kit, while over 8800 DMRs overlapped with DMRs present in the EPIC BeadChip. Data analyzed with BaseSpace MethylSeq and BaseSpace MethylKit Apps. Results may vary based on sample type or experimental conditions.

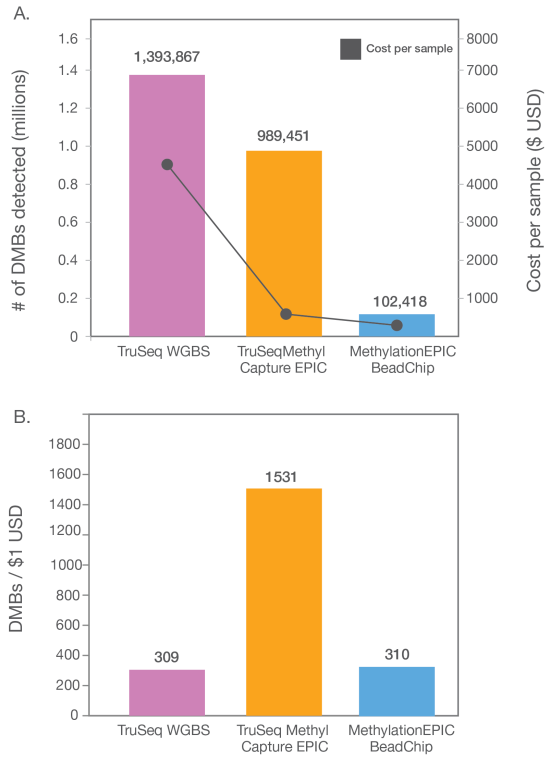

Figure 4: TruSeq Methyl Capture EPIC Library Prep Kit Cost-Effectiveness — **A)** DMBs Detected. The TruSeq Methyl Capture EPIC Kit provided over 70% of the same DMBs as WGBS, for a significantly lower cost per sample. WGBS, targeted Methyl-Seq, and methylation array analysis were performed on the HCC1187 tumor-normal pair according to the manufacturers' instructions for TruSeq DNA Methylation Library Prep, TruSeq Methyl Capture EPIC Library Prep, and Infinium MethylationEPIC BeadChip kits. Differences in methylation at individual CpG sites were counted with the BaseSpace MethylKit App. **B)** DMBs Identified Per Dollar. The TruSeq Methyl Capture EPIC kit delivered significantly more DMBs per research dollar compared to WGBS and the EPIC BeadChip. Data analyzed with BaseSpace MethylSeq and BaseSpace MethylKit Apps. Results may vary based on sample type or experimental conditions.

Targeted Methyl-Seq Provides More Comprehensive Data

The TruSeq Methyl Capture EPIC Library Prep Kit delivers a more comprehensive level of coverage across a region of interest compared to methylation arrays.[7] At individual CpG sites, the information provided by Methyl-Seq and methylation arrays is largely equivalent (Figure 5A). However, arrays cannot detect the methylation status of other CpG sites within a region of interest, without specific array probes. By offering base-by-base coverage across the entire region of interest, Methyl-Seq provides the higher resolution needed to distinguish between epigenomic profiles that may appear similar based on array data, but that show significant differences in methylation status at the nucleotide level (Figure 5).

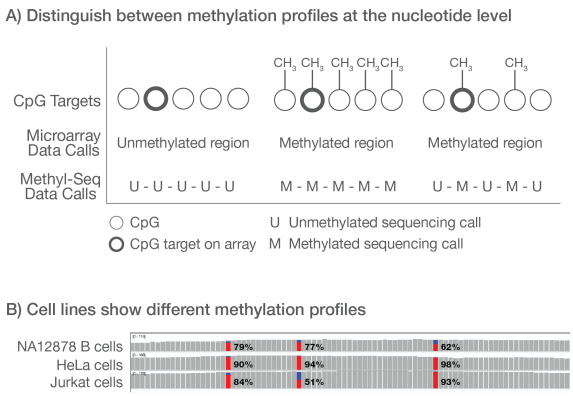

Figure 5: High-Resolution Data with Methyl-Seq — **A)** Methyl-Seq libraries deliver nucleotide-level resolution across the entire region of interest. **B)** TruSeq Methyl Capture EPIC libraries were prepared with 3 different cell lines to show heterogeneity of methylation across a target region. Each grey bar represents a single nucleotide, red indicates % reads methylated, and blue indicates % of reads unmethylated. The cell lines show distinct methylation profiles across the region and different percentages of methylated reads at the same nucleotide position. Data analyzed with MethylKit BaseSpace App and visualized with BaseSpace Integrative Genomics Viewer.[16]

