



HAL
open science

Réponse du méthylome suite à l'exposition au froid chez une espèce à génome complexe: le maïs (*Zea mays* ssp. *mays*)

Zeineb Achour

► To cite this version:

Zeineb Achour. Réponse du méthylome suite à l'exposition au froid chez une espèce à génome complexe: le maïs (*Zea mays* ssp. *mays*). Génomique, Transcriptomique et Protéomique [q-bio.GN]. Université Paris Saclay (COMUE), 2018. Français. NNT: 2018SACLS102 . tel-02130625

HAL Id: tel-02130625

<https://theses.hal.science/tel-02130625>

Submitted on 16 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Réponse du méthylome suite à l'exposition au froid chez une espèce à génome complexe : le maïs (*Zea mays ssp. mays*)

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°567 Sciences du Végétal : du gène à l'écosystème
Spécialité de doctorat : Biologie

Thèse présentée et soutenue le 15 mai 2018, à Gif-sur-Yvette, par

Zeineb ACHOUR

Composition du Jury :

Etienne BUCHER Directeur de recherche, INRA (UMR IRHS)	Rapporteur
Christoph GRUNAU Professeur, Université de Perpignan (UMR IHPE)	Rapporteur
Martin CRESPI Directeur de recherche, CNRS (IPS2)	Président
Mélanie JUBAULT Maître de conférence, AgroCampusOuest (UMR IGEPP)	Examinatrice
Leandro QUADRANA Chargé de recherche, CNRS (IBENS)	Examinateur
Clémentine VITTE Chargée de recherche, CNRS (UMR GQE-Le Moulon)	Directrice de thèse
Catherine GIAUFFRET Chargée de recherche, INRA (UR AgrolImpact)	Invitée



Remerciements

Avec ces remerciements je voudrais m'adresser à toute personne qui m'a aidée, d'une manière ou d'une autre, à mener à terme ce défi qui a duré 3 ans et demi, une expérience inoubliable, pleine d'enrichissements sur le plan scientifique comme sur le plan humain et personnel, des moments de partage et de belles rencontres.

Mes premiers remerciements s'adressent à ma directrice de thèse Clémentine VITTE, de m'avoir donné l'opportunité de concrétiser mon souhait de réaliser une thèse en biologie fondamentale malgré mon background principalement agronomique. Merci pour ses explications claires en épigénomiques et en bio-informatiques au démarrage de ma thèse et tout le long. Merci pour ses conseils et pour les discussions intéressantes, pour sa disponibilité et pour ses nombreuses relectures et corrections de mon manuscrit. Merci d'avoir été à l'écoute et de m'avoir soutenue surtout à la fin pendant les moments de doute. Un grand merci !

Je remercie vivement Etienne BUCHER et Christoph GRUNAU d'avoir accepté d'être les rapporteurs de ma thèse. Je remercie également Mélanie JUBAULT, Leandro QUADRANA et Martin CRESPI d'avoir accepté d'être les examinateurs pour mon travail. Merci aussi à Catherine GIAUFFRET d'avoir accepté l'invitation pour participer à mon jury de thèse.

Je tiens aussi à remercier les membres de mes deux comités de suivi pour les doctorants, Catherine GIAUFFRET, Angélique DELERIS, Jean-Philippe PICHON, Stéphane MAURY, Marianne DELARUE et Frédéric CHOULET. Merci d'avoir fait le déplacement pour le comité malgré la distance, merci pour les discussions intéressantes/bienveillantes que nous avons eues et les idées que vous m'avez données. Merci également à la direction de l'école doctorale pour l'organisation de formations intéressantes, ainsi que leurs encouragement et soutien pour les doctorants. Je pense en particulier à Marianne DELARUE et Martine FOURNIER. Merci à l'association DOC' en herbe d'organiser des rencontres scientifiques et aussi des moments de convivialité entre doctorants.

Merci à la mission universitaire de Tunisie de m'avoir donné une bourse qui m'a permis de réaliser ma thèse. Merci au projet 'Amaizing' d'avoir financé mes travaux de recherche.

Pendant ma thèse j'ai eu l'opportunité de collaborer avec d'autres laboratoires et d'avoir de belles rencontres. Un grand merci à tous ces gens qui m'ont beaucoup aidé pour la réalisation de plusieurs étapes pendant ma thèse :

UR AgrolImpact, Estrées-Mons, Péronne

A Catherine GIAUFFRET pour le temps que tu m'as consacré pour m'expliquer les réponses physiologiques du maïs au froid, pour m'avoir appris à faire des dissections sur des plantules de maïs ainsi que les réflexions sur le dispositif expérimental !

UE GCIE, Estrées-Mons

A Dominique RABIER, Hélène SELLIER, Benoit DECAUX et Jean-François HU pour votre aide technique sur l'installation des deux cultures, les prélèvements et les croisements, et le suivi des conditions de la serre et de la chambre climatisée. Pour tout ce que vous m'avez appris sur la gestion de la culture du maïs en serre : installation, prélèvements, autofécondations, croisement !

Biogemma, Chappes

A Magalie LEVEUGLE, Hervé DUBORJAL et Jean-Philippe PICHON pour les différentes réunions et discussions pour la réalisation du BS_seq, du SeqCapEpi et du mapping. Merci pour les explications sur certains aspects méthodologiques et bioinformatiques. Merci d'avoir fait au mieux pour respecter les délais avec ma thèse malgré vos agendas assez chargés !

IPS2, Orsay

A Véronique BRUNAUD, Christine PAYSANT-LE ROUX, Ludivine SOUBIGOU-TACONNAT, José CAIUS pour la réalisation de la partie paillasse jusqu'au mapping des reads pour l'analyse du mRNAseq. Merci d'avoir fourni ces données assez rapidement pour que je puisse les intégrer dans ma thèse !

Réseau RTP3E pour avoir organisé des formations pour les jeunes débutants en épigénétiques et des journées thématiques intéressantes.

J'arrive maintenant au **Moulon** :

Merci à tous les membres de mon équipe DyGAP ; Maud TENAILLON pour sa bonne humeur et son dynamisme ! A ma collègue du bureau Karine ALIX pour son enthousiasme, ses conseils et son soutien ! A Johann JOETS, pour les explications et son aide sur la réalisation des analyses bio-informatiques. A Domenica MANICACCI pour les explications statistiques. A Agnès ROUSSELET et Martine Le Guilloux pour les extractions d'ADN et d'ARN. Agnès ROUSSELET et Hélène CORTI merci pour les déplacements à Mons avec moi. A Philippe BRABANT, Pierre GERARD, Amandine CORNILLE.

A Arnaud DESBIEZ-PIAT, je te souhaite une bonne continuation pour ta thèse.

A Natalia MARTINEZ, merci pour ta bonne humeur et ton soutien ; merci de m'avoir écoutée quand je ne disais pas forcément de choses intéressantes. Je te souhaite bon courage pour la fin de ta thèse. Coraje vas a llegar !

Je tiens à remercier toute l'équipe administrative du moulon. Merci à Valérie LESPINAS, à Rozenn LE GUYADER, à Cyril MARTIN, Khalid OUSKOU, Stéphanie AVODAGBE, Frédéric HUYNH et Djibril POUYE. Merci d'avoir toujours trouvé des solutions rapides pour la gestion des trucs de dernières minutes !

Un immense merci à tous les gens qui m'ont aidée et soutenue, au sein du Moulon, ne serait-ce que par un petit mot d'encouragement !

Le Moulon n'est pas seulement un lieu de travail agréable, c'est un lieu de belles rencontres aussi. Merci à Valérie C. pour sa sympathie et pour ses délicieux gâteaux qui ne manquent pas pour les différentes fêtes. Merci à tous les jeunes pour la bonne ambiance et l'humour qui permet de décompresser quand l'esprit voit trouble ! A Anthony, Marianyela, Natalia, Christophe, Jean-Tristan, Héloïse, Charlotte, Dorian, Pierre et Yannick. Merci à Gaëlle, Emma, Julie, Alice et Emmanuelle pour les séances de gym ! A Margaux-Alison et Fabien pour leur soutien en fin de thèse !

Mes derniers remerciements s'adressent à ma famille. Merci à mes parents, ma sœur, mon frère et toute ma famille qui ont su me soutenir malgré la distance ! Merci d'avoir cru en moi. Merci à ma belle-famille également pour leur encouragement et leur soutien pendant la dernière ligne droite et spécialement à ma belle-mère pour ma part de gâteau que tu m'envoyais quand je ne pouvais pas être là !

A Mehdi pour sa présence, sa patience, son écoute et sa compréhension particulièrement pendant l'étape de rédaction. Merci d'avoir supporté et bien géré des moments de stress intense. Merci de savoir changer ma mauvaise humeur quand ce n'était pas évident de le faire. Merci pour plein d'autres choses ...

Enfin, merci encore une fois à tous ceux qui m'ont aidée, épaulée et encouragée pour finir ma thèse !

