

**Univerzita Karlova**  
**Přírodovědecká fakulta**

Studijní program: Biologie

Studijní obor: Biologie



**Jan Polák**

Role sekvenčního kontextu v metylaci DNA  
Role of sequence context in DNA methylation

Bakalářská práce

Školitel: RNDr. Lukáš Fischer Ph.D.

Praha 2019



### *Poděkování*

Chtěl bych především poděkovat svému školiteli Lukáši Fischerovi za jeho trpělivost a nakažlivý entuziasmus.

### **Prohlášení:**

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze 15. 8 2019

Podpis

# Obsah

Abstrakt.....	2
Abstract.....	3
Seznam použitých zkratk.....	4
1 Úvod.....	5
2 O chromatinu.....	5
3 Obecné funkce a způsoby značení.....	6
3.1.1 Histonové varianty.....	7
3.1.2 Histonové modifikace.....	7
4 Metylce DNA.....	8
4.1. Výskyt v genomu.....	8
4.1.1 Euchromatin.....	8
4.1.2 Heterochromatin.....	9
4.2. Remodelační komplexy a interpretační proteiny.....	9
4.2.1 Remodelační komplexy.....	9
4.2.2 Interpretační proteiny.....	10
4.3. Udržování a de novo metylace.....	10
4.4. Cytosinové kontexty zdálky.....	11
4.5. DNA metyltransferázy.....	11
4.5.1 MET1.....	11
4.5.2 Rodina CMT.....	12
4.5.3 RdDM a DRM2.....	13
4.6. SUVH rodina proteinů.....	14
4.6.1 SUVH4.....	15
4.6.2 SUVH5.....	15
4.6.3 SUVH6.....	15
4.6.4 SUVH2 a SUVH9.....	16
4.7. Cytosinové kontexty zblízka.....	16
4.7.1 CG.....	16
4.7.2 Obecně o non-CG.....	16
4.7.2.1 CHG.....	16
4.7.2.2 CHH.....	17
5 Demetylce.....	18
5.1. DNA demetylce.....	18
5.2. Demetylce histonů.....	19
6 Závěr.....	19
Použitá literatura.....	20

## Abstrakt

Metylace DNA v cytosinu je důležitá epigenetická značka, která se podílí na regulaci genové exprese, umlčování transponovatelných elementů a spoluurčuje celkový stavu chromatinu. Cytosin se v DNA může vyskytovat ve třech kontextech: CG, CHG a CHH (kde H zastupuje C, A, nebo T). *Arabidopsis thaliana* (a rostliny obecně) disponuje molekulárním aparátem, kterým je schopna metylovat DNA v cytosinu ve všech kontextech. Tento aparát má dva úkoly: udržovat již zavedené metylační značky a v případě potřeby vytvářet nové. Udržování metylace se pro jednotlivé kontexty liší svými mechanismy. K metylaci symetrických kontextů CG a CHG se využívá informace z již metylovaného vlákna. Asymetrické kontexty CHH a i CHG k metylaci využívají informaci z další epigenetické značky: modifikace histonů, konkrétně dimetylace histonu H3 na lysinu 9. Takto vzniká zpětnovazebná smyčka mezi metylací histonů a DNA. Metylace CHH pak také může být naváděna pomocí malých RNA komplementárních k danému místu. Mechanismus navádění metylačního enzymu pomocí malých RNA je také využíván při zavádění nových mDNA značek, a to v případě všech kontextů. Nejvíce je metylován kontext CG, který kromě velkého zastoupení v heterochromatinu je přítomný i v konstitutivně přepisovaných oblastech dlouhých středně intenzivně přepisovaných genů.

**Klíčová slova:** *Arabidopsis*, epigenetika, chromatin, metyltransferáza, metylace histonu

## Abstract

Cytosine methylation of DNA is a pivotal epigenetic mark, which contributes to the regulation of the gene expression, silencing of transposable elements, and co-defines chromatin state. There are three cytosine contexts: CG, CHG and CHH (where H stands for C, A, or T). *Arabidopsis thaliana* (and plants in general) has an arsenal of molecular mechanisms capable of cytosine methylation in all of its contexts. That said, there are two tasks at hand: maintaining of pre-existing methylation and if need be, creating new methylated spots. The actual process of maintaining of the methylation depends on the cytosine context. Methylation of symmetrical contexts of CG and CHG can utilize the information about the methylation pattern from the second DNA strand. The asymmetrical context of CHH, and also CHG need to look for this information elsewhere: in the methylation of the lysine 9 of H3 histone. This creates a self-reinforcing loop and a crosstalk between two epigenetic mechanisms. Maintenance of methylation of CHH is also navigated by small RNA complementary to the locus in question. This mechanism of enzyme navigating by RNA is also used in establishing a new methylated site for all of the contexts. CG methylation is most prevalent in both heterochromatin and euchromatin. It also has a special function of increasing transcription processivity of long, constitutively expressed genes.

**Keywords:** *Arabidopsis*, epigenetics, chromatin, methyltransferase, histone methylation

## Seznam použitých zkratek

AGO	ARGONAUTE
BAH	bromo-adjacent homology
CD	chromo domain
CLSY1	SNF2 domain-containing protein CLASSY1
CMT	CHROMOMETHYLASE
DCL	DICER-LIKE PROTEIN
DDM1	DECREASED DNA METHYLATION1
DME	TRANSCRIPTIONAL ACTIVATOR DEMETER
DML	TRANSCRIPTIONAL ACTIVATOR DEMETER-LIKE2
DRD1	DEFECTIVE IN RNA DIRECTED DNA METHYLATION1
DRM2	DOMAIN REARRANGED METHYLASE2
gbM	gene-body methylation
HDA6	HISTONE DEACETYLASE
IBM1	INCREASE IN BONSAI METHYLATION
TE	transposable element
TRD	target recognition domain
TSS	transcription side start
SAR	SET and RING
SHH1	SAWADEE HOMEODOMAIN HOMOLOG1
SUVH	SUPPRESSOR OF VARIATION HOMOLOG
RdDM	RNA-directed DNA methylation
ROS1	REPRESSOR OF SILENCING
HDA6	HISTOE DEACETYLASE
MEMs	DNA methylation monitoring sequence
MET1	METHYLTRANSFERASE1
PolIV	RNA polymeraza IV
PolV	RNA polymeraza V
VIM	VARIANT IN METHYLATION
ZMET	ZEA METHYLTRANSFERASE

# 1 Úvod

Snaha vysvětlit fenomény odchylovající se od Mendelových zákonů, např. rodičovský imprinting, fenotyp mezidruhových kříženců, paramutace a umlčování transgenů, přispěla k vytvoření pojmu epigenetiky. Exprese genu je podmíněna genetickou složkou, tedy přítomností a podobou daného genu (včetně promotoru) v genomu, ale také epigenetickou složkou, tedy mechanismy, které genovou expresi ovlivňují jinak, než nukleotidovými sekvencemi. Některé tyto vlivy jsou dokonce dědičné. Mezi hlavní epigenetické mechanismy patří složky ovlivňující stav chromatinu: histony, jejich varianty a modifikace, další interagující proteiny a metylace DNA.

