

Reduced Representation Bisulfite Sequencing in Maize

Fei-Man Hsu¹, Chung-Ju Rachel Wang²,* and Pao-Yang Chen²,*

¹Graduate School of Frontier Sciences, the University of Tokyo, Chiba, Japan; ²Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan

*For correspondence: rwang@gate.sinica.edu.tw; paoyang@gate.sinica.edu.tw

[Abstract] DNA methylation is an epigenetic modification that regulates plant development (Law and Jacobsen, 2010). Whole genome bisulfite sequencing (WGBS) is a state-of-the-art method for profiling genome-wide methylation patterns with single-base resolution (Cokus *et al.*, 2008). However, for an organism with a large genome, *e.g.*, the 2.1 Gb genome of maize, WGBS may be very expensive. Reduced representation bisulfite sequencing (RRBS) has been developed in mammalian studies (Smith *et al.*, 2009). By digesting the genome with *MspI* with a size selection range of approximately 40-220 bp, CG-rich regions covering only ~1% of the human genome can be specifically sequenced. However, unlike mammalian genomes, plant genomes do not exhibit clear CpG islands. Therefore the original RRBS protocol is not suitable for plants. Accordingly, we developed an *in silico* pipeline to select specific enzymes to generate a region of interest (ROI)-enriched, *e.g.*, promoter-enriched, reduced representation genome in plants (Hsu *et al.*, 2017). By digesting the maize genome with *MseI* and selecting 40-300 bp segments, we sequenced about one-fourth of the maize genome while preserving 84.3% of the promoter information. The protocol has been successfully established in maize and can be broadly used in any genome. Our *in silico* pipeline is combined with the RRBS library preparation protocol, allowing for the computational analysis and experimental validation.

Keywords: Bisulfite sequencing, DNA methylation, Epigenetics, Maize, Methylome

[Background] DNA methylation is a heritable epigenetic modification that plays an important role in many developmental processes of animals, plants and fungi by regulating gene expression and the chromatin structure (Law and Jacobsen, 2010). WGBS is a genome-wide scale method for profiling DNA methylation at single-base resolution, although high sequencing costs are required to achieve sufficient coverage (Cokus *et al.*, 2008). In mammals, RRBS has been developed to specifically sequence CG-dense regions, *e.g.*, CpG islands (Smith *et al.*, 2009). In this protocol, we aimed to adapt RRBS for plants. To be specific, we developed an *in silico* pipeline (Figure 1) to performed enzyme selection by targeting specific genome regions to generate RRBS methylomes and provided an experimental validation protocol.

Our method has been successfully established in the maize genome, one of the major global crops, which has a 2.1 Gb genome (Hsu *et al.*, 2017). In addition to its large size, 85% of the maize genome consists of various repetitive sequences (Schnable *et al.*, 2009). This feature could cause multiple mapping, *i.e.*, many short reads from sequencing, which need to be discarded. These characteristics make targeted bisulfite sequencing (BS-seq) more cost-effective. mCHH islands were found to be

located upstream of transcription start sites (TSSs) with higher methylation level (Gent *et al.*, 2013; Li *et al.*, 2015). We therefore aimed to perform promoter-enriched RRBS in maize, and an *in silico* pipeline was developed for enzyme selection.

Our *in silico* pipeline currently has 85 pre-installed restriction enzymes. Users can easily append more enzymes. A genome FASTA file, refFlat annotation file, and repeat gff3 annotation file are the required input files to run this pipeline. ROIs, including promoters, exons, introns, splicing sites, repeats, UTRs and intergenic regions are pre-selected for enrichment analysis. As soon as the input files are prepared, users can run our pipeline to select ideal enzymes by typing simple commands. Users can also verify the prediction by performing RRBS library construction following the experimental protocol provided and performing sequencing.

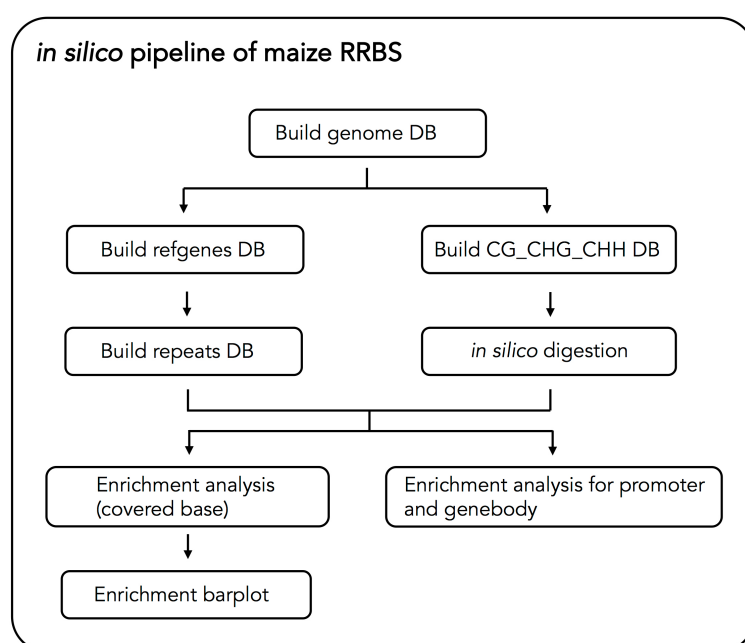


Figure 1. Flowchart of maize RRBS *in silico* pipeline

Materials and Reagents

1. Pipette tips
Note: Low retention tips are recommended.
2. 1.5 ml microcentrifuge tubes (Eppendorf, catalog number: 0030108051)
3. PCR tubes (Thermo Fisher Scientific, Applied Biosystems™, catalog number: A30588)
4. Clean razor blade
5. Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Invitrogen™, catalog number: Q32850) (used to quantify double strand DNA in combination with Qubit Fluorometer)
6. *Mse*I (New England Biolabs, catalog number: R0525S) (used to digest maize genomic DNA, CutSmart Buffer is included)

Note: The enzyme could be replaced according to the in silico enzyme selection result in

Procedure A.