Liste des abréviations

ABA : Abscissic Acid
Ac/Ds : Activator/Dissociation
Adh : Alcohol dehydrogenase
ADN : Acide DésoxyriboNucléique
ADNc : Acide DésoxyriboNucléique complémentaire
AFLP : Amplification Fragment Length Polymorphism
AGO : ARGONAUTE
ARN : Acide RiboNucléique
BAH : Bromo Adjacent Homology
BNS : Bonsaï
bZIP : basic region leucine Zipper Protein
C : Cytosine
CDK : Cyclin Dependent Kinase
CENH3 : Centromeric Histone3
CMT : Chromo Methyltransferase
CNI : Carbon/Nitrogen Insensitive
Cnr : Colorless non-ripening
CO2 : Dioxyde de carbone
CTCF : CCCTC-binding factor
Cyc : Cycloidea
dCas : dead CRISP associated
DCL : Dicer-like nuclease
DDM : Decrease in DNA Methylation
DMCs : Differentially Methylated Cytosines
DML : DEMETER-LIKE
DMRs : Differentially Methylated Regions
Dnmt : DNA methyltransferases
DRM : Domain Rearranged Methyltransferase
dsRNA : double strand RiboNucleic Acid
epiRIL : epigenetic Recombinant Inbred Line
gbM : Gene body Methylation
GO : Gene Ontology
GWAS : Genome Wide Association Studies
HDAC : histone deacetylase
HpaII : enzyme de restriction obtenue à partir de *Haemophilus*
IAP : intracisternal A-particle
IBM : Increased in Bonsai Methylation
InDels : Insertions/Deletions
Kyp : Kryptonite
Ler : Landsberg erecta

LINE : Long Interspersed Nuclear Element
IncARN : Long non coding Acide RiboNucléique
LTR : Long Terminal Retrotransposon
MBD : Methyl Binding Domain
McrBC : Methylcytosine restriction enzyme
MeDIP : Methylation DNA Immuno-Precipitation
MET : Methyltransferase
MspI : Methylation specific PCR
NaCl : Chlorure de Sodium
NRPD : Nuclear RNA Polymerase D
osDNMT : Oryza sativa DNA methyltransferases
osMET : Oryza sativa Methyltransferase1
PCR : Polymerase Chain Reaction
Pol : ARN Polymérase
qRT-PCR : quantitative Reverse Transcription Polymerase Chain Reaction
QTL : Quantitative Trait Loci
RdDM: RNA directed DNA Methylation
RDR : RNA-dependent RNA polymerase
RILs : Recombinant Inbred Lines
RMR : Required to Maintain Repression
ROS : Repressor Of Silencing
RT-PCR : Reverse Transcriptase Polymerase Chain Reaction
SINE : Short Interspersed Nuclear Element
siRNA : small interfering RiboNucleic Acid
SNP: Single Nucleotide Polymorphism
SUVH : SUppressor of Variegation 3-9-Homologous
T : Thymine
ET: Element Transposable
TET : Ten Eleven Translocation
TSS : Transcription Start Site
TTS : Transcriptional Termination Sites
U : Uracil
VIM : Variant in Methylation
WGBS : Whole Genome Bisulfite Sequencing
ZmCBF : Zea mays Core Binding Factor
ZmCDPK : Zea mays calcium dependent protein kinase
ZmCLC-D : Zea mays Chloride Channel D
ZmCPK : Calcium-Dependent Protein kinase
ZmDREB : Zea mays Dehydration Responsive Element Binding Protein
ZmMPK : Zea mays Mitogen Activated Protein Kinase
ZmRCI : Zea mays Rare Cold Inducible
ZmRLK : Zea mays Receptor Like Protein kinase

Table des matières

Introduction générale.....	1
Chapitre I. Introduction.....	3
I. Le froid, une contrainte abiotique pour le maïs.....	3
1. Notion de contrainte et de stress.....	3
2. Le maïs, une plante tropicale adaptée au climat tempéré.....	4
3. Réponse physiologique du maïs au froid	5
II. Organisation et fonctions de l'épigénome.....	7
1. La chromatine organise l'information génomique.....	7
1.1. Structure et fonctions de la chromatine	7
1.2. Bases moléculaires des états chromatinien.....	11
2. Mécanistique de la méthylation de l'ADN.....	13
2.1. Mise en place.....	13
2.2. Maintenance.....	15
3. Profil de méthylation des différentes composantes génomiques	18
3.1. Méthylation des gènes	18
3.2. Méthylation des éléments transposables	22
3.3. Méthylation des centromères.....	25
3.4. Méthylation le long des chromosomes	26
III. Origines de la variation du méthylome	29
1. Origine génétique.....	29
2. Régulation développementale	33
3. Réponse du méthylome aux conditions environnementales.....	34
3.1 Réponse sous contrainte	34
3.1.1. Analyses ciblées sur des locus choisis	34
3.1.2. Analyses à l'échelle du génome	35
3.1.3. Réponse en association avec d'autres marques de la chromatine	36
3.2. Maintien et transmission de la réponse induite	38
IV. Le maïs, un modèle d'étude pour analyser l'impact de l'environnement sur le méthylome ..	40
1. Organisation du génome	40

2. Organisation du méthylome.....	45
3. Spécificité des voies de régulation de la méthylation.....	45
4. Variabilité intra-spécifique du génome et du méthylome.....	46
5. Impact des contraintes abiotiques sur la méthylation chez le maïs.....	49
6. Le maïs : qualités et contraintes.....	49
V. Analyse du méthylome : défis et techniques dédiées.....	49
VI. Présentation des objectifs de la thèse.....	55
Préambule. Vue d'ensemble de l'expérimentation mise en place pour répondre aux grands axes de la thèse.....	57
1. Choix des géotypes et des conditions de culture.....	57
2. Mise en place du dispositif expérimental.....	58
3. Prélèvements.....	59
Chapitre II. Le froid prolongé induit des variations de la méthylation de l'ADN chez une lignée de maïs sensible au froid.....	61
1. Contexte et objectifs.....	61
2. Achour <i>et al.</i> en préparation pour soumission dans <i>Genome Biology</i> : Prolonged cold in maize triggers genome-wide hypermethylation and local strong methylation variation.....	62
3. Conclusion et perspectives.....	116
Chapitre III. Dynamique de la variation de la méthylation de l'ADN au cours du développement et entre deux générations.....	119
I. Contexte et objectifs.....	119
II. Choix du matériel végétal à prélever.....	120
1. Prélèvements dédiés à l'analyse de la maintenance des niveaux de méthylation au cours du développement.....	120
2. Prélèvements dédiés à l'analyse de la maintenance des niveaux de méthylation dans la descendance (génération G1).....	122
III. Suivi des variations de méthylation pour le géotype sensible B73.....	123
1. Sélection des régions génomiques à analyser.....	123
2. Choix des échantillons à analyser.....	127
3. Analyses à réaliser à partir des données.....	128
Chapitre IV. Comparaison de la réponse du méthylome au froid entre trois géotypes contrastés pour leur réponse phénotypique.....	133
I. Contexte et objectifs.....	133
II. Résultats préliminaires et discussion.....	135

Chapitre V. Discussion générale et perspectives	141
Annexe 1. Analyse bibliographique des régions génomiques impliquées dans la réponse aux contraintes abiotiques chez le maïs.....	148
Annexe 2. Détail des dispositifs expérimentaux.....	153
Annexe 3. Mise au point de la détection de régions différentiellement méthylées entre échantillons de maïs après un stress abiotique	163
Annexe 4. Formations suivies au cours de ma thèse et valorisation de mes travaux	173
Références bibliographiques.....	175

Introduction générale

Au cours de ces vingt dernières années, la génomique est passée de l'analyse d'un génome de petite taille (157Mb, Arabidopsis Genome Initiative, 2000) à la comparaison de nombreux génomes de plus grande taille (par exemple ceux du genre *Oryza* (267 à 389 Mb, Stein *et al.* (2018)). La comparaison des génomes des différentes espèces a permis de mettre en évidence une dynamique importante des génomes en terme de taille, de ploïdie, de composition en gènes et en éléments transposables. La comparaison d'individus de la même espèce a également permis de révéler la nature extrêmement plastique du génome sur des échelles de temps courtes, sous forme de variants nucléotidiques (SNP), de petits indels, de variants structuraux de plus grande taille (duplications, délétions, inversions), et d'insertion différentielle d'éléments transposables. Seulement 8 ans après la séquence génomique d'*Arabidopsis thaliana*, un méthylome de ce génome a été séquencé (Cokus *et al.*, 2008, Lister *et al.*, 2008), et des méthylomes de nombreuses espèces ont été caractérisés depuis. Aujourd'hui, caractériser le génome et plusieurs méthylomes d'une espèce non modèle comme le pommier est réalisable en quelques années (Daccord *et al.*, 2017).

La caractérisation conjointe des génomes et des méthylomes de plusieurs espèces a permis de caractériser les profils de méthylation des gènes et des éléments transposables, et de mieux comprendre le lien qui existe entre la structure du génome et la composition du méthylome. Par exemple, la méthylation des éléments transposables est très différente entre des génomes où gènes et éléments transposables sont souvent associés, comme chez le maïs, et des génomes où les éléments transposables sont essentiellement localisés dans les régions péricentromériques, comme *A. thaliana*. En parallèle, le développement de mutants d'*A. thaliana* a permis de caractériser les grandes voies impliquées dans la mise en place et le maintien de la méthylation sur les différents composants du génome. Actuellement, les travaux de recherche s'orientent vers la caractérisation du maintien et des originalités de ces voies entre espèces (Niederhuth *et al.*, 2016), afin de faire le lien entre des organisations (épi)génomiques plus complexes et les mécanismes de régulation de la méthylation sous-jacents.

La méthylation de l'ADN est une marque de la chromatine sensible aux signaux de l'environnement. En particulier, elle peut être modifiée par des contraintes abiotiques. A l'inverse des mammifères chez lesquels la méthylation est réinitialisée au cours de la reproduction, chez les plantes les variations de la méthylation de l'ADN sont plus facilement héritées entre générations (Quadrona & Colot, 2016). En conséquence, des modifications de la méthylation induite par une contrainte abiotique, si elles sont maintenues après retour à des conditions « normales », pourraient être transmises à la descendance, générant ainsi des « épiallèles ». Ces épiallèles pourraient être sélectionnés et ainsi participer à l'adaptation des individus à leur environnement.

La caractérisation de nouveaux épiallèles induits par des contraintes environnementales nécessite la mise en place de dispositifs contrôlés complexes, l'analyse parallèle de plusieurs individus et de plusieurs générations, et la génération de méthylomes de grande qualité permettant de détecter des signaux faibles. De plus, l'épigénome étant une combinaison de multiples facteurs régulateurs (ARN non codants, méthylation de l'ADN, modifications post-traductionnelle des histones), il peut être

intéressant d'analyser simultanément plusieurs de ces marques afin de déterminer si les changements de méthylation détectés sont liés à une réponse directe, ou indirecte, à la contrainte.