V této práci bych chtěl popsat současnou představu o tom, jakým způsobem *Arabidopsis thaliana* navozuje, udržuje a mění metylaci DNA v cytosinu. Při této snaze pokládám za nutné pro lepší výklad i nastínit funkce, jaké metylace DNA u *Arabidopsis* zastává, a také popsat ostatní mechanismy, které do stavu a procesu metylace DNA zasahují. Děje související se stavem chromatinu jsou totiž úzce provázané do té míry, že při popisování jednoho z mechanismů (zde tedy metylace DNA) se nelze vyhnout ostatním jevům v chromatinu. Proto nejprve představím všechny možné součásti, které dohromady tvoří velmi jemně vyladěný systém. Získám tak pojmový aparát, pomocí kterého budu moci vysvětlit roli sekvenčního kontextu v metylaci DNA.

## 2 O chromatinu

Chromatin byl původně definován jako barvitelná struktura v buněčném jádře. Skládá se z molekuly DNA a asociovaných proteinů. Základní jednotka chromatinu je nukleosom, oktamer složený z proteinového jádra dvojic histonových proteinů. Nukleosom se ovinut molekulou DNA o délce 147 bp. Jednotlivé nukleosomy jsou propojeny krátkým úsekem DNA, které mohou být asociované s histonem H1. Ten také propojuje centrální histony a tvoří tak vyšší řád organizace, chromatinové vlákno. Oblasti s H1 jsou kompaktnější, méně přístupné a transkripčně neaktivní (Grant-Downton a Dickinson 2005).

Chromatin se tradičně dělí na lépe barvitelný a pozorovatelný heterochromatin a hůře barvitelný a pozorovatelný euchromatin. Tento efekt je způsobený složením: v kondenzovanějším heterochromatinu je DNA ovinuta těsněji, euchromatin je naopak rozvolněnější, oblasti počátku transkripce (TSS) jsou pak charakteristické přítomností nestabilních nukleosomů. Kromě vizuálního efektu při barvení se liší hlavně funkcí:



heterochromatin vykazuje menší aktivitu transkripce a rekombinace a obsahuje více repetitivních sekvencí.

V *Arabidopsis thaliana* byly popsány čtyři hlavní stavy chromatinu: heterochromatin, „Polycomb“ (chromatin s vývojově umlčenými geny vazbou Polycomb repressive komplexu), transkripčně aktivní chromatin a intergenické oblasti. Byly provedeny i podrobnější analýzy zastoupení jednotlivých histonů, jejich variant a modifikací, metylace DNA a genomických elementů (TE, typ RNA, promotory, introny atd.), podle kterých se dá chromatin rozředit na základě společných vlastností do 9-12 stavů: chromatin v okolí počátku transkripce, chromatin v okolí počátku kódující sekvence, chromatin v okolí konce transkripce, chromatin dlouhých genů a intronů. Dále heterochromatin bohatý na CG a heterochromatin bohatý na AT. Dále tři typy „Polycomb“ chromatinu. (Sequeira-Mendes et al. 2014).

### **3 Obecné funkce a způsoby značení**

Pro detailnější rozlišení stavu a potažmo funkce chromatinu je nutné sledovat značky, které na chromatinu molekulárně interpretují interagující proteiny zajišťující základní funkce chromatinu včetně enzymů modifikujících DNA a histony a faktorů remodelujících chromatin. Mezi tato značení patří posttranslační modifikace histonů, začlenění variant histonů do nukleosomů a metylace DNA. V euchromatinu je DNA lépe přístupná molekulárním mechanismům, které jsou asociované s DNA (aparátem transkripce, rekombinace, replikace atd.), a obsahuje méně repetitivních oblastí. Procesy spojené s heterochromatinovou DNA (například přepis nebo metylace DNA) vyžadují často aktivitu specifických remodelačních komplexů, které upraví strukturu chromatinu (nukleosomů) a DNA tím zpřístupní. Uspořádání nukleosomů kondenzovaného chromatinu totiž blokuje vazbu transkripčních faktorů k DNA. Remodelační komplexy jsou rozmanitá skupina proteinů a často fungují v součinnosti s enzymem, kterému DNA zpřístupňují (vlastní remodelační proteiny mají metyltransferázy a vlastní transkripční faktory). To, jak je DNA přístupná, se odvozuje od zastoupení histonových variant a modifikací (níže). Stav chromatinu je tak klíčový pro všechny děje spojené s DNA a jeho součástí tvoří komplexní, vzájemně propojenou a dynamickou souhru regulačních procesů a výstup této souhry je fenotyp buňky, buněčných populací a celého organismu (Loidl 2004, Grant-Downton a Dickinson 2005, Roudier et al. 2009).

### 3.1.1 Histonové varianty

Jádro oktameru histonů v nukleosomu je tvořeno dvojicemi histonů H2A, H2B, H3, H4, a tyto oktamery jsou propojeny histonem H1. Jednotlivé histony se vyskytují ve více variantách, které spolu s posttranslačními modifikacemi histonů vytváří „epigenetický histonový kód“. Například rDNA chromatin v jádru obsahuje výlučně H1 histony, jiná místa mohou obsahovat jednu ze tří variant H1-1, H1-2, které jsou rozmístěny různě po chromozómech či H1-3, který chybí v repetitivních sekvencích (Grant-Downton a Dickinson 2005).

Histon H3 může být v nukleosomu nahrazen variantami, nejčastěji H3.1 a H3.3, které se liší stavbou a funkcí. H3.1 je exprimován během DNA replikace a je charakteristický pro transkripčně neaktivní chromatin, H3.3 během S-fáze buněčného cyklu. H3.3 se vyskytuje u konců transkripce konstitutivně exprimovaných genů a jeho zastoupení pozitivně koreluje s metylací exonů exprimovaných genů (tzv. gene-body methylation, viz dále). Knock-out mutace všech tří genů kódujících H3.3 je lethální, není tedy zastupitelný jinými variantami histonu H3. Přítomnost H3.3 umožňuje aktivitu DNA metyltransferáz, pravděpodobně proto, že brání vazbě H1 histonu a tím skládání chromatinu, které brání aktivitě metyltransferáz (Wollmann et al. 2017).

H2A histon má např. variantu H2A.Z, který zabraňuje vysoké kondenzaci chromatinu, a vyskytuje se u 5' konce přepisovaných úseků, kde s H3.3 vytváří nestabilní nukleosomy (Jin a Felsenfeld 2007, Zilberman et al. 2008).