7. Agarose

Note: For the gel size selection, we recommend low melting point agarose.

8. TBE buffer (Thermo Fisher Scientific, Invitrogen™, catalog number: 15581044)

9. AMPure XP (Beckman Coulter, catalog number: A63881)

10. Absolute ethanol

Note: To dilute, use DNase/RNase-free distilled water.

11. 1 M Tris-HCl (Thermo Fisher Scientific, Invitrogen™, catalog number: 15568025)

Note: To dilute, use DNase/RNase-free distilled water.

12. NEBuffer 2 (New England Biolabs, catalog number: B7002S)

13. Klenow Fragment (3'→5' exo-) (New England Biolabs, catalog number: M0212S) (used to end-repair and A-tail DNA)

14. T4 DNA Ligase (New England Biolabs, catalog number: M0202S, T4 DNA ligation buffer is included) (used to ligate sequencing adaptors)

15. 100-bp DNA ladder

16. QIAquick Gel Extraction Kit (QIAGEN, catalog numbers: 28704, 28706)

17. EpiTect Fast Bisulfite Conversion Kit (QIAGEN, catalog number: 59824)

Note: Keep this kit at an appropriate temperature according to the manufacturer's guide.

18. TruSeq DNA LT Sample Prep Kit (Illumina, catalog number: FC-121-2001)

19. PfuTurbo DNA Polymerase (Agilent Technologies, catalog number: 600250)

20. DNase/RNase-free distilled water (Thermo Fisher Scientific, Invitrogen™, catalog number: 10977015)

21. Deoxynucleotide (dNTP) Solution Set (New England Biolabs, catalog number: N0446S)

22. MinElute PCR Purification Kit (QIAGEN, catalog number: 28004)

23. 10 mM dNTP mix (see Recipes)

24. 5x RRBS dNTP mix (see Recipes)

Equipment

1. Pipettes:

- a. Thermo Fisher Scientific, Thermo Scientific™, model: P5000, catalog number: 4641110N
- b. Thermo Fisher Scientific, Thermo Scientific™, model: P1000, catalog number: 4641100N
- c. Thermo Fisher Scientific, Thermo Scientific™, model: P300, catalog number: 4641090N
- d. Thermo Fisher Scientific, Thermo Scientific™, model: P200, catalog number: 4641080N
- e. Thermo Fisher Scientific, Thermo Scientific™, model: P100, catalog number: 4641070N
- f. Thermo Fisher Scientific, Thermo Scientific™, model: P20, catalog number: 4641060N
- g. Thermo Fisher Scientific, Thermo Scientific™, model: P10, catalog number: 4641030N
- h. Thermo Fisher Scientific, Thermo Scientific™, model: P2, catalog number: 4641010N

2. Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Invitrogen, model: Qubit® 2.0 Fluorometer, catalog number: Q32866)
3. Heating block (Eppendorf, model: ThermoMixer comfort, catalog number: 5355000011)
4. Incubator (Only Science, Firstek, model: S300)
5. Thermocycler
 - a. Thermo Fisher Scientific, Applied Biosystems™, model: GeneAmp™ PCR System 9700, catalog number: 4413750
 - b. Bio-Rad Laboratories, model: T100™ Thermal Cycler, catalog number: 1861096
6. Centrifuge
 - a. Eppendorf, model: Centrifuge MiniSpin® plus, catalog number: 5453000011
 - b. Eppendorf, model: Centrifuge 5424, catalog number: 5424000410
 - c. GYROZEN, model: 1730R
7. Magnetic stand
 - a. Thermo Fisher Scientific, DynaMag™-2 Magnet, catalog number: 12321D
 - b. Thermo Fisher Scientific, Applied Biosystems™, DynaMag™-PCR Magnet, catalog number: 492025

Software

1. System requirements: Linux/Unix or Mac OS, python 2.7+, R 3.2.0
2. Python and R scripts and a tutorial are available at
https://gitlab.com/fmhsu0114/maize_RRBS

Procedure

The complete pipeline contains 1. *in silico* enzyme selection, followed by 2. RRBS library preparation. The resulted library is to be sequenced to generate DNA methylation profiles at genomic regions of interest.

A. *In silico* enzyme selection

1. Clone the GitLab project from: https://gitlab.com/fmhsu0114/maize_RRBS.
2. Download the genome FASTA file of your interest, *i.e.*, maize_genome.fa, and use the following command to build the genome database:

```
$ python Build_Gnome.py AGPv3.fa
```

2 output files (genome.shelve and log-genome.txt) will be generated.
3. Download the annotation refFlat file, *i.e.*, refFlat.txt, and use the following command to build the gene annotation database:

```
$ python Build_refgene_shelve.py refFlat.txt log-genome.txt
```

1 output file (refgenes.shelve) will be generated.
4. Download the annotation gff3 file and grep 'repeats' into an individual file, *i.e.*, repeats.gff3, and

use the following command to build the repeat annotation database:

```
$ python Build_repeats_Gnome.py repeats.gff3, log-genome.txt
```

1 output file (repeat.shelve) will be generated.

