Epigenetics
Made Simple™

Zymo-Seq Methyl Spike-in Control

Standards for DNA methylation analysis workflows

Highlights

- **Accurate quantification:** Reliable calculation of bisulfite conversion efficiency post bisulfite library prep
- **Precise calibration points:** Six amplicons with 0, 10, 25, 50, 75, and 100% methylation levels allow for a standard curve and robust data normalization
- **Versatile application:** Compatible with various species (except *E. coli*) and bisulfite sequencing library preparation methods

Catalog Numbers:
D5500

Scan with your smart-phone camera to
view the online protocol/video.

Table of Contents

Product Contents

Product Name	D5500 (25 prep)	Storage Temperature
Zymo-Seq Methyl Spike-in Control	25 µl	-20 ºC

Specifications

- **Source –** Double-stranded synthetic amplicons derived from the *E. coli* genome.

- **Complete reference sequences –**
  The reference genome of *E.coli* strain K-12 substrain MG1655 can be accessed at the following web address:
  https://www.ncbi.nlm.nih.gov/genome/167?genome_assembly_id=161521
  Amplicon sequences can also be found in Appendix I.

- **DNA concentration –** 60 pg/µl

- **Methylation levels –** Table 1 shows the expected methylation levels of the amplicons in **Zymo-Seq Methyl Spike-in Control**.

  Table 1**:** Methylation levels of each amplicon.

Amplicon #	Methylation Level
1	0%
2	10%
3	25%
4	50%
5	75%
6	100%

Product Description

Zymo-Seq Methyl Spike-in Control is a mixture of six unique double-stranded synthetic amplicons (180-200 bp) derived from the *E. coli* K12 genome. Each amplicon represents a different CpG methylation level ranging from 0% to 100% and can be identified based on its distinct sequence composition. Amplicons with different methylation levels are generated by mixing an *in vitro* methylated amplicon with the unmethylated amplicon of the same sequence at the specified ratio.

Zymo-Seq Methyl Spike-in Control serves as an *in situ* control for genome-wide and whole genome bisulfite sequencing library preparations from DNA of any species, except *E. coli*. Since the control is well-defined, it can be used to evaluate various aspects of the library preparation and bioinformatics pipeline. It is ideal for evaluating efficiency of the bisulfite-conversion reaction because CpH sites (H = A, T, C) are not methylated. This is especially useful for evaluating bisulfite conversion efficiency in samples known to have methylation in CpH context (for example, plants, pluripotent stem cells, etc.). The control can also be used to assess possible errors in bioinformatics analysis, such as trimming, alignment, or methylation calling. Since **Zymo-Seq Methyl Spike-in Control** is comprised of six amplicons with defined methylation levels, it can be used to validate bioinformatics pipeline calibration by demonstrating a strong correlation between the observed and expected methylation levels of the amplicons (Figure 1).
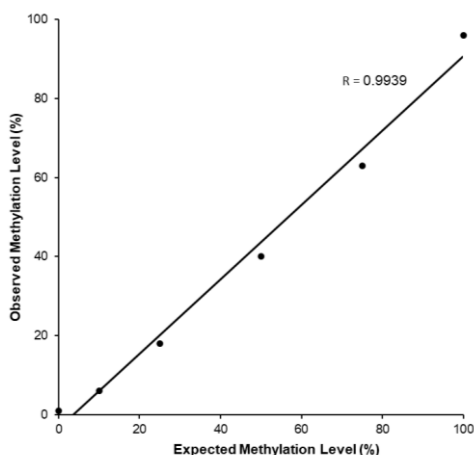

Figure 1. **The observed methylation levels of the amplicons in Zymo-Seq Methyl Spike-in Control strongly correlate with the expected methylation levels.** An RRBS library was prepared from 300 ng of human genomic DNA with 60 pg of **Zymo-Seq Methyl Spike-in Control**. Following library preparation and sequencing of the library, the reads were aligned to the **Zymo-Seq Methyl Spike-in Control** reference genome. The methylation level of each cytosine in CpG context was calculated as the number of reads reported as a C divided by the total number of reads reported as a C or T. The methylation level of each amplicon is reported as the average methylation level of all CpG sites sequenced in the amplicon. The correlation between the observed versus expected methylation level was 0.9939. Each dot on the graph represents an amplicon in **Zymo-Seq Methyl Spike-in Control**.

Expected methylation levels of the amplicons are as follows: amplicon 1 – 0%, amplicon 2 – 10%, amplicon 3 – 25%, amplicon 4 – 50%, amplicon 5 – 75%, amplicon 6 – 100%.

Protocol

1. Use the guidelines below to determine when to add the **Zymo-Seq Methyl Spike-in Control** to a DNA sample for an NGS library preparation:

    a) Reduced Representation Bisulfite Sequencing (RRBS): Add the **Zymo-Seq Methyl Spike-in Control** prior to digestion with MspI.

    b) Conventional WGBS (DNA adapterized prior to bisulfite conversion): Add the **Zymo-Seq Methyl Spike-in Control** after DNA fragmentation and before the end-repair step.

    c) Post-bisulfite library prep (DNA adapterized after bisulfite conversion): Add the **Zymo-Seq Methyl Spike-in Control** prior to bisulfite conversion.