### 3.1.2 Histonové modifikace

Z nukleosomu vystupují N konce histonů, které jsou tak vystaveny modifikačním enzymům. Histonové posttranslační modifikace jsou kovalentní změny v aminokyselinách. Prominentní jsou modifikace histonů H3 a H4. Mezi tyto modifikace patří zejména místně specifické acetylace a metylace lysinových zbytků (dále také ubiquitinylace a sumoylace lysinu, metylace argininu a fosforylace serinu a threoninu). Modifikace histonů způsobují strukturální a funkční změnu chromatinu a modifikované histony jsou považovány za znaky epigenetického histonového kódu, pomocí kterého lze zjistit, v jakém stavu se chromatin nachází. Tento kód je interpretován proteiny s chromodomény, které se váží na metylovaný lysin, a s bromodomény, které se váží na acetylovaný lysin (Loidl 2004, Grant-Downton a Dickinson 2005).

Pro transkripčně aktivní místa je u *Arabidopsis* charakteristická kombinace H3K4me2, H3K4me3, H3K36me3 a H3Ac. Pro dlouhé geny H3K4me1, H2Bub a H3K36me3 (které pravděpodobně zvyšují procesivitu transkripce). Pro umlčené geny v „Polycomb chromatinu“ je charakteristická značka H3K27me3. Pro heterochromatin jsou charakteristické H3K9me2 a H3K27me1 (Sequeira-Mendes et al. 2014).

Metylace histonů je zprostředkována histon metyltransferázami. Metylace H3K9 je zprostředkována enzymy z rodiny SUVH. Tvoří propojení mezi dvěma epigenetickými mechanismy, metylací DNA a H3K9me2 (Roudier et al. 2009).

## 4 Metylace DNA

### 4.1. Výskyt v genomu

#### 4.1.1 Euchromatin

Metylace DNA je další epigenetická značka, která ovlivňuje strukturu a přístupnost chromatinu. Mezi její hlavní funkce patří regulace genové exprese. U genů může docházet k metylaci promotorů, což má obvykle za důsledek umlčení exprese daného genu (o výjimce níže). Metylační značka v promotoru inhibuje expresi buď přímo tak, že brání vazbě transkripčního aktivátoru nebo vytváří vazbu s transkripčním represorem, nepřímo pak tak, že navádí histon metyltransferázy a ty následně zprostředkují změnu struktury chromatinu. *Arabidopsis* však metylací promotoru reguluje expresi mála genů (v poměru s jinými druhy) a mutantní rostliny neschopné DNA metylace jsou životaschopné (v porovnání s rýží a kukuřicí, pro které je ztráta metylace letální). Lze na ní proto efekt mutací v metylačním aparátu dobře zkoumat. Expresi genu také brání metylace oblasti konce transkripce (Zhang et al. 2018).

Další oblast genu, která je u krytosemenných rostlin metylovaná, je přímo v exonech konstitutivně exprimovaných genů, tzv. gene-body methylation (gbM). Geny s gbM jsou delší s více exony, evolučně konzervovanější a důležitější (tím se také vysvětluje nižší frekvence fixace mutací, protože jsou více zatíženy selekcí). Jedna z možných funkcí gbM je, že se mDNA značka podílí na definování rozhraní exonů a intronů a brání inkorporaci H2A.Z, který brání silné transkripci. GbM snad také zabraňuje zahájení transkripce v pseudopromotorech uvnitř genu (Zilberman et al. 2008, Takuno a Gaut 2012).

Další metylovaná oblast genu jsou introny některých genů, které obsahují retrotransposony. Při demethylaci těchto oblastí dochází k předčasné terminaci transkripce a špatnému splicingu. Dále jsou metylovány kratší transposony, které se vyskytují poblíž aktivně přepisovaných genů.

#### **4.1.2 Heterochromatin**

Nejvíce metylovaného cytosinu je v transponovatelných elementech a brání tak jejich přepisu polymerázou II, šíření a množení. Transponovatelné elementy v heterochromatinu jsou charakteristické vyšším zastoupením H3K9me2 histonů než ty v euchromatinu (Zemach et al. 2013). Dále je metylovaný pericentromerický heterochromatin a repetitivní sekvence. Právě v heterochromatinu je pro zavádění metylace nutná součinnost metyltransferáz a jejich remodelačních komplexů (Zemach et al. 2013).

### **4.2. Remodelační komplexy a interpretační proteiny**

#### **4.2.1 Remodelační komplexy**

DDM1 remodeluje chromatin s obsahem histonu H1 a umožňuje aktivitu DNA metyltransferáz. Je nutný pro metylaci cytosinu heterochromatinových transponovatelných elementů. Patří mezi ně Snf2 rodina remodelačních enzymů (např. DDM1), které hydrolyzují ATP, a energii využívají k pohybu podél DNA a mění stavbu nukleosomu, aby tak zpřístupnily DNA. Ztráta DDM1 způsobí sníženou metylaci heterochromatinu a nárůst přepisu transpononů (Zemach et al. 2013).

DRD1 je také z třídy Snf2 a remodeluje chromatin pro přístup Pol V působící v RdDM dráze (dráha, která využívá malých RNA k navádění lokusů určených k metylaci. viz níže), přítomnost DRD1 je nutnou podmínkou pro přepis daného lokusu Pol V (viz níže). Remodelaci ale nezvládá v oblastech s H1 histonem (Zemach et al. 2013). Dalším remodelačním proteinem v RdDM je CLSY1, který upravuje chromatin pro Pol IV, což umožňuje přepis vedoucí k produkci hc-siRNA (viz níže) (Law et al. 2011).

Remodelační komplexy mohou mít opačnou funkci, tedy kondenzaci chromozomu a umlčení transponovatelných elementů. V *Arabidopsis* má tuto funkci v pericentrickém chromatinu

proteinová rodina MORC. Patrně není propojená s metylací DNA a uvádím ji jenom jako příklad mechanismu umlčování jiného než metylace (Moissiard et al. 2012).

### 4.2.2 Interpretační proteiny

Epigenetická informace v chromatinu je interpretována různými druhy proteinů. Patří mezi ně například rodina chromometyltransferáz (CMT), která kromě své domény s metyltransferázovou aktivitou má i chromodoménu a BAH doménu. Pomocí té se umí vázat k H3K9me2 a tato vazba podmiňuje DNA metylační aktivitu CMT. Mezi chromatin interpretující proteiny patří také SHH1, který se váže k H3K9me2 svou SAWADEE doménou a přivádí do daného lokusu Pol IV, jednu z RNA polymeráz RdDM dráhy (Law et al. 2011, Law et al. 2013).

Jsou i proteiny, které interpretují metylační stav molekul DNA, a navádějí enzymy k místům metylace. Souhrně se označují jako MBD (Zemach a Grafi 2007). Mezi DNA interpretační proteiny patří VIM a SUVH rodina proteinů. VIM a SUVH mají společné to, že pro interpretaci DNA disponují SRA doménou, kterou interagují a vybírají sekvenci nukleotidů obsahující metylovaný cytosin. VIM se váže k sekvencím C\*G i k C\*HG (Woo et al. 2007). Členové lépe popsané SUVH rodiny se účastní *de novo* i udržovací metylace a někteří jsou zároveň histon metyltransferázami, což z nich činí multifunkční enzymy stojící na křižovatce epigenetických (metylových) značek na molekulách DNA a histonů (Johnson et al. 2007), a proto jim bude věnována samostatná kapitola. Nejprve je nutné ale představit další mechanismy, abych mohl lépe popsat, co se kde propojuje.

