



An Optimized Method for the Construction of a DNA Methylome from Small Quantities of Tissue or Purified DNA from *Arabidopsis* Embryo

Hyunjin Yoo^{1,2}, Kyunghyuk Park^{1,2}, Jaehoon Lee^{1,2}, Seunga Lee¹, and Yeonhee Choi^{1,*}

¹Department of Biological Sciences, Seoul National University, Seoul 08826, Korea, ²These authors contributed equally to this work.

*Correspondence: yhc@snu.ac.kr
<https://doi.org/10.14348/molcells.2021.0084>
www.molcells.org

DNA methylation is an important epigenetic mechanism affecting genome structure, gene regulation, and the silencing of transposable elements. Cell- and tissue-specific methylation patterns are critical for differentiation and development in eukaryotes. Dynamic spatiotemporal methylation data in these cells or tissues is, therefore, of great interest. However, the construction of bisulfite sequencing libraries can be challenging if the starting material is limited or the genome size is small, such as in *Arabidopsis*. Here, we describe detailed methods for the purification of *Arabidopsis* embryos at all stages, and the construction of comprehensive bisulfite libraries from small quantities of input. We constructed bisulfite libraries by releasing embryos from intact seeds, using a different approach for each developmental stage, and manually picking single-embryo with microcapillaries. From these libraries, reliable *Arabidopsis* methylome data were collected allowing, on average, 11-fold coverage of the genome using as few as five globular, heart, and torpedo embryos as raw input material without the need for DNA purification step. On the other hand, purified DNA from as few as eight bending torpedo embryos or a single mature embryo is sufficient for library construction when RNase A is treated before DNA extraction. This method can be broadly applied to cells from different tissues or cells from other model organisms. Methylome construction can be achieved using a minimal amount of input material

using our method; thereby, it has the potential to increase our understanding of dynamic spatiotemporal methylation patterns in model organisms.

Keywords: bisulfite sequencing library, DNA methylation, embryo, methylome

INTRODUCTION

DNA methylation involves the chemical modification of cytosine bases by the addition of methyl groups at 5th-position carbon, forming 5-methyl cytosine (5mC). This methylation is critical for normal development in both plants and mammals as it is associated with gene regulation, genomic imprinting, and the silencing of transposable elements (TE) in both kingdoms (Zeng and Chen, 2019). DNA methylation may also act as a biomarker for cell age and identity, and for various diseases, including cancers (Bell et al., 2019; Kim and Costello, 2017; Kim et al., 2021; Levenson, 2010; Locke et al., 2019; Salas et al., 2018). The DNA methylation landscape is not only cell-type specific but also dynamic, as global reprogramming of DNA methylation occurs during gamete formation and embryo development in mammals (Zeng and Chen, 2019). Conversely, seed plants do not seem to have such global reprogramming, and DNA methylation patterns

Received 9 April, 2021; revised 25 May, 2021; accepted 8 June, 2021; published online 23 August, 2021

eISSN: 0219-1032

©The Korean Society for Molecular and Cellular Biology.

©This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>.

are maintained over multiple trans-generations, particularly in CpG-rich regions (Hofmeister et al., 2017; Picard and Gehring, 2017). However, recent methylome data has revealed a global and gradual increase in methylation at CHH sites, where H represents A, C, or T residues, during plant embryo development (Bouyer et al., 2017; Kawakatsu et al., 2017; Papareddy et al., 2020). These data suggest a role for CHH methylation in the control of embryo development.

DNA methylation research has accelerated with the development of whole-genome and single-cell bisulfite sequencing (Chatterjee et al., 2012; Clark et al., 2017; Karemaker and Vermeulen, 2018; Krueger et al., 2012; Li et al., 2011; Smallwood et al., 2014). Bisulfite treatment of DNA converts cytosine to uracil by hydrolytic deamination, while methyl-cytosine remains unaffected. Polymerase chain reaction (PCR) amplification of bisulfite treated DNA leads to the conversion of uracil to thymine (CT conversion) by DNA polymerases (Frommer et al., 1992). Sequencing of these PCR products allows the identification of cytosine to thymine substitutions when compared to original sequences, identifying regions of un-methylated DNA, with the presence of cytosine in a sequence identifying regions of DNA methylation. Thus, bisulfite sequencing allows the discrimination of methyl-cytosine from unmethylated cytosine in the genome at single base resolution.

DNA methylation patterns, and the extent to which DNA is methylated, are frequently associated with gene and TE expression. These can differ depending on cell type and developmental stage and it is, therefore, important to isolate and enrich samples based on these criteria. In *Arabidopsis*, methylation profiles remain poorly understood, despite a wealth of omics data. This is, in part, due to difficulties in cell- and developmental stage-specific sampling resulting from the structural complexity and the small size of the *Arabidopsis* genome compared with that of animal model organisms, and difficulties in culturing cells. Furthermore, the isolation of specific cell-types often needs specialized protocols to obtain an appropriate amount of sample (Deal and Henikoff, 2011; Ibarra et al., 2012; Park et al., 2016). Despite recent advances in the understanding of methylome changes in whole seeds (Bouyer et al., 2017; Kawakatsu et al., 2017; Lin et al., 2017), methylation dynamics during embryo development in *Arabidopsis* are not fully understood. Recently, Papareddy et al. (2020) examined methylome dynamics at CHH context using dissected embryos. However, their methodology is labor- and time-intensive as it requires purified DNA from more than 50 embryos in all stages as an input for methylome construction. In order to reduce the amount of this input, and simplify the method while maintaining data quality, we adapted cutting edge methods of the construction of mammalian single-cell methylome libraries (Allen et al., 2006; Luo et al., 2018; Yu et al., 2017) for use in *Arabidopsis* embryos. We identified the conditions necessary for embryo isolation and degradation of early embryo cell walls to allow their use as inputs for methylome construction. From globular to torpedo stage embryos, we discovered that as few as five raw globular embryos are sufficient as input material for methylome construction without the need for DNA purification. For later stage embryos, we found that a sufficient amount of DNA could be

purified from as few as eight bending torpedo embryos or a single mature embryo for methylome construction using our optimized protocols. By using a series of suitable tests to minimize the amount of input material while still ensuring its quality, this study provides detailed conditions for optimal sample preparation and the generation of high-quality libraries with on average 11-fold genome coverage.

MATERIALS AND METHODS

Plant growth conditions

Arabidopsis thaliana, Col-*gl* (Columbia-glabrous) ecotype was used for embryo isolation. Plants were grown on soil in an environmentally controlled room at 22°C under long photoperiods (16-h light/8-h dark) with cool white fluorescent light (100 $\mu\text{mole}/\text{m}^2/\text{s}$).

