PORTLAND PRESS

Review Article

Decoding the sorghum methylome: understanding epigenetic contributions to agronomic traits

Ulduz Vafadarshamasbi[1], Emma Mace[2,3], David Jordan[2] and Peter A. Crisp[1]

[1]School of Agriculture and Food Sciences, The University of Queensland, Brisbane, QLD 4072, Australia; [2]Centre for Crop Science, Queensland Alliance for Agriculture and Food Innovation, Hermitage Research Facility, Warwick, QLD 4370, Australia; [3]Agri-Science Queensland, Department of Agriculture and Fisheries (DAF), Hermitage Research Facility, Warwick, QLD 4370, Australia

Correspondence: Peter A. Crisp (p.crisp@uq.edu.au)

OPEN ACCESS

DNA methylation is a chromatin modification that plays an essential role in regulating gene expression and genome stability and it is typically associated with gene silencing and heterochromatin. Owing to its heritability, alterations in the patterns of DNA methylation have the potential to provide for epigenetic inheritance of traits. Contemporary epigenomic technologies provide information beyond sequence variation and could supply alternative sources of trait variation for improvement in crops such as sorghum. Yet, compared with other species such as maize and rice, the sorghum DNA methylome is far less well understood. The distribution of CG, CHG, and CHH methylation in the genome is different compared with other species. CG and CHG methylation levels peak around centromeric segments in the sorghum genome and are far more depleted in the gene dense chromosome arms. The genes regulating DNA methylation in sorghum are also yet to be functionally characterised; better understanding of their identity and functional analysis of DNA methylation machinery mutants in diverse genotypes will be important to better characterise the sorghum methylome. Here, we catalogue homologous genes encoding methylation regulatory enzymes in sorghum based on genes in *Arabidopsis*, maize, and rice. Discovering variation in the methylome may uncover epialleles that provide extra information to explain trait variation and has the potential to be applied in epigenome-wide association studies or genomic prediction. DNA methylation can also improve genome annotations and discover regulatory elements underlying traits. Thus, improving our knowledge of the sorghum methylome can enhance our understanding of the molecular basis of traits and may be useful to improve sorghum performance.

Introduction

In higher plants, cytosine methylation at the 5′ carbon in genomic DNA (5mC) is a conserved chromatin modification that plays a pivotal role in the regulation of gene expression, transgene silencing, genomic stability, imprinting and inactivation of transposable elements (TEs) [1–4]. DNA methylation can be inherited through meiotic and mitotic divisions in plant cells [5] and is thus an epigenetic factor due to its heritable role in genome maintenance [6] and gene regulation [7,8]. DNA methylation states are mostly stably inherited in a mendelian fashion. However, methylation variants can cause variability in phenotypic behaviour due to spontaneous methylation changes that cause unexpected non-mendelian inheritance, as is the case in paramutation [9]. Variation in the patterns of DNA methylation may produce novel epialleles that may contribute to plant improvement and adaptation [10]. Profiles of DNA methylation can also be applied to refine genome annotations or identify putative *cis*-regulatory elements, as has been demonstrated in *Sorghum bicolor* (sorghum) [11,12].

Sorghum is an important cereal due to its adaptation to biotic and abiotic stress [13–15], and its usage in food, alcoholic beverages, animal feed and in biofuel production [16]. Its close relationship to the other major crops such as maize [17], sugarcane, and pearl millet [18] also makes sorghum an

attractive focus for cereal genome research. Compared with *Arabidopsis* or even to other grasses like maize and rice, in sorghum there is more limited information about the regulation and natural variation for patterns of DNA methylation. In this review, we highlight what is known about the sorghum methylome, including similarities and differences to its close relative maize. Better elucidating methylation signatures in the sorghum genome may also be useful for sorghum improvement.

Detection and distribution of DNA methylation in the sorghum genome

Detection of DNA methylation

In plants, DNA methylation occurs in three sequence forms including CG, CHG and CHH, where H can be A, C, or T [19]. CG and non-CG contexts both have significant potential to be methylated [20]. Methods for DNA methylation detection include those based on bisulfite conversion, enzymatic digests that use combinations of methylation sensitive and insensitive restriction enzymes, and affinity-based immunoprecipitation methods [21–24]. Whole-genome bisulfite DNA sequencing or 'WGBS' is the gold-standard approach, which generates a base-pair resolution DNA methylation map [25]. WGBS based on chemical conversion can damage DNA and results in DNA degradation, so an enzymatic-based alternative was developed that uses 10–11 translation (TETs) dioxygenases to oxidise methylated cytosines (5-mCs), and convert unmethylated regions to uracils by catalytic polypeptide-like 3A (APOBEC3A) treatment [26]. Owing to the high cost of whole-genome profiling in large genomes, numerous cheaper alternatives have been developed, including sequence-capture approaches, array-based approaches and reduced-representation strategies analogous to genotype-by-sequence methods for genotyping [27,28]. However, these methods have significant limitations and progress in this field would be accelerated by the development of more economical profiling methods.

Distribution of DNA methylation

Plant species show extensive variation in both levels and patterns of 5′ cytosine methylation [29] (Figure 1A). Previous studies have profiled DNA methylation in sorghum [12,29–31], maize [6,32,33], rice [34] and many other species [35,36]. By profiling DNA methylation levels genome-wide, regions in the genome can be grouped into different domains based on the types and levels of DNA methylation, including regions with high CHH, CG, and CHG ('High CHH'), regions with high CG and CHG only ('Heterochromatin'), high CG only ('CG-only' or gene body methylation), intermediate, and regions with low methylation in all contexts ('Unmethylated'). For maize, sorghum, and rice the relative distribution of methylation domains vary; CG and CHG methylation occur in heterochromatic regions and are enriched at TEs and intergenic regions but depleted in genes and comprise 79% of the large genome of maize and 47% in sorghum, and only 26% in the smaller genome of rice (Figure 1A,B). These species vary in genome size, 389 Mb, 730 Mb, and >2 Gb in rice, sorghum, and maize, respectively (Figure 1A). Rice maintains ∼10% CHH methylation, the highest of these species, which is largely due to the RNA directed DNA methylation (RdDM) pathway [8]. The maize genome is methylated least at CHH sites with <1%. Rice has a higher rate of unmethylated regions at 43% compared with sorghum (27%) and maize (8%) (Figure 1B); however, in terms of absolute base pairs, the unmethylated space in each genome is remarkably similar, ∼125–132 Mb (Figure 1A).

Comparing sorghum to maize, rice and *Arabidopsis*, in each genome, gene density is highest in the chromosome arms and lower towards the pericentromeric regions, although relative gene density is much, much higher in the chromosome arms for sorghum compared with maize, rising to over 100 genes per megabase (Figure 2A), and higher again for Arabidopsis, with over 300 gene per megabase. For sorghum rice and *Arabidopsis*, the relative density of TEs and repeats displays an inverse pattern to gene density (Figure 2B). In contrast, TE density in maize is the highest of all species and is high throughout the chromosome. Likewise, the average DNA methylation levels along a chromosome varies between species (Figure 2C). For all, CG has the highest levels of methylation, followed by CHG, and CHH is the least methylated form. For all species, CHH has a low-level distribution pattern across the chromosome but is highest in rice, and also high in the pericentromeric regions of *Arabidopsis*. For maize, with its high abundance of TEs, CG and CHG are maintained at high levels right across the chromosome, while the proportion of CG- and CHG-methylated sites peaks around centromeric segments in sorghum, rice and *Arabidopsis* and are far more depleted in the gene dense chromosome arms (Figure 2C).
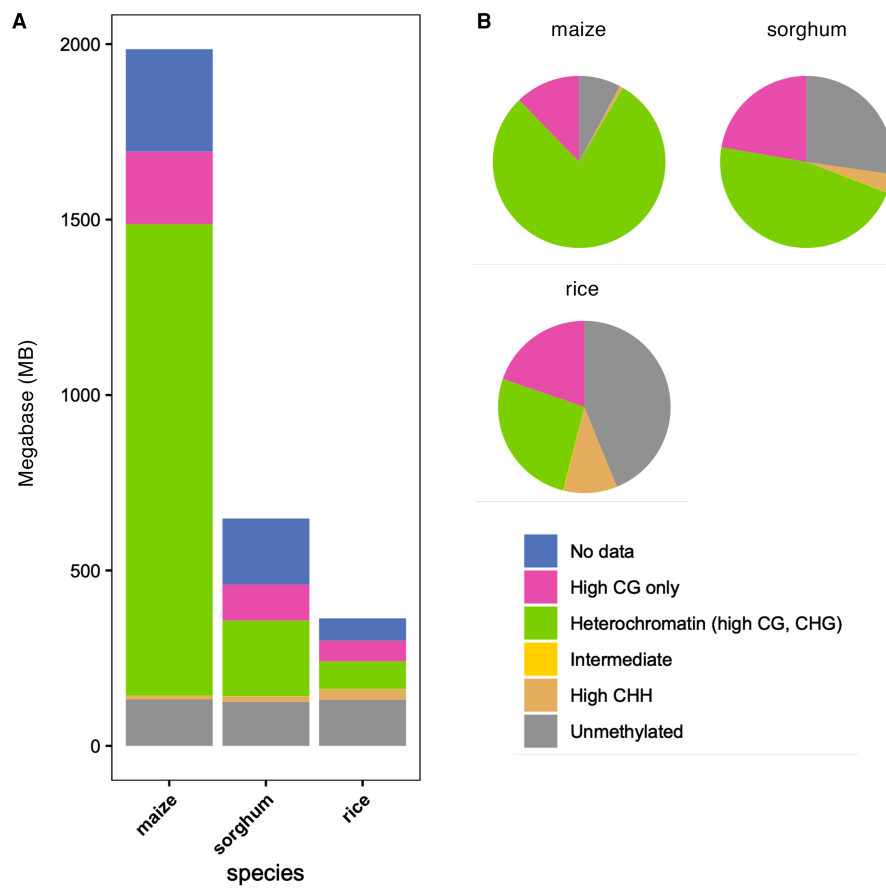
Figure 1. Comparison of DNA methylation domains between genomes.
(**A**) The total base pairs of each methylation domain in the maize, sorghum, rice genomes. Data sourced from Crisp et al. [11];
(**B**) the relative abundance of different methylation domains in the maize, sorghum and rice genomes, expressed as a percent
of each genome that can be analysed (excluding unmappable regions and regions with no data).