### 4.3. Udržování a *de novo* metylace

Metylace DNA se dá rozdělit na dva základní procesy: udržování již existujícího vzoru metylace DNA při/po replikaci a nová metylace předtím nemetylovaných oblastí (tzv. *de novo*). Již zavedenou metylaci DNA je třeba udržovat především kvůli udržení transponovatelných elementů v umlčeném stavu, ale udržovací metylace také pomáhá zachovávat identitu buněk při diferenciaci. Způsob, jakým se udržuje metylace, se odvozuje od toho, v jakém sekvenčním kontextu a v jakém typu chromatinu se nachází metylovaný cytosin. *De novo* metylace umožňuje buňce vytvořit si metylační vzor podle aktuálních

požadavků, čemuž napomáhá i schopnost metylované cytosiny aktivně odstraňovat. K tomu jsou buňky vybaveny odpovídajícím molekulárním arsenálem. *De novo* metylační značení rostliny zavádějí pro aktivní transponovatelné elementy, exogeny a také během vývoje, kdy se epigenetické úpravy podílejí na buněčné diferenciaci (Law a Jacobsen 2010). Dále se zavádějí během odpovědi na stres a je možné, že tak fungují jako adaptační mechanismus (Zhang et al. 2018).

#### **4.4. Cytosinové kontexty zdálky**

Cytosin se v genomu vyskytuje v sousedství různých nukleotidů, tedy v různých kontextech. Rostliny dokážou, na rozdíl od savců, metylovat cytosin ve všech možných kontextech, ale liší se mechanismy, kterými to dělají. Kontexty mohou být symetrické (polosymetrické), jako je to v případě CG a CHG (kde H je A, T nebo C), pro které enzymy udržovací metylace mohou využít informaci z metylovaného vlákna k její replikaci na nemetylovaném, a asymetrické, jako je to v případě CHH (ale do značné míry i CHG), kde jsou metylační enzymy naváděny buď modifikací histonů (H3K9me2), nebo pomocí malých RNA (sRNA) komplementárních k danému úseku DNA. RNA také navádí enzymy k metylaci *de novo* ve všech kontextech. Metylaci cytosinu je udržována místně specificky (především v závislosti na stavu chromatinu) v jednotlivých kontextech distinktivními mechanismy. Ty se ale často navzájem ovlivňují (Wendte a Schmitz 2018).

#### **4.5. DNA metyltransferázy**

Metylová skupina se v DNA rostlin přidává na cytosin, konkrétně na jeho 5' pozici prostřednictvím DNA metyltransferáz a donorem metylové skupiny je S-adenosyl-L-methionin (Zhang et al. 2018).

##### **4.5.1 MET1**

MET1 je metyltransferáza specifická pro CG kontext a je to homolog lidské DNMT1. Udržovací CG metylaci provádí bezprostředně po replikaci. Chromatinový remodelační komplex DDM1 je nutný pro metylaci CG v některých heterochromatinových lokusech. VIM1-VIM3 snadno rozeznávají symetrický C\*G/CG a navádí k ní MET1. Mutanti *vim1*, *vim2*, *vim3* individuálně metylaci CG vykazují beze změny, trojitý mutant ji však ztrácí (Woo 2008, Stroud et al. 2013). MET1 má také roli v metylaci CCG, kde nejdříve metyluje

CC\*G/C\*GG, a to rekrutuje SUVH5 a SUVH6, který tuto sekvenci umí rozeznat a vytvoří histonové značky H3K9me2. Ty poté rekrutují CMT3, a ta dometyluje i zbylý vnější cytosin na C\*C\*G/C\*GG (Yaari et al. 2015, Gouil a Baulcombe 2016). MET1 přímo interaguje s HDA6 a mutant v této histonové deacetyláze *hda6* vykazuje pokles CG metylace v některých lokusech, což napovídá, že MET1 pro svou aktivitu HDA6 v těchto lokusech vyžaduje (Wendte a Schmitz 2018).

#### 4.5.2 Rodina CMT

CMT jsou evolučně stará rodina DNA metyltransferáz rostlin. Mají CHROMO a BAH domény, které se váží k H3K9me2 a tato vazba podmiňuje jejich aktivitu. Dále mají metyltransferázovou aktivitu na cytosinu (Du 2013). Chromometylázy rostlin tvoří dvě oddělené skupiny, které vznikly pravděpodobně v rámci celogenomové duplikační události (WGD) u evolučního kořene krytosemenných rostlin. *Arabidopsis* disponuje CMT3 a CMT1, které se řadí do jedné skupiny, a CMT2, která je z druhé skupiny. Jejich funkce je udržování cytosinové metylace transposonů a dalších repetitivních sekvencí v pericentromerických oblastech. U *Arabidopsis* cílí na kontexty CHG (v případě CMT3) a CHH (v případě CMT2). Funkce těchto chromometyláz se v menší míře překrývá (Gouil a Baulcombe 2016). Členové CMT rodiny u jiných druhů mohou mít jinou specifitu pro dané kontexty, např. kukuřice má CMT pouze z první skupiny, ZMET2 a ZMET5, ale ty metylují CHG i CHH. Aktivita CMT má součinnost s histonovými metyltransferázami SUVH4/5/6 a demetylázou IBM1 (Jiao et al. 2011, Bewick et al. 2017, Wendte a Schmitz 2018).

CMT3 váže nukleosomy s H3K9me2, a tak propojuje modifikaci tohoto histonu s CHG sekvencí (Bewick et al. 2017). H3K9me2 však zřejmě vzniká i v nežádoucích místech. To by mohlo způsobit i nežádoucí metylaci CHG v přepisovaných genech. Zde to řeší histonová demetyláza IBM1 (viz níže), která H3K9me2 zase demetyluje jako chybně vloženou značku. Když to však IBM1 nestihne, H3K9me2 rekrutuje CMT3 a ta může nametylovat CGG/CCG a hrozí, že se zafixuje jako CG metylace aktivitou MET1 i po odstranění H3K9me2 IBM1. Mutant *ibm1* si časem zametyluje aktivně přepisované geny i v kontextu mCHG (Bewick et al. 2016, Wendte a Schmitz 2018). Na CMT3 ortologu kukuřice ZMET2 byl popsán mechanismus, kterým rozpoznává metylované histony. Ukazuje se, že oproti mononukleosomu s preferencí váže dinukleosomy s H3K9me, jeden svou BAH doménou a druhý svou CD. Methyltransferázová doména, která se nachází uprostřed, poté s preferencí metyluje propojovací DNA. U BAH domény se předpokládá alosterická regulace. ZMET2 tak

snad rozpoznává i vyšší organizaci chromatinu, a vyšší počet nukleosomů za sebou tak možná zvyšuje i aktivitu ZMET2 (Stoddard et al. 2019).