Pollination

After 24 h of emasculation (Park et al., 2016), fully matured stamens were picked from open flowers with tweezers. Pollen was rubbed onto the emasculated stigma for fertilization, under a dissection microscope. Pollinated plants were incubated in the growth room until they reached a suitable stage for sampling; 4 days after pollination (DAP4) for globular, DAP5 for heart, DAP7 for torpedo, DAP9 for bending torpedo, and DAP12 for mature green stage embryos (Fig. 1).

Preparation of globular stage embryos

Siliques at DAP4 were dissected for globular stage embryo preparation, and seeds were collected together in 1.5 ml tubes (seeds from 3-5 siliques per tube) with 50-100 μl of isolation buffer (1 \times TE buffer; 10 mM Tris-HCl [pH 8.0], 1 mM EDTA [pH 8.0]). After brief centrifugation, seeds were ground gently with a pestle, which allows intact globular em-

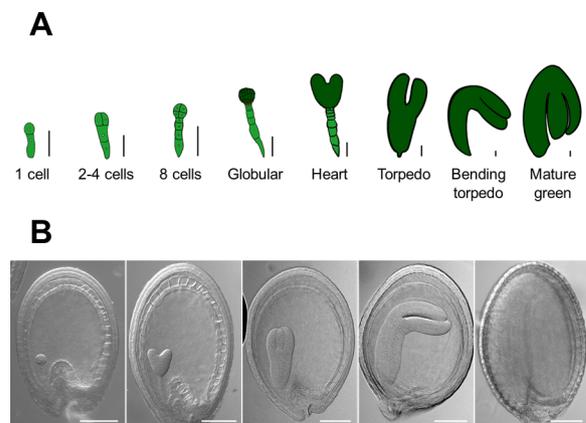


Fig. 1. The developmental stages of *Arabidopsis* embryos. (A) The developmental stages of *A. thaliana* embryos. Vertical scale bars = 50 μm . (B) Differential interference contrast images of embryos within seeds used in this study. From left to right: globular, heart, torpedo, bending torpedo and mature green stage. Scale bars = 100 μm .

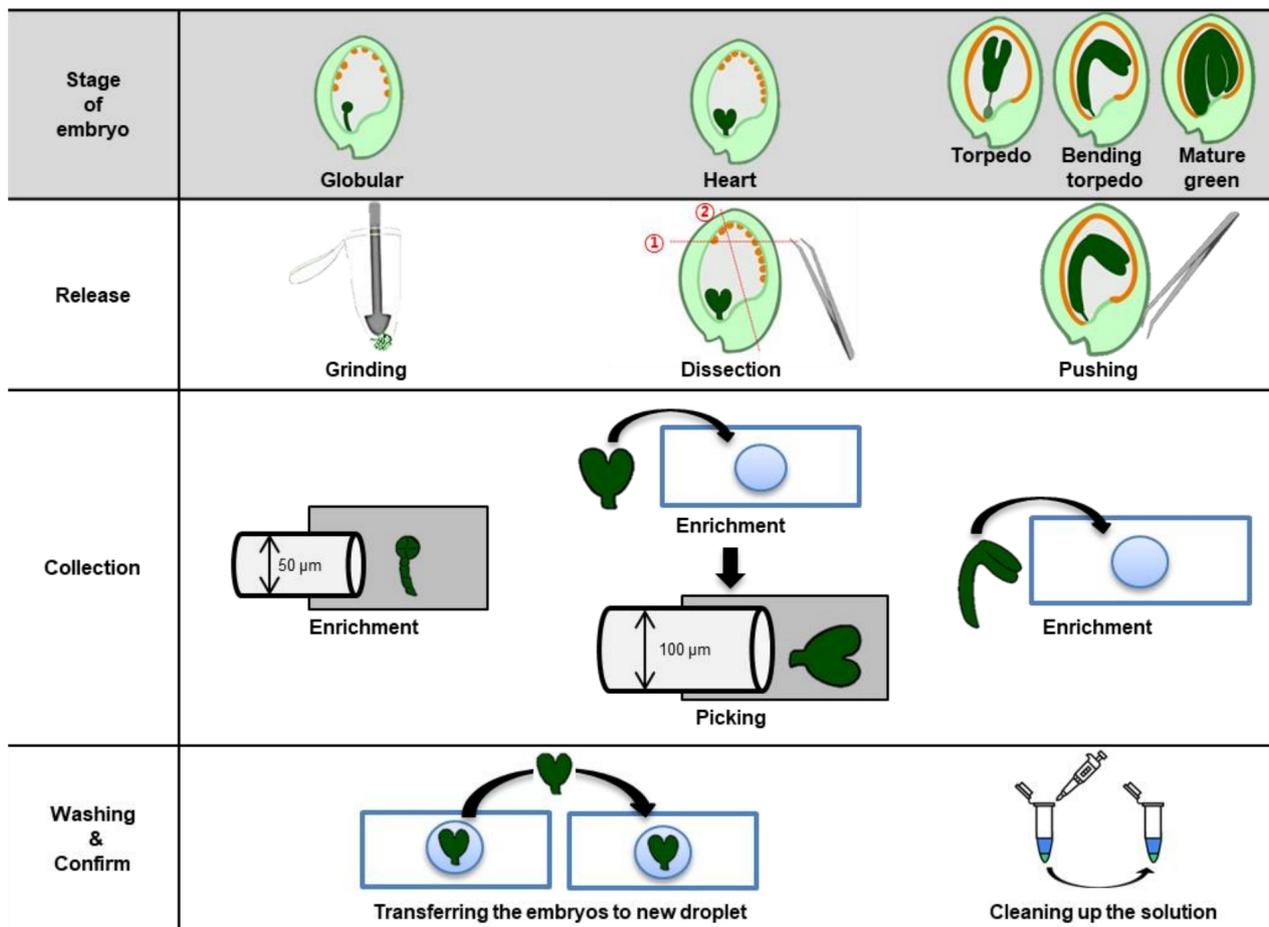


Fig. 2. A summary of sampling methods according to developmental stage. Three steps (release, collection, and washing and confirm) were optimized for each embryo stage. The heart stage embryo was dissected for sampling first horizontally then vertically. The diameters of S shaped glass micropipettes used were 50 μm and 100 μm for globular and heart stage embryos, respectively. The washing step was conducted repeatedly until embryos are purified.

bryos to be released from the seeds (Fig. 2).

Preparation of heart stage embryos

Siliques at DAP5 were dissected in 50 μl isolation buffer for heart stage embryo preparation. The dissection method was modified from a previous study (Xiang et al., 2011), with incisions made horizontally at the upper part of the seed, and then vertical cuts were made at the chalazal side, to avoid embryo damage (Fig. 2). Endosperm cells were dissected out to expose the embryo. Since heart stage embryos appear transparent, reflected light from the dissection microscope was used to visualize embryos, which were released using tweezers. Released embryos were finally moved to a new droplet of isolation buffer using a pipet for enrichment.