5. Build the genome CG, CHG and CHH sites database with the following command:

```
$ python Build_CG_CHG_CHH_Gnome.py genome.shelve log-genome.txt
```

1 output file (CG_CHG_CHH.shelve) will be generated.

6. Perform *in silico* digestion with the provided enzyme list in the GitLab project (enzyme_cutting_sites.txt) with the following command:

```
$ python Enzyme_digestione.py
```

fragment_info (per MSRE and per size range) for each enzyme and a Summary-enzyme-coverage.txt file will be generated.

7. Place the output files from 1.6 into one directory ./fragment/ and run the enrichment analysis with the following command:

```
$ python RRBS_fragments_enrichment.py repeats.shelve, refgenes.shelve
```

log₂_fold_enrichment (per MSRE and per size range) files will be generated (see Table 1).

Table 1. Fold enrichment values in log₂ scale of the *MseI* 40-300-bp fragments

	ALL	Promoter	5' UTR	Exon	Intron	3' UTR	Splice_site	Repeats	Intergenic
Genome	206,065,361	37,355,117	12,574,937	39,031,481	77,413,200	14,101,473	267,900	1,720,406,932	1,879,909,493
Genome (percentage)	1.000	0.018	0.006	0.019	0.038	0.007	0.000	0.835	0.912
RRBS	565,511,128	13,622,681	2,846,911	6,149,536	26,638,124	4,442,513	83,686	452,360,346	511,727,677
RRBS (percentage)	1.000	0.024	0.005	0.011	0.047	0.008	0.000	0.800	0.905
Log₂FC	0.000	0.41	-0.278	-0.801	0.326	0.199	0.187	-0.062	-0.012

8. Move all of the log₂_fold_enrichment files into one directory ./output/ and plot the enrichment bar charts with the following command:

```
$ Rscript Barchart_enrichment.R
```

See Figure 2 to Figure 4 as examples.

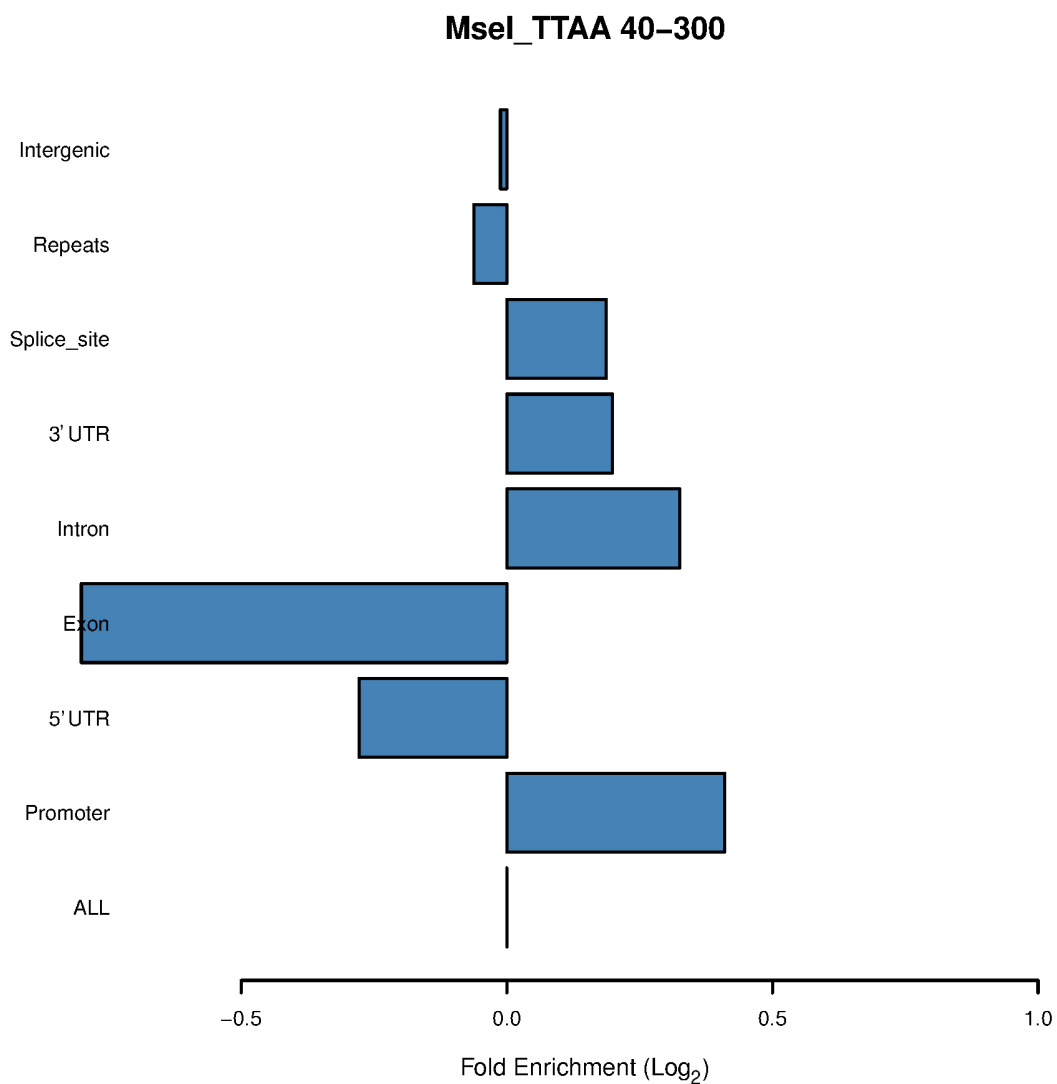


Figure 2. Bar chart of the *MseI* 40-300-bp fragments shows enrichment in promoters