Mon projet de thèse s'inscrit dans ce cadre scientifique et dans un contexte contemporain où la contribution des épialèles à l'adaptation est une question majeure. En utilisant le maïs comme modèle d'étude, les travaux de recherche décrits dans ce manuscrit ont pour objectifs de comprendre la réponse du méthylome au froid ainsi que d'étudier la transmission de cette réponse au cours du développement et des générations suivantes. Ces questions seront abordées dans les différents chapitres et leur réponse sont à différents niveaux d'avancement. Mes résultats apportent une vision nouvelle quant à l'implication de la structure du génome dans le modelage du méthylome en réponse à une contrainte abiotique.

Du point de vue théorique, mon travail de thèse m'a permis d'acquérir des connaissances, notamment sur la composition et la structure des génomes végétaux, les mécanismes impliqués dans la mise en place et la maintenance de la méthylation de l'ADN, et la réponse physiologique aux contraintes abiotiques. Au niveau technique, j'ai pu apprendre à mettre en place toutes les étapes, de la réflexion sur le plan expérimental à l'analyse des données. Ceci m'a permis de me former à différents aspects de la biologie du maïs comme la culture en milieu contrôlé, la dissection des méristèmes, la mise en place des prélèvements, la réalisation de croisements et d'autofécondations, et à l'analyse de données de BS-seq, pour lesquelles j'ai appris à programmer en bash, en python, et en R. J'ai aussi appris à analyser mes données en les confrontant aux données déjà disponibles dans la bibliographie.

Le manuscrit que je présente ici est divisé en cinq chapitres, précédés d'une introduction générale et suivis de quatre annexes. Le chapitre 1 correspond à l'introduction. Elle présente d'abord le contexte scientifique de l'étude du froid chez le maïs. Je présente ensuite les caractéristiques moléculaires et fonctionnelles de la chromatine ainsi que les facteurs à l'origine de modifications du méthylome. Par la suite, j'expose les caractéristiques propres au maïs, puis j'expose les enjeux relatifs à l'analyse de la méthylation de l'ADN chez les plantes. Je présente enfin la problématique de mon sujet de thèse ainsi que les objectifs abordés au cours de ce travail. Le préambule présente le choix des génotypes utilisés pour ce travail et les questions qui ont conduit au choix du dispositif expérimental. Le chapitre 2 est présenté sous forme d'un article en cours de préparation. Il présente les résultats de l'étude comparative des méthylomes entre plantes « stressées » après séjour au froid et « non stressées » pour une lignée sensible au froid, B73. Le chapitre 3 a pour objectif le suivi des variations de la méthylation de l'ADN au cours du développement et dans la descendance chez B73. Le chapitre 4 pose la question de la relation entre les variations du méthylome et du phénotype et présente des résultats préliminaires sur la comparaison de la réponse du méthylome de trois génotypes de maïs (B73, F2, et F331) ayant une réponse phénotypique contrastée au froid. Le chapitre 5 présente une discussion générale qui souligne l'intérêt de mon travail et présente les perspectives qu'il reste à aborder pour mieux comprendre la réponse du méthylome au froid et son application en amélioration variétale. L'annexe 1 présente une revue bibliographique des gènes de réponse au froid chez le maïs, que j'ai réalisée au cours de ma thèse. L'annexe 2 présente en détail l'installation de l'expérimentation des trois génotypes de maïs. L'annexe 3 présente la méthode que j'ai développée pour la détection des régions différentiellement méthylées (DMR) utilisée dans les chapitres 2, 3 et 4. L'annexe 4 présente la valorisation de mes travaux de recherche dans des conférences nationales et internationales ainsi que les formations que j'ai suivies pendant ma thèse.

Chapitre I.

Introduction

I. Le froid, une contrainte abiotique pour le maïs

1. Notion de contrainte et de stress

La notion de stress en biologie a été définie par assimilation à la physique, suite à l'observation de la déformation d'un fil en acier sur lequel est positionnée une charge. Celui-ci développe une tension (« strain »), dont la mesure correspond à la déformation du fil, qui peut être irréversible (plastique) ou réversible (élastique). La force de la tension exercée est proportionnelle à la taille de la charge et de sa position sur le fil. La traduction en français rend confuse la terminologie du 'stress'. Par abus de langage, on appelle « stress » la contrainte appliquée. Cependant, le stress est en fait la réponse suite à une situation de contrainte. Ainsi, la contrainte provoque le « stress » (Levitt, 1972).

Une contrainte est définie comme un ensemble de conditions environnementales induisant une réduction en termes de croissance et par rapport à celle de l'individu dans des conditions optimales de croissance. Ainsi, le stress (le « strain » ou la « déformation ») peut également être défini comme toute perturbation de la croissance réduisant la production de biomasse en conditions limitantes de ressources (Cramer *et al.*, 2011, Evan, 2004, Jackson, 1986). La contrainte peut être soit de nature abiotique, d'origine physique (par exemple la sécheresse, la température, les blessures) ou chimique (par exemple la salinité, le pH du sol, la présence de métaux lourds), soit de nature biotique (agressions par les pathogènes, compétition entre plantes). En réponse à ces contraintes, la plante induit des modifications physiologiques et biochimiques permettant sa survie. Mais cela a un coût, qui réduit la croissance. Par exemple, la baisse importante de l'assimilation photosynthétique du CO₂ atmosphérique est observée chez le maïs en condition de stress froid. La réponse entraîne la fermeture des stomates, puis la réduction des activités enzymatiques. Le ralentissement du transfert des électrons puis la dégradation du photosystème Ils sont également observés en cas de stress sévère. Ceci se manifeste par une chlorose au niveau des feuilles (Long & Spence, 2013, Fracheboud *et al.*, 1999). Le mécanisme de synthèse de phytohormones peut aussi être affecté en réponse au stress. Par exemple, la teneur en acide abscissique peut atteindre un niveau 10 à 30 fois supérieur à la normale chez des plantes de maïs soumises à un stress hydrique (Leung *et al.*, 2012).

La réponse à une contrainte environnementale peut se manifester sous forme de tolérance ou de sensibilité. La tolérance correspond à la capacité des plantes à supporter les contraintes auxquelles elles sont exposées. Par exemple, par l'élévation de la viscosité du cytoplasme des cellules, la protection des enzymes et des membranes par certains antioxydants et osmoprotectants, et par la modification de la composition phospholipidique des membranes cellulaires. Un génotype sensible est incapable de développer des mécanismes de défenses face aux agressions auxquelles il fait face (Calatayud *et al.*, 2013). A titre d'exemple, en réponse au froid chez le maïs, les génotypes tolérants sont capables d'accumuler plus d'acide abscissique et de manière plus rapide que les génotypes sensibles (Riva-Roveda & Périlleux, 2015).

La réponse d'une plante aux contraintes varie dans le temps (immédiate ou tardive) et dans l'espace (plante entière, un organe, un tissu, un type cellulaire). Elle est contrôlée par un réseau complexe guidé par de multiples voies de signalisation (Blum, 2015, Maron *et al.*, 2016). Une même contrainte peut induire différentes réponses suivant le génotype, l'organe affecté, le moment d'exposition des plantes à la contrainte, ou son intensité. Inversement, différentes contraintes peuvent engendrer une réponse physiologique similaire. Par exemple, la fermeture des stomates est observée en cas de stress froid ou de sécheresse.

2. Le maïs, une plante tropicale adaptée au climat tempéré

Le maïs (*Zea mays ssp. mays*) a été domestiqué à partir de la téosinte, sous-espèce *Zea mays ssp. parviglumis*, il y a environ 9000 ans dans la vallée du Balsas au Mexique (Matsuoka *et al.*, 2002, Piperno *et al.*, 2009). Après la phase initiale de domestication, l'Homme a répandu le maïs vers le Nord et vers le Sud du continent Américain (Doebley *et al.*, 1988). Ceci s'est traduit par l'adaptation des populations traditionnelles à des climats variés, et notamment les climats tempérés des latitudes extrêmes : Argentine/Chili au Sud et Nord des USA/Canada au Nord. L'adaptation aux températures basses et aux périodes de jours courts s'est essentiellement accompagnée par la sélection de maïs de floraison précoce (Malvar *et al.*, 1997, Trzcinska-Danielewicz *et al.*, 2009). Elle a conduit à un raccourcissement du cycle de vie de la plante. Cependant, le maïs reste particulièrement sensible à la température et il est considéré comme une espèce sensible au froid exigeant des températures relativement élevées pour assurer la germination (>10°C) et la croissance végétative (>15°C) (Riva-Roveda & Périlleux, 2015). Alors que l'arrivée du maïs dans les régions de basses altitudes dans le sud de l'Amérique (Argentine, Chili) est datée à 6000 ans (Piperno *et al.*, 2009), son expansion dans les régions de hautes altitudes des Andes est plus récente (il y a 4000 ans) (Grobman *et al.*, 2012, Perry *et al.*, 2006). La transition à partir des basses altitudes (zone de domestication) vers les zones de hautes altitudes s'est accompagné par une adaptation à la baisse de température, à l'augmentation en radiations UV et à la réduction partielle de la pression des gaz atmosphériques (Körner, 2007). Cette adaptation a été réalisée par plusieurs populations de façon convergente, suite aux différents événements d'introductions dans les régions de hautes altitudes (Takuno *et al.*, 2015).

Enfin, l'introduction du maïs en Europe à la fin du XV^{ème} siècle s'est faite à partir de plusieurs origines américaines (Rebourg *et al.*, 2003). En particulier, le matériel adapté au climat tempéré (cornés du Nord ou « Northern Flints » issu du Nord des USA/Canada) a été introduit en Europe du Nord entre 1492 et 1539, et le matériel tropical des Caraïbes (issu d'Amérique du Sud, Doebley *et al.*, 1988) a été introduit en Espagne par Christophe Colomb en 1493. Des similitudes entre des variétés italiennes traditionnelles et des variétés d'Amérique du sud suggèrent une troisième introduction du maïs en Europe à travers l'Italie (Brandenburg *et al.*, 2017). L'hybridation de ces deux types de matériel, suivie de sélection locale, a conduit à une diversification induisant l'émergence d'un groupe génétique spécifique au maïs Européen, le groupe de cornés Européens (« European Flints ») (Camus-Kulandaivelu *et al.*, 2006). L'ensemble de ce parcours est retracé dans la Figure I.1.