Collection of the globular and heart stage embryo

Samples in 10-20 μl isolation buffer were transferred onto glass slides for isolation of embryos using a manual micro-pipetting device (MPP-200B Micro Pick and Place Manipulator System B without Microscope; Nepa Gene, Japan) under an inverted fluorescent imager (ZOE fluorescent cell imager,



Fig. 3. The isolation of released embryo using micro-glass pipette. Isolation of a late heart to early torpedo embryo (orange arrows) using a micro-pipetting device under an inverted fluorescent imager.

1450031; Bio-Rad, USA) (Fig. 3). For heart stage embryos, gentle tapping of the microcapillary was used to position the axis of the embryos towards the microcapillary, allowing access to the embryo for suctioning (Fig. 2). Isolated globular and heart stage embryos were subsequently washed; five or six droplets of isolation buffer (~10 μl each) were placed on the glass slide and embryos were moved between these droplets (Fig. 2). This washing step was repeated until all debris, accumulated during dissection process, was removed.

For manual isolation, the lid of a flat cap PCR tube was cut (Cat. No.137-211C; WATSON, Japan), as shown in [Supplementary Fig. S1](#). Pure embryos were transferred onto this cap in 5-10 μ l of isolation buffer. Sample purity was confirmed under the microscope, before the PCR tube was reassembled with the lid containing the purified embryos. The PCR tubes were centrifuged for 30 s to 1 min and stored at -80°C . Different diameter sizes of micro glass pipette were used according to the embryo type and size: 50 μ m for globular stage embryo and 100 μ m for heart stage embryo ([Fig. 2](#)) (1-GT50S-6, 50 μ m; 1-GT100S-6, 100 μ m; Nepa Gene).

Preparation and isolation of torpedo to mature green stage embryos

Sample preparation was conducted using DAP7, DAP9, and DAP12 seeds for isolation of torpedo, bending torpedo and mature green stage embryo, respectively. At these later stages, embryos are more easily detectable due to their greener color and larger size compared with globular and heart stage embryos. On slide glass, seeds were stabbed and pushed using tweezers to release the embryos, leaving seed coat and endosperm clumps behind. Next, the released embryos were collected in 20 μ l of isolation buffer. A sufficient number of embryos were transferred to 1.5 ml tubes and washed in an isolation buffer. After several washes, purified embryos were transferred to a new 1.5 ml tube and stored at -80°C .

Cell wall degradation test

Samples were thawed at room temperature and incubated at 95°C for 30 s in a water bath, before re-freezing in liquid nitrogen. This freeze-thaw cycle was repeated five times. Samples were centrifuged briefly and vortexed for 1 min for globular embryos or 5 min for the heart embryos. For torpedo embryos, vortex time was extended and sample degradation was assessed visually. After degradation, M-digestion buffer (2 \times , D5021-9; Zymo Research, USA) and 1 μ l of proteinase K (D3001-2-5; Zymo Research) were added to the PCR tube lids (up to 20 μ l total) and the samples were centrifuged and vortexed briefly. Samples were incubated at 50°C for 30 min, and vortexed as described above for each embryo stage. After a further brief centrifugation, samples were ready for CT conversion, which is the first step in the construction of bisulfite sequencing library.

DNA purification from bending torpedo and mature green embryos with pre-RNase A treatment

Samples were ground with a pestle in 50 μ l of TE, to which 1 μ l of DNase-free RNase A (20-40 mg/ml, R4642-50MG; Merck, Germany) was added before incubation at 37°C for 30 min with the pestle. DNA preparation from pre-RNase A-treated sample was conducted as described previously ([Allen et al., 2006](#)). DNA was precipitated after overnight incubation of isopropanol at -20°C for maximal yield.

Construction of bisulfite sequencing libraries using the snmC-seq2 method

EZ DNA Methylation-DirectTM kit (D5020; Zymo Research) was used for CT conversion. CT conversion reagent (130 μ l) was added to 20 μ l of purified DNA from bending torpedo to ma-

ture green embryos or freeze-thawed raw samples of globular-torpedo stage embryo (see above). CT conversion was performed according to the manufacturer's instructions with two exceptions: to remove washing solution completely, an additional washing step followed by centrifugation (13,000g) was added; to maximize elution, samples were incubated for 5 min after the addition of M-Elution buffer. Sequencing library construction was conducted by the snmC-seq2 method ([Luo et al., 2018](#)), with two modifications: the temperature was gradually increased (8°C for 4 min then 16°C to 37°C , ramping rate 0.1°C per second, and finally 37°C for 30 min) instead of increasing it abruptly (4°C for 5 min and 25°C for 5 min, then 37°C for 60 min) during random-primed DNA synthesis; and all clean-up procedures were conducted using SPRI beads (Sera-Mag SpeedBeads Magnetic Carboxylate Modified; Merck) after each enzyme reaction. SPRI beads (0.8 \times) were added to the samples, suspended and incubated on a DynaMagTM-PCR Magnet (492025; Invitrogen, USA) for 5 min until the solution was clear (approximately 5 min), before being washed with 150 μ l of fresh 80% EtOH. Washing was repeated twice more and beads were dried after removal of the solution. DNA was eluted from beads by vortexing and incubation for 5 min in EB buffer (19086; Qiagen, Germany). Finally, the supernatant was transferred to a new tube.

Construction of bisulfite sequencing libraries with the Zymo Pico kit

The Pico Methyl-SeqTM Library Prep Kit (D5456; Zymo Research) was used, following the manufacturer's instructions. Lightning Conversion Reagent (130 μ l) was added to 20 μ l of purified DNA from bending torpedo to mature green embryos (the product of DNA purification with pre-RNase A treatment) or to freeze-thaw treated raw samples of globular-torpedo stage embryos (see above). The remainder of the procedure was performed according to the manufacturer's instructions with two exceptions: to remove the washing solution completely, an additional wash step followed by centrifugation (13,000g) was added; and to maximize elution, samples were incubated for 5 min after the addition of M-Elution buffer. To avoid excessive primer dimerization, half of the amount of LibraryAmp primers proposed by the manufacturer was used.

WGBS data processing and the complexity of methylome libraries

All Sequencing procedures were performed using the HiSeqXten platform (Macrogen, Korea). Paired end reads (150 bp) were generated. All reads were trimmed (10 bp for 5'end and 5 bp for 3'end) using Trim galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), and low quality and short reads (<70 bp) were removed using Trimmomatic (<http://www.usadellab.org/cms/?page=trimmomatic>). Reads were mapped to the *Arabidopsis* TAIR10 genome by hisat2 (<http://daehwankimlab.github.io/hisat2/>) using Bismark (<https://www.bioinformatics.babraham.ac.uk/projects/bismark/>) under the option `-hisat2 -local`. PCR duplicates were removed and methylation levels were extracted using the Bismark toolset (`deduplicate_bismark` and `bismark_methylation_extractor`, respectively). Read counts for all cytosine methylation contexts (CpG, CHG, and CHH) were calcu-

lated by how many reads were mapped on each context (1 bp resolution). Methylation levels were basically calculated by dividing the counts of methylated cytosine by the number of cytosine (meC + C). Except for genomic features, we used 50 bp windowed average having more than each three cytosine context with at least five reads.