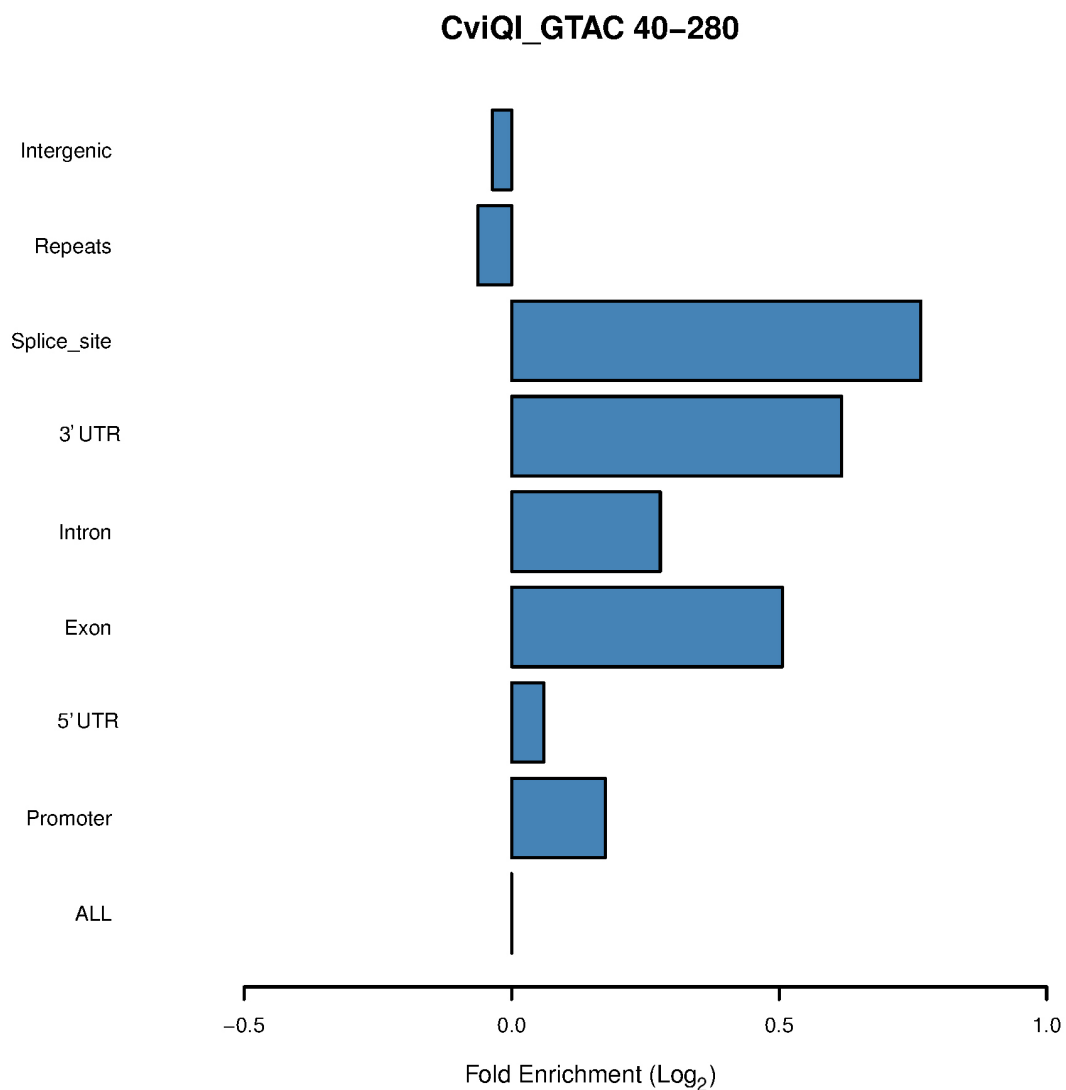


Figure 3. Bar chart of the CviQI 40-280-bp fragments shows enrichment in exons, introns, UTRs and splice sites