Exceptional Coverage and Reproducibility

To demonstrate the exceptional performance of the TruSeq Methyl Capture EPIC Library Prep Kit, 5 different cell types were sequenced and subsampled to a range of coverage levels (Figure 6). Based on the results of these and other internal studies, Illumina recommends a sequencing depth of 11 Gb (55 million paired-end, 2 × 100 bp reads) for the generation of 30–50× coverage to perform methylation calling. At this sequencing depth, typically > 90% of bases are sequenced at ≥ 10× coverage (the recommended coverage level for methylation calling). The sequencing results also show excellent reproducibility among samples run in triplicate and high correlation with samples assayed via WGBS and the MethylationEPIC BeadChip (Table 4).

Table 4: Correlation of TruSeq Methyl Capture EPIC Methylation Calls

Sample	Pearson's Correlation
Methyl Capture EPIC HCC1187 tumor DNA library in triplicate	0.98
Methyl Capture EPIC HCC1187 normal DNA library in triplicate	0.99
Methyl Capture EPIC Coriell NA12878 library in triplicate	0.97
Methyl Capture EPIC NA12878 library vs WGBS NA12878 library	0.97
Methyl Capture EPIC NA12878 library vs MethylationEPIC Array NA12878	0.96

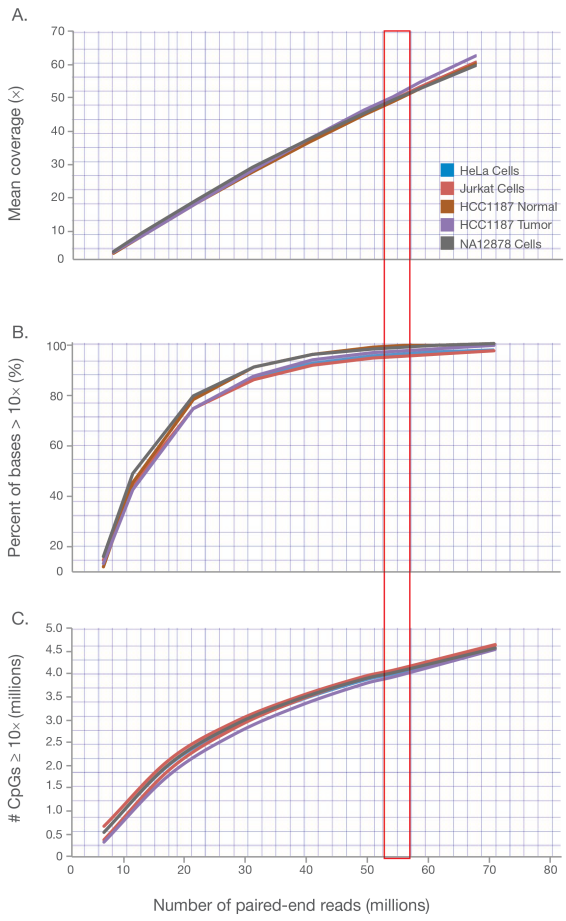

Figure 6: Coverage and Reproducibility for TruSeq Methyl Capture EPIC Libraries — Five cell lines were sequenced with the TruSeq Methyl Capture EPIC Library Prep Kit and sampled to various read depths. The red window shows recommended number of reads to obtain > 90% of bases at ≥ 10× coverage. Data analyzed with BaseSpace MethylSeq App. **A)** At a read depth of 55 million reads, all samples displayed 40–50× mean coverage levels. **B)** At 55 million reads, > 90% of bases were covered at ≥ 10×. **C)** At 55 million reads, ≥ 3.7 million CpGs were covered at ≥ 10×.