RESULTS

Early globular- to torpedo-stage embryos can be used as inputs for the construction of methylomes without DNA purification

During embryogenesis, progressive changes to pattern formation become established. Embryonic patterns in early globular to heart embryos in particular, change more rapidly than late-stage embryos (Laux et al., 2004). However, studies into the epigenetic states in *Arabidopsis* have mainly concentrated on late stage embryos or seeds (Bouyer et al., 2017; Kawakatsu et al., 2017), mainly because of the relative ease with which methylome libraries in these stages are generated. However, to understand the comprehensive DNA methylation roadmap during embryogenesis, the development of methods tailored to each developmental stage is required. While methods for the construction of methylome libraries from single mammalian cells have been developed (Clark et al., 2017; Smallwood et al., 2014), it is difficult to apply these methods to *Arabidopsis* without extensive optimization due to the relatively small genome size of this model organism. Moreover, early embryos are technically far more difficult to isolate than late-stage embryos as they are transparent, consisting of a small number of cells that are buried deep within seeds. Therefore, we first optimized the isolation method of each stage of embryos to avoid contamination and provide methylome libraries with 5-20× coverage.

First, we applied seed dissection to release early embryos.

However, due to the small size of embryo and the strong connection between suspensor and seed coat, globular embryos were difficult to release. We, therefore, adopted a previously described seed grinding method (Raissig et al., 2013), which allowed us to collect a number of globular embryos sufficient for downstream processing. This method similarly allowed for the release of earlier-stage embryos from before DAP4 (Fig. 4A). However, embryos without intact suspensors were also released (Fig. 4A).

Larger heart stage embryos can be seen under a dissection microscope. These were dissected from seeds in an isolation buffer using a protocol modified from previously described (Raissig et al., 2013), as illustrated in Fig. 2. However, this method produced debris, which resulted in contamination. Therefore, the released embryos were purified and enriched in a fresh isolation buffer after dissection. Embryos in the globular stage were taken out of the seeds with gentle pestle grinding, without breaking the embryo (detailed in the Materials and Methods section). Pure globular and heart embryos were enriched using different sized microcapillaries (Fig. 2).

For the construction of bisulfite library, DNAs extracted from the tissues or cells are treated with chemicals for the CT conversion. However, we found that DNA extraction starting from a small number of cells and tissues failed to obtain sufficient amount of DNA. Accordingly, subsequent libraries were not generated. Therefore, we decided to use the raw embryos rather than extracted DNAs from those as an input for the following CT conversion. To do so, the intact cell wall and nuclear membrane require degradation for the DNA to be exposed. Therefore, we examined freeze-thaw cycles—freezing in liquid nitrogen and heating to 95°C to enable embryo degradation. Cell wall degradation started from the 3rd cycle in both globular and heart embryos (Fig. 4B). Due to their larger size and complex structure, the cell walls of torpedo embryos were not degraded by this simple freeze-thaw method. We, therefore,

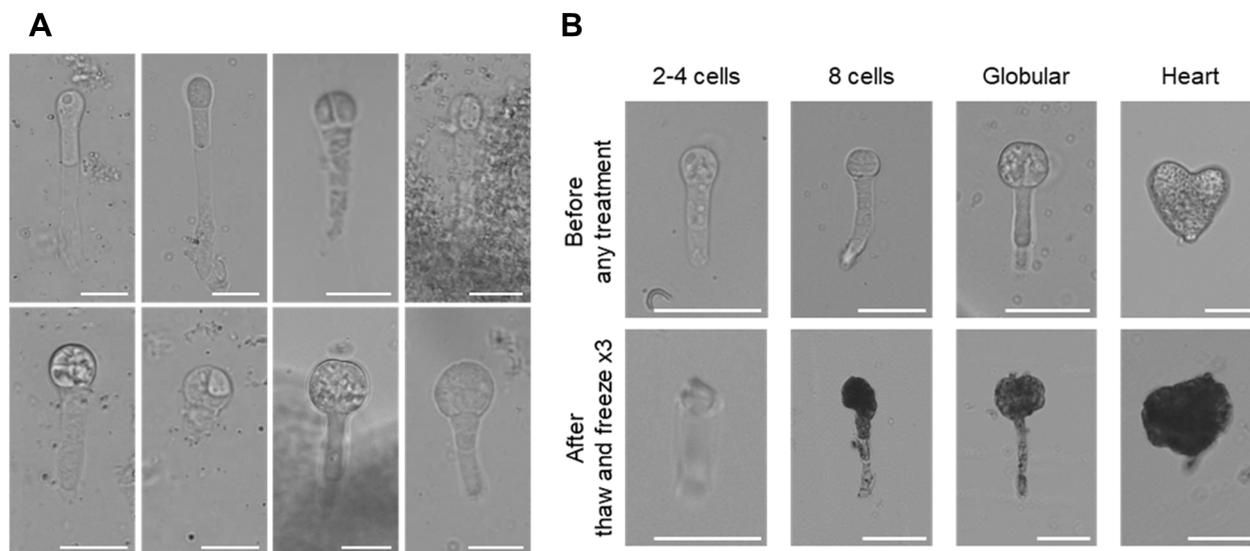


Fig. 4. Released embryos from intact seeds. (A) The release of globular embryos from seeds by gentle grinding. Embryos consisting of 1-4 cells are shown in the upper panel. Eight-cell to early globular stage embryos are shown in the lower panel. Scale bars = 25 μm . (B) Before and after 3 cycles of freeze-thaw without vortexing. Cell wall degradation is observed in the lower panel. Scale bars = 50 μm .

examined the effectiveness of two further treatments (vortexing for 30 min with or without freeze-thaw). The combined vortex and freeze-thaw treatments greatly improved torpedo embryo degradation (Table 1).