Molecular mechanisms regulating DNA methylation

DNA methylation in plant genomes occurs with the addition of a methyl group (CH_3) to cytosine residues,
and is regulated by multiple complementary pathways and enzymes depending on the sequence context [1,37].
The components of the methylation pathways have been extensively studied in the model plant *Arabidopsis*
[38,39] and also in maize [40] and rice [41,42]. In *Arabidopsis*, these components have been knocked out indi-
vidually and in combination to reveal many important insights into methylation in each sequence context [39].
However, in larger crop genomes, mutants in many components have not been obtained despite attempts, sug-
gesting these genes are indispensable [40], which has hampered functional analysis. In rice, CRISPR/Cas dis-
ruption of each methyltransferases individually led to viable plants except for MET1 loss of function [42]. Most
homologous genes encoding methylation regulatory enzymes in maize have been identified [43]. Here, we
review the sorghum gene homologues based on genes identified in *Arabidopsis*, maize, and rice; however, func-
tional studies in sorghum are an important area for future research.

Initiating and maintaining methylation

Cytosine-5′ DNA methyltransferases (C5′-MTases) catalyse the addition of methyl groups with a complemen-
tary role for 'de novo' and 'maintenance' in regulating methylation [29]. C5′-MTases consist of three main
enzyme groups: (1) DOMAINS REARRANGED METHYLTRANSFERASES (DRMs) that catalyse *de novo*
methylation in all three methylation contexts through RNA-directed DNA methylation (RdDM) [44]; (2)
METHYLTRANSFERASES (METs) that are responsible for the maintenance of CG methylation at each round
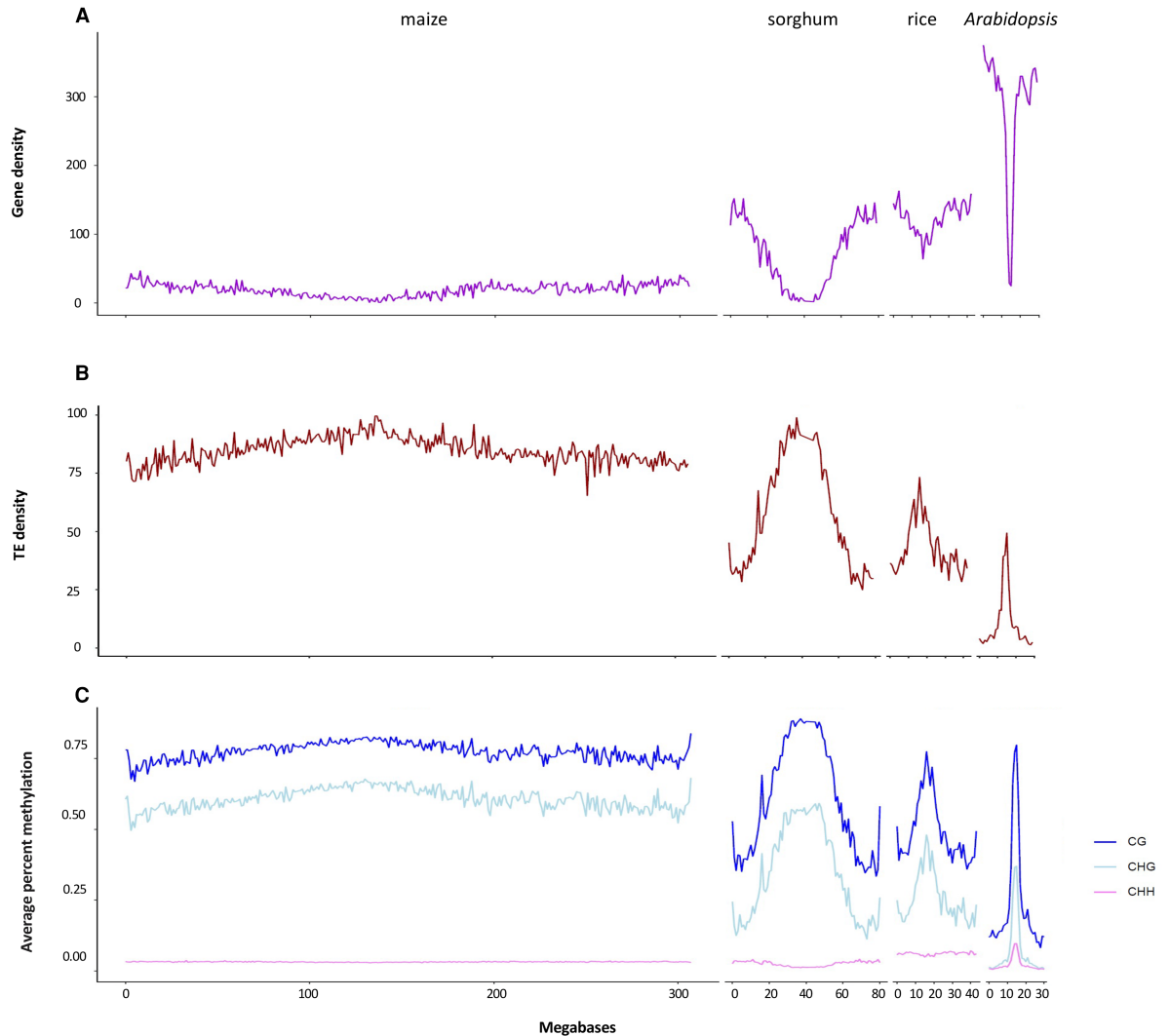
Figure 2. Distribution of genes, TEs and DNA methylation across maize, sorghum, rice, and *Arabidopsis* chromosome 1.
(**A**) Gene density expressed as the total number of genes per megabase across chromosome 1 in each species; (**B**) distribution of TEs across chromosome 1 in each species expressed as percent of base pairs annotated as TE; (**C**) the average level of CG, CHG and CHH methylation per megabase. Methylation levels were determined per 100 bp tile and then averaged into megabase tiles across chromosomes for each species. Each chromosome is set to the same scale to illustrate the different size of maize (B73), sorghum (Btx623), rice (Nipponbare), and *Arabidopsis* (Col-0) chromosomes. Gene annotations sourced from Phytozome13, methylation data for maize, sorghum and rice sourced from Crisp et al. [11] and *Arabidopsis* from Crisp et al. [129]; TE annotations generated using EDTA [130] and the maize TE annotation generated using EDTA sourced from Hufford et al. [131].

of DNA replication [45]; and (3) CHROMOMETHYLASES (CMTs) that catalyse the maintenance of CHG and CHH methylation [46,47].

De novo DNA methylation is guided by small interfering RNAs (siRNAs) and scaffold RNAs via the RdDM pathway [19,38]. The siRNA precursors and scaffold RNAs are transcribed by two plant specific RNA polymerase II (Pol II)-related enzymes including Pol IV and Pol V [48,49]. In the 'canonical' RdDM pathway, Pol IV produces a single-stranded RNA (ssRNA). The ssRNA is reverse-transcribed by RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) to produce a double stranded RNA (dsRNA), which is finally cleaved to 24 nucleotide (nt) siRNAs [41]. During canonical RdDM, Pol IV and Pol V are recruited to target loci by other heterochromatic marks and co-factors, for example SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) that binds

histone H3 lysine 9 methylation (H3K9me2) [50,51]. The initiation of RdDM has been more enigmatic; however, recent findings suggest that the self-reenforcing cycle of RdDM is initiated by an AGO4 clade ARGONAUTE guided by an siRNA to a locus independent of pre-existing DNA methylation [52].