Convenient, Hassle-Free Kit

The TruSeq Methyl Capture EPIC Library Prep Kit contains nearly all the reagents needed to produce high-quality Methyl-Seq libraries, including bisulfite conversion reagents. It is one of the most inclusive kits on the market, requiring only 2 separate catalog numbers (Table 5). This convenient, consolidated kit provides the indexed adapters, enrichment probes, and purification beads as well as all the reagents needed for the initial bisulfite conversion. With only 500 ng input DNA (prebisulfite conversion), the kit supports fast and easy preparation of up to 48 high-quality targeted Methyl-Seq libraries. The total assay time of the TruSeq Methyl Capture EPIC Library Prep Kit is less than 2 days, which is faster than the 3- and 5-day workflows of competitor kits (Table 5). With bead-based enrichment and clean-up steps, the protocol is easily adapted to high-throughput procedures. Additionally, the kit is part of a fully integrated sequencing solution (Figure 7).

Table 5: Comparison of Contents in Various Targeted Methyl-Seq Kits

Reagents Included	TruSeq Methyl Capture EPIC	Competitor X Methyl-Seq Kit[12]	Competitor Y Methyl-Seq Kit[13]
Number of Vendors Required	2	4	3
Number of Part Numbers Required	2	5	5
Bisulfite Conversion Reagents	✓	×	×
End Repair and 3' A-Tailing Reagents	✓	×	×
Adapter Indexes	✓	×	×
Purification Reagents	✓	×	×
Oligo Pool	✓	✓	✓
Polymerase for Enrichment Amplification	×	×	×
Library Prep Total Assay Time	2 days	3 days	5 days

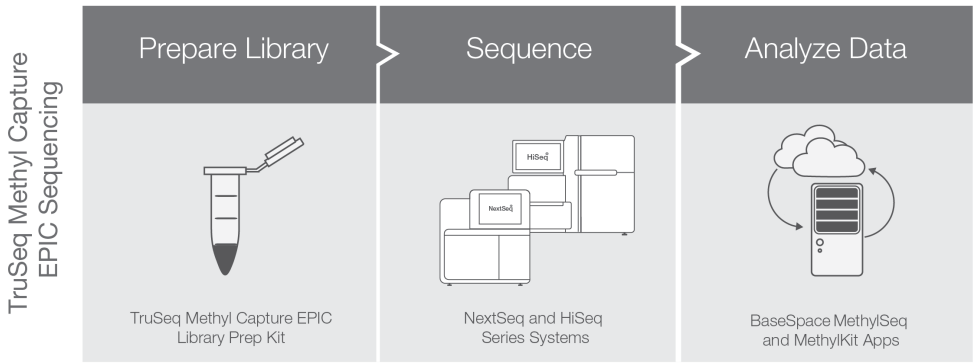

Figure 7: TruSeq Methyl Capture EPIC Sequencing Workflow — The TruSeq Methyl Capture EPIC Library Prep Kit is part of an integrated workflow that includes library preparation, sequencing, and data analysis.