Pestle grinding and DNA purification with RNase A pre-treatment in bending-torpedo stage embryos

The freeze-thaw cycle method described above cannot be applied to embryos from the bending-torpedo stage. These embryos, which have a greater cell number than globular-

to torpedo-stage embryos with rigid cell walls, were not degraded by freeze-thawing. Therefore, DNA purification is required for embryos in this stage. We first optimized the previously described DNA purification method (Allen et al., 2006), initially examining whether these relatively small embryos (roughly 0.5 mm in diameter) compared to other plant tissues such as leaves could be ground with a pestle (Table 2). Under a microscope, we saw consistent degradation of embryos from bending torpedo stage to mature green stage, but not torpedo embryos (Table 2). A general DNA purification method (Allen et al., 2006) with 10 bending torpedo embryos from DAP 9 did not yield detectable DNA. Interestingly, omitting RNase A treatment, a step shown to result in DNA loss due to the DNA binding activity of RNase A (Dona and Houseley, 2014), we were able to obtain approximately 200 pg of DNA from a single bending torpedo embryo (Table 2). As we were concerned that omitting RNase A treatment completely could lead to contamination, we next examined whether pre- and post-treatment with RNase A affected DNA yield. Pre-treatment with RNase A yielded sufficient DNA without RNA contamination from 8-10 bending torpedo embryos or even from a single mature embryo for library construction (Table 3). Post-treatment with RNase A starting from the same amounts of embryos failed to do so (Table 3). Our method was further used for the purification of DNA from small tissue samples, such as a single seedling sample

Table 1. Verification of cell wall degradation

Treatment	1st trial	2nd trial	3rd trial
No treatment	10/9	12/10	12/11
Vortex only (30 min)	10/4	11/5	12/3
Thaw and freeze 3 cycles + vortex (30 min)	10/-	12/-	12/-

Values are presented as input sample amount/output sample amount.

Changes detected after freeze-thaw and vortexing of embryos. Numbers indicate the number of intact torpedo stage embryos detected and were counted under a microscope. Shrunken embryos were counted as degraded embryos.

Table 2. Analysis of pestle grinding as a method of extracting DNA from purified embryos

Method	Sample input	Pestle grinding	DNA (pg/ μ l)	Total (ng)
CTAB without RNase treatment	1 Torpedo	X	-	-
	1 Torpedo	X	-	-
	1 Torpedo	O	-	-
	1 Torpedo	O	-	-
	1 Bending torpedo	O	11	0.22
	3 mg leaf	O	7,150	143
	8 mg leaf	O	6,200	124
	Non control	-	-	-

DNA can be extracted from single bending torpedo-stage embryos by pestle grinding followed by CTAB DNA extraction without RNase treatment. Embryo degradation by pestle grinding was examined under a microscope. CTAB, cetyltrimethylammonium bromide.

Table 3. Advantage of DNA purification with pre-RNase A treatment for small amounts of sample

Method	Sample input	RNase A treatment	Total DNA (ng)	RNA (pg/ μ l) ^a
CTAB	8 Bending torpedo	None	7.3	2,100
	10 Bending torpedo	Before elution	-	-
	6 Bending torpedo	Before elution	-	-
	8 Bending torpedo	Post-extraction	-	-
	8 Bending torpedo	Pre-extraction	1.8	-
	10 Bending torpedo	Pre-extraction	7.1	-
	1 Mature green	Pre-extraction	5.4	-
	1 Mature green	Pre-extraction	5.8	-

RNase A pre-treatment assists DNA purification from few bending torpedo embryos and single mature embryos, without RNA contamination.

CTAB, cetyltrimethylammonium bromide.

^aRNA concentrations of less than 20 pg/ μ l cannot be detected.

at 13 days after germination or hundreds of protoplasts. This modified protocol for the purification of DNA from small amounts of input sample would allow the examination of genome heterogeneity and the epigenome of individual seedlings originating from the same mother in a near future.

Two complementary methods for the construction of DNA methylation libraries

The use of post-bisulfite adaptor tagging and whole genome amplification in the construction of single-cell methylomes has greatly facilitated epigenome studies (Clark et al., 2017; Miura and Ito, 2015; Smallwood et al., 2014). Using this method, the heterogeneity and stochasticity of DNA methylation has been examined in single cells (Parry et al., 2021; Smallwood et al., 2014). However, the small genome size and difficulties in culturing plants has made the construction of single-cell methylomes challenging. To overcome these difficulties, we adapted a method of mammalian single cell methylome construction for use in *Arabidopsis*. We optimized two complementary methods for the generation of bisulfite sequencing libraries that use post-bisulfite adaptor tagging and whole genome amplification; the Pico-kit method and the snmC-seq2 method (Luo et al., 2018). Methylome construction with Pico-kit produced fewer over-represented read sequences than the snmC-seq2 method. This may be due to the unique adaptor reaction step in the snmC-seq2 method (Table 4), which reduces the generation of primer dimers in a library by tagging with a second adaptor without primers, but generates over-represented sequences of oligo A at the beginning of reverse reads. Since this adaptation step can

Table 4. Analysis of two complementary methylome construction methods

Method	Over-represented sequences	Description
Pico-kit	0/25 libraries	
snmC-seq2	9/31 libraries	Adaptase reaction step generates over-represented oligo A sequences

The Pico-kit method resulted in fewer over-represented sequences than the snmC-seq2 method. Numbers indicate the number of libraries containing over-represented sequences among all libraries made by each method. Over-represented sequences were confirmed using the FastQC program.

Table 5. Data quality analysis

Sample	Globular	Heart	Torpedo	Bending torpedo	Mature green
No. of libraries	7 (3P + 4S)	3 (2P + 1S)	6 (2P + 4S)	2 (2P)	5 (4P + 1S)
Coverage (average)	6×	15×	10×	22×	14×
Total coverage (×)	41×	44×	61×	44×	70×
CT conversion rate (%) (average)	99.2	99.5	99.5	99.1	99.4

Information about bisulfite sequencing data collected from the five stages of the embryo development. CT conversion rates were calculated using the un-methylated plastid genome.

The minimum number of samples for making one library used in this study was five for globular embryos.

P, Pico-kit method; S, snmC-seq2 method.

extend up to 50 base pair, these synthetic sequences caused a failure in the mapping of reverse reads. Libraries produced by snmC-seq2 showed more complexity than those produced by the Pico-kit method, with snmC-seq2 libraries covering more regions of the genome than Pico-kit libraries. Peri-centromeric regions, in particular, were mapped more frequently using snmC-seq2 (Fig. 5). These data suggest that the use of a combination of both the Pico-kit and snmC-seq2 library construction methods can generate a reliable and enriched dataset.

Analysis of data quality and application of the method

Data quality was analyzed with reference to the epigenome consortiums guide (Roadmap Epigenomics; <http://www.roadmapepigenomics.org/>) (Roadmap Epigenomics Consortium et al., 2015), with reads providing no less than 30-fold coverage with at least two replicates per sample. This criterion was achieved with embryos from all stages of samples with a robust CT conversion ratio (Table 5). These data suggest that good quality libraries could be constructed from small amounts of sample with our modified method and pipeline. For example, as few as five globular embryos were able to generate a reliable library (Table 5). The detailed statistic information such as the number of input embryos, total number of sequenced reads, % mapping efficiency, % deduplication, number of leftover reads, coverage, CT con-

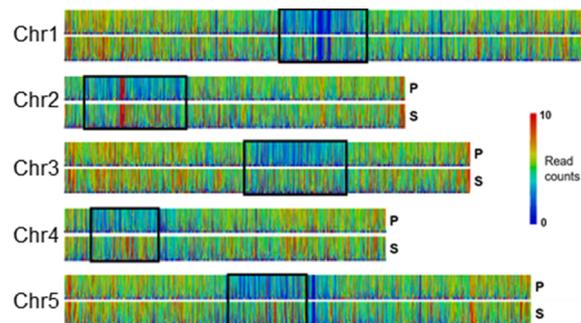


Fig. 5. Comparison of the snmC-seq2 and Pico-kit methods of library construction. Comparison of library complexity produced by two different bisulfite sequencing methods. The snmC-seq2 (each lower panel) method has better coverage (more read counts), particularly at peri-centromeric regions (black squares), than the Pico-kit method (upper panels). P, Pico-kit method; S, snmC-seq2 method.

version ratio, % CG, CHG, and CHH of each library is in the [Supplementary Dataset 1](#).