In plants, the RdDM pathway recruits DRMs for maintenance of CHH methylation [38,53]. In rice, OsDRM2 plays a major role in CHH methylation [8]. The siRNAs binding to ARGONAUTE 4 (AGO4) or AGO6 and scaffold RNAs activate the enzyme DOMAINS REARRANGED METHYLTRANSFERASES 2 (DRM2) for DNA methylation at the target loci [53,54]. CMT2 in *Arabidopsis* also plays a role in the maintenance of CHH methylation [46]; while maize lacks a CMT2 homologue, sorghum does have a homologue, SbCMT2 (Sobic.009G083900). In the maize genome, DRM-like genes ZMET3 and ZMET7 are homologous to *Arabidopsis* DRM1 and DRM2, and ZMET6 that is related to DRM3 [40]. Homologues of DRM1/DRM2, and DRM3 in sorghum are SbDRM2 (Sobic.003G124000) and SbDRM3 (Sobic.009G032200) (Table 1).

Following DNA replication, the methylation patterns for CG, CHG, and CHH are maintained by different mechanisms [38]. DNA METHYLTRANSFERASE 1 (MET1) is responsible for maintenance of CG sites in *Arabidopsis* [38]. The rice genome encodes two redundant MET1 genes, OsMET1-1 and OsMET1-2 that maintain methylation at CG sites, with higher levels of OsMET1-2 expression [55]. The maize genome encodes two tandemly duplicated MET1-like genes that are both in the monocot MET1b clade [56], including ZmMET1b-1 (Zm00001eb301630) and ZmMET1b-2 (Zm00001eb301620) [43]. In sorghum, there is a homologue of rice OsMET1a, SbMET1a (Sobic.001G055700) and a homologue of the maize ZmMET1b duplicates, SbMET1b (Sobic.002G056000) (Table 1).

The methylation of CHG sites is maintained by chromomethylases (CMTs) in a reinforcing loop and requires the activity of a chromodomain, which recognises specific histone modifications [41]. In this loop, histone H3 lysine 9 di-methylation (H3K9me2) is catalysed by the H3K9 methyltransferases KRYPTONITE (KYP)/SUPPRESSOR OF VARIEGATION 3–9 HOMOLOGUE 4/5/6 (SUVH4/5/6) at genomic regions [57,58]. *Arabidopsis* has three CMT genes: CMT1, CMT2 and CMT3 [4], although CHG methylation relies on the chromomethylases CMT3 rather than CMT2 [47,57] and CMT1 function is unknown. There are in-paralogs of CMT3 in maize ZMET2 and ZMET5 that catalyse the maintenance of CHG sites [40,59], that have a close relationship with CMTs in other monocots but the maize genes are considered a separate subclade and named ZMETs . Based on the genealogical relationship between CMTs and ZMETs, it is suggested that they have been derived from an ancestral whole-genome duplication among angiosperms [60]. Maize ZMET2 seems to contribute to the maintenance of CHG as a main functional homologue with CMT3, while its paralog ZMET5 is expressed to lesser degree [40,59]. ZMETs in sorghum also arose from two independent duplications, which might have been derived from Poaceae-specific duplication events [61]. However, there are differences between genes and amino acids of ZMETs in sorghum compared with maize [61]; the sorghum homologues are SbCMT3a (Sobic.004G197400) and SbCMT3b (Sobic.006G214000) (Table 1). In rice, three CMT genes including OsCMT2, OsCMT3a, and OsCMT3b have been identified; however OsCMT3a takes charge of maintaining CHG sites as the main functional homologue of CMT3 that contributes in the expression of genes and TEs, while OsCMT3b seems to contribute at minor level [42,62]. The chromomethylase CMT2 recognises H3K9me2 within deep heterochromatin enriched with TEs to methylate CHH [63]. Rice, unlike maize, does have a functional CMT2 homologue, OsCMT2 [62], which is also homologous to sorghum SbCMT2 (Sobic.009G083900) (Table 1).

The maintenance of DNA methylation is also regulated by chromatin remodellers [41]. DECREASE IN DNA METHYLATION1 (DDM1) is a nucleosome remodeller, and is required to maintain DNA methylation in *Arabidopsis* especially at TEs [46]. The two homologues of DDM1 in maize (ZmDDM1a and ZmDDM1b), are highly expressed with similar patterns including in the embryo [40,64] and play a critical role for CHG methylation, and to a lesser extent for CGs in heterochromatic regions. DDM1 also contributes to methylation of CHH islands, located in euchromatic regions in interaction with ZmAGO4 [65,66]. In sorghum, DDM homologues are SbDDM1a (Sobic.002G021200) and SbDDM1b (Sobic.001G109700) (Table 1).

Demethylation

In addition to *de novo* methylation and maintenance, demethylation can play a critical role in plants, for example, controlling fruit ripening, in gametophytes, and in plant aging and senescence [67,68]. In plants, demethylation events occur through the base excision repair pathway, which itself is an essential defense pathway for the genome and performs the replacement of 5-methylcytosine with cytosine in an active demethylation process [69]. Passive demethylation also contributes to demethylation, where 5-methylcytosine is diluted

Table 1 Methylation and Demethylation genes in sorghum

Gene		Sorghum gene	Sorghum identifier	Homologues		
				Maize	Rice	Arabidopsis [75]
Methylation	MET1b	SbMET1b	Sobic.002G056000	ZmMET1b-1 (Zm00001eb301630) and ZmMET1b-2 (Zm00001eb301620) [43]	OsMET1-2 (LOC_Os07g08500) [55]	AtMET1 (AT5G49160)
	MET1a	SbMET1a	Sobic.001G055700		OsMET1-1 (LOC_Os03g58400) [55]	AtMET1 (AT5G49160)
	CMT3/ Zmet	SbCMT3a	Sobic.004G197400	ZMET5 (Zm00001d002330 v4 *missing in maize v5) [40]	OsCMT3a (LOC_Os10 g01570) [41,62]	AtCMT3 (AT1G69770)
	CMT3/ Zmet	SbCMT3b	Sobic.006G214000	ZMET5 (Zm00001d002330) and ZMET2 (Zm00001d026291) [40]	OsCMT3b (LOC_Os03g12570) [41]	AtCMT3 (AT1G69770)
	DDM1a	SbDDM1a	Sobic.002G021200	CHR101 (ZmDDM1A; Zm00001eb117870)	OsDDM1a (LOC_Os09g27060) [8]	AtDDM1 (AT5G66750)
	DDM1b	SbDDM1b	Sobic.001G109700	CHR101 ZmDDM1A; Zm00001eb117870) or CHR106 (ZmDDM1B; Zm00001eb055640) [40,64]	OsDDM1b (LOC_Os03g51230) [8]	AtDDM1 (AT5G66750)
	CMT2	SbCMT2	Sobic.009G083900		OsCMT2 (LOC_Os05g13780) [46]	AtCMT2 (AT4G19020)
	DRM1/ DRM2	SbDRM2	Sobic.003G124000	ZMET3 (Dmt103; Zm00001eb404130), ZMET7 (Dmt107; Zm00001eb000820) [40]	OsDRM1a (LOC_Os11g01810); OsDRM1b (LOC_Os12g01800); OsDRM2 (LOC_Os03g02010) [41]	AtDRM1 (AT5G15380)
	DRM3	SbDRM3	Sobic.009G032200	ZMET6 (Dmt106; Zm00001eb354730) [40]	OsDRM3 (LOC_Os05g04330) [42,76]	AtDRM1 (AT5G15380)
Demethylation	ROS1	SbROS1c	Sobic.008G085300	DNG102 (Zm00001eb241240) [43,74]	OsROS1a (LOC_Os01g11900) OsROS1b-d (LOC_Os05g37410, LOC_Os05g37350, LOC_Os02g29230) [73,77]	AtROS1 (AT2G36490); AtDML2 (AT3G10010)
	ROS1	SbROS1b	Sobic.004G149800	DNG101 (Zm00001eb202980) [43,74]	As above	As above
	ROS1	SbROS1a	Sobic.009G155900	DNG103 (Zm00001eb289030) [43,74]	As above	As above
	DML3	SbDML3	Sobic.006G224100	DNG105 ZmDML3 (Zm00001eb241310) [73]	OsDML3a-b (LOC_Os04g28860; LOC_Os02g29380) [41]	AtDML3 (AT4G34060)

To identify putative genes in sorghum associated with methylation and demethylation, the homologous genes previously annotated for Arabidopsis, maize and rice genomes were identified through literature searches (as cited) and each pairwise aligned to the sorghum genome using the BLAST tool (NCBI).

from the genome during DNA replication [4]. In *Arabidopsis*, four genes have been identified as 5-methylcytosine DNA glycosylases (DNGs) belonging to three distinct orthology groups, the REPRESSOR OF SILENCING 1 (ROS1), DEMETER (DME) and two DEMETER-Likes, DML2 and DML3 [70]. ROS1 plays an

essential role in deactivation of methylated cytosines and inhibits transcriptional silencing of both transgenic and endogenes loci in the *Arabidopsis* genome [71]. It is also reported that there is a dynamic counteraction between the RdDM and DNA demethylation for ROS1 expression, so that RdDM targets methylation in the promoter of ROS1, but in another way, prevents ROS1 hypermethylation by inducing ROS1 transcription that leads to active DNA demethylation [71,72]. Monocots lack DME orthologs but do have ROS1 and DML orthologs [73]. Two homologues in the maize genome for ROS1; DNG101 and DNG103 have been reported [43]. Sorghum genes of the DNG family include SbROS1a, SbROS1b, and SbROS1c that are homologous to DNG103 (Zm00001eb289030), DNG101 (Zm00001eb202980), and DNG102 (Zm00001eb241240) in maize [43,74] (Table 1). The sorghum genome also encodes a homologue to ZmDML3 [73], SbDML3 (Sobic.006G224100), that is a DME-like gene, (Table 1). However, the functional activity of these sorghum demethylation pathway genes needs to be experimentally explored.