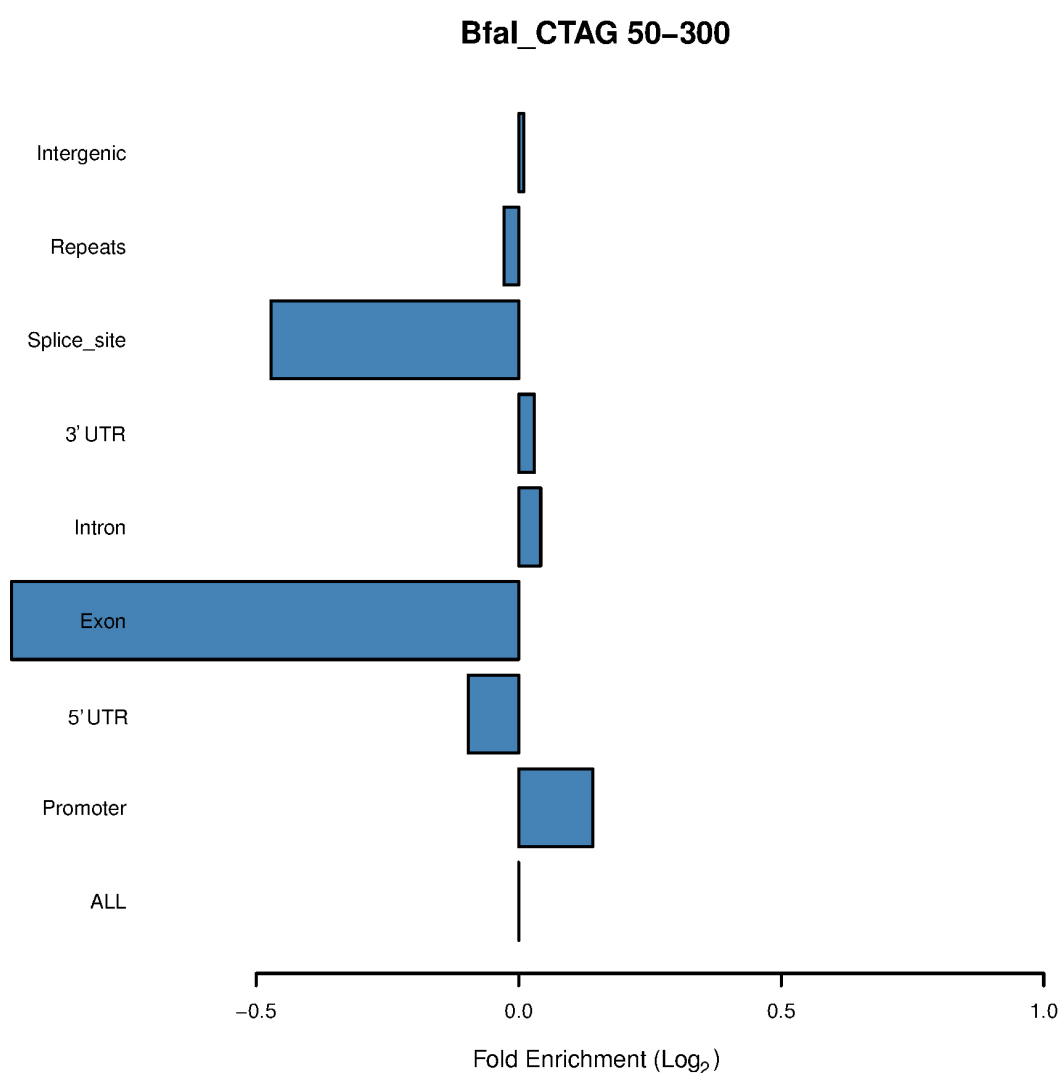


Figure 4. Bar chart of the *Bfal* 50-300-bp fragments shows enrichment in promoters

- Use the following command to calculate the covered promoter and genebody region:

```
$ python RRBS_fragments_enrichment_gene_level_per_MRE_Size.py -i Enzyme -s size_range -c 1
```

This script has 3 parameters:

-i: restriction enzyme

-s: size selection range, e.g., 40-280

-c: number of cytosines per promoter/genebody to be included

An output file, `Fragment-Enrichment-genebody-pmt-Enzyme-size_range-1.txt`, will be generated (Table 2 and Table 3). This file contains the percentages of the promoters and genebody covered in the reduced representation (RR) genome.

Note: In Table 2 and Table 3, Pct refers to percentage; Pmt refers to promoter.

Table 2. Promoter and genebody covered in the *MseI* 40-300-bp fragments

Enzyme	Cutting_Site	Size_Selection	Genome_Size	Fragment_Covered	Fragment_Covered_Pct	Total_Genes	Covered_Genebody	Covered_Genebody_Pct	Log ₂ _Fold_Enrichment_Genebody	Covered_Pmt	Covered_Pmt_Pct	Log ₂ _Fold_Enrichment_Pmt
<i>MseI</i>	TTAA	40-300	2,060,653,601	565,511,128	0.274	38,653	29,219	0.756	1.462	32,600	0.843	1.620

Table 3. Promoter and genebody covered in the *Bfal* 50-300-bp fragments

Enzyme	Cutting_Site	Size_Selection	Genome_Size	Fragment_Covered	Fragment_Covered_Pct	Total_Genes	Covered_Genebody	Covered_Genebody_Pct	Log ₂ _Fold_Enrichment_Genebody	Covered_Pmt	Covered_Pmt_Pct	Log ₂ _Fold_Enrichment_Pmt
<i>Bfal</i>	CTAG	50-300	2,060,653,601	472,960,164	0.230	38,653	30,256	0.783	1.770	28,906	0.748	1.704

B. RRBS library preparation

1. Use Qubit Fluorometric Quantification with Qubit dsDNA BR assay to quantify maize genomic DNA according to the manufacturer's instructions.

Note: To acquire high quality double strand genomic DNA, we recommend spin column-based DNA extraction kit, e.g., QIAGEN DNeasy Plant Mini kit. If Pheno-Chloroform method is used, we suggest another round of spin-column based purification.