To further verify the quality of libraries, we compared methylation levels of our libraries with the previously published libraries (Bouyer et al., 2017; Papareddy et al., 2020). We chose heart stage and mature green stage libraries because these two libraries from other groups were generated from the embryos with the most similar, although not exactly the same, developmental stages to our libraries. CG, CHG and CHH methylation of the whole genome and genic regions show almost the same levels (Fig. 6, Global and Gene), indicating that our libraries recapitulated the previously published results, and supporting that our methods using small amounts of input sample are reliable. CG and CHG methylation in TE show similar levels, but interestingly, CHH methylation in our libraries shows slightly reduced levels than the previous published ones (Fig. 6, TE). We assume that this might be due to the slightly different developmental stages of the samples used to generate each library. It has been known that CHH methylation levels are being increased as seeds mature (Kawakatsu et al., 2017; Lin et al., 2017). Therefore, our samples for the heart and mature green embryos are likely in slightly earlier stages than the embryos from other groups. Accordingly, we observed consistent increase of methylation levels in TE regions of our libraries during embryogenesis (Supplementary Fig. S2). Our libraries also represent the same methylation patterns with the previous published libraries in individual genes, such as *FWA* and *AT3G17400* (Supplementary Fig. S3).

We also compared the similarity among the libraries generated from the same stage. In general, library replicates from one specific stage are in the similar positions in the CHH PCA plot (Supplementary Fig. S4), indicating that each library in the same stage generates similar methylation result. The library complexity plot, where the x-axis is the number of sequenced reads and y-axis is the number of uniquely mapped and deduplicated reads, showed generally proportional in our libraries (Supplementary Fig. S5). All these analyses support the idea that our methods using smaller amounts of input than the previous methods can generate comparable and decent libraries.

Overall, our procedure uses sampling methods tailored to each stage of embryos (Figs. 2 and 3), with cell wall degra-

dition methods without DNA purification for earlier stage embryos (Fig. 4B), and DNA extraction methods with pre-RNase A treatment for later stage embryos, which largely reduces the amount of input sample required compared with common DNA extraction methods (Allen et al., 2006). These methods and pipelines are generally applicable to other tissues from plants, or other species when sample size is limited (Fig. 7).

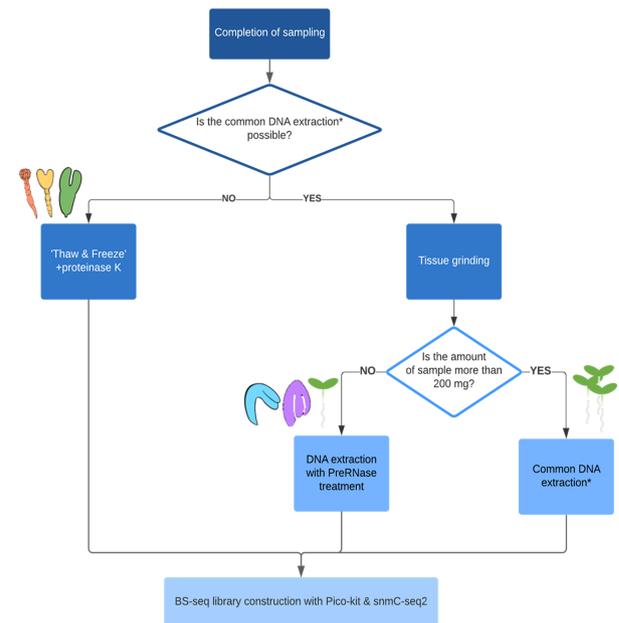


Fig. 7. Strategies of bisulfite sequencing (BS-seq) library construction. Input tissues are degraded by freeze-thaw with proteinase K incubation, or by grinding with a pestle, according to the size and the shape of each sample. If the weight of the sample is under 200 mg, we perform DNA extraction with pre-RNase A treatment. The cetyltrimethylammonium bromide (CTAB) method (Allen et al., 2006) is generally used for genomic DNA extraction (identified with an asterisk).

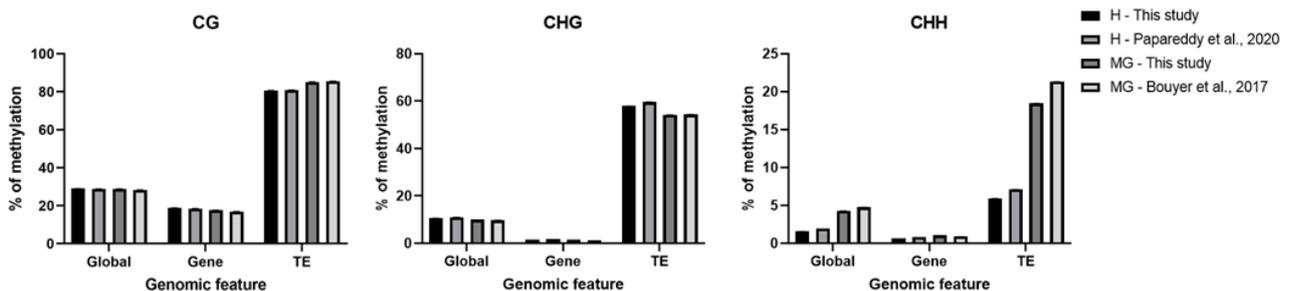


Fig. 6. The comparison of our libraries to the previously published data. In order to show the data reproducibility, our embryo libraries of the heart and mature green stages were compared to the libraries from Papareddy et al. (2020) for the heart stage and from the Bouyer et al. (2017) for the mature green stage. CG, CHG, and CHH methylation levels are shown in global, genic and TE regions.