Opportunities and challenges for crop improvement

Variation in DNA can underpin phenotypic traits of agronomic relevance [7,78,79]. In this section, we highlight areas that require a better understanding in sorghum in order to leverage the methylome for crop improvement, and discuss avenues for implementation. It is not yet clear if some traits might be more amenable to improvement via epigenetic approaches; however, it will be valuable to investigate this for key targets for agronomic improvement. Example key traits include, yield and yield components such as grain size and number; abiotic stress tolerance, including cold tolerance at emergence, heat tolerance during flowering and drought tolerance; and grain quality and digestibility.

Heritability

Patterns of symmetric DNA methylation can be faithfully maintained and inherited through cell divisions and even CHH methylation can be propagated via the RdDM pathway. This means that patterns of methylation are highly stable for much of the life cycle of plants, especially during vegetative development. That said, tissue specific differences in methylation are reported, including in sorghum [30,80]. Epigenetic reprogramming causes changes in DNA methylation states during gametogenesis [81], and in turn affects reproductive lineages. Notably, the agronomically valuable cereal endosperm is one tissue where DNA methylation levels are reduced compared with other tissues, which has been observed in rice [82], maize [83] and sorghum [80].

The heritability of DNA methylation states can be defined by the three classes of epiallele: obligatory, facilitated, and pure [84]. Obligatory epialleles are a direct result of *cis-* and *trans-*acting sequence variation. As an example, an insertion of a transposon can be associated with epigenetic silencing of neighbouring coding sequences. Facilitated epialleles are also linked to genetic variation but can act independently; an example is where a transposon causes methylation in a nearby gene but then is segregated out [5,84]. Facilitated epigenetic variation could occur due to either local or cis-acting genetic variants or trans-acting variants [85,86]. Lastly, pure epialleles are entirely independent of genetic variation, and are generated by the activity (or failure) of the DNA methylation/demethylation pathways [84]. Changes in methylation for epialleles have a potential to persist and pass onto the next generations as epigenetic alleles independent of the underlying DNA sequence, and could influence phenotypic variation in traits making them of interest to applied breeding programs [87]. Importantly, only obligatory and cis-acting facilitated epialleles might be already captured by genetic marker-based (e.g. SNP) selections schemes, while epialleles generated by trans-acting variants and pure epialleles are likely to escape genetic marker-based selection [86].

For a particular genomic region, methylation levels can be categorised as non-methylated, low, partial, or high-methylated, and the heritability of these states range from high stability to entire instability [86,88]. Epigenetic variation derived from trans-acting genetic factors and genetic variants may cause a paramutation, which is an interaction between alleles for a specific region/locus that potentially alter mendelian inheritance and lead to a heritably altered gene expression state through mitosis [89]. Epigenetic alleles or 'epialleles' can drive changes in phenotype across generations in the absence of any DNA sequence changes (SNPs) [90]. Spontaneous or random errors in the maintenance of methylation lead to 'epimutation' and generate single methylation polymorphisms (SMPs) including CG-, CHG-, CHH-SMPs [91], or differentially methylated regions (DMRs) [92]. Spontaneous epimutation rates are estimated to be similar across species [93], with SMPs in the range 1×10^{-4} to 10^{-5} sites/haploid/generation, which is far higher than SNPs (1×10^{-8} to 10^{-9}) [86]. In contrast with SMPs, the level of methylation across larger regions is more stable (1×10^{-5} to 10^{-6} sites/haploid/generation) [94,95]. Even in the face of environmental stress, DNA methylation is highly stable over generations

[94]. DMRs among genotypes are very heritable, although they have lower stability in comparison with SNPs [86,92], which can be due to spontaneous backward epimutation in CG sites [96].

Natural variation

Comparisons of DNA methylation from different varieties or genotypes within a species have identified extensive natural variation. However, for most species and especially for crop species, profiling large populations is very limited. Extensive variation has been identified in *Arabidopsis* [91,97], the only species to have been profiled on a population scale. Consistent with findings in *Arabidopsis*, smaller-scale studies in maize, soybean, rice, wheat and brachypodium [86] point to extensive variation, including variants associated with domestication. Likewise, sequence capture profiling of ∼0.2% of the hexaploid wheat genome [98] identified diversity associated with agronomic traits among 104 varieties [99]. Profiling of 263 diverse maize varieties found many epialleles that were associated with gene expression changes and metabolic traits [6]. Importantly, around half of the trait associated regions with methylation polymorphisms identified in this study would not have been identified by profiling SNPs. The extent and spectrum of natural variation within sorghum is an exciting area for future research.

Inducing variation

Future efforts that seek to use epialleles for sorghum improvement could take advantage of several important sources of induced variation including chemical treatments, tissue culturing and genetic and epigenetic mutation. For example, new epialleles for the gene *pericarp color1* (*p1*) derived from tissue culture in maize, showed loss of pigmentation [100]. Also, induced epigenetic changes during tissue culture in rice were reported to be inherited in regenerated plants [101], and can also lead to activation of the Tos17 retrotransposon [102–105]. Tissue culture induced changes in DNA methylation have also been observed in sorghum using MSAP profiling [106]. Mutation in MET1 or DDM1 globally decreases CG methylation [7], which can be used to generate epigenetic recombinant inbred lines (epiRILs). epiRILs have been very useful in investigating the epigenetic basis of traits, although populations in large genome crop species are yet to be generated [5]. Chemicals such as zebularine can also be used to generate populations of epi-mutagenised plants, which has been applied to wheat to select for variation in flowering time and spike morphology [107]. Going forward, the use of epigenome editing will be a powerful alternative means to precisely generate or investigate epialleles [2]. Some studies have also found that the environment can induce changes in methylation. For example, a study on root system architecture found methylation variation associated with environmental stresses [108]; other studies have found that DNA methylome is stable under stresses [94].

Use of epigenetic variation in trait dissection

The most straightforward application of DNA methylation in breeding would be to use methylation marks for selection of epi-haplotypes associated with phenotypes of interest [109]. The impact of such an approach would require that epialleles were frequent and explained significant amounts of phenotypic variation and were stably transmitted [110,111]. Data from a range of crops suggests this may indeed be the case. Epigenetic variants have been shown to contribute to agronomic traits such as dwarf phenotypes in rice [112,113], drought and salt tolerance in rice [114], increase in protein and decrease in oil contents in *Brassica napus* [115]. *NMR19-4* was reported as a stable environmental-associated epiallele that negatively regulates the expression of a PHEOPHYTIN PHEOPHORBIDE HYDROLASE (PPH), which may result in variations in leaf senescence [116]. Thus, epigenome-wide association studies (EWAS) could complement genome association studies [117,118]. Several DMRs in *Arabidopsis* are epiQTLs that control flowering time and root length, and accounting for between 60–90% stable inheritance, which could be used as epi-markers for selection trials [119]. EWAS has been used in maize to identify genomic regions with differential methylation associated with traits that would be missed by SNP based GWAS alone [6].

Genetic mapping and GWAS studies play a crucial role in our ability to understand the genetic architecture of complex traits and identify markers for major genes that can be used in marker-assisted selection. Yet, sometimes these markers fall outside of genes. Chromatin profiling and the patterns of DNA methylation can be used to refine genome annotations and identify functional genomic elements [120]. In sorghum, DNA methylation profiles have been used to identify high confidence gene models from draft annotations [121]. Identifying the unmethylated portion of the genome is also a powerful method to identify putative *cis*-regulatory elements and expressed genes and has been successfully applied to sorghum [11]. This use of

methylation in this way can aid in better understanding of the sorghum genome and the molecular basis of traits including quantitative trait loci (QTLs), which can contribute to crop improvement molecular breeding strategies.

Use of epigenetic variation to enhance genomic selection

DNA methylation signatures also have potential to be incorporated in genomic selection schemes [122]. For example, in *Arabidopsis* DNA methylation data has been used to predict plant height [123]. Inclusion of epialleles in genomic selection models would require research to identify the best way to exploit this data and integrate it into genomic selection models, however, it will be very interesting to see if this can improve predictive power. In addition, prediction for performance of F1 hybrids remains challenging; genetic information is not sufficient for predicting heterosis perfectly, epigenome profiles may improve genome prediction due to epigenetic contributions to heterosis [124,125]. Thus, epigenetic modelling has the potential to fill current gaps and enhance our ability to predict phenotypic behaviour [126].