2. Aliquot 1 µg of maize genomic DNA into a 1.5 ml tube, add 5 µl of 10x CutSmart Buffer, 1 µl of MseI and ddH₂O to a final volume of 50 µl. Mix carefully with a pipette and incubate at 37 °C on the heating block for 4 h.

Note: MseI used here is designed for promoter-enriched RRBS in maize. For different genomes (organisms), researchers should conduct 'in silico enzyme selection' and replace MseI with the predicted enzyme, and modify the digestion time, buffer and temperature according to the manufacturer's instructions.

3. [Optional] Take 5 µl for gel electrophoresis (100 ng, 1.5%, 0.5x TBE) and add 5 µl ddH₂O to the enzyme digestion product.

Note: This step is to confirm the enzyme digestion efficiency.

4. Post-digestion AMPure beads purification

- a. Add 150 µl of well-inverted AMPure XP beads solution to 50 µl of DNA, vortex for 10 sec, briefly centrifuge and incubate at RT for 5 min.

Note: When vortexing solutions with AMPure XP beads, avoid high vortexing strength causing bubbles. After vortexing, use microcentrifuge to spin down quickly to remove liquids from the lid and the wall of the tube.

- b. Place the tube on the magnetic stand for 3 min.
- c. Remove the supernatant.
- d. Add 200 µl of 80% EtOH without disturbing the beads and incubate at RT for 30 sec. Remove the EtOH.
Note: Do not remove the tube from the magnetic stand.
- e. Repeat the washing Step B4d.
- f. Air-dry the beads for 5 min.
- g. Remove the tube from the magnetic stand.
- h. Add 32 µl of 10 mM Tris-HCl and pipet up and down to mix thoroughly. Incubate at RT for 2 min.
- i. Place the tube on the magnetic stand for 3 min.
- j. Transfer the supernatant to a PCR tube.

5. End repair and A-tailing

- a. Add 6 µl of 10x NEBuffer 2, 12 µl of 5x RRBS dNTP mix and 3 µl of 3'→5' Klenow exo- to the purified DNA and add ddH₂O to a final volume of 60 µl.
- b. Incubate using thermocycler at 30 °C for 20 min and 37 °C for 20 min, then place the tube on ice.

6. DNA purification with AMPure XP beads
 - a. Add 180 μ l of well-inverted AMPure XP beads to the 60 μ l of DNA, vortex for 10 sec, briefly centrifuge and incubate at RT for 5 min.
 - b. Place the tube on the magnetic stand for 3 min.
 - c. Remove the supernatant.
 - d. Add 200 μ l of 80% EtOH without disturbing the beads and incubate at RT for 30 sec. Remove the EtOH.
Note: Do not remove the tube from the magnetic stand.
 - e. Repeat the washing Step B6d.
 - f. Air-dry the beads for 5 min.
 - g. Remove the tube from the magnetic stand.
 - h. Add 24 μ l of 10 mM Tris-HCl and pipet up and down to mix thoroughly. Incubate at RT for 2 min.
 - i. Place the tube on the magnetic stand for 3 min.
 - j. Transfer the supernatant to a PCR tube.
7. Ligate with a TruSeq methylated Illumina adaptor
 - a. Add 3 μ l of 10x T4 DNA ligation buffer, 2 μ l of adaptor and 1 μ l of T4 DNA ligase to the 24 μ l end-repaired and A-tailed DNA to a final volume of 30 μ l.
 - b. Incubate at 16 °C overnight in a thermocycler.
8. Gel size selection (Figure 5)
 - a. Load all the adapter-ligated DNA to 1.5% 0.5x TBE gel (15 μ l per well, in total 2 wells) for gel electrophoresis with a 100-bp DNA ladder.
 - b. Use a clean razor blade for each sample to cut 40-300-bp fragments and purify the DNA using the QIAGEN Qiaquick Gel Extraction Kit following the manufacturer's instructions. Elute DNA twice with a total volume of 34 μ l of EB buffer. Transfer the solution to a PCR tube. Take 1 μ l for Qubit quantification (optional).

Note: One should avoid UV light during gel size selection. Since the adaptor sequence is about 120 bp, we cut 160-420-bp fragments from the gel. For each well in the previous step, use one QIAGEN Qiaquick Gel Extraction reaction. For the first round of elution, add 10 μ l of EB buffer to each MinElute spin column, and collect the 20 μ l of the solution in a new tube. Then use another 14 μ l of EB buffer to elute the remaining DNA from each spin column one by one.

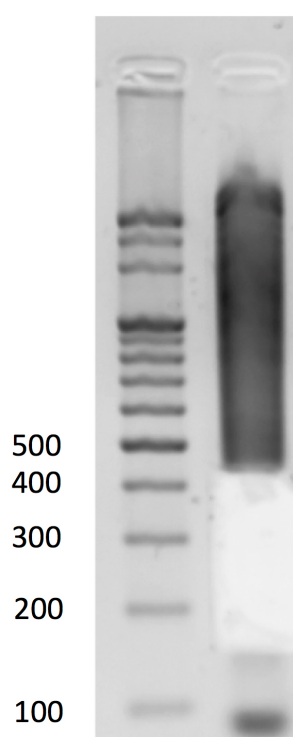


Figure 5. Gel size selection

9. Bisulfite conversion

Note: Confirm that the reagents in the EpiTect Fast Bisulfite Conversion Kit are ready to use.