DISCUSSION

Global gene expression profiles and epigenetic data, including DNA methylation, have been studied extensively with the development of whole genome and single-cell methods (Chatterjee et al., 2012; Clark et al., 2017; Karemaker and Vermeulen, 2018; Krueger et al., 2012; Li et al., 2011; Rich-Griffin et al., 2020; Smallwood et al., 2014; Stuart and Satija, 2019). Although these technical advances have been of great benefit, cell- or tissue-type specific epigenome profiling in *Arabidopsis* has remained challenging, mainly because it has a smaller genome size than mammals. In addition, the structural complexity and variation of each tissue type has made data interpretation difficult, especially if the input sample is a mixture of cell types. For example, seeds comprise three parts: seed coat, embryo, and endosperm. Accordingly, DNA methylation patterns differ greatly between the tissue types, with endosperm more hypomethylated than the embryos (Hsieh et al., 2009; Ibarra et al., 2012). Therefore, without dissection of the seeds, unique and distinct tissue-specific profiles could be missed. Therefore, there is a demand for single cell- and tissue-specific epigenomic profiling from a limited number of cells, and the concomitant optimization of protocols.

Here, we introduced an optimized pipeline for the construction of embryo methylomes, to aid the investigation of cell- and tissue-type specific genetic and epigenetic changes during *Arabidopsis* embryogenesis. We have developed effective methods for embryo release from seeds in a developmental stage-dependent manner, followed by manual embryo picking with micro glass pipettes (Figs. 2 and 3). The methods were optimized for single-cell DNA methylation library construction. We successfully generated libraries from early-stage embryos, using as few as five globular sage embryos as a raw input sample along with cell wall degradation step (Fig. 4, Table 5). We found that this is a very effective way to generate decent libraries when the input samples are limited. Our method enables to avoid DNA extraction step which normally requires enough amounts of sample. We also found that for bending torpedo embryos, a DNA purification step is essential and better for library construction than using raw samples (Table 2).

We further found that pre-RNase A treatment enables the purification of sufficient DNA from as few as eight bending torpedo embryos or a single mature embryo, to construct more than one methylome (Table 3). Based on the information of unexpected DNA binding activity of RNase A (Dona and Houseley, 2014), we have tested and verified DNA binding activity of RNase A and the consequent DNA loss (Tables 2 and 3). Indeed, we could purify DNA from only a single torpedo embryo or 100 leaf cells without RNase treatment. However, the minimum amount of input tissue was increasing to 8 folds with RNase A treatment. Nevertheless, since the remaining RNAs after DNA extraction might inhibit following experimental procedures, we had to optimize and develop the method with RNase A treatment to minimize the amount of input sample. We assumed that RNase A treatment before exposure of naked DNA by simultaneous RNase A treatment and nuclei lysis might help to reduce the chance of unexpect-

ed DNA binding activity of RNase A than the conventional DNA extraction methods which conducts RNase A treatment after DNA concentration. We have tested this (Tables 2 and 3) and developed an enhanced method by rearranging the timing of RNase A treatment during DNA extraction procedure. This method made it possible to reduce the necessary amount of plant tissue for DNA extraction without RNA contamination (Tables 2 and 3). Discovering other RNase enzymes with better specificity for RNA degradation would help to reduce the amount of input. Or simply, we may develop the more sensitive DNA extraction methods from small amounts of input by testing the minimal amount of RNase A enzyme to degrade RNA molecules with minimizing DNA loss in detail.

To compare methylome data from different methods we used, evaluation of the data quality is required. The comparison of library quality obtained by cell wall degradation methods (globular, heart, torpedo) and pre-RNase A DNA extraction methods (bending torpedo, mature green), and the conventional DNA extraction methods for methylome construction can be evaluated by the CT conversion ratio and mapping ratio (Table 5, Supplementary Dataset 1). CT conversion of all of our data is over 99%. This means that all of our data is sufficient to interpret the methylation values as they are. In addition, except for the fifth methylome of DAP4 globular embryos, the mapping ratio of our methylome data is 37%-62% (Supplementary Dataset 1). This means that our libraries show not only reliable quality of methylome data but also better than referenced single-cell bisulfite sequencing libraries from mammals. Furthermore, consistent with the previous methylome data from *Arabidopsis* (Kawakatsu et al., 2017; Lin et al., 2017), our methylome recapitulated those data in that global CG methylation levels are well maintained whereas methylation levels are increased during embryo development (Fig. 6, Supplementary Fig. S2). Individual gene loci also showed similar methylation patterns in data from ours and other groups (Supplementary Fig. S3). Our methylome data that show very similar global CG methylation levels to the previously published data in heart and mature green stage are from the methods of cell wall degradation and pre-RNase DNA extraction, respectively. Therefore, we can compare all of our data in parallel.

Furthermore, we optimized two complementary methods of methylome construction (Pico-kit and snmC-seq2 method), with heterochromatic regions mapped by snmC-seq2 libraries, while Pico-kit libraries produced fewer over-represented read sequences (Fig. 5, Table 4). We recommend using both methods in tandem for each target sample.

Our methods make the construction of DNA methylation libraries accessible by further reducing the current minimum amount of input material, to our knowledge at least in plants, while retaining data quality. We tailored the pipeline for each stage of embryo development (Fig. 7). The methods presented here have the advantage that each pipeline protocol can be applied to other tissues from plants, or even to other species, since our method provides information on what to consider for each stage, from sampling to library construction.

Seed plants have comprehensive and versatile DNA methylation systems and patterns of methylation are tightly regu-

lated. Therefore, the development of methods to study these patterns to aid our understanding of genome heterogeneity and the methylome are of considerable importance.

Note: Supplementary information is available on the Molecules and Cells website (www.molcells.org).

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (2020R1A2C2009382) to Y.C., J.L., and S.L. were supported by the Stadelmann-Lee Scholarship Fund, Seoul National University.

AUTHOR CONTRIBUTIONS

H.Y., K.P., J.L., and Y.C. conceived and designed the pipeline. H.Y. and J.L. performed embryo isolation experiments. K.P. performed methylome construction. K.P. and S.L. performed methylome library optimization. K.P. and J.L. processed and analyzed the data. H.Y., K.P., J.L., S.L., and Y.C. wrote the manuscript. Y.C. supervised the project.

CONFLICT OF INTEREST

The authors have no potential conflicts of interest to disclose.

ORCID

Hyunjin Yoo <https://orcid.org/0000-0002-5847-9232>
Kyunghyuk Park <https://orcid.org/0000-0002-3141-5087>
Jaehoon Lee <https://orcid.org/0000-0001-6087-4410>
Seunga Lee <https://orcid.org/0000-0001-8488-439X>
Yeonhee Choi <https://orcid.org/0000-0002-2796-5262>

REFERENCES

Allen, G.C., Flores-Vergara, M.A., Krasynanski, S., Kumar, S., and Thompson, W.F. (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat. Protoc.* *1*, 2320-2325.

Bell, C.G., Lowe, R., Adams, P.D., Baccarelli, A.A., Beck, S., Bell, J.T., Christensen, B.C., Gladyshev, V.N., Heijmans, B.T., Horvath, S., et al. (2019). DNA methylation aging clocks: challenges and recommendations. *Genome Biol.* *20*, 249.

Bouyer, D., Kramdi, A., Kassam, M., Heese, M., Schnittger, A., Roudier, F., and Colot, V. (2017). DNA methylation dynamics during early plant life. *Genome Biol.* *18*, 179.