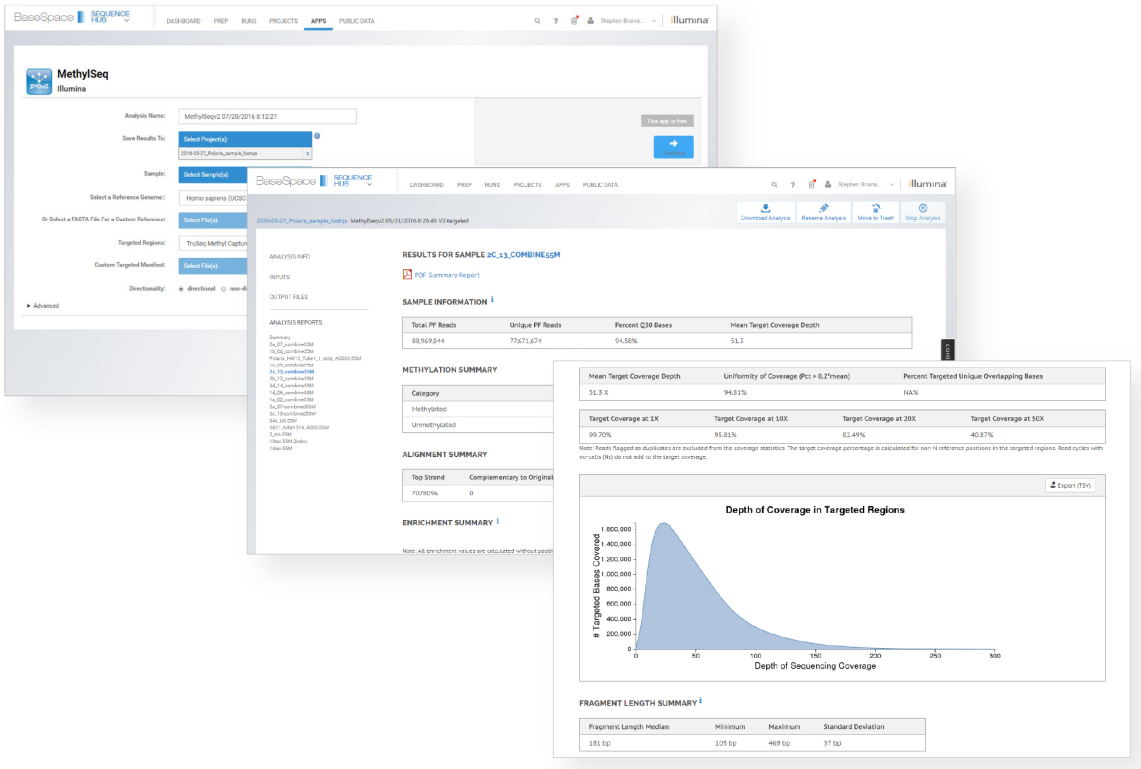

Figure 8: BaseSpace MethylSeq App— After sequencing is complete, the MethylSeq BaseSpace App aligns reads to the genome and calls methylation at all cytosines. The TruSeq Methyl Capture EPIC panel is available for selection by drop-down menu to enable calculation of targeted coverage metrics. The MethylSeq App also reports a coverage graph showing depth of sequencing coverage vs. number of targeted bases covered.

User-Friendly Data Analysis Solution

Methylation-Specific Apps

As part of the comprehensive Methyl-Seq solution, sequence data can be transferred seamlessly into the BaseSpace Sequence Hub and analyzed with BaseSpace Apps. BaseSpace Sequence Hub provides click-and-go access to a broad range of user-friendly apps including the MethylSeq (Figure 8) and MethylKit Apps (Figure 9). The MethylSeq App uses Bismark[17] and Bowtie2[18] to map bisulfite-treated sequencing reads to the genome of interest and performs methylation calls, while the MethylKit App analyzes sequencing data for differences in methylation between samples. BaseSpace Methyl-Seq Apps support targeted data and common epigenomics analysis tasks such as methylation calling, analysis of differential methylation between samples, and categorization of significant methylation regions. BaseSpace Methyl-Seq Apps are designed for push-button ease of use for any researcher, regardless of bioinformatics experience.

Biological Interpretation Apps

Illumina also offers the BaseSpace Correlation Engine—a tool which contains data sets from over 18,000 public studies, enriching sequencing results with functional information. After read alignment and methylation calling are completed, BaseSpace Correlation Engine can be used to identify disease mechanisms, drug targets, and prognostic or predictive biomarkers associated with sequencing results and cited in peer-reviewed literature. The Ingrative Genomics Viewer (IGV) BaseSpace App enables visualization of regions of interest across multiple samples simultaneously. BaseSpace Sequence Hub and its growing library of apps enable biologists to keep pace with research and manage data easily.