- a. Add RNase-free water to the DNA solution to a final volume of 40 μ l, mix with a pipette and aliquot 20 μ l to another PCR tube.
- b. Add 35 μ l of DNA Protect Buffer and 85 μ l of a Bisulfite Solution to each tube to a final volume of 140 μ l and mix thoroughly with gentle pipetting.
- c. Run in a thermocycler with the following program:

95 °C	5 min
60 °C	20 min
95 °C	5 min
60 °C	20 min

Then hold at 20 °C. Lid temperature: 100 °C.

10. Briefly centrifuge the PCR tubes containing the bisulfite reaction with microcentrifuge to remove the liquid from the tube lid and transfer to clean 1.5 ml microcentrifuge tubes. For each tube, add 310 μ l of freshly prepared Buffer BL containing 10 μ g/ml carrier RNA, mix thoroughly with pipetting and then centrifuge briefly.
11. Add 250 μ l of 96-100% EtOH, gently pipet for 10 sec and then centrifuge briefly.
12. Transfer all of the solution to MinElute DNA spin columns with collection tubes. Centrifuge at maximum speed for 1 min and discard the flow-through.
13. Washing

Add 500 μ l Buffer BW to each spin column, centrifuge at maximum speed for 1 min and discard

the flow-through.

14. Desulfonation

- a. Add 500 µl of Buffer BD to each spin column and incubate for 15 min at RT.
- b. Centrifuge at maximum speed for 1 min and discard the flow-through.
- c. Add 500 µl Buffer BW to each spin column, centrifuge at maximum speed for 1 min and discard the flow-through.
- d. Repeat the washing Step B14c.
- e. Add 250 µl of 96-100% EtOH to each spin column and centrifuge at maximum speed for 1 min.
- f. Place each spin column into a clean 1.5 ml microcentrifuge tube, open the lid and incubate using a 60 °C heating block to evaporate the liquid.
- g. Elute the DNA by adding 15 µl of Buffer EB onto the center of the spin column membrane, close the lid and incubate at RT for 1 min.
- h. Centrifuge at 15,000 x g for 1 min.

15. PCR amplification

- a. Add 2.5 µl Illumina TruSeq PCR primer cocktail, 0.5 µl of 10 mM dNTP, 5 µl of 10x Pfu Turbo Buffer, 0.5 µl of Pfu Turbo DNA polymerase and 26.5 µl of ddH₂O to the eluted DNA solution to a final volume of 50 µl. Mix thoroughly and centrifuge briefly.
- b. Run in a thermocycler with the following program:

1	95 °C	2 min	} 10-12 cycles
2	95 °C	15 sec	
3	63 °C	15 sec	
4	72 °C	30 sec	
5	72 °C	2 min	
6	4 °C	hold	

16. DNA purification with AMPure XP beads

- a. Add 150 µl of well-inverted AMPure XP beads to the 50 µl of DNA, vortex for 10 sec, briefly centrifuge and incubate at RT for 5 min.
- b. Place the tube on the magnetic stand for 3 min.
- c. Remove the supernatant.
- d. Add 200 µl of 80% EtOH without disturbing the beads and incubate at RT for 30 sec. Remove the EtOH.
Note: Do not remove the tube from the magnetic stand.
- e. Repeat the washing Step B16d.
- f. Air-dry the beads for 5 min.
- g. Remove the tube from the magnetic stand.
- h. Add 22 µl of 10 mM Tris-HCl and pipet up and down to mix thoroughly. Incubate at RT for 2 min.
- i. Place the tube on the magnetic stand for 3 min.
- j. Transfer the supernatant to a 1.5 ml microcentrifuge tube.

17. Quantify the final library concentration with the Qubit assay.
18. Sequence the RRBS library with Illumina paired-end sequencing. Any regular length of pair-end reads, e.g., 100-250 bp, is fine.

Data analysis

1. Our *in silico* pipeline creates a table and bar chart for each enzyme with its associated range of size selection. Figure 1 is a flowchart of this pipeline, and users could see input and output files of each step. Figure 2 is a bar chart showing the enrichment of each potential ROI from approximately 40-300 bp size-selected *MseI*-digested DNA. The X-axis is in log₂ scale, so values > 0 can be considered to be an enrichment of these *MseI*-digested DNA fragments. Table 1 includes the values used to plot Figure 2. From Figure 2 and Table 1, we can see that the *MseI*-digested 40-300-bp fragments are enriched in promoters. As an alternative, Figure 3 shows the result from another enzyme (*CviQI*) 40-280 bp, with fragments more enriched in exons, splice sites, introns and UTRs.
2. Different enzymes produce similar enrichment values. Figure 2 and Figure 4 suggest that *MseI* 40-300-bp and *Bfal* 50-300-bp fragments are enriched in promoters. Table 2 and Table 3 show the statistics of the promoters and genebody percentages covered. For the *MseI* RR genome, by sequencing 566 Mbp of the 2.1 Gb genome, 84.3% of promoters are included. As for the *Bfal* RR genome, sequencing 23.0% of the genome includes 74.8% of the promoters. Users need to determine the tradeoff between the cost of 4.4% of the genome and 8.5% of the promoters.