Chatterjee, A., Stockwell, P.A., Rodger, E.J., and Morison, I.M. (2012). Comparison of alignment software for genome-wide bisulphite sequence data. *Nucleic Acids Res.* *40*, e79.

Clark, S.J., Smallwood, S.A., Lee, H.J., Krueger, F., Reik, W., and Kelsey, G. (2017). Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat. Protoc.* *12*, 534-547.

Deal, R.B. and Henikoff, S. (2011). The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nat. Protoc.* *6*, 56-68.

Dona, F. and Houseley, J. (2014). Unexpected DNA loss mediated by the DNA binding activity of ribonuclease A. *PLoS One* *9*, e115008.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA

strands. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 1827-1831.

Hofmeister, B.T., Lee, K., Rohr, N.A., Hall, D.W., and Schmitz, R.J. (2017). Stable inheritance of DNA methylation allows creation of epigenotype maps and the study of epiallele inheritance patterns in the absence of genetic variation. *Genome Biol.* *18*, 155.

Hsieh, T.F., Ibarra, C.A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009). Genome-wide demethylation of *Arabidopsis* endosperm. *Science* *324*, 1451-1454.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N., Machlicova, A., Nishimura, T., et al. (2012). Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* *337*, 1360-1364.

Karemaker, I.D. and Vermeulen, M. (2018). Single-cell DNA methylation profiling: technologies and biological applications. *Trends Biotechnol.* *36*, 952-965.

Kawakatsu, T., Nery, J.R., Castanon, R., and Ecker, J.R. (2017). Dynamic DNA methylation reconfiguration during seed development and germination. *Genome Biol.* *18*, 171.

Kim, M. and Costello, J. (2017). DNA methylation: an epigenetic mark of cellular memory. *Exp. Mol. Med.* *49*, e322.

Kim, M.J., Lee, H.J., Choi, M.Y., Kang, S.S., Kim, Y.S., Shin, J.K., and Choi, W.S. (2021). UHRF1 induces methylation of the TXNIP promoter and down-regulates gene expression in cervical cancer. *Mol. Cells* *44*, 146-159.

Krueger, F., Kreck, B., Franke, A., and Andrews, S.R. (2012). DNA methylome analysis using short bisulfite sequencing data. *Nat. Methods* *9*, 145-151.

Laux, T., Wurschum, T., and Breuninger, H. (2004). Genetic regulation of embryonic pattern formation. *Plant Cell* *16* Suppl, S190-S202.

Levenson, V.V. (2010). DNA methylation as a universal biomarker. *Expert Rev. Mol. Diagn.* *10*, 481-488.

Li, W., Liu, H., Cheng, Z.J., Su, Y.H., Han, H.N., Zhang, Y., and Zhang, X.S. (2011). DNA methylation and histone modifications regulate de novo shoot regeneration in *Arabidopsis* by modulating WUSCHEL expression and auxin signaling. *PLoS Genet.* *7*, e1002243.

Lin, J.Y., Le, B.H., Chen, M., Henry, K.F., Hur, J., Hsieh, T.F., Chen, P.Y., Pelletier, J.M., Pellegrini, M., Fischer, R.L., et al. (2017). Similarity between soybean and *Arabidopsis* seed methylomes and loss of non-CG methylation does not affect seed development. *Proc. Natl. Acad. Sci. U. S. A.* *114*, E9730-E9739.

Locke, W.J., Guanzone, D., Ma, C., Liew, Y.J., Duesing, K.R., Fung, K.Y.C., and Ross, J.P. (2019). DNA methylation cancer biomarkers: translation to the clinic. *Front. Genet.* *10*, 1150.

Luo, C., Rivkin, A., Zhou, J., Sandoval, J.P., Kurihara, L., Lucero, J., Castanon, R., Nery, J.R., Pinto-Duarte, A., Bui, B., et al. (2018). Robust single-cell DNA methylome profiling with snmC-seq2. *Nat. Commun.* *9*, 3824.

Miura, F. and Ito, T. (2015). Highly sensitive targeted methylome sequencing by post-bisulfite adaptor tagging. *DNA Res.* *22*, 13-18.

Papareddy, R.K., Paldi, K., Paulraj, S., Kao, P., Lutzmayer, S., and Nodine, M.D. (2020). Chromatin regulates expression of small RNAs to help maintain transposon methylome homeostasis in *Arabidopsis*. *Genome Biol.* *21*, 251.

Park, K., Frost, J.M., Adair, A.J., Kim, D.M., Yun, H., Brooks, J.S., Fischer, R.L., and Choi, Y. (2016). Optimized methods for the isolation of *Arabidopsis* female central cells and their nuclei. *Mol. Cells* *39*, 768-775.

Parry, A., Rulands, S., and Reik, W. (2021). Active turnover of DNA methylation during cell fate decisions. *Nat. Rev. Genet.* *22*, 59-66.

Picard, C.L. and Gehring, M. (2017). Proximal methylation features associated with nonrandom changes in gene body methylation. *Genome Biol.* *18*, 73.

Raissig, M.T., Gagliardini, V., Jaenisch, J., Grossniklaus, U., and Baroux,

- C. (2013). Efficient and rapid isolation of early-stage embryos from *Arabidopsis thaliana* seeds. *J. Vis. Exp.* (76), 50371.
- Rich-Griffin, C., Stechemesser, A., Finch, J., Lucas, E., Ott, S., and Schafer, P. (2020). Single-cell transcriptomics: a high-resolution avenue for plant functional genomics. *Trends Plant Sci.* 25, 186-197.
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317-330.
- Salas, L.A., Wiencke, J.K., Koestler, D.C., Zhang, Z., Christensen, B.C., and Kelsey, K.T. (2018). Tracing human stem cell lineage during development using DNA methylation. *Genome Res.* 28, 1285-1295.
- Smallwood, S.A., Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., Andrews, S.R., Stegle, O., Reik, W., and Kelsey, G. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* 11, 817-820.
- Stuart, T. and Satija, R. (2019). Integrative single-cell analysis. *Nat. Rev. Genet.* 20, 257-272.
- Xiang, D., Venglat, P., Tibiche, C., Yang, H., Risseuw, E., Cao, Y., Babic, V., Cloutier, M., Keller, W., Wang, E., et al. (2011). Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in *Arabidopsis*. *Plant Physiol.* 156, 346-356.
- Yu, B., Dong, X., Gravina, S., Kartal, O., Schimmel, T., Cohen, J., Tortoriello, D., Zody, R., Hawkins, R.D., and Vijg, J. (2017). Genome-wide, single-cell DNA methylomics reveals increased non-CpG methylation during human oocyte maturation. *Stem Cell Rep.* 9, 397-407.
- Zeng, Y. and Chen, T. (2019). DNA methylation reprogramming during mammalian development. *Genes (Basel)* 10, 257.