Use of epigenetic variation to modify meiotic recombination

Meiotic recombination has a critical influence on genetic gain in plant breeding programs. In some situations, high recombination can be favourable, such as when it is necessary to reduce the impact of linkage drag while in other situations lower recombination may be helpful in retaining favourable linkage blocks. Changes in methylation in both CG [127] and non-CG [128] contexts in pericentromeric heterochromatin in *Arabidopsis* has been shown to increase meiotic recombination. It is conceivable that if methylation can be manipulated in sorghum, that could be used to favourably change recombination rate to enhance progress in breeding programs, for example by enabling access to variation in regions of low recombination.

Perspectives

- DNA methylation is a chromatin modification that may provide an untapped layer of information on top of DNA sequence.

- This information is heritable, therefore, DNA methylation may provide a mechanism of epigenetic inheritance of trait variation that could be helpful in sorghum crop improvement. The sorghum DNA methylome and the genes regulating DNA methylation are understudied.

- It will be critical to determine the extent of variation for DNA methylation and the heritability of these epialleles in sorghum because this information could be very useful in selection, genomic prediction, pre-breeding and research, and complement current genetic approaches.

Abbreviations

AGO4, ARGONAUTE 4; CMTs, chromomethylases; DDM1, DECREASE IN DNA METHYLATION1; DME, DEMETER; DMRs, differentially methylated regions; DRM2, DOMAINS REARRANGED

METHYLTRANSFERASES 2; EWAS, epigenome-wide association studies; MET1, METHYLTRANSFERASE 1; SMPs, single methylation polymorphisms; TEs, transposable elements.

References

1   Stroud, H., Hale, C.J., Feng, S., Caro, E., Jacob, Y., Michaels, S.D. et al. (2012) DNA methyltransferases are required to induce heterochromatic re-replication in Arabidopsis. *PLoS Genet.* **8**, e1002808 https://doi.org/10.1371/journal.pgen.1002808

2   Gallego-Bartolomé, J. (2020) DNA methylation in plants: mechanisms and tools for targeted manipulation. *New Phytol.* **227**, 38–44 https://doi.org/10.1111/nph.16529

3   Ahn, J., Heo, S., Lee, J. and Bang, D. (2021) Introduction to single-cell DNA methylation profiling methods. *Biomolecules* **11**, 1013 https://doi.org/10.3390/biom11071013

4   Zhang, H., Lang, Z. and Zhu, J.-K. (2018) Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell. Biol.* **19**, 489–506 https://doi.org/10.1038/s41580-018-0016-z

5   Niederhuth, C.E. and Schmitz, R.J. (2014) Covering your bases: inheritance of DNA methylation in plant genomes. *Mol. Plant.* **7**, 472–480 https://doi.org/10.1093/mp/sst165

6   Xu, J., Chen, G., Hermanson, P.J., Xu, Q., Sun, C., Chen, W. et al. (2019) Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. *Genome Biol.* **20**, 1–16 https://doi.org/10.1186/s13059-018-1612-0

7   Springer, N.M. and Schmitz, R.J. (2017) Exploiting induced and natural epigenetic variation for crop improvement. *Nat. Rev. Genet.* **18**, 563–575 https://doi.org/10.1038/nrg.2017.45

8   Tan, F., Zhou, C., Zhou, Q., Zhou, S., Yang, W., Zhao, Y. et al. (2016) Analysis of chromatin regulators reveals specific features of rice DNA methylation pathways. *Plant Physiol.* **171**, 2041–2054 https://doi.org/10.1104/pp.16.00393

9   Weigel, D. and Colot, V. (2012) Epialleles in plant evolution. *Genome Biol.* **13**, 249 https://doi.org/10.1186/gb-2012-13-10-249

10 Varotto, S., Tani, E., Abraham, E., Krugman, T., Kapazoglou, A., Melzer, R. et al. (2020) Epigenetics: possible applications in climate-smart crop breeding. *J. Exp. Bot.* **71**, 5223–5236 https://doi.org/10.1093/jxb/eraa188

11 Crisp, P.A., Marand, A.P., Noshay, J.M., Zhou, P., Lu, Z., Schmitz, R.J. et al. (2020) Stable unmethylated DNA demarcates expressed genes and their cis-regulatory space in plant genomes. *Proc. Natl Acad. Sci. U.S.A.* **117**, 23991–24000 https://doi.org/10.1073/pnas.2010250117

12 Olson, A., Klein, R.R., Dugas, D.V., Lu, Z., Regulski, M., Klein, P.E. et al. (2014) Expanding and vetting sorghum bicolor gene annotations through transcriptome and methylome sequencing. *Plant Genome.* **7**, lantgenome2013.08.0025 https://doi.org/10.3835/plantgenome2013.08.0025

13 Mace, E.S., Rami, J.-F., Bouchet, S., Klein, P.E., Klein, R.R., Kilian, A. et al. (2009) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput diversity array technology (DArT) markers. *BMC Plant Biol.* **9**, 13 https://doi.org/10.1186/1471-2229-9-13

14 Hariprasanna, K. and Patil, J.V. (2015) Sorghum: Origin, classification, biology and improvement. In *Sorghum Molecular Breeding* (Madhusudhana, R., Rajendrakumar, P. and Patil J. eds.), pp. 3–20, Springer India, New Delhi

15 Ananda, G.K.S., Myrans, H., Norton, S.L., Gleadow, R., Furtado, A. and Henry, R.J. (2020) Wild sorghum as a promising resource for crop improvement. *Front. Plant Sci.* **11**, 1108 https://doi.org/10.3389/fpls.2020.01108

16 Gnansounou, E., Dauriat, A. and Wyman, C.E. (2005) Refining sweet sorghum to ethanol and sugar: economic trade-offs in the context of North China. *Bioresour. Technol.* **96**, 985–1002 https://doi.org/10.1016/j.biortech.2004.09.015

17 Swigonová, Z., Lai, J., Ma, J., Ramakrishna, W., Llaca, V., Bennetzen, J.L. et al. (2004) Close split of sorghum and maize genome progenitors. *Genome Res.* **14**, 1916–1923 https://doi.org/10.1101/gr.2332504

18 Kellogg, E.A. (1998) Relationships of cereal crops and other grasses. *Proc. Natl Acad. Sci. U.S.A.* **95**, 2005–2010 https://doi.org/10.1073/pnas.95.5.2005

19 Nie, W.-F. (2021) DNA methylation: from model plants to vegetable crops. *Biochem. Soc. Trans.* **49**, 1479–1487 https://doi.org/10.1042/BST20210353

20 Mahmood, A.M. and Dunwell, J.M. (2019) Evidence for novel epigenetic marks within plants. *AIMS Genet.* **6**, 70–87 https://doi.org/10.3934/genet.2019.4.70

21 Susan, J.C., Harrison, J., Paul, C.L. and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22**, 2990–2997 https://doi.org/10.1093/nar/22.15.2990

22 Bock, C., Tomazou, E.M., Brinkman, A.B., Müller, F., Simmer, F., Gu, H. et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat. Biotechnol.* **28**, 1106–1114 https://doi.org/10.1038/nbt.1681

23 Harris, R.A., Wang, T., Coarfa, C., Nagarajan, R.P., Hong, C., Downey, S.L. et al. (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat. Biotechnol.* **28**, 1097–1105 https://doi.org/10.1038/nbt.1682

24 Morrison, J., Koeman, J.M., Johnson, B.K., Foy, K.K., Beddows, I., Zhou, W. et al. (2021) Evaluation of whole-genome DNA methylation sequencing library preparation protocols. *Epigenetics Chromatin* **14**, 28 https://doi.org/10.1186/s13072-021-00401-y

25 Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D. et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**, 215–219 https://doi.org/10.1038/nature06745

26 Vaisvila, R., Ponnaluri, V.K.C., Sun, Z., Langhorst, B.W., Saleh, L., Guan, S. et al. (2021) Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. *Genome Res.* **31**, 1280–1289 https://doi.org/10.1101/gr.266551.120

27 Kurdyukov, S. and Bullock, M. (2016) DNA methylation analysis: choosing the right method. *Biology* **5**, 3 https://doi.org/10.3390/biology5010003

28 Paun, O., Verhoeven, K.J.F. and Richards, C.L. (2019) Opportunities and limitations of reduced representation bisulfite sequencing in plant ecological epigenomics. *New Phytol.* **221**, 738–742 https://doi.org/10.1111/nph.15388

29 Niederhuth, C.E., Bewick, A.J., Ji, L., Alabady, M.S., Kim, K.D., Li, Q. et al. (2016) Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* **17**, 194 https://doi.org/10.1186/s13059-016-1059-0

30 Turco, G.M., Kajala, K., Kunde-Ramamoorthy, G., Ngan, C.-Y., Olson, A., Deshphande, S. et al. (2017) DNA methylation and gene expression regulation associated with vascularization in sorghum bicolor. *New Phytol.* **214**, 1213–1229 https://doi.org/10.1111/nph.14448