2. Thaw the **Zymo-Seq Methyl Spike-in Control** on ice. Mix by pipetting and spin down.

3. Determine the amount of the **Zymo-Seq Methyl Spike-in Control** to be added to DNA samples for library preparation. Recommended amounts are specified in the table below:

DNA Input for Library Prep	Zymo-Seq Methyl Spike-in Control Amount
100 – 500 ng	60 pg
50 – 99 ng	30 pg
20 – 49 ng	10 pg
< 20 ng	< 10 pg[1]

4. Depending on the amount of the **Zymo-Seq Methyl Spike-in Control** needed per DNA sample, dilute the 60 pg/µl **Zymo-Seq Methyl Spike-in Control** to either 30 pg/µl or 10 pg/µl in TE buffer[2].

5. Add 1 µl of the diluted **Zymo-Seq Methyl Spike-in Control** to each DNA sample at the appropriate step in an NGS library preparation protocol.

---

[1] The appropriate amount of the **Zymo-Seq Methyl Spike-in Control** may need to be determined experimentally.

[2] Aliquots of 60 pg/µl and 30 pg/µl dilution of the **Zymo-Seq Methyl Spike-in Control** can be stored at -20° C.

Appendices

Appendix I: Reference Amplicon Sequences
In the sequences below, all CpG sites are in bold, and MspI recognition sites are highlighted in grey.

Amplicon 1 (methylation level – 0%)

GATTGCTATTAGTGCCT**CCG**GTGCTTTGCCAGACACATATAGTAGC
CTGCCTG**CG**TTACCTT**CG**CTGGAAGGGCTGA**CG**GTA**CG**CAAGCTG
CAACTCTCTATGGACC**CG**ATGCTGACTATGATGGGGATGCAGATGC
TAATGGAGAAATATGG**CG**ATCAGG**CG**ATGG**CCG**GGATGGATCACA
G

Amplicon 2 (methylation level – 10%)

CAT**CG**TTGA**CG**AACACA**CCG**GT**CG**TACCATGCAGGGC**CG**T**CG**CTG
GTC**CG**ATGGTCTGCACCAGGCTGTGGAAG**CG**AAAGAAGGTGTGCA
GATCCAGAA**CG**AAAACCAAA**CG**CTGGCTTGACTCACCTTCCAGAAC
TACTTC**CG**TCTGTATGAAAAACTGG**CG**GGGATGA**CCG**GTACTGCTG
ATAC**CG**AAG

Amplicon 3 (methylation level – 25%)

GGTGAAAGTT**CCG**GATGGCAC**CG**TTGATCCATTT**CG**TCTGAC**CG**CA
GCAAACATGCTGGATGCCCAAAGAACA**CG**GTGC**CG**TTATCCTTAC**CG**
CTCATGAAGTCA**CG**GGGCTGATT**CG**TGAAGG**CGCG**A**CG**GTGTG**CG**
GTGTT**CG**TGTA**CG**TAACCATCTCA**CCG**G**CG**AAACTCAG

Amplicon 4 (methylation level – 50%)

GGATTGCTTTCCAGTT**CCG**GTA**CG**GTCAGCACTTTGG**CG**CAGG**CG**A
TATTCAGTTCAGCAAAG**CG**TTCTTTTACTTC**CGCG**ATGGTATGTGT**C
G**CCAGCCAGGCATGA**CG**TTTCTCTTCAACCAGTGGGC**CG**TAAGGG
CATTGACTA**CG**GTGGATAAGCTGAGTGCCTT**CCG**GGATTTCTGG**CG**
TG

Amplicon 5 (methylation level – 75%)

CAGG**CG**GGTAAGATTTG**CCG**GTTCTTGCTGATACAGCAAT**CG**TGCC
AGA**CG**CAGAGCTTCAGCCTTGTTA**CG**GGT**CG**CCA**CG**CTGACAGCAT
AA**CG**CTCCTCAAGCATTTCATTGG**CG**GGGAGGGTGG**CG**AGCAGTTT
TTG**CG**CTG**CG**T**CG**TACTGACCTTCTTTATACAGCAC**CCG**GTAG**CG**T**C
GCG**CCAACAACATACTGG**CG**

<u>Amplicon 6 (methylation level – 100%)</u>

CATACCAGGGAAAAAT C**CG**G CCTC**CG**CAGCA**CG**AAGCAGAGTG**CG**
AACTATCAGAAATTT**CG**CTTCAGTAT**CG**GCCCATGCCATGGCTGC**C
G**AAAGAAATCCCCACAGCAGTGTTGT**CG**TACCAATCCAGGTTCTGG
CCCCCAGTTTG**CG**CATCAAAAGATT**CG**C**CG**GAACACCCAGAAAC

Appendix II: Zymo-Seq Methyl Spike-in Control Analysis
<u>Reference index preparation</u>
Prepare the **Zymo-Seq Methyl Spike-in Control** reference index with
Bismark (v0.23.0). Navigate to the working directory where the **Zymo-
Seq Methyl Spike-in Control** FASTA is stored, and run the following
command:

```
bismark_genome_preparation .
```

<u>Trimming</u>
The FASTQ files containing the reads should then be trimmed with
optimal trimming parameters to remove any adapter sequence, low
quality reads, or artifacts introduced by the library prep methodology
used.