CMT2 metyluje nesymetrický kontext CHH v heterochromatinu. S preferencí metyluje CAA a CTA (Gouil a Baulcombe 2016). Umí metylovat i místa v heterochromatinu s H1, která ji zpřístupňuje remodelační komplex DDM1 (Zemach et al. 2013). Oproti CMT3, která nemá preferenci k počtu metylovaných značek na lysin, váže CMT2 s preferencí H3K9me2 oproti H3K9me1, a pravděpodobně je to způsobeno rozdíly v BAH a chromodoménách (Stroud et al. 2014, Du et al. 2015).

CMT1 u *A. thaliana* je pravděpodobně pseudogen, pokud má nějakou funkci, zatím nebyla objevena (Bewick et al. 2017).

### 4.5.3 RdDM a DRM2

RdDM Pomocí této dráhy se *de novo* metylují všechny kontexty. Také má roli v udržování CHH. Pokud rostlina ztratí metylaci cytosinu, může ji znovu získat tímto mechanismem. Klíčový pro tuto dráhu je přepis daného lokusu vedoucí ke vzniku malých RNA, které navádí metyltransferázu DRM2. Nejvíce cílových sekvencí RdDM je v evolučně mladších transposonech. RdDM se také podílí v odpovědi na stres, kdy rostlina epigeneticky mění genovou expresi, a při tvorbě epialel a imprintingu (Matzke a Mosher 2014). Pro RdDM je nutný přepis daného lokusu, a tudíž umlčuje pouze lokusy, které jsou přístupné RNA polymeráze V, např. transponovatelné elementy, aktivní exogeny a aktivní patogeny. Je velmi přesná a dokáže metylovat lokusy i v těsné blízkosti přepisovaných genů. Má klíčovou roli např. v umlčování LTR retrotransposonů *Gypsy*, které se vyskytují v pericentrickém heterochromatinu (Zemach et al. 2010).

RdDM můžeme dělit na dva sledy událostí: i) produkce naváděcích siRNA a ii) vlastní metylace. Na počátku kanonické RdDM dráhy je chromatin remodelován komplexem CLSY1 a Pol IV je přivedena k lokusu s H3K9me2 proteinem SHH1, který propojuje histonové modifikace s aktivitou Pol IV. Pol IV produkuje ncRNA, které slouží jako templát pro RNA dependentní RNA polymerasu 2 (RDR2), která produkuje dsRNA. Tyto dsRNA jsou rozstříhané dicer-like proteiny (DCL3) na 24-nukleotidové siRNA, které se váží především na AGO4 a AGO6 proteiny. Při *de novo* RdDM mohou být zdrojem malých RNA i transkripty Pol II (Cuerda-Gil and Slotkin 2016). Paralelně je RNA polymeráza V naváděna SUVH2 a 9 na metylovanou DNA. Přístup k ní (rozvolnění chromatinu) zajišťuje DRD1, který opět remodeluje chromatin a PolV tak může přepisovat lokusy určené pro metylaci a produkuje „scaffold RNA“, na které se vážou komplexy siRNA-AGO. Výsledný komplex scaffoldRNA-



siRNA-AGO poté navádí DRM2, která metyluje přepisované vlákno daného lokusu (Zhong et al. 2014, Thao et al. 2013, Matzke a Mosher 2014).

DRM2 metyltransferáza RdDM dráhy je homolog lidské DNMT3. Má rozpoznávací doménu (target recognition domain, TRD) a katalytickou doménu, kde probíhá přenos metylové skupiny. Substrátová DNA je umístěna do žlábků mezi těmito doménami. Cytosin určený k metylaci se vykroutí směrem k aktivnímu místu. V kanonické RdDM je DRM2 naváděna 24nt siRNA na homologní lokusy, kde zároveň H1 nebrání H3K9me2 k vázání s SHH1. Kromě udržovací metylace je to také *de novo* metyltransferázou. Pro svou aktivitu vyžaduje přítomnost svého efektorového proteinu rodiny AGO, který váže siRNA na daném locusu. Metyluje vlákno DNA, které bylo přepsáno Pol V (Zhong et al. 2014).

*Arabidopsis* je zvláštní v tom, že rostliny s nefunkční Pol IV a V (mutanti *nrpe1* a *nrpd1*) mají téměř nepozměněný vývoj. Analogický mutant rýže či kukuřice vykazuje velké vývojové poruchy. Vysvětluje se to mezidruhovými rozdíly v epigenetickém působení transposonů na genovou expresi. *Arabidopsis* má pro rostliny netypicky nízké zastoupení retrotransposonů (Kim 2017), přitom právě tyto elementy nejvíce ovlivňují expresi sousedních genů. Narušení RdDM u *Arabidopsis* má malý vliv na geny ovlivněné LINE/L1 elementy, velký naopak na ty ovlivněné RathE1. Je to však dáno tím, kde se dané transposony vyskytují, než v jejich vlastnostech (Gouil a Baulcombe 2016).

## 4.6. SUVH rodina proteinů

SUVH je rodina proteinů se SET and RING (SRA) doménou. Patří mezi ně tři histonové metyltransferázy, které propojují metylaci DNA a metylaci histonu H3. SUVH4 (který má synonymní označení KRYPTONITE, KYP) a SUVH5/6 jsou H3K9 metylázy (v trojitě mutantu *suvh4/suvh5/suvh6* chybí H3K9me2). Jsou schopné interpretovat metylační značky na DNA a používají je ke svému navádění. Svým produktem, H3K9me2, poté zase umožňují aktivitu CMT3, CMT2 a RdDM a vytvářejí tak zpětnovazebnou smyčku mezi metylovanou DNA a H3K9me2. Proteiny této rodiny mají různou afinitu k jednotlivým metylovaným cytosinovým kontextům (Li et al. 2018). Zároveň je jejich metyltransferázová aktivita podmíněna udržováním non-CG DNA metylace, protože trojitý mutant *drm2, cmt2* a *cmt3* ztrácí i metylaci H3K9 (Stroud et al. 2014).

Analýza vazby SUVH6 na metylovanou DNA ukazuje, že SUVH mají SRA doménu, která svým žlábkem může hydrofilně vázat sekvenci prvních čtyř nukleotidů nemetylovaného vlákna. Tu rozpoznává vodíkovými vazbami s fosfátovými skupinami. Také vytvoří

vodíkovou vazbu přímo s bázemi prvních dvou nukleotidů metylovaného vlákna a jejich palindromem na nemetylovaném. Metylovaný cytosin je vykroucen směrem do kapsi v SRA doméně. Vlákno je stabilizováno skrz malý DNA žlábek smyčky SRA domény, váže se tam glycin té smyčky s nyní nespárovanou G<sup>c</sup> bází.