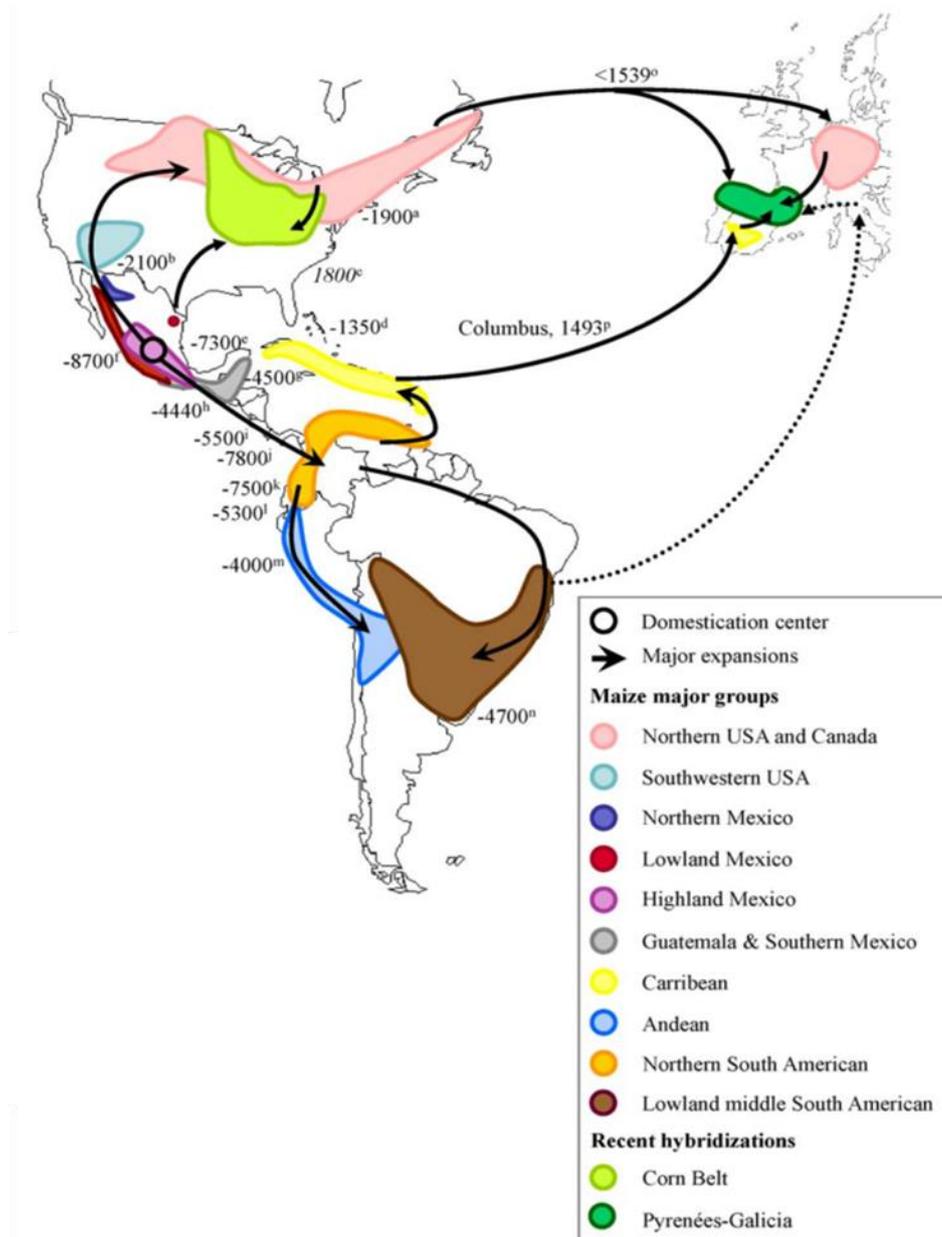


Figure I.1 : Diffusion du maïs en Amérique et en Europe (Tenaillon & Charcosset, 2011)

Les flèches montrent les différentes migrations du maïs depuis son centre de domestication (Mexique) en Amérique et les 3 introductions en Europe. Les dates négatives présentent des nombres d'années calculés par rapport au présent. Les dates positives correspondent à des années. Les couleurs présentent les groupes majeurs du maïs.

3. Réponse physiologique du maïs au froid

La réponse du maïs au froid est liée à trois composantes : la sévérité du froid, la durée du froid, et le stade de développement de la plante au moment de son exposition au froid. En termes de température, on distingue le stress « modéré » (10-15 °C) et le stress « sévère » (2-10 °C) (Leipner & Stamp, 2009). Certains auteurs décrivent également un état de stress « doux » (réduction de la photosynthèse et du développement) pour des températures de 12-17°C (Marocco *et al.*, 2005). En termes de durée, on distingue le froid « court » (quelques jours) et le froid « prolongé » (plusieurs semaines). En termes de développement, on distingue les stades de germination, d'hétérotrophie et

d'autotrophie. Une étape sensible est le passage de la plantule à l'état d'autotrophie entre les stades 3 et 5 feuilles visibles, qui peut compromettre le fonctionnement du méristème et qui est également importante pour la transition florale pour le matériel précoce (Hund *et al.*, 2004, yGreaves, 1996). Au niveau racinaire, le froid (arrosage avec de l'eau à 6°C et mesure phénotypique après 24h) affecte la biomasse racinaire, qui peut provoquer la chute de la conductance hydraulique au niveau des racines et, par suite, un stress hydrique (Melkonian *et al.*, 2004). Une modification du système reproducteur est observée essentiellement quand la plante est soumise au froid au moment de la transition florale. Le froid induit alors une diminution du nombre de ramifications (Bechoux *et al.*, 2000, Louarn *et al.*, 2010) et l'avortement de l'épi primaire, qui est alors remplacé par une languette foliacée (Lejeune & Bernier, 1996). La combinaison de ces trois composantes (sévérité, durée, stade de développement) induit une réponse physiologique complexe. Certaines combinaisons ont un impact similaire sur la plante. Par exemple, un froid modéré de longue durée ou un froid sévère de courte durée déclenchent chez la plante les mêmes mécanismes de protection. Par contre, un stress sévère de longue durée rend irréversibles les dégâts cellulaires occasionnés (Leipner & Stamp, 2009).

Le maïs présente une variabilité génétique de réponse au froid. Par exemple, la comparaison de 49 lignées de maïs appartenant à des groupes génétiques diverses pour leur réponse à un traitement froid de 15 jour/3°C nuit au stade 7 feuilles a révélé une réponse physiologique contrastée pour différents paramètres comme l'échange de CO₂, la teneur en chlorophylle, la composition du Photosystème II, la conductance des feuilles ou le poids sec (Lee *et al.*, 2002a). Une autre étude a montré que chez des génotypes sensibles, un froid de 6 °C pendant 48h sur des plantules au stade 3 feuilles visibles modifie la balance rédox et induit ainsi des changements au niveau du chloroplaste, ce qui inhibe la photosynthèse et provoque la sénescence (Foyer *et al.*, 2002). Chez des génotypes tolérants, le froid (6 °C pendant 48 h au stade 3 feuilles visibles) module les propriétés de l'apoplasme et la composition lipidique des membranes cellulaires (Nguyen *et al.*, 2009). Les génotypes tolérants possèdent une activité malate déshydrogénase plus importante que les génotypes sensibles (Leipner & Mayer, 2008). De plus, le froid induit l'expression de gènes codant des protéines de défense telles que les kinases, les phosphatases, ainsi que des protéines impliquées dans de nombreuses voies de régulation et de biosynthèse d'hormones connues par leur réponse aux contraintes abiotiques (notamment les phénylpropanoïdes, l'acide salicylique, l'acide abscissique, l'acide jasmonique), ainsi que des gènes codant pour des protéines impliquées dans les divisions cellulaires, la structure cellulaire, le métabolisme secondaire, le métabolisme des lipides, le métabolisme des carbohydrates, la signalisation cellulaire, des protéines de dégradation, de stockage et de développement (Foyer & Noctor, 2005, Janda *et al.*, 1999, Janowiak *et al.*, 2002, Maiti & Satya, 2014, Marocco *et al.*, 2005). D'un autre côté, des approches de transcriptomique globale ont également montré la modification de l'expression de gènes impliqués dans la synthèse des anthocyanes, le métabolisme des acides aminés et des lipides, dans le rythme circadien, et dans le métabolisme de la photosynthèse (Nguyen *et al.*, 2009, Sobkowiak *et al.*, 2014, Trzcinska-Danielewicz *et al.*, 2009).

Plusieurs études ont permis de détecter la présence de gènes de réponse au froid, ainsi que la réponse de certains éléments transposables. Les régions génomiques de réponses au froid ainsi qu'à d'autres contraintes abiotiques chez le maïs sont résumées dans le tableau 1 de l'annexe 1. Par exemple, les gènes codant pour l'alcool déshydrogénase (*Adh1*) (Kusano *et al.*, 1995), pour des protéines kinases (*ZmCDPK1*, *ZmMAPKKK*, *ZmMPK17*) (Kusano *et al.*, 1995, Pan *et al.*, 2012, Yang *et al.*, 2010) et pour des facteurs de transcription (*ZmDREB1*) (Riva-Roveda *et al.*, 2016).

Enfin, l'analyse phénotypique des plantes semées au champ précocement (début avril) dans le nord de l'Europe ont révélé que pendant les années froides, les maïs sensibles montrent une modification de la phénologie, de l'activité photosynthétique, de la croissance et par suite du rendement (Andrews *et al.*, 1995, Giauffret *et al.*, 1995). Bien que ces plantes soient soumises au froid au début de leur cycle de développement, une différence d'accumulation de biomasse est visible au stade pré-floraison, suggérant un impact prolongé du froid sur la croissance, en particulier la croissance de feuilles qui n'étaient pas développées au moment de l'exposition au froid. L'origine moléculaire de ce retard de croissance n'est pas encore connue. L'impact à long terme pourrait être dû au maintien d'un signal métabolique *via* des réseaux de régulation dont les acteurs amont seraient touchés en début du cycle de développement. Une autre hypothèse serait une perturbation de la régulation chromatinienne de certains gènes au niveau du méristème, qui serait ensuite maintenue à travers les divisions cellulaires.