31   Wang, H., Feng, Q., Zhang, M., Yang, C., Sha, W. and Liu, B. (2010) Alteration of DNA methylation level and pattern in sorghum (Sorghum bicolor L.) pure-lines and inter-line F1 hybrids following low-dose laser irradiation. *J. Photochem. Photobiol. B* **99**, 150–153 https://doi.org/10.1016/j.jphotobiol.2010.03.011

32   Xu, G., Lyu, J., Li, Q., Liu, H., Wang, D., Zhang, M. et al. (2020) Evolutionary and functional genomics of DNA methylation in maize domestication and improvement. *Nat. Commun.* **11**, 5539 https://doi.org/10.1038/s41467-020-19333-4

33   Qian, Y., Hu, W., Liao, J., Zhang, J. and Ren, Q. (2019) The dynamics of DNA methylation in the maize (Zea mays L.) inbred line B73 response to heat stress at the seedling stage. *Biochem. Biophys. Res. Commun.* **512**, 742–749 https://doi.org/10.1016/j.bbrc.2019.03.150

34   Wang, W., Qin, Q., Sun, F., Wang, Y., Xu, D., Li, Z. et al. (2016) Genome-Wide differences in DNA methylation changes in Two contrasting rice genotypes in response to drought conditions. *Front. Plant Sci.* **7**, 1675 https://doi.org/10.3389/fpls.2016.01675

35   Laanen, P., Saenen, E., Mysara, M., Van de Walle, J., Van Hees, M., Nauts, R. et al. (2021) Changes in DNA methylation in *Arabidopsis thaliana* plants exposed over multiple generations to gamma radiation. *Front. Plant Sci.* **12**, 532 https://doi.org/10.3389/fpls.2021.611783

36   Shen, Y., Zhang, J., Liu, Y., Liu, S., Liu, Z., Duan, Z. et al. (2018) DNA methylation footprints during soybean domestication and improvement. *Genome Biol.* **19**, 1–14 https://doi.org/10.1186/s13059-018-1516-z

37   Gahlaut, V., Zinta, G., Jaiswal, V. and Kumar, S. (2020) Quantitative epigenetics: a new avenue for crop improvement. *Epigenomes* **4**, 25 https://doi.org/10.3390/epigenomes4040025

38   Law, J.A. and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 https://doi.org/10.1038/nrg2719

39   Stroud, H., Greenberg, M.V.C., Feng, S., Bernatavichute, Y.V. and Jacobsen, S.E. (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the arabidopsis methylome. *Cell* **152**, 352–364 https://doi.org/10.1016/j.cell.2012.10.054

40   Li, Q., Eichten, S.R., Hermanson, P.J., Zaunbrecher, V.M., Song, J., Wendt, J. et al. (2014) Genetic perturbation of the maize methylome. *Plant Cell* **26**, 4602–4616 https://doi.org/10.1105/tpc.114.133140

41   Lanciano, S. and Mirouze, M. (2017) DNA methylation in rice and relevance for breeding. *Epigenomes* **1**, 10 https://doi.org/10.3390/epigenomes1020010

42   Hu, D., Yu, Y., Wang, C., Long, Y., Liu, Y., Feng, L. et al. (2021) Multiplex CRISPR-Cas9 editing of DNA methyltransferases in rice uncovers a class of non-CG methylation specific for GC-rich regions. *Plant Cell* **33**, 2950–2964 https://doi.org/10.1093/plcell/koab162

43   Noshay, J.M, Crisp, P.A and Springer, N.M. (2018) The Maize Methylome. In *Compendium of Plant Genomes* (Bennetzen, J., Flint-Garcia, S., Hirsch, C. and Tuberosa R. eds.), pp. 81–96, Springer International Publishing, Cham

44   Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M. et al. (2003) Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr. Biol.* **13**, 2212–2217 https://doi.org/10.1016/j.cub.2003.11.052

45   Saze, H., Scheid, O.M. and Paszkowski, J. (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat. Genet.* **34**, 65–69 https://doi.org/10.1038/ng1138

46   Zemach, A., Kim, M.Y., Hsieh, P.-H., Coleman-Derr, D., Eshed-Williams, L., Thao, K. et al. (2013) The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205 https://doi.org/10.1016/j.cell.2013.02.033

47   Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L. et al. (2013) Non-CG methylation patterns shape the epigenetic landscape in arabidopsis. *Nat. Struct. Mol. Biol.* **21**, 64–72 https://doi.org/10.1038/nsmb.2735

48   Haag, J.R. and Pikaard, C.S. (2011) Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nat. Rev. Mol. Cell. Biol.* **12**, 483–492 https://doi.org/10.1038/nrm3152

49   Wierzbicki, A.T., Haag, J.R. and Pikaard, C.S. (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* **135**, 635–648 https://doi.org/10.1016/j.cell.2008.09.035

50   Zhang, H., He, X. and Zhu, J.-K. (2013) RNA-directed DNA methylation in plants: where to start? *RNA Biol.* **10**, 1593–1596 https://doi.org/10.4161/rna.26312

51   Law, J.A., Du, J., Hale, C.J., Feng, S., Krajewski, K., Palanca, A.M.S. et al. (2013) Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**, 385–389 https://doi.org/10.1038/nature12178

52   Sigman, M.J., Panda, K., Kirchner, R., McLain, L.L., Payne, H., Peasari, J.R. et al. (2021) An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate DNA methylation. *Nat. Plants* **7**, 1461–1474 https://doi.org/10.1038/s41477-021-01008-7

53   Zhong, X., Du, J., Hale, C.J., Gallego-Bartolome, J., Feng, S., Vashisht, A.A. et al. (2014) Molecular mechanism of action of plant DRM *de novo* DNA methyltransferases. *Cell* **157**, 1050–1060 https://doi.org/10.1016/j.cell.2014.03.056

54   Zhang, H., Ma, Z.-Y., Zeng, L., Tanaka, K., Zhang, C.-J., Ma, J. et al. (2013) DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. *Proc. Natl Acad. Sci. U.S.A.* **110**, 8290–8295 https://doi.org/10.1073/pnas.1300585110

55   Hu, L., Li, N., Xu, C., Zhong, S., Lin, X., Yang, J. et al. (2014) Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proc. Natl Acad. Sci. U.S.A.* **111**, 10642–10647 https://doi.org/10.1073/pnas.1410761111

56   Thomas, M., Pingault, L., Poulet, A., Duarte, J., Throude, M., Faure, S. et al. (2014) Evolutionary history of methyltransferase 1 genes in hexaploid wheat. *BMC Genomics* **15**, 922 https://doi.org/10.1186/1471-2164-15-922

57   Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S. et al. (2001) Requirement of CHROMOMETHYLASE3 for maintenance of cpXpG methylation. *Science* **292**, 2077–2080 https://doi.org/10.1126/science.1059745

58   Jackson, J.P., Lindroth, A.M., Cao, X. and Jacobsen, S.E. (2002) Control of cpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556–560 https://doi.org/10.1038/nature731

59   Papa, C.M., Springer, N.M., Muszynski, M.G., Meeley, R. and Kaeppler, S.M. (2001) Maize chromomethylase Zea methyltransferase2 is required for CpNpG methylation. *Plant Cell* **13**, 1919–1928 https://doi.org/10.1105/TPC.010064

60   Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E. et al. (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature.* **473**, 97–100 https://doi.org/10.1038/nature09916

61   Bewick, A.J., Niederhuth, C.E., Ji, L., Rohr, N.A., Griffin, P.T., Leebens-Mack, J. et al. (2017) The evolution of CHROMOMETHYLASES and gene body DNA methylation in plants. *Genome Biol.* **18**, 1–13 https://doi.org/10.1186/s13059-017-1195-1

62 Cheng, C., Tarutani, Y., Miyao, A., Ito, T., Yamazaki, M., Sakai, H. et al. (2015) Loss of function mutations in the rice chromomethylase OsCMT3a cause a burst of transposition. *Plant J.* **83**, 1069–1081 https://doi.org/10.1111/tpj.12952

63 Gouil, Q. and Baulcombe, D.C. (2016) DNA methylation signatures of the plant chromomethyltransferases. *PLoS Genet.* **12**, e1006526 https://doi.org/10.1371/journal.pgen.1006526

64 Long, J.C., Xia, A.A., Liu, J.H., Jing, J.L., Wang, Y.Z., Qi, C.Y. et al. (2019) Decrease in DNA methylation 1 (DDM1) is required for the formation of CHH islands in maize. *J. Integr. Plant Biol.* **61**, 749–764 https://doi.org/10.1111/jipb.12733

65 Li, Q., Gent, J.I., Zynda, G., Song, J., Makarevitch, I., Hirsch, C.D. et al. (2015) RNA-directed DNA methylation enforces boundaries between heterochromatin and euchromatin in the maize genome. *Proc. Natl Acad. Sci. U.S.A.* **112**, 14728–14733 https://doi.org/10.1073/pnas.1514680112