SUVH4/5/6 metylují H3K9 pomocí katalytické kazety pre-SET/SET/post-SET na C-konci bez allosterických regulací. Je pro to zřejmě potřeba kofaktor navázaný mez SET a post-SET. V nepřítomnosti substrátu má SET doména autoinhibiční mechanismus - umí zablokovat své vazebné místo pro histon H3 interakcí se svou „two-helix bundle“ doménou (Li et al. 2018).

#### 4.6.1 SUVH4

Mutace v genu pro SUVH4 se projeví poklesem hladiny metylace CWG (kde W je A nebo T) až o polovinu více, než CCG. SUVH4 má v rámci CHG silnou preferenci pro C\*AG/C\*TG, C\*AT/ATC a C\*C\*G/C\*GG, a také pro hemimetylovaný C\*G/CG. Nízkou preferenci má pro CC\*G/C\*GG a pro plně metylovaný C\*G/C\*G (Li et al. 2018). U nukleosomů s histonem H3K9me2 metylovaným SUVH4 v blízkosti jak CAG/CTG, tak CGG H3K9me2 rekrutuje CMT3, a ta pak metyluje nejen CAG/CTG, ale i CCG. V osamělých CCG kontextech je ale pozorována nižší úroveň metylace CCG, protože SUVH4 k nim má nižší afinitu oproti SUVH5/6 (Gouil a Baulcombe 2016).

#### 4.6.2 SUVH5

SUVH5 se spolu se SUVH6 zástupně váží na kontext CCG. SUVH5 se nejslaběji váže k C\*AT/ATG, středně k C\*AG/CTG zato má preferenci pro C\*C\*G/C\*GG a CC\*G/CGG, a také C\*G/C\*G i (slaběji) C\*G/CG. Jako jediný z rodiny nemá SUVH5 autoinhibiční mechanismus (Li et al. 2018). Pokusy s odpovědí rostlin na světelné podmínky ukazují, že SUVH5 přispívá k procesu klíčení. Dimetylací H3K9 umlčuje expresi transkripčních faktorů genů navozujících dormanci semen, kyseliny abscisové spolu s jejími receptory a katabolických genů pro gibereliny (Gu et al. 2019).

#### 4.6.3 SUVH6

SUVH6 se váže stejně silně ke všem metylovaným kontextům, má tedy zřejmě redundantní funkci pro SUVH4 i SUVH5 ve všech kontextech. Kromě toho se s preferencí váže k přepisovaným repetitívám. SUVH6 má ve struktuře N-terminální svazek tvořený dvěma  $\alpha$  helixy („two-helix bundle“), který interaguje s nemetylovaným vláknem DNA. Je umístěn

uprostřed a po stranách jsou naproti sobě domény SRA a kazeta pre-set/SET/post-SET (Li et al. 2018).

#### **4.6.4 SUVH2 a SUVH9**

SUVH2 a SUVH9 nejsou metyltransferázy, chybí jim post-SET doména, která je spolu se SET aktivním místem enzymu. Jsou to ale významné komponenty RdDM dráhy. Pomocí svých SRA domén se váží k mDNA a asociují s DDR komplexem a fungují jako adaptorové proteiny navádějící Pol V do RdDM lokusů (Liu et al. 2014).

### **4.7. Cytosinové kontexty zblízka**

V této kapitole bych chtěl shrnout výskyt cytosinu v genomu a jeho interakci s metylačním aparátem. Epigenomové (metylomové) mapy *Arabidopsis thaliana* odhalují, že metylovaný cytosin se vyskytuje nejčastěji v CG kontextu (55%), dále CHG (23%) a CHH (22%) (Millar et al. 2008)

#### **4.7.1 CG**

CG metylace se v genomu vyskytuje nejvíce, a to v transposonech, repetitivních sekvencích, i v genových oblastech. Pokud jsou konstitutivně přepisované geny metylovány, je to výlučně v kontextu CG (popřípadě vnitřního cytosinu v CCG). Transposony jsou metylovány v CG i non-CG (Zabet et al. 2017). Metylace je replikována semikonzervativně - je udržována vazbou VIM1-3 na hemimetylovanou DNA, což aktivuje metyltransferázu MET1. Heterochromatinová oblast kolem centromer má velké zastoupení C\*G. Metylace CG v různých subkontextech (myslí se tím třeba CGA atd.) nejsou stejně zastoupeny (rozdíl je ale velmi malý ve srovnání s non-CG). U *Arabidopsis* je nižší zastoupení metylace v CGT, tento rozdíl se projevuje více v genech a transposonech (Gouil a Baulcombe 2016).

#### **4.7.2 Obecně o non-CG**

CHG a CHH se v genomu často vyskytují spolu (Millar et al. 2008). Tyto kontexty nemohou být udržovány stejně jako CG. Rostliny to dělají tak, že vytvořily pozitivní zpětnovazebnou smyčku mezi metylací H3K9me2 a non-C\*G (Li et al. 2018).

##### **4.7.2.1 CHG**

CHG hladina metylace se liší podle stavu chromatinu daného lokusu, rozdíl je mezi euchromatinem a heterochromatinem a i mezi jednotlivými subkontexty, v heterochromatinu je nejméně zastoupen C\*CG (o 20-50%). V euchromatinu *Arabidopsis* jsou metylované CHG

zastoupeny málo a přítomné rozdíly v subkontextech (jsou-li) jsou nerozeznatelné (Gouil a Baulcombe 2016). Udržování metylace CHG je do značné míry závislé na CMT3. Ke ztrátě CHG metylace také dochází z 85% u *met1* a *vim1/2/3* mutantech v místech, kde nedochází k obnovení CG metylace (Stroud et al. 2013).

CCG kontext po metylaci vypadá takto C\*C\*G/C\*GG a může být zároveň chápán jako CG i jako CCG kontext (Li et al. 2018). CCG je zvláštní i v jiných ohledech. Má ze všech CHG nejmenší hladinu metylace. Mutant *met1* téměř úplně přestane vykazovat metylaci CCG. Mutant *cmt3* má ten samý efekt, ale sníží navíc i úroveň metylace ostatních CHG. Mutant *suvh4* sníží disproporčně C\*WG, C\*C\*G jen málo. Methylace CCG je tedy MET1 i CMT3 dependentní (Gouil a Baulcombe 2016). Pro metylace vnějšího cytosinu je nutná metylace vnitřního, tedy kde není metylován vnitřní, není ani vnější. Neznamená to ale, že bude vnější všude metylován. Tam kde je metylován vnitřní cytosin, je vnější metylován ve 40 %. CCG kontext se tedy může *in vivo* z hlediska metylace cytosinu vyskytovat ve třech ze čtyř teoreticky možných podob: CCG, CC\*G, C\*C\*G (Gouil a Baulcombe 2016), výsledky z naší laboratoře ale ukazují, že v průběhu de novo RdDM naopak dochází k přednostní metylaci vnějšího cytosinu a přechodně se tedy vyskytuje i varianta C\*CG (Příbylová et al. submitted). C\*C\*G jsou častěji asociovány s H3K9me2. CC\*G se nachází převážně v genech a nevyskytuje se společně s H3K9me2. U některých body-metylated genů je to působením IBM1, která H3K9me2 demetyluje. Jsou tak uchráněny proti metylaci vnějšího cytosinu díky přerušení regulační smyčky s CMT3. Dále jsou méně často v transposonech (Zabet et al. 2017).