Summary

By querying millions of methylation sites across multiple samples, the TruSeq Methyl Capture EPIC Library Prep Kit delivers high-impact results for a fraction of the cost of TruSeq WGBS. With content that encompasses the Infinium MethylationEPIC BeadChip, researchers can move seamlessly between array and sequencing technologies, taking advantage of the low price point of the MethylationEPIC BeadChip for large-scale studies and diving deep on specific samples or regions of interest with the TruSeq Methyl Capture EPIC Library Prep Kit. By combining the power of NGS with targeted bisulfite sequencing, the TruSeq Methyl Capture EPIC Library Prep Kit enables scientists who study cancer and other complex diseases to accelerate their research and gain new insights into the workings of the epigenome.

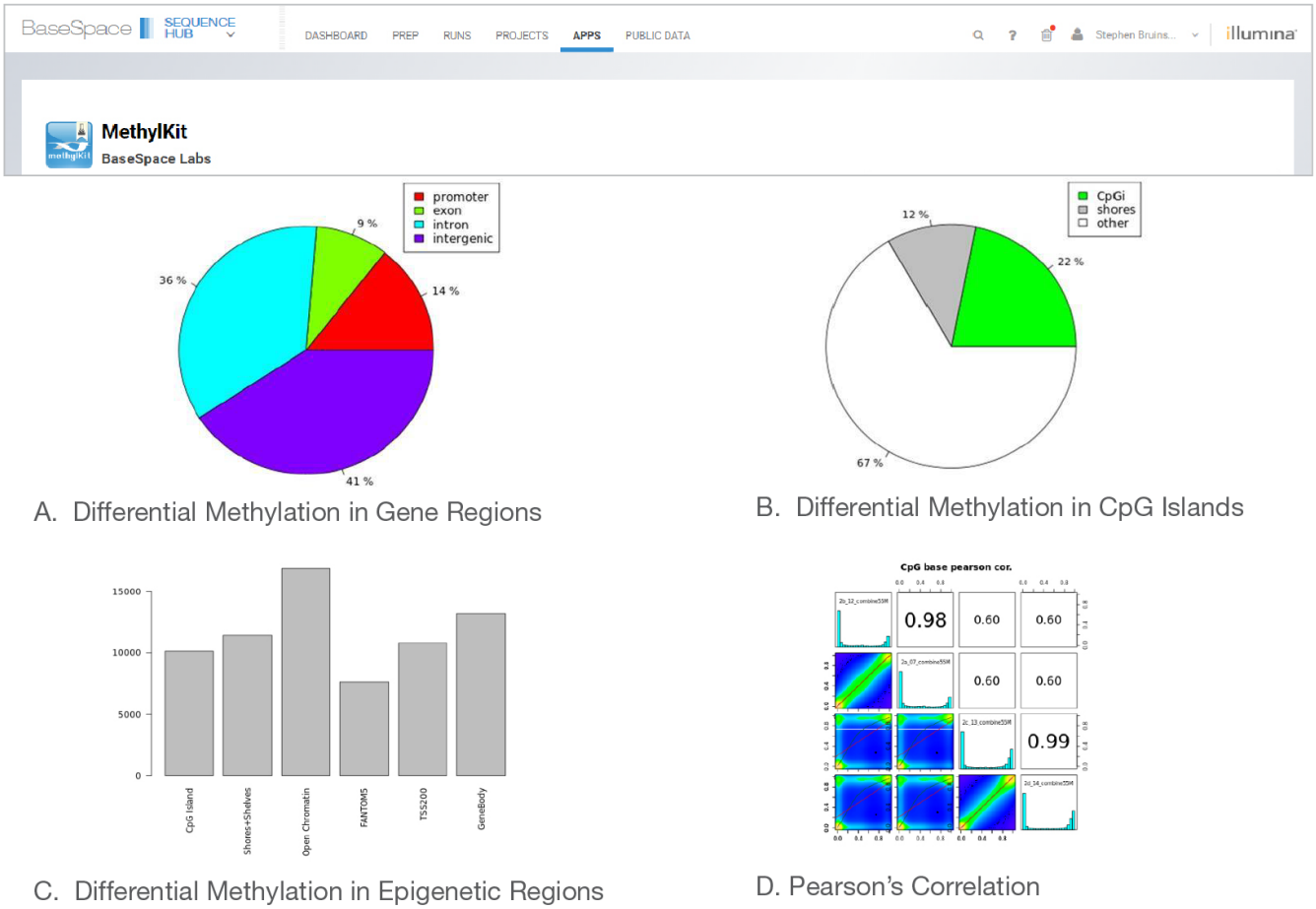

Figure 9: BaseSpace MethylKit App— After alignment using MethylSeq, the MethylKit BaseSpace App enables differential methylation calling between 2 samples. HCC1187 tumor-normal samples were sequenced by TruSeq Methyl Capture EPIC and methylation was compared using the MethylKit BaseSpace App. DMBs are sorted by MethylKit according to category relative to **A)** gene regions and **B)** CpG Islands. **C)** DMRs are also counted and binned into various epigenetic categories and **D)** Pearson's correlation calculated.

Ordering Information

Product	Catalog No.
TruSeq Methyl Capture EPIC Library Prep Kit, 48 samples	FC-151-1003
TruSeq Methyl Capture EPIC Library Prep Kit, 12 samples	FC-151-1002
KAPA HiFi HotStart Uracil+ with ReadyMix (2×)	KK2801

Order library prep kits at www.Illumina.com, and PCR reagents at www.kapabiosystems.com.

Learn More

To learn more about the TruSeq Methyl Capture EPIC Library Prep Kit, visit
www.illumina.com/methylcapture

For more on the TruSeq DNA Methylation Kit and WGBS, visit
www.illumina.com/products/truseq-dna-methylation.html

To learn more about the Infinium MethylationEPIC BeadChip, visit
www.illumina.com/products/infinium-methylation-EPIC-array.html

References

1. Vijayakumar NT, Judy MV. Autism spectrum disorders: Integration of the genome, transcriptome and the environment. *J Neurol Sci*. 2016; 364:167-76.