II. Organisation et fonctions de l'épigénome

1. La chromatine organise l'information génomique

1.1. Structure et fonctions de la chromatine

Dans les cellules eucaryotes, l'ADN, localisé dans le noyau, est organisé autour de protéines histones, l'ensemble formant la chromatine. Cette structure peut prendre plusieurs états de condensation, permettant ainsi à la fois de compacter l'ADN à une taille compatible avec celle du noyau et de réguler l'état d'ouverture, et donc d'accessibilité, de l'ADN (Qiu, 2006).

L'observation cytologique du noyau interphasique a révélé la présence de territoires fonctionnels de niveau de compaction différente : l'hétérochromatine, une forme condensée, visible sous forme de régions denses, peu transcrite ; et l'euchromatine, une forme peu condensée et permettant la transcription (Figure I.2A). L'hétérochromatine peut être divisée en deux classes selon le degré et la durée de compaction : l'hétérochromatine constitutive, condensée de façon constitutive et localisée au niveau des régions péri-centromériques, des renflements hétérochromatiniens et des télomères, et l'hétérochromatine facultative qui peut adopter un état de compaction transitoire et réversible suivant le type cellulaire et l'état d'expression des gènes (Fransz & de Jong, 2011, Roudier *et al.*, 2009, Eichten *et al.*, 2014, Wang *et al.*, 2009). Cette compaction transitoire permet une lecture sélective de l'information stockée dans le noyau, à différents moments du développement ou de l'interaction avec l'environnement (Figure I.2B).

L'unité de base de la chromatine est le nucléosome (Figure I.3A), autour duquel sont enroulées environ 147 pb d'ADN. La succession des nucléosomes et de l'ADN nu confère à la chromatine une structure dite en « collier de perles » visible en interphase (Figure I.3B). Le nucléosome consiste en un octamère de quatre protéines histones (H2A, H2B, H3 et H4), chacune constituée d'un domaine globulaire et d'une « queue » spécifique (Figure I.3C). Ces « queues », sont formées d'acides aminés basiques, dépassent vers l'extérieur et peuvent ainsi subir des modifications post-traductionnelles qui modifient leurs caractéristiques chimiques (notamment acétylation, ubiquitination, méthylation ou phosphorylation). Ces modifications jouent un rôle dans la compaction de la chromatine (Figure I.2A). Un « code histone » a été défini, afin de répertorier les différents types de modifications qui peuvent être ajoutées sur les différents acides aminés des différentes protéines histones (Figure I.3, D et E). En plus des quatre protéines histones localisées au niveau du nucléosome, l'histone H1 (« linker histone »)

peut se lier à la molécule de l'ADN entre les nucléosomes (Gibney & Nolan, 2010, Qiu, 2006). Il existe également des variants d'histone ayant des propriétés particulières. Par exemple, au niveau des centromères, une partie des histones H3 est remplacée par le variant CENH3, qui permet le recrutement de protéines telles que CENP-C et par suite la formation des kinétochores (Van Hooser *et al.*, 2001). Ce variant ne possède pas de résidus Lys-9 et Lys-27 (Bailey *et al.*, 2013) ce qui empêche la présence de certaines modifications post-traductionnelles dans ces régions génomiques. Enfin, la méthylation de l'ADN consiste en la présence d'un groupement méthyl en position 5 de la cytosine. On parle alors de 5-méthyl-cytosine (5-mC) (Figure 1.4A). Ce groupement n'interfère pas avec l'appariement C : G des deux brins de l'ADN (Figure 1.4B). Par contre, il modifie les propriétés biochimiques de l'ADN et permet ainsi sa régulation. L'ensemble des cytosines méthylées dans un génome constitue le méthylome. Celui-ci montre une symétrie entre les deux brins de l'ADN (Figure 1.5). Cette modification chimique peut être gagnée ou perdue, ce qui permet le passage de cytosine vers 5-methylcytosine et réciproquement. L'acquisition du groupement méthyl a lieu de façon active et la perte peut se faire de manière active ou de manière passive (Gehring & Henikoff, 2007, Lister *et al.*, 2008).

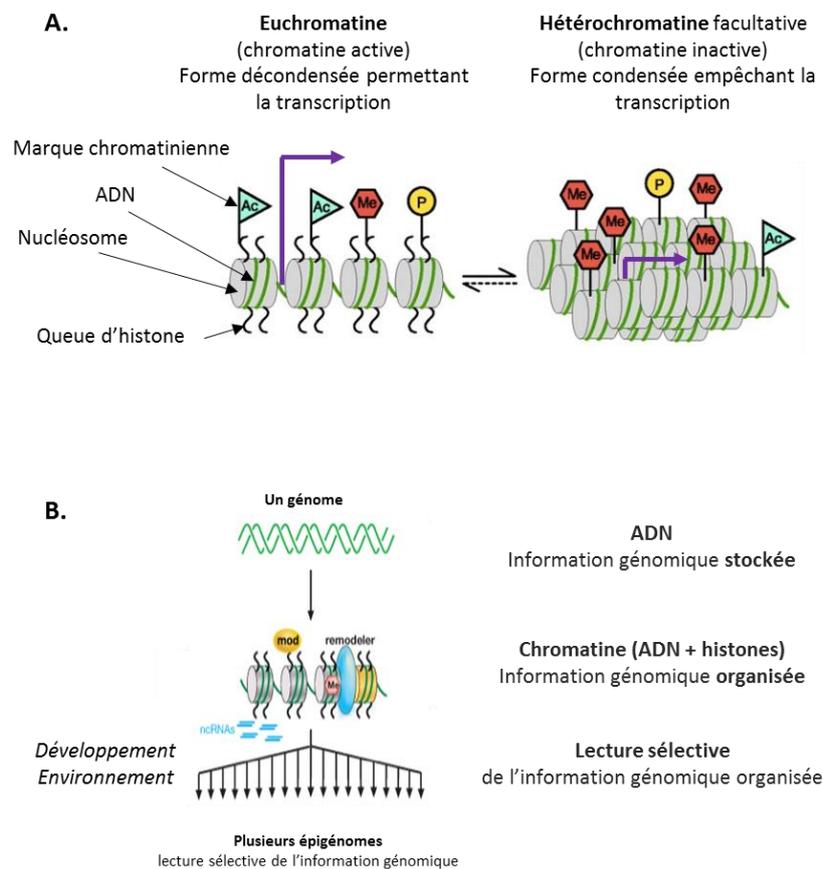


Figure 1.2: La chromatine organise l'information génomique

A. Schéma des deux structures majeures de la chromatine: l'euchromatine (à gauche) et l'hétérochromatine (à droite). Le passage d'un état à un autre est réversible, et est médié par des modifications des marques chromatinienne. Les flèches violettes indiquent le niveau de transcription (Me: méthylation, Ac: Acétylation, P: Phosphorylation) (Adapté de Delgado-Cuello, 2011). **B.** Schéma illustrant la lecture sélective de l'information codée par le génome, via l'organisation de la chromatine au cours du développement et en réponse à l'environnement (Adapté de Allis *et al.* (2007)).