#### 4.7.2.2 CHH

CHH je metylován v méně kondenzovaném chromatinu bez H1 histonu pomocí RdDM. (Zemach et al. 2013). Zde je metylace stejně zastoupena ve všech subkontextech. Metyltransferáza dráhy RdDM DRM2 je naváděna sRNA a ne proteiny, jako CMT rodina, která je naváděna H3K9me2, na který jsou metylové značky vnášené prostřednictvím SUVH. Mutant *nrpe1* (podjednotka Pol V) nedokáže přepisovat lokusy určené pro metylaci, což narušuje RdDM, a má snížené zastoupení metylace ve všech subkontextech v chromozomových raménkách. Je však pozorovatelný rozdíl i na hranici pericentromerického heterochromatinu.

Jen malá část metylovaného CHH se však nachází v euchromatinu, většina se nachází v kondenzovanějším heterochromatinu s H1, kde je metylace zprostředkována remodelací pomocí DDM1 a následnou metylací prostřednictvím CMT2 (Zemach et al. 2013).

U *Arabidopsis* je v heterochromatinu nejvíce metylovaný subkontext CAA a CTA (35-40 %), oproti tomu CCC a CCT pouze do 8 %. V pericentromerickém heterochromatinu je také nejvíce zastoupen celkově. V roli metyltransferázy je zde CMT2, která má preferenci k CAA a CTA, a tak vzniká ona disproporce. Mutant *cmt2* vykazuje celkově nižší metylaci. Oblasti na hranici pericentromerického heterochromatinu jsou ovlivněny i RdDM, protože se zde hromadí siRNA z euchromatinu (Gouil a Baulcombe 2016).

## 5 Demethylace

Demethylace DNA probíhá samovolně nebo katalyzovaně. Zpětnovazebná smyčka mezi několika chromatinovými modifikacemi je robustní mechanismus, který zajišťuje stálé umlčení parazitických DNA elementů. Jeho nechtěným vedlejším produktem může být ale rozšíření metylací do oblastí, které umlčeny být nemají. Tyto špatně zanesené značky jsou proto aktivně odstraňovány (Wendte a Schmitz 2018).

### 5.1. DNA demethylace

K DNA demethylaci může docházet pasivně vlivem umlčení exprese metyltransferáz při vývoji samičích gametofytů nebo *in vitro* při nedostatku donoru metylových skupin, nebo aktivně působením enzymů.

*Arabidopsis* má DNA demethylázy vegetativních pletiv ROS1, DML2 a DML3, a DME, která je aktivní při vývoji gametofytů. Ve vegetativní buňce pylového zrna způsobuje přepis transposonů. Z těchto transkriptů jsou vyrobeny malé RNA, které putují do spermatických buněk, kde indukují DNA metylaci pomocí RdDM. DME je také aktivní před oplodněním centrální buňky, ze které vzniká endosperm (Zhang et al. 2018).

ROS1 je glykosyláza, která demetyluje cytosin mechanismem „base excision repair“ – tedy odstraňuje celou metylovanou bázi, místo níž je opravným mechanismem doplněn nemetylovaný cytosinový nukleotid. Aktivita ROS1 je regulována koordinačně regulační smyčkou: metylace promotoru genu *ROS1* (tzv. MEM sekvence) umožňuje jeho transkripci na rozdíl od ostatních promotorů, kde metylace transkripci umlčuje. Upstream od MEMs je transposon ze skupiny helitronů, který zřejmě láká metyltransferázy. Methylace promotoru ROS1 je zajištěna aktivitou *de novo* metylace RdDM. Vysoká aktivita RdDM (či nízká aktivita demetylačních enzymů) tak vede ke zvýšené aktivitě ROS1 a udržuje tak rovnováhu

v metylaci. Naopak nízká aktivita RdDM snižuje expresi ROS1 a tak brání přílišné hypometylaci. ROS1 má největší aktivitu u transposonů v blízkosti genů, které tím chrání před ektopickým efektem metyltransferáz (Williams et al. 2015).

## 5.2. Demetylace histonů

IBM1 je histonová demetyláza. Její exprese je umlčena v mutantech *met1-3*, protože exprese *IBM1* vyžaduje metylaci CG v intronu tohoto genu. IBM1 zřejmě konstantně odstraňuje nepatřičné represní značky H3K9me2, které se objevují v přepisovaných oblastech vlivem chybné „off-target“ metylace histonů, tak jak bylo popsáno v případě vzniku „gene-body metylace“. Po nastolení mutace *imbl* nepatřičné H3K9me2 značky způsobují rychlou metylaci CHG a umlčení genu. IBM1 je tak nutná v obraně rostliny proti ustanovení ektopické CHG a CHH metylace v přepisovaných genech, které by mohlo narušit fungování transkripce (Bewick et al. 2016)

## 6 Závěr

Arabidopsis (a potažmo) rostliny má dynamický molekulární aparát k nakládání se svým chromatinem. Umožňuje jí plasticky reagovat na aktuální podmínky, ve kterých se nachází, změnit svůj expresní profil, a po odeznění vnějšího impulzu se zpět navrátit do základního stavu, či snad vytvořit nový, lépe adaptovaný na dané prostředí. Zároveň jí umožňuje konzervativně udržovat stav chromatinu, který je nutný pro správné fungování genomu. Metylace cytosinu je klíčová součást tohoto aparátu, která recipročně ovlivňuje celý epigenom. Přihlédnutí k tomu, v jakém kontextu se metylovaný cytosin nachází, dobře napovídá, jakou má konkrétní metylační značka funkci. K vytvoření celkového obrazu epigenetické souhry je nutné popsat dílčí mechanismy, které se jí účastní. Jejich zmapování také povede k možnosti předvídat chování celého systému při pozměnění jeho částí a naopak zhodnotit, v jakém stavu se rostlina nachází. To může mít využití při plánování dalších experimentů, ale zejména nám to pomůže lépe pochopit fungování rostlinného organismu.