66 Long, J., Liu, J., Xia, A., Springer, N.M. and He, Y. (2021) Maize decrease in DNA methylation 1 targets RNA-directed DNA methylation on active chromatin. *Plant Cell* **33**, 2183–2196 https://doi.org/10.1093/plcell/koab098

67 Zhu, J.-K. (2009) Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* **43**, 143–166 https://doi.org/10.1146/annurev-genet-102108-134205

68 Ogneva, Z.V., Dubrovina, A.S. and Kiselev, K.V. (2016) Age-associated alterations in DNA methylation and expression of methyltransferase and demethylase genes in *Arabidopsis thaliana. Biol. Plant* **60**, 628–634 https://doi.org/10.1007/s10535-016-0638-y

69 Roldán-Arjona, T., Ariza, R.R. and Córdoba-Cañero, D. (2019) DNA base excision repair in plants: an unfolding story With familiar and novel characters. *Front. Plant Sci.* **10**, 1055 https://doi.org/10.3389/fpls.2019.01055

70 Liu, R. and Lang, Z. (2020) The mechanism and function of active DNA demethylation in plants. *J. Integr. Plant Biol.* **62**, 148–159 https://doi.org/10.1111/jipb.12879

71 Gong, Z., Morales-Ruiz, T., Ariza, R.R., Roldán-Arjona, T., David, L. and Zhu, J.K. (2002) ROS1, a repressor of transcriptional gene silencing in arabidopsis, encodes a DNA glycosylase/lyase. *Cell* **111**, 803–814 https://doi.org/10.1016/S0092-8674(02)01133-9

72 He, X.-J., Chen, T. and Zhu, J.-K. (2011) Regulation and function of DNA methylation in plants and animals. *Cell Res.* **21**, 442–465 https://doi.org/10.1038/cr.2011.23

73 Zemach, A., Yvonne Kim, M., Silva, P., Rodrigues, J.A., Dotson, B., Brooks, M.D. et al. (2010) Local DNA hypomethylation activates genes in rice endosperm. *Proc. Natl Acad. Sci. U.S.A.* **107**, 18729–18734 https://doi.org/10.1073/pnas.1009695107

74 Li, J., Li, C. and Lu, S. (2018) Systematic analysis of DEMETER-like DNA glycosylase genes shows lineage-specific Smi-miR7972 involved in SmDML1 regulation in Salvia miltiorrhiza. *Sci. Rep.* **8**, 7143 https://doi.org/10.1038/s41598-018-25315-w

75 Cheng, C.-Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S. and Town, C.D. (2017) Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* **89**, 789–804 https://doi.org/10.1111/tpj.13415

76 Henderson, I.R., Deleris, A., Wong, W., Zhong, X., Chin, H.G., Horwitz, G.A. et al. (2010) The *de novo* cytosine methyltransferase DRM2 requires intact UBA domains and a catalytically mutated paralog DRM3 during RNA-directed DNA methylation in *Arabidopsis thaliana. PLoS Genet.* **6**, e1001182 https://doi.org/10.1371/journal.pgen.1001182

77 Ono, A., Yamaguchi, K., Fukada-Tanaka, S., Terada, R., Mitsui, T. and Iida, S. (2012) A null mutation of ROS1a for DNA demethylation in rice is not transmittable to progeny. *Plant J.* **71**, 564–574 https://doi.org/10.1111/j.1365-313X.2012.05009.x

78 Raju, S.K.K. and Niederhuth, C.E. (2018) Epigenetic Diversity and Application to Breeding. In *Advances in Botanical Research* (Mirouze, M., Bucher, E. and Gallusci, P. eds.), pp. 49–86, Academic Press, Oxford, UK

79 Crisp, P.A., Bhatnagar-Mathur, P., Hundleby, P., Godwin, I.D., Waterhouse, P.M. and Hickey, L.T. (2021) Beyond the gene: epigenetic and cis-regulatory targets offer new breeding potential for the future. *Curr. Opin. Biotechnol.* **73**, 88–94 https://doi.org/10.1016/j.copbio.2021.07.008

80 Zhang, M., Xu, C., von Wettstein, D. and Liu, B. (2011) Tissue-specific differences in cytosine methylation and their association with differential gene expression in sorghum. *Plant Physiol.* **156**, 1955–1966 https://doi.org/10.1104/pp.111.176842

81 Kawashima, T. and Berger, F. (2014) Epigenetic reprogramming in plant sexual reproduction. *Nat. Rev. Genet.* **15**, 613–624 https://doi.org/10.1038/nrg3685

82 Rodrigues, J.A., Ruan, R., Nishimura, T., Sharma, M.K., Sharma, R., Ronald, P.C. et al. (2013) Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. *Proc. Natl Acad. Sci. U.S.A.* **110**, 7934–7939 https://doi.org/10.1073/pnas.1306164110

83 Waters, A.J., Makarevitch, I., Eichten, S.R., Swanson-Wagner, R.A., Yeh, C.-T., Xu, W. et al. (2011) Parent-of-origin effects on gene expression and DNA methylation in the maize endosperm. *Plant Cell* **23**, 4221–4233 https://doi.org/10.1105/tpc.111.092668

84 Richards, E.J. (2006) Inherited epigenetic variation — revisiting soft inheritance. *Nat. Rev. Genet.* **7**, 395–401 https://doi.org/10.1038/nrg1834

85 Melquist, S. and Bender, J. (2004) An internal rearrangement in an arabidopsis inverted repeat locus impairs DNA methylation triggered by the locus. *Genetics* **166**, 437–448 https://doi.org/10.1534/genetics.166.1.437

86 Noshay, J.M. and Springer, N.M. (2021) Stories that can't be told by SNPs; DNA methylation variation in plant populations. *Curr. Opin. Plant Biol.* **61**, 101989 https://doi.org/10.1016/j.pbi.2020.101989

87 Kumar, S. and Mohapatra, T. (2021) Dynamics of DNA methylation and Its functions in plant growth and development. *Front. Plant Sci.* **12**, 596236 https://doi.org/10.3389/fpls.2021.596236

88 Schmitz, R.J., Schultz, M.D., Lewsey, M.G., O'Malley, R.C., Urich, M.A., Libiger, O. et al. (2011) Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334**, 369–373 https://doi.org/10.1126/science.1212959

89 Melquist, S., Luff, B. and Bender, J. (1999) Arabidopsis PAI gene arrangements, cytosine methylation and expression. *Genetics* **153**, 401–413 https://doi.org/10.1093/genetics/153.1.401

90 Sarkies, P. (2020) Molecular mechanisms of epigenetic inheritance: possible evolutionary implications. *Semin. Cell Dev. Biol.* **97**, 106–115 https://doi.org/10.1016/j.semcdb.2019.06.005

91 Schmitz, R.J., Schultz, M.D., Urich, M.A., Nery, J.R., Pelizzola, M., Libiger, O. et al. (2013) Patterns of population epigenomic diversity. *Nature* **495**, 193–198 https://doi.org/10.1038/nature11968

92 Johannes, F. and Schmitz, R.J. (2019) Spontaneous epimutations in plants. *New Phytol.* **221**, 1253–1259 https://doi.org/10.1111/nph.15434

93 Shahryary, Y., Symeonidi, A., Hazarika, R.R., Denkena, J., Mubeen, T., Hofmeister, B. et al. (2020) Alphabeta: computational inference of epimutation rates and spectra from high-throughput DNA methylation data in plants. *Genome Biol.* **21**, 260 https://doi.org/10.1186/s13059-020-02161-6

94   Ganguly, D.R., Crisp, P.A., Eichten, S.R. and Pogson, B.J. (2017) The arabidopsis DNA methylome is stable under transgenerational drought stress. *Plant Physiol.* **175**, 1893–1912 https://doi.org/10.1104/pp.17.00744

95   Hofmeister, B.T., Denkena, J., Colomé-Tatché, M., Shahryary, Y., Hazarika, R., Grimwood, J. et al. (2020) A genome assembly and the somatic genetic and epigenetic mutation rate in a wild long-lived perennial *Populus trichocarpa*. *Genome Biol.* **21**, 1–27 https://doi.org/10.1186/s13059-020-02162-5

96   Quadrana, L. and Colot, V. (2016) Plant transgenerational epigenetics. *Annu. Rev. Genet.* **50**, 467–491 https://doi.org/10.1146/annurev-genet-120215-035254

97   Kawakatsu, T., Huang, S.-S.C., Jupe, F., Sasaki, E., Schmitz, R.J., Urich, M.A. et al. (2016) Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell* **166**, 492–505 https://doi.org/10.1016/j.cell.2016.06.044

98   Gardiner, L.-J., Quinton-Tulloch, M., Olohan, L., Price, J., Hall, N. and Hall, A. (2015) A genome-wide survey of DNA methylation in hexaploid wheat. *Genome Biol.* **16**, 1–15 https://doi.org/10.1186/s13059-014-0572-2

99   Gardiner, L.-J., Joynson, R., Omony, J., Rusholme-Pilcher, R., Olohan, L., Lang, D. et al. (2018) Hidden variation in polyploid wheat drives local adaptation. *Genome Res.* **28**, 1319–1332 https://doi.org/10.1101/gr.233551.117