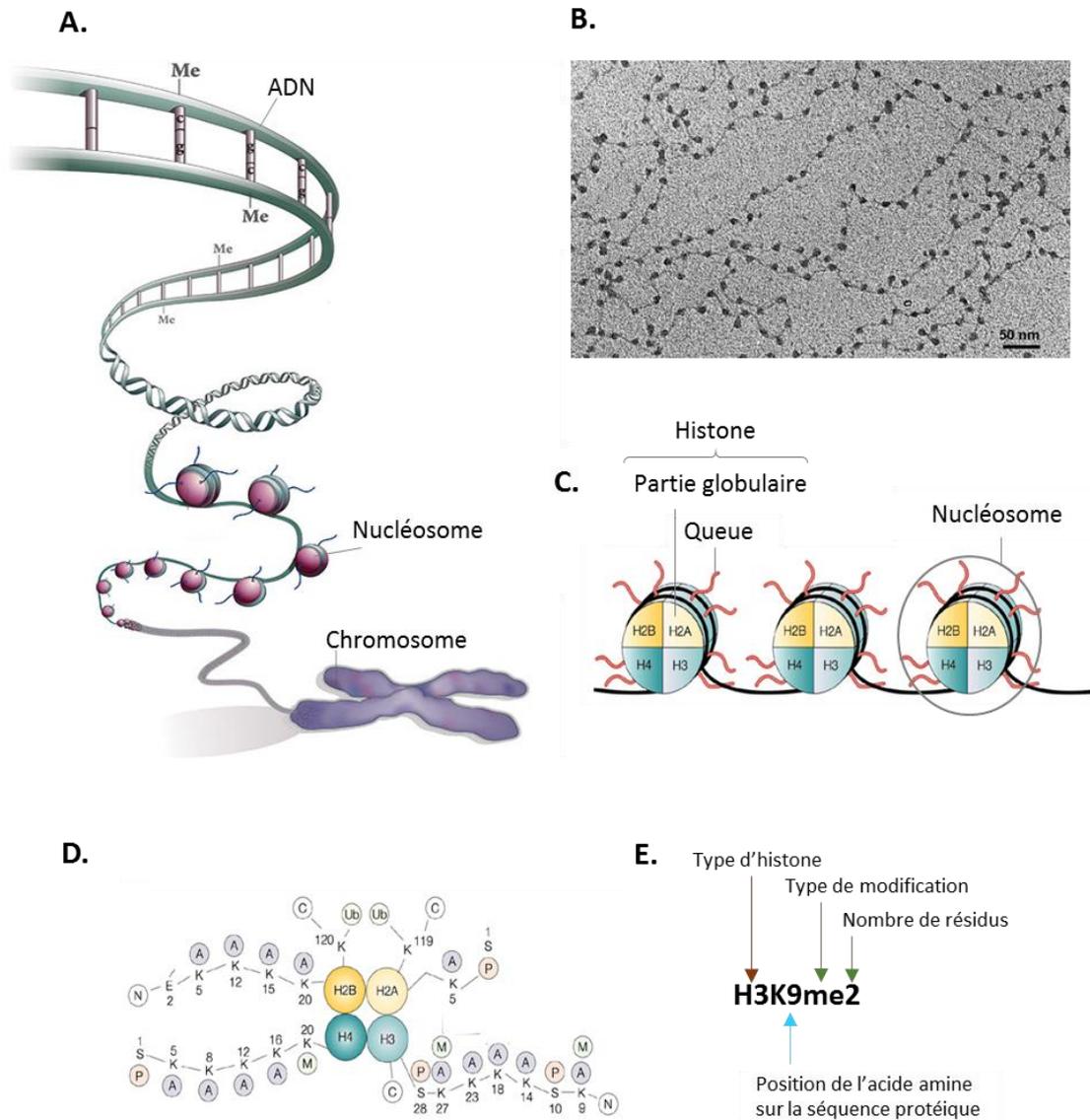


Figure 1.3 : Organisation de la chromatine

- A.** Schéma de la structure de l'ADN au chromosome, illustrant l'enroulement de l'ADN autour des nucléosomes et la présence de méthylation (Me) sur la molécule d'ADN (Adapté de Qiu (2006)).
- B.** Organisation de l'ADN en interphase, illustrant la structure en « collier de perles » générée par l'enroulement de la molécule d'ADN autour des nucléosomes.
- C.** Schéma de l'enroulement de l'ADN autour des nucléosomes, des structures formées de quatre protéines histones (H2A, H2B, H3 et H4), chacune possédant un partie globulaire et une queue (Marks *et al.*, 2001).
- D.** Détail de la composition en acides aminés des différentes queues d'histone, et des modifications qui peuvent avoir lieu sur celles-ci (une même protéine histone peut subir différents types et nombre de modifications) (Marks *et al.*, 2001).
- E.** Exemple de nomenclature d'une modification post-traductionnelle d'une protéine histone, ici une diméthylation (me2) de la lysine en position 9 (K9) de l'histone 3 (H3).

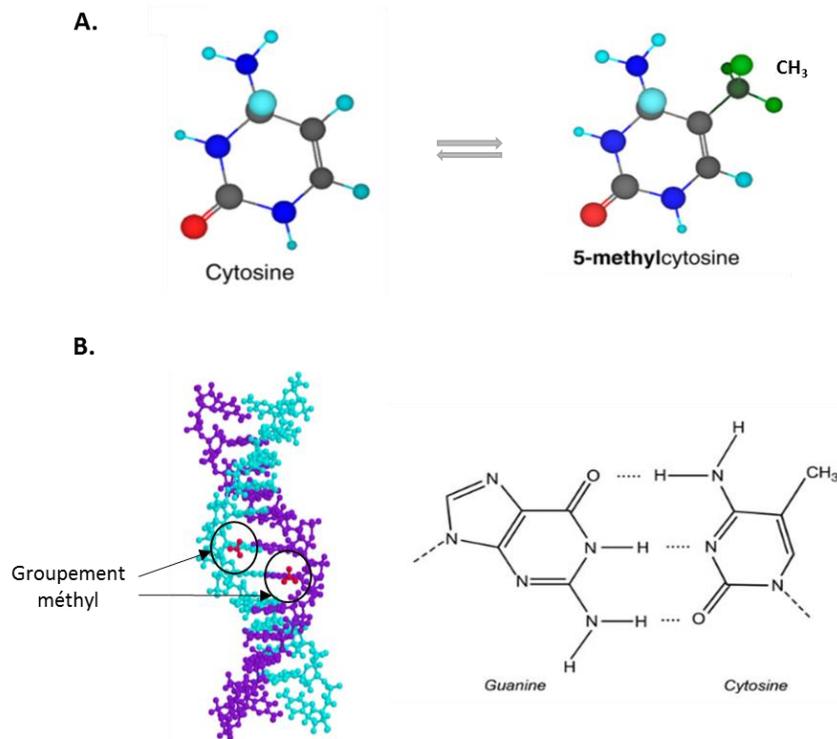


Figure I.4 : Présentation biochimique de la méthylation des cytosines

A. Biochimie de la méthylation des cytosines montrant le groupement méthyl en position 5 dans la 5-méthylcytosine (Prokhortchouk & Defosse, 2008). **B.** Position extérieure des groupes méthyl, qui ne perturbent pas les ponts hydrogènes entre les bases cytosine et guanine (Gibney & Nolan, 2010, Prokhortchouk & Defosse, 2008).

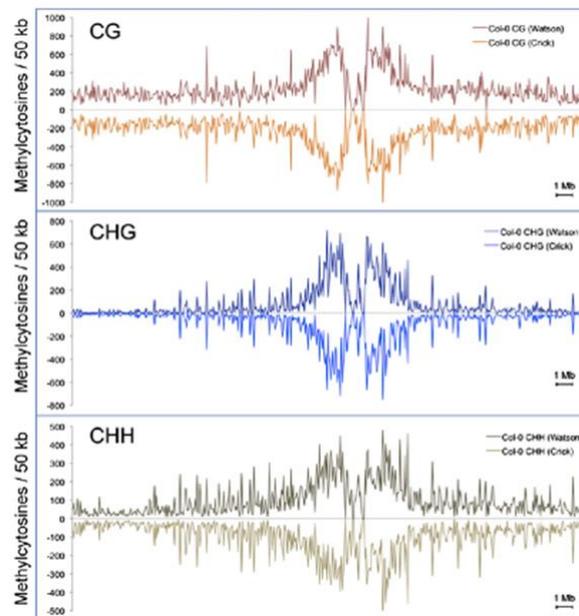


Figure I.5 : Profil de la méthylation le long d'un chromosome d'*Arabidopsis thaliana*, illustrant la symétrie observée entre les deux brins d'ADN (Lister *et al.*, 2008)

Les profils des cytosines des trois contextes sont représentés sur des fenêtres de 50kb. Le profil du brin Watson est représenté par des valeurs positives, celui du brin Crick par des valeurs négatives.

1.2. Bases moléculaires des états chromatinien

La méthylation de l'ADN et les modifications post-traductionnelles des histones constituent les « marques de la chromatine », et jouent un rôle important dans la dynamique de la chromatine et son accessibilité à la machinerie de transcription (Figure I.6) (Qiu, 2006, Lacoste & Côté, 2003, Lauria & Rossi, 2011, Roudier *et al.*, 2009). Les modifications post-traductionnelles des histones ont différents effets selon le type de modification chimique, et le nombre des modifications. Par exemple, la présence d'un groupement tri-méthyl sur la lysine 4 de l'histone 3 (H3K4me3) au niveau du TSS est associée à l'activation transcriptionnelle des gènes (Lauria & Rossi, 2011, Roudier *et al.*, 2009) (Figure I.6).

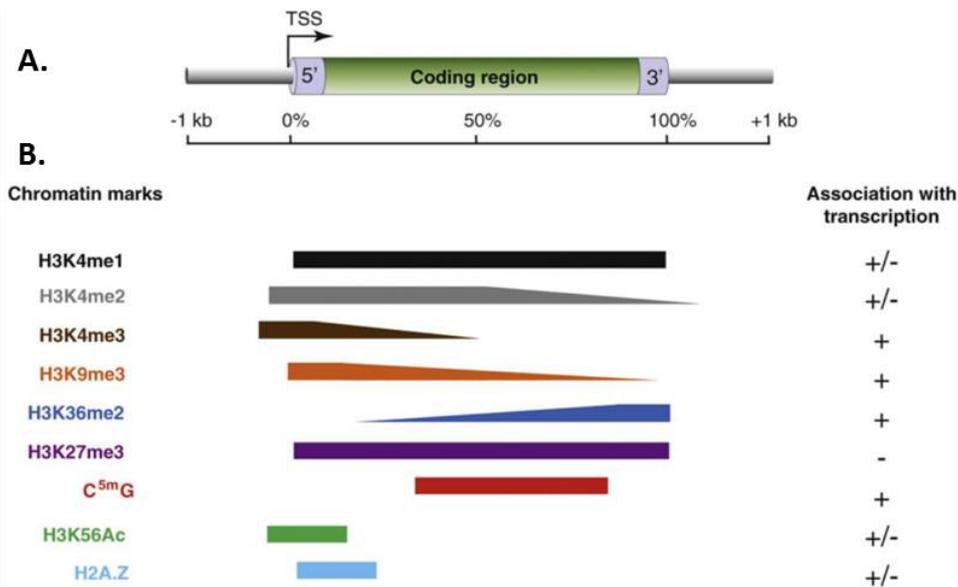


Figure I.6 : Schéma récapitulant le lien entre marques chromatinien et transcription (Roudier *et al.*, 2009).

A. Structure du gène, illustrant la position du début de transcription (« Transcription Start Site », ou TSS), des UTR (5' et 3') et de la région codante (en vert). **B.** Enrichissement des différentes marques chromatinien dans les différentes parties du gène, et relation avec le niveau de transcription. (+) et (-) indiquent respectivement l'état actif ou non de la transcription

La combinaison des différentes marques permet de catégoriser la chromatine en quatre états majeurs. Chaque état est caractérisé par l'abondance particulière d'un sous ensemble de ces marques, et correspond à un sous-ensemble de régions génomiques similaires en termes de composition génomique et de niveau d'expression. Chez *Arabidopsis thaliana*, ces quatre classes correspondent aux gènes exprimés, aux gènes réprimés par les protéines de type polycomb, aux régions hétérochromatiques (notamment riches en éléments transposables), et en régions d'euchromatine silencieuse (Figure I.7). La première est enrichie en marques H3K4me2/3, H3K36me3 et H2Bub et H3K9me3, et présente un niveau moyen de méthylation des cytosines. La deuxième est enrichie en H3K4me2 et H3K27me2/3 et présente une absence de méthylation de l'ADN. La troisième est riche en H3K9me2, H3K27me1/2 et H4K20me1 et présente un fort niveau de méthylation de l'ADN. La quatrième classe n'est pas enrichie en une marque particulière. Ces états chromatinien décrits chez *A. thaliana* sont conservés chez d'autres organismes pluricellulaires comme le nématode (*Caenorhabditis elegans*) (Gerstein *et al.*, 2010, Liu *et al.*, 2011), la drosophile (*Drosophila melanogaster*) (Kharchenko *et al.*, 2011, Roy *et al.*, 2010) et l'Homme (*Homo sapiens*) (Ernst & Kellis,

2010). Les états chromatinien dépendent des protéines de remodelage de la chromatine et des histones transférases permettant la régulation de l'expression des gènes au cours du développement. Parmi les marques chromatinien, la méthylation de l'ADN joue un rôle important dans la régulation des états chromatinien relatifs aux éléments transposables et aux gènes (Figure I.7). Les états chromatinien sont régulés à l'échelle de la composition génomique (gènes, éléments transposables) ce qui permet la régulation indépendante de chacun des composants qui se succèdent le long du génome (Figure I.8).