## Použitá literatura

- Bewick AJ, Ji L, Niederhuth CE *et al.*, 2016. On the origin and evolutionary consequences of gene body DNA methylation. *Proceedings of the National Academy of Sciences* **113**, 9111–9116.
- Bewick AJ, Niederhuth CE, Ji L *et al.*, 2017. The evolution of CHROMOMETHYLASES and gene body DNA methylation in plants. *Genome Biology* **18**, 1–13.
- Du J, Zhong X, Bernatavichute Y V. *et al.*, 2012. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* **151**, 167–180.
- Du J, Johnson LM, Jacobsen SE, Patel DJ, 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nature Reviews Molecular Cell Biology* **16**, 519–532.
- Gouil Q, Baulcombe DC, 2016. DNA Methylation Signatures of the Plant Chromomethyltransferases. *PLoS Genetics* **12**, 1–17.
- Grant-Downton RT, Dickinson HG, 2005. Epigenetics and its implications for plant biology. 1. The epigenetic network in plants. *Annals of Botany* **96**, 1143–1164.
- Grant-Downton RT, Dickinson HG, 2006. Epigenetics and its implications for plant biology 2. The „epigenetic epiphany“: Epigenetics, evolution and beyond. *Annals of Botany* **97**, 11–27.
- Gu D, Ji R, He C *et al.*, 2019. Arabidopsis Histone Methyltransferase SUVH5 Is a Positive Regulator of Light-Mediated Seed Germination. *Frontiers in Plant Science* **10**, 1–13.
- Jiao Y, Wickett NJ, Ayyampalayam S *et al.*, 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* **473**, 97–100.
- Jin C, Felsenfeld G, 2007. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes and Development* **21**, 1519–1529.
- Johnson LM, Bostick M, Zhang X *et al.*, 2007. The SRA Methyl-Cytosine-Binding Domain Links DNA and Histone Methylation. *Current Biology* **17**, 379–384.
- Kim NS, 2017. The genomes and transposable elements in plants: are they friends or foes? *Genes and Genomics* **39**, 359–370.
- Law JA, Jacobsen SE, 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics* **11**, 204–220.
- Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE, 2011. SHH1, a Homeodomain protein required for DNA Methylation, as well as RDR2, RDM4, and Chromatin remodeling factors, associate with RNA Polymerase IV. *PLoS Genetics* **7**.
- Law JA, Du J, Hale CJ *et al.*, 2013. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**, 385–389.

- Li X, Harris CJ, Zhong Z *et al.*, 2018. Mechanistic insights into plant SUVH family H3K9 methyltransferases and their binding to context-biased non-CG DNA methylation. *Proceedings of the National Academy of Sciences* **115**, E8793–E8802.
- Liu Z, Shao C, Zhang C *et al.*, 2014. The SET Domain Proteins SUVH2 and SUVH9 Are Required for Pol V Occupancy at RNA-Directed DNA Methylation Loci. **10**.
- Loidl P, 2004. A plant dialect of the histone language. *Trends in Plant Science* **9**, 84–90.
- Matzke MA, Mosher RA, 2014. RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15**, 394–408.
- Millar AH, Tonti-Filippini J, O'Malley RC *et al.*, 2008. Highly Integrated Single-Base Resolution Maps of the Epigenome in Arabidopsis. *Cell* **133**, 523–536.
- Moissiard G, Cokus SJ, Cary J *et al.*, 2012. MORC family ATPases required for heterochromatin condensation and gene silencing. *Science* **336**, 1448–1451.
- Příbylová A, Čermák V, Tyč D, Fischer L, 2019. Detailed INSIGHT INTO THE dynamics of the initial phases of de novo RNA-directed DNA methylation in plant cells. *submitted to Epigenetics & Chromatin*.
- Roudier F, Teixeira FK, Colot V, 2009. Chromatin indexing in Arabidopsis: an epigenomic tale of tails and more. *Trends in Genetics* **25**, 511–517.
- Sequeira-Mendes J, Araguez I, Peiro R *et al.*, 2014. The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *The Plant Cell* **26**, 2351–2366.
- Stoddard CI, Feng S, Campbell MG *et al.*, 2019. A Nucleosome Bridging Mechanism for Activation of a Maintenance DNA Methyltransferase. *Molecular Cell* **73**, 73-83.e6.
- Stroud H, Greenberg MVC, Feng S, Bernatavichute Y V, Jacobsen SE, 2012. Resource Comprehensive Analysis of Silencing Mutants Reveals Complex Regulation of the Arabidopsis Methylome. *Cell* **152**, 352–364.
- Stroud H, Greenberg MVC, Feng S, Bernatavichute Y V., Jacobsen SE, 2013. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**, 352–364.
- Stroud H, Do T, Du J *et al.*, 2014. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nature Structural and Molecular Biology* **21**, 64–72.
- Takuno S, Gaut BS, 2012. Body-methylated genes in arabidopsis thaliana are functionally important and evolve slowly. *Molecular Biology and Evolution* **29**, 219–227.
- Wada Y, Ohya H, Yamaguchi Y, Koizumi N, Sano H, 2003. Preferential de Novo Methylation of Cytosine Residues in Non-CpG Sequences by a Domains Rearranged DNA Methyltransferase from Tobacco Plants. *Journal of Biological Chemistry* **278**, 42386–42393.
- Wendte JM, Schmitz RJ, 2018. Specifications of Targeting Heterochromatin Modifications in Plants. *Molecular Plant* **11**, 381–387.



- Wendte JM, Zhang Y, Ji L *et al.*, 2019. Epimutations are associated with CHROMOMETHYLASE 3-induced de novo DNA methylation. *eLife* **8**, 1–27.
- Williams BP, Pignatta D, Henikoff S, Gehring M, 2015. Methylation-Sensitive Expression of a DNA Demethylase Gene Serves As an Epigenetic Rheostat. *PLoS Genetics* **11**, 1–18.
- Wollmann H, Stroud H, Yelagandula R *et al.*, 2017. The histone H3 variant H3.3 regulates gene body DNA methylation in *Arabidopsis thaliana*. *Genome Biology* **18**, 1–10.
- Woo HR, Pontes O, Pikaard CS, Richards EJ, 2007. Protein Required for Centromeric Heterochromatinization. *Genes & Development* **1**, 267–277.
- Yaari R, Noy-Malka C, Wiedemann G *et al.*, 2015. DNA METHYLTRANSFERASE 1 is involved in <sup>m</sup>CG and <sup>m</sup>CCG DNA methylation and is essential for sporophyte development in *Physcomitrella patens*. *Plant Molecular Biology* **88**, 387–400.
- Zabet NR, Catoni M, Prischi F, Paszkowski J, 2017. Cytosine methylation at CpCpG sites triggers accumulation of non-CpG methylation in gene bodies. *Nucleic Acids Research* **45**, 3777–3784.
- Zemach A, Grafi G, 2007. Methyl-CpG-binding domain proteins in plants: interpreters of DNA methylation. *Trends in Plant Science* **12**, 80–85.
- Zemach A, McDaniel IE, Silva P, Zilberman D, 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–919.
- Zemach A, Kim MY, Hsieh PH *et al.*, 2013. The arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205.
- Zhang H, Lang Z, Zhu JK, 2018. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology* **19**, 489–506.
- Zhong X, Du J, Hale CJ *et al.*, 2014. Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* **157**, 1050–1060.
- Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S, 2008. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456**, 125–129.