2. Teschendorff AE, Menon U, Gentry-Maharaj A, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res*. 2010; 20:440-6.
3. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet*. 2014; 383:1990-8.
4. Rakyan VK, Down TA, Balding DJ, and Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet*. 2011; 12:529-41.
5. McCabe MT, Brandes JC, Vertino PM. Cancer DNA methylation: molecular mechanisms and clinical implications. *Clin Cancer Res*. 2009;15:3927-37.
6. Hon GC, Hawkins RD, Caballero OL, et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res*. 2012; 22:246-58.

7. Lee EJ, Luo J, Wilson JM, Shi H. Analyzing the cancer methylome through targeted bisulfite sequencing. *Cancer Lett.* 2013; 340:171-8.
8. Masser DR, Stanford DR, and Freeman WM. Targeted DNA methylation analysis by next-generation sequencing. *J Vis Exp.* 2015; (96).
9. ENCODE Consortium (www.genome.gov/10005107/encode-project/). Accessed June 2016.
10. FANTOM Consortium (fantom.gsc.riken.jp/). Accessed June 2016.
11. Epigenomics Project Roadmap (www.roadmapepigenomics.org/). Accessed June 2016.
12. Agilent SureSelect MethylSeq target regions (earray.chem.agilent.com/suredesign) Accessed 14 June 2016.
13. Roche Nimblegen CpGiant target regions (sequencing.roche.com/products/nimblegen-seqcap-target-enrichment/seqcap-epi-system/seqcap-epi-cpgiant-enrichment-kit.html) Accessed 14 June 2016.
14. BaseSpace MethylSeq App (www.illumina.com/informatics/research/sequencing-data-analysis-management/basespace/basespace-apps/methylseq-1927928.html). Accessed July 2016.
15. BaseSpace MethylKit App (www.illumina.com/informatics/research/sequencing-data-analysis-management/basespace/basespace-apps/methylkit-1933933.html). Accessed July 2016.
16. BaseSpace Integrative Genomics Viewer (www.illumina.com/informatics/research/sequencing-data-analysis-management/basespace/basespace-apps/integrative-genomics-viewer-1886885.html). Accessed July 2016.
17. Krueger F and Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics.* 2011; 27:1571–1572.
18. Langmead B and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods.* 2012; 9: 357-359.

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

For Research Use Only. Not for use in diagnostic procedures.

 M-GL-01451. Current as of 16 November 2022.

illumina®