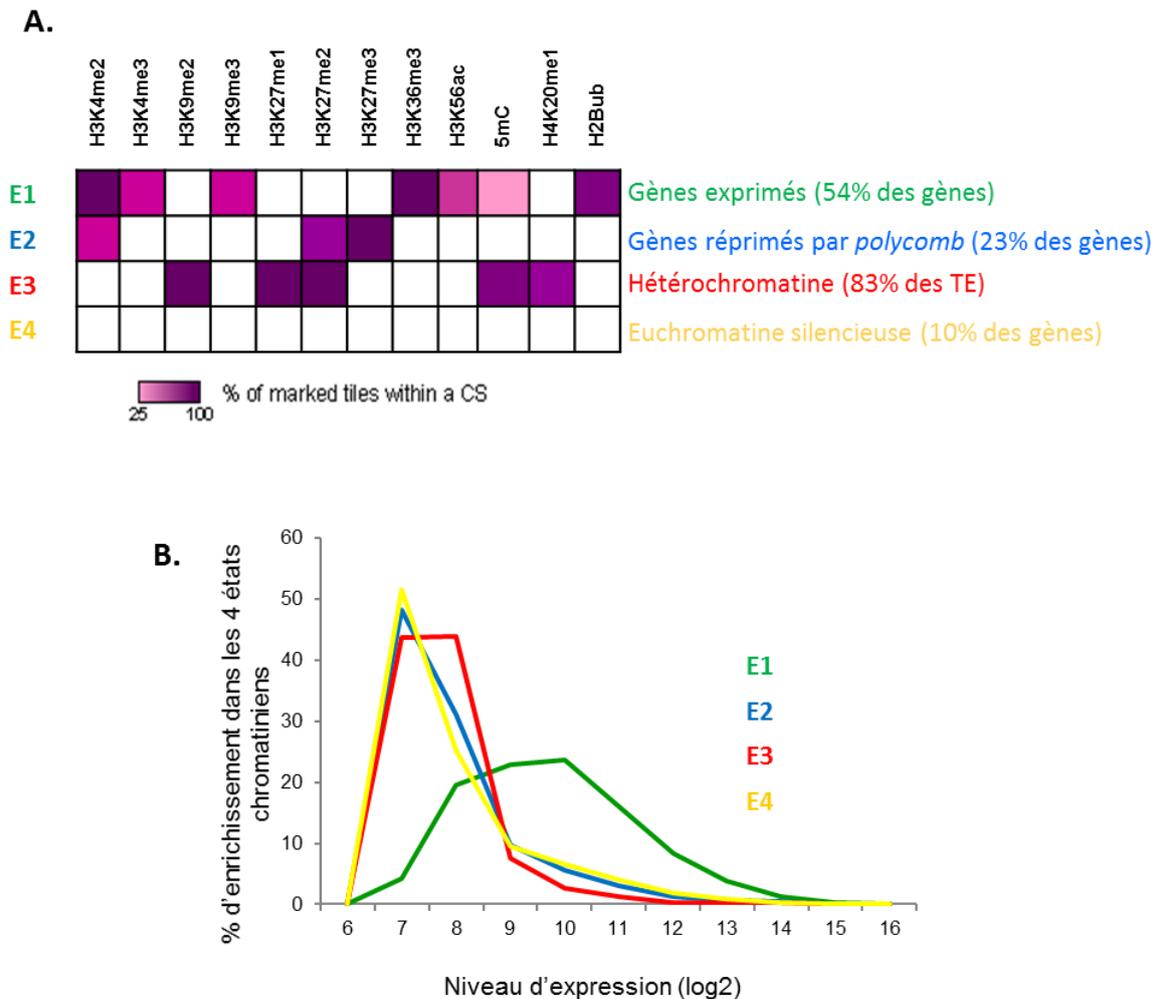


Figure I.7 : Description des différents états chromatinien chez *A.thaliana* (Roudier *et al.*, 2011)

A. Enrichissement des différentes marques de la chromatine dans les quatre états chromatinien. Les marques sont représentées en haut et les différents états chromatinien à gauche (de E1 à E4). A droite est représentée la catégorie fonctionnelle qui a été faite à partir de l'annotation de la localisation de ces états, présentée entre parenthèses. **B.** Présentation des niveaux d'expression des différents états chromatinien proposés, illustrant que seul l'état 1 (E1) est enrichi en séquences exprimées.

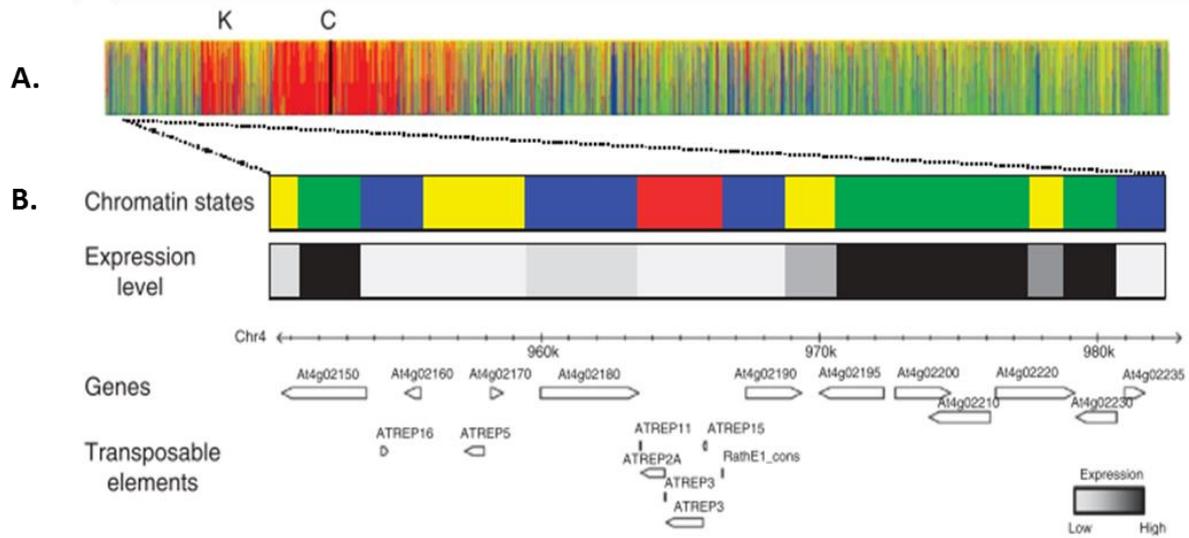


Figure I.8 : Succession des états chromatinien le long du génome (Roudier *et al.*, 2011)

A. Succession des différents états chromatinien le long du chromosome 4 (la position du centromère est indiquée par un C, celle du renflement chromatinien ou « knob » par un K).

B. Zoom sur une région du génome, présentant de haut en bas : la succession des états chromatinien, les niveaux d'expression (blanc : faible, noir : fort) et l'annotation des gènes et des éléments transposables. Les couleurs des états chromatinien sont celles de la figure I.7.

2. Mécanistique de la méthylation de l'ADN

2.1. Mise en place

Le mécanisme de méthylation s'opère en deux grandes étapes : l'établissement (ou méthylation *de novo*) et la maintenance. Les facteurs impliqués dans cette voie ont principalement été décrits chez la plante modèle *A.thaliana*, grâce à l'analyse de profils de méthylation chez différents mutants (Matzke & Mosher, 2014, Law & Jacobsen, 2010). La méthylation *de novo* est définie par l'insertion de nouveaux groupements méthyl (CH₃) sur de nouvelles cytosines du génome. Ce processus s'opère en dehors de la réplication de l'ADN contrairement au mécanisme de maintenance, et est médié par la voie « RNA-dépendant DNA Methylation » (RdDM) dont les principaux acteurs sont les petits ARN (Matzke & Mosher, 2014). Cette voie est responsable de l'établissement de la méthylation sur les trois contextes de cytosines (CG, CHG et CHH).

La voie RdDM canonique (Matzke & Mosher, 2014, Cuerda-Gil & Slotkin, 2016), est initiée par la transcription par l'ARN polymérase IV (PolIV). Les transcrits sont ensuite associés à la protéine RDR2 afin de générer des transcrits double brin (« dsRNA ») qui sont ensuite coupés en des ARN de 24 nucléotides (« siRNA ») par l'action de la protéine DCL3. Ces derniers sont ensuite pris en charge par AGO4 (ou son homologue proche AGO6). En parallèle, le locus cible est transcrit par PolIV, ce qui génère un transcrit non codant qui resterait associé au locus cible. La complémentarité de séquence entre le petit ARN chargé dans AGO4/6 et cet ARN généré par PolIV permet le recrutement de méthyltransférase DRM2, ce qui conduit à la méthylation de l'ADN du locus (Figure I.9A).