100 Rhee, Y., Sekhon, R.S., Chopra, S. and Kaeppler, S. (2010) Tissue culture-induced novel epialleles of a Myb transcription factor encoded by pericarp color1 in maize. *Genetics* **186**, 843–855 https://doi.org/10.1534/genetics.110.117929

101 Stroud, H., Ding, B., Simon, S.A., Feng, S., Bellizzi, M., Pellegrini, M. et al. (2013) Plants regenerated from tissue culture contain stable epigenome changes in rice. *eLife* **2**, e00354 https://doi.org/10.7554/eLife.00354

102 Liu, Z.L., Han, F.P., Tan, M., Shan, X.H., Dong, Y.Z., Wang, X.Z. et al. (2004) Activation of a rice endogenous retrotransposon Tos17 in tissue culture is accompanied by cytosine demethylation and causes heritable alteration in methylation pattern of flanking genomic regions. *Theor. Appl. Genet.* **109**, 200–209 https://doi.org/10.1007/s00122-004-1618-8

103 Cheng, C., Daigen, M. and Hirochika, H. (2006) Epigenetic regulation of the rice retrotransposon Tos17. *Mol. Genet. Genomics* **276**, 378–390 https://doi.org/10.1007/s00438-006-0141-9

104 Ding, Y., Wang, X., Su, L., Zhai, J., Cao, S., Zhang, D. et al. (2007) SDG714, a histone H3K9 methyltransferase, is involved in Tos17 DNA methylation and transposition in rice. *Plant Cell* **19**, 9–22 https://doi.org/10.1105/tpc.106.048124

105 La, H., Ding, B., Mishra, G.P., Zhou, B., Yang, H., Bellizzi, M.R. et al. (2011) A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon Tos17 and promotes its transposition in rice. *Proc. Natl Acad. Sci. U.S.A.* **108**, 15498–15503 https://doi.org/10.1073/pnas.1112704108

106 Zhang, M., Xu, C., Yan, H., Zhao, N., von Wettstein, D. and Liu, B. (2009) Limited tissue culture-induced mutations and linked epigenetic modifications in F hybrids of sorghum pure lines are accompanied by increased transcription of DNA methyltransferases and 5-methylcytosine glycosylases. *Plant J.* **57**, 666–679 https://doi.org/10.1111/j.1365-313X.2008.03719.x

107 Finnegan, E.J., Ford, B., Wallace, X., Pettolino, F., Griffin, P.T., Schmitz, R.J. et al. (2018) Zebularine treatment is associated with deletion of FT-B1 leading to an increase in spikelet number in bread wheat. *Plant Cell Environ.* **41**, 1346–1360 https://doi.org/10.1111/pce.13164

108 Gladman, N., Hufnagel, B., Regulski, M., Liu, Z., Wang, X., Chougule, K. et al. (2021) Sorghum root epigenetic landscape during limiting phosphorus conditions. *bioRxiv* https://doi.org/10.1101/2021.05.25.445633

109 Latutrie, M., Gourcilleau, D. and Pujol, B. (2019) Epigenetic variation for agronomic improvement: an opportunity for vegetatively propagated crops. *Am. J. Bot.* **106**, 1281–1284 https://doi.org/10.1002/ajb2.1357

110 Danchin, E., Pocheville, A., Rey, O., Pujol, B. and Blanchet, S. (2019) Epigenetically facilitated mutational assimilation: epigenetics as a hub within the inclusive evolutionary synthesis. *Biol. Rev. Camb. Philos. Soc.* **94**, 259–282 https://doi.org/10.1111/brv.12453

111 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 https://doi.org/10.1186/s13059-017-1288-x

112 Chen, X. and Zhou, D.-X. (2013) Rice epigenomics and epigenetics: challenges and opportunities. *Curr. Opin. Plant Biol.* **16**, 164–169 https://doi.org/10.1016/j.pbi.2013.03.004

113 Miura, K., Agetsuma, M., Kitano, H., Yoshimura, A., Matsuoka, M., Jacobsen, S.E. et al. (2009) A metastable DWARF1 epigenetic mutant affecting plant stature in rice. *Proc. Natl Acad. Sci. U.S.A.* **106**, 11218–11223 https://doi.org/10.1073/pnas.0901942106

114 Garg, R., Narayana Chevala, V.V.S., Shankar, R. and Jain, M. (2015) Divergent DNA methylation patterns associated with gene expression in rice cultivars with contrasting drought and salinity stress response. *Sci. Rep.* **5**, 14922 https://doi.org/10.1038/srep14922

115 Long, Y., Xia, W., Li, R., Wang, J., Shao, M., Feng, J. et al. (2011) Epigenetic QTL mapping in Brassica napus. *Genetics* **189**, 1093–1102 https://doi.org/10.1534/genetics.111.131615

116 He, L., Wu, W., Zinta, G., Yang, L., Wang, D., Liu, R. et al. (2018) A naturally occurring epiallele associates with leaf senescence and local climate adaptation in arabidopsis accessions. *Nat. Commun.* **9**, 460 https://doi.org/10.1038/s41467-018-02839-3

117 Johannes, F., Porcher, E., Teixeira, F.K., Saliba-Colombani, V., Simon, M., Agier, N. et al. (2009) Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet.* **5**, e1000530 https://doi.org/10.1371/journal.pgen.1000530

118 Can, S.N., Nunn, A., Galanti, D., Langenberger, D., Becker, C., Volmer, K. et al. (2021) The EpiDiverse plant epigenome-wide association studies (EWAS) pipeline. *Epigenomes* **5**, 12 https://doi.org/10.3390/epigenomes5020012

119 Cortijo, S., Wardenaar, R., Colomé-Tatché, M., Gilly, A., Etcheverry, M., Labadie, K. et al. (2014) Mapping the epigenetic basis of complex traits. *Science* **343**, 1145–1148 https://doi.org/10.1126/science.1248127

120 Crisp, P.A., Noshay, J.M., Anderson, S.N. and Springer, N.M. (2019) Opportunities to Use DNA methylation to distil functional elements in large crop genomes. *Mol. Plant* **12**, 282–284 https://doi.org/10.1016/j.molp.2019.02.006

121 Hu, Z., Olatoye, M.O., Marla, S. and Morris, G.P. (2019) An integrated genotyping-by-sequencing polymorphism Map for over 10,000 sorghum genotypes. *Plant Genome* **12**, 180044 https://doi.org/10.3835/plantgenome2018.06.0044

122 Kakoulidou, I., Avramidou, E.V., Baránek, M., Brunel-Muguet, S., Farrona, S., Johannes, F. et al. (2021) Epigenetics for crop improvement in times of global change. *Biology* **10**, 766 https://doi.org/10.3390/biology10080766

123 Hu, Y., Morota, G., Rosa, G.J.M. and Gianola, D. (2015) Prediction of plant height in *Arabidopsis thaliana* using DNA methylation data. *Genetics* **201**, 779–793 https://doi.org/10.1534/genetics.115.177204

124 Schnable, P.S. and Springer, N.M. (2013) Progress toward understanding heterosis in crop plants. *Annu. Rev. Plant Biol.* **64**, 71–88 https://doi.org/10.1146/annurev-arplant-042110-103827

125 Shen, H., He, H., Li, J., Chen, W., Wang, X., Guo, L. et al. (2012) Genome-wide analysis of DNA methylation and gene expression changes in two Arabidopsis ecotypes and their reciprocal hybrids. *Plant Cell* **24**, 875–892 https://doi.org/10.1105/tpc.111.094870

126 Uzun, Y., Wu, H. and Tan, K. (2021) Predictive modeling of single-cell DNA methylome data enhances integration with transcriptome data. *Genome Res.* **31**, 101–109 https://doi.org/10.1101/gr.267047.120

127 Yelina, N.E., Lambing, C., Hardcastle, T.J., Zhao, X., Santos, B. and Henderson, I.R. (2015) DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. *Genes Dev.* **29**, 2183–2202 https://doi.org/10.1101/gad.270876.115

128 Underwood, C.J., Choi, K., Lambing, C., Zhao, X., Serra, H., Borges, F. et al. (2018) Epigenetic activation of meiotic recombination near centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Res.* **28**, 519–531 https://doi.org/10.1101/gr.227116.117

129 Crisp, P.A., Ganguly, D.R., Smith, A.B., Murray, K.D., Estavillo, G.M., Searle, I. et al. (2017) Rapid recovery gene downregulation during excess-light stress and recovery in Arabidopsis. *Plant Cell* **29**, 1836–1863 https://doi.org/10.1105/tpc.16.00828

130 Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J.R.A., Hellinga, A.J. et al. (2019) Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275 https://doi.org/10.1186/s13059-019-1905-y

131 Hufford, M.B., Seetharam, A.S., Woodhouse, M.R., Chougule, K.M., Ou, S., Liu, J. et al. (2021) *De novo* assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science* **373**, 655–662 https://doi.org/10.1126/science.abg5289