<u>Alignment</u>
Since the **Zymo-Seq Methyl Spike-in Control** was applied *in situ*, the
entire trimmed FASTQ file for each sample should be aligned to the
reference using Bismark (v0.23.0). If the library is non-directional, the --
non_directional flag can be added to the command line.

```
bismark <path_to_reference_index> -1 <input_read1> -2
<input_read2>
```

The overall alignment should be low since a very small amount of the
control has been spiked in. However, these reads should be sufficient for
high coverage of the control sequences.

<u>Evaluating Bisulfite Conversion Efficiency</u>
Zymo-Seq Methyl Spike-in Control provides an unbiased and accurate
way to calculate bisulfite conversion efficiency. A common method to
evaluate bisulfite conversion efficiency is by determining the percentage
of cytosines in non-CpG context that are converted into thymines.
However, this approach is not applicable when working with sample
types known to have CpH (H = A, T, or C) methylation, such as
embryonic stem cells, pluripotent stem cells, oocytes, neurons, glial cells,
and plant DNA.

Since **Zymo-Seq Methyl Spike-in Control** is manufactured to have methylation only in a CpG context, non-CpG sites of the amplicons in the control are guaranteed to be unmethylated. This ensures that bisulfite conversion efficiency is calculated correctly when examining non-CpG sites in **Zymo-Seq Methyl Spike-in Control**.

Once Bismark finishes running, alignment report in TXT format should be used to calculate the bisulfite conversion efficiency using the following formula:
%bs = [1- (Total methylated C's in CHG context + Total methylated C's in CHH context) / (Total methylated C's in CHG context + Total methylated C's in CHH context + Total unmethylated C's in CHG context + Total unmethylated C's in CHH context)] * 100%

Methylation Level Analysis
There may be a discrepancy between the observed and the expected methylation levels of the amplicons in **Zymo-Seq Methyl Spike-in Control**. The amplicons are enzymatically methylated *in vitro*, and it has been observed that the amplicon expected to be 100% methylated is 96 – 99% methylated. Variation can also be introduced by artifacts in the bioinformatics pipeline. It is recommended to overlook the discrepancy between the observed and expected methylation levels of the amplicons if there is overall strong correlation as demonstrated earlier in Figure 1, and if there is an obvious trend of increasing methylation levels from amplicon 1 to amplicon 6. Table 2 shows an example of the typical result of analyzing expected versus observed methylation levels of the amplicons in the control.

Table 2: Example of expected versus observed methylation levels of amplicons in **Zymo-Seq Methyl Spike-in Control**.

Amplicon #	Expected Methylation Level (%)	Observed Methylation Level (%)
1	0	0
2	10	6
3	25	18
4	50	40
5	75	63
6	100	96

Frequently Asked Questions

Q: What percentage of reads are assigned to the Zymo-Seq Methyl Spike-in Control in sequencing?

A: Less than 0.5% of the total reads goes into the Zymo-Seq Methyl Spike-in Control.

Ordering Information

Product Description	Catalog No.	Size
Zymo-Seq Methyl Spike-in Control	D5500	25 Preps.
Zymo-Seq RRBS Library Kit	D5460 D5461	24 Preps. 48 Preps.
Zymo-Seq WGBS Library Kit	D5465	24 Preps.
Human HCT116 DKO Non-Methylated DNA	D5014-1	5 µg/20 µl
Human HCT116 DKO Methylated DNA	D5014-2	5 µg/20 µl
Bisulfite-Converted Universal Methylated Human DNA Standard	D5015	1 µg/50 µl
ZymoTaq™ Premix	E2003 E2004	50 Rxns. 200 Rxns.
ZymoTaq™ qPCR Premix	E2054 E2055	50 Rxns 200 Rxns.
EZ DNA Methylation™ Kit	D5001 D5002	50 Rxns 200 Rxns.
EZ DNA Methylation-Gold™ Kit	D5005 D5006	50 Rxns 200 Rxns.
EZ DNA Methylation-Direct™ Kit	D5020 D5021	50 Rxns 200 Rxns.
EZ DNA Methylation-Lightning™ Kit	D5030 D5031	50 Rxns 200 Rxns.

100% satisfaction guarantee on all Zymo Research products, or your money back.

Zymo Research is committed to simplifying your research with quality products and services. If you are dissatisfied with this product for any reason, please call 1(888) 882-9682.

The *BEAUTY* of *SCIENCE* is to Make Things *SIMPLE*®