Conclusion

1. We present an *in silico* pipeline to perform ROI-directed RRBS in maize with an experimentally verifiable prediction. The most critical steps for users to run our pipeline are: 1) Preparing the input files; 2) Choosing an optimal enzyme to show clear enrichment of the selected ROIs (see Figure 2).
2. Input files can often be downloaded from public databases, such as Ensembl (Zerbino *et al.*, 2017). It is important to note that our pipeline requires an individual repeat annotation file in .gff3 format. In addition to the bar charts and table (Figures 2 and 3, Table 1), when judging which size-selection region is optimal, the total RR genome size and read length of the sequencing platform should also be considered. For example, if two RR genomes show 2x enrichment in promoters and one covers 50% of the whole genome while the other covers 25%, the tradeoff between sequencing cost and quantity of information should be considered.
3. For experimental validation, size selection and bisulfite conversion are the most critical steps. In this protocol, we used gel size selection to ensure that the fragment length could be visualized. As an alternative, AMPure XP beads can also be used. To ensure efficient bisulfite conversion, 0.01% sheared lambda phage DNA could be pooled into original genomic DNA prior to enzyme digestion. Since lambda phage DNA lacks methylated cytosine, non-conversion rate can be

estimated by calculating the methylation level of lambda phage DNA to evaluate bisulfite conversion efficiency.

Recipes

1. 10 mM dNTP mix
10 mM dATP
10 mM dTTP
10 mM dCTP
10 mM dGTP
2. 5x RRBS dNTP mix
200 μ M dATP
200 μ M dTTP

Acknowledgments

This study was supported by funding from the Institute of Plant and Microbial Biology, Academia Sinica and grants obtained from Taiwan Ministry of Science and Technology (106-2633-B-001-001) to P.-Y.C. We thank Yi-Jing Lee for providing the equipment information. The authors declare that they have no competing financial interests.

References

1. Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M. and Jacobsen, S. E. (2008). [Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning](#). *Nature* 452(7184): 215-219.
2. Gent, J. I., Ellis, N. A., Guo, L., Harkess, A. E., Yao, Y., Zhang, X. and Dawe, R. K. (2013). [CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize](#). *Genome Res* 23(4): 628-637.
3. Hsu, F. M., Yen, M. R., Wang, C. T., Lin, C. Y., Wang, C. R. and Chen, P. Y. (2017). [Optimized reduced representation bisulfite sequencing reveals tissue-specific mCHH islands in maize](#). *Epigenetics Chromatin* 10(1): 42.
4. Law, J. A. and Jacobsen, S. E. (2010). [Establishing, maintaining and modifying DNA methylation patterns in plants and animals](#). *Nat Rev Genet* 11(3): 204-220.
5. Li, Q., Gent, J. I., Zynda, G., Song, J., Makarevitch, I., Hirsch, C. D., Hirsch, C. N., Dawe, R. K., Madzima, T. F., McGinnis, K. M., Lisch, D., Schmitz, R. J., Vaughn, M. W. and Springer, N. M. (2015). [RNA-directed DNA methylation enforces boundaries between heterochromatin and euchromatin in the maize genome](#). *Proc Natl Acad Sci U S A* 112(47): 14728-14733.
6. Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang,

J., Fulton, L., Graves, T. A., Minx, P., Reilly, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S. M., Belter, E., Du, F., Kim, K., Abbott, R. M., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S. M., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Kohlberg, S., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L., Wei, S., Kumari, S., Faga, B., Levy, M. J., McMahan, L., Van Buren, P., Vaughn, M. W., Ying, K., Yeh, C. T., Emrich, S. J., Jia, Y., Kalyanaraman, A., Hsia, A. P., Barbazuk, W. B., Baucom, R. S., Brutnell, T. P., Carpita, N. C., Chaparro, C., Chia, J. M., Deragon, J. M., Estill, J. C., Fu, Y., Jeddelloh, J. A., Han, Y., Lee, H., Li, P., Lisch, D. R., Liu, S., Liu, Z., Nagel, D. H., McCann, M. C., SanMiguel, P., Myers, A. M., Nettleton, D., Nguyen, J., Penning, B. W., Ponnala, L., Schneider, K. L., Schwartz, D. C., Sharma, A., Soderlund, C., Springer, N. M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T. K., Yang, L., Yu, Y., Zhang, L., Zhou, S., Zhu, Q., Bennetzen, J. L., Dawe, R. K., Jiang, J., Jiang, N., Presting, G. G., Wessler, S. R., Aluru, S., Martienssen, R. A., Clifton, S. W., McCombie, W. R., Wing, R. A. and Wilson, R. K. (2009). [The B73 maize genome: complexity, diversity, and dynamics](#). *Science* 326(5956): 1112-1115.

7. Smith, Z. D., Gu, H., Bock, C., Gnirke, A. and Meissner, A. (2009). [High-throughput bisulfite sequencing in mammalian genomes](#). *Methods* 48(3): 226-232.
8. Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C. G., Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O. G., Janacek, S. H., Juettemann, T., To, J. K., Laird, M. R., Lavidas, I., Liu, Z., Loveland, J. E., Maurel, T., McLaren, W., Moore, B., Mudge, J., Murphy, D. N., Newman, V., Nuhn, M., Ogeh, D., Ong, C. K., Parker, A., Patricio, M., Riat, H. S., Schuilenburg, H., Sheppard, D., Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Zadissa, A., Frankish, A., Hunt, S. E., Kostadima, M., Langridge, N., Martin, F. J., Muffato, M., Perry, E., Ruffier, M., Staines, D. M., Trevanion, S. J., Aken, B. L., Cunningham, F., Yates, A. and Flicek, P. (2017). [Ensembl 2018](#). *Nucleic Acids Res.*