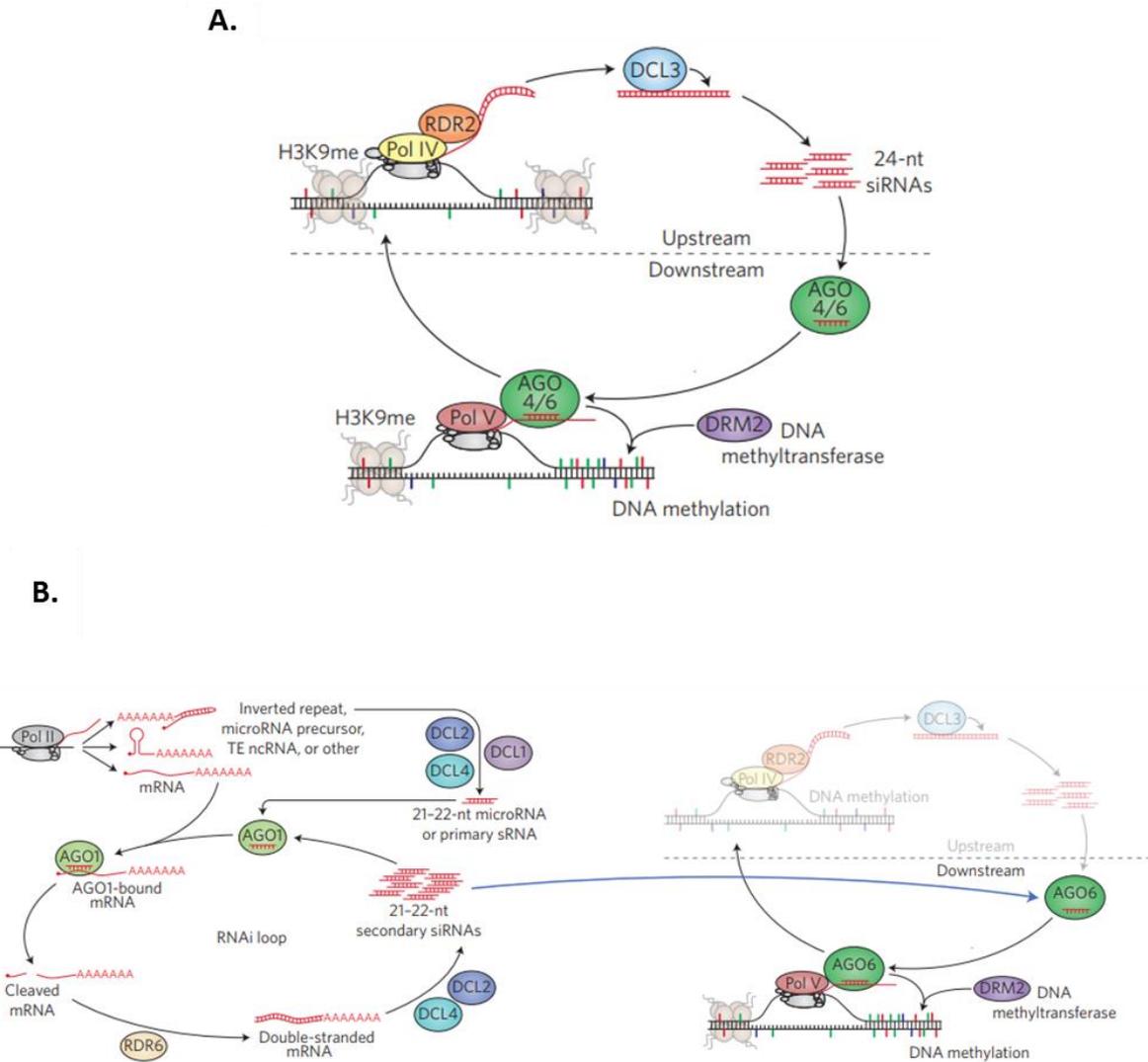


Figure I.9 : Mise en place de la méthylation *de novo* par voie RdDM (d'après Cuerda-Gil and Slotkin (2016))

A. Voie RdDM canonique. PolIV interagit avec RDR2 pour produire les ARN double brin (dsRNA). Ces derniers sont coupés en des ARN de 24 nucléotides (siRNA) par DCL3. Ces petits ARN sont chargés sur AGO4/6, et vont cibler les transcrits produits par PolV. L'interaction entre AGO4/6 et les transcrits de la PolV induit la méthylation de l'ADN par l'action de la méthyltransférase DRM2. Les couleurs sur le brin d'ADN représentent les cytosines de contexte CG (rouge), CHG (bleu) et CHH (vert). « Upstream » : avant production des siRNA. « Downstream » : après génération des siRNA.

B. Voie RdDM non canonique. Cette voie fait aussi appel à la production de siRNA, mais ceux-ci sont générés par l'action de PolII et RDR6. Les siRNA produits sont ensuite pris en charge par AGO6 identiquement à ce qui est décrit en A.

Ce mécanisme de RdDM se basant au début sur l'activité de la PolIV constitue la voie canonique de la RdDM. Cependant, cette voie PolIV-dépendante fonctionne pour la régulation de l'expression de ET endogène au génome, mais elle n'est pas responsable de la régulation d'un ADN exogène (par exemple viral) (Herr *et al.*, 2005). Elle pourrait donc correspondre à une voie de maintenance de la méthylation

dans les régions hétérochromatiques déjà méthylées, plutôt qu'à une vraie voie permettant la mise en place de la méthylation *de novo* (revu dans(Cuerda-Gil & Slotkin, 2016).

Une voie RdDM faisant intervenir la polymérase II (PolII) et RDR6 au lieu de PolIV/RDR2 a été décrite, et permet de méthyler des régions nouvelles. Dans cette voie non canonique, une reconnaissance entre l'ADN et la Pol II entraîne la production des transcrits d'environ 45 nucléotides (Blevins *et al.*, 2015, Law *et al.*, 2013, Law *et al.*, 2011, Ye *et al.*, 2016, Zhai *et al.*, 2015). Ces transcrits sont par la suite convertis en ARN double brin (dsRNAs) par RDR6. Ces derniers sont coupés en petits ARN (siRNA) par DCL2 et DCL4 pour produire des ARN de 22 et 21 nucléotides. Ces petits ARN, au lieu d'être pris en charge par AGO1 comme dans la voie PTGS pour cliver l'ARN cible, sont pris en charge par AGO6, avec un devenir similaire à celui de la voie canonique, et induisant la méthylation du locus générant un transcrite PolIV (revu dans Cuerda-Gil and Slotkin (2016)) (Figure I.9B). Cette voie non canonique pourrait constituer la première mise en place de la méthylation à un locus non méthylé. La voie RdDM canonique interviendrait en second lieu pour renforcer le signal nouvellement initié.

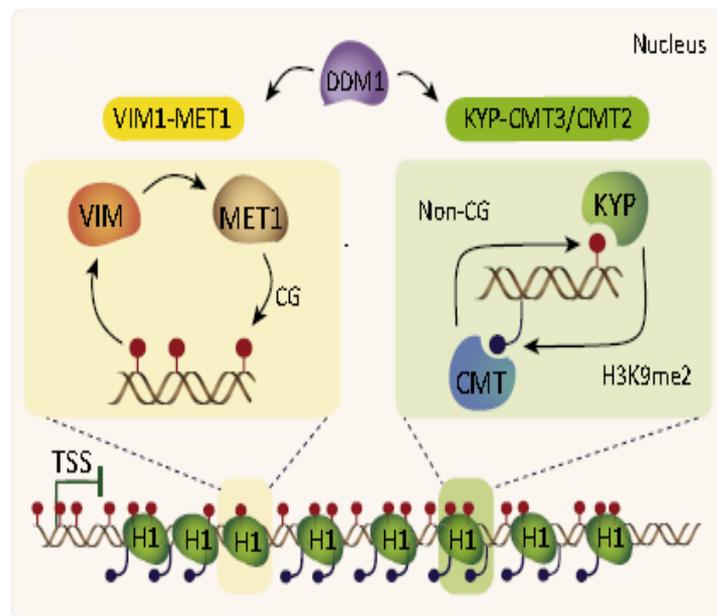


Figure I.10 : Maintenance de la méthylation de l'ADN, (d'après Kim and Zilberman (2014))

Schéma du mécanisme de maintenance de la méthylation de l'ADN pour le contexte CG (à gauche, faisant appel à l'action de VIM1 et MET1) et dans le contexte CHG (à droite, faisant appel à la marque d'histone H3K9me2, l'action de KYP et de CMT3/2).

2.2. Maintenance

Une fois établie, la méthylation de l'ADN est conservée à travers les divisions cellulaires. Contrairement aux mammifères où la méthylation de l'ADN est maintenue par une seule méthyltransférase (DNMT1), les plantes possèdent plusieurs méthyltransférases qui permettent une maintenance différente des cytosines selon leur contexte (CG, CHG ou CHH). Cette maintenance a lieu lors de la réplication de l'ADN, et repose sur l'existence de cytosines méthylées sur le brin parent. Chez *A. thaliana*, la méthylation des cytosines en contexte symétrique CG est essentiellement maintenue par MET1 qui reconnaît les CG hémiméthylés générés lors de la réplication, avec l'aide de la protéine VIM1 (Law &

Jacobsen, 2010) (Figure I.10). La méthylation des cytosines en contexte CHG est intimement liée à la marque H3K9me2 et est maintenue par l'action d'une boucle d'auto-renforcement : les histones methyltransferases KYP, SUVH5 et SUVH6 se lient aux sites méthylés CHG et catalysent la méthylation de la lysine 9 de l'histone H3 (H3K9me2). La protéine CMT3 (et partiellement CMT2) se lient au chromodomaine de cette marque et méthylent en retour l'ADN (Stroud *et al.*, 2014, Stroud *et al.*, 2013) (Figure I.10). La méthylation CHH est asymétrique (Figure I.11) et après la réplication de l'ADN, il n'existe pas de cytosine pour guider la maintenance de la méthylation. Celle-ci est maintenue essentiellement par la voie RdDM canonique faisant intervenir PolIV/RDR2/AGO4 pour les éléments transposables courts et les extrémités des éléments transposables longs, ou par l'action de CMT2, essentiellement pour la partie interne des éléments transposables longs (Stroud *et al.*, 2013). La maintenance de la méthylation CHH et CHG dans les régions hétérochromatiques peut également impliquer des remodeleurs de la chromatine comme DDM1 qui facilite l'activité de CMT2 et DRD1 (Law & Jacobsen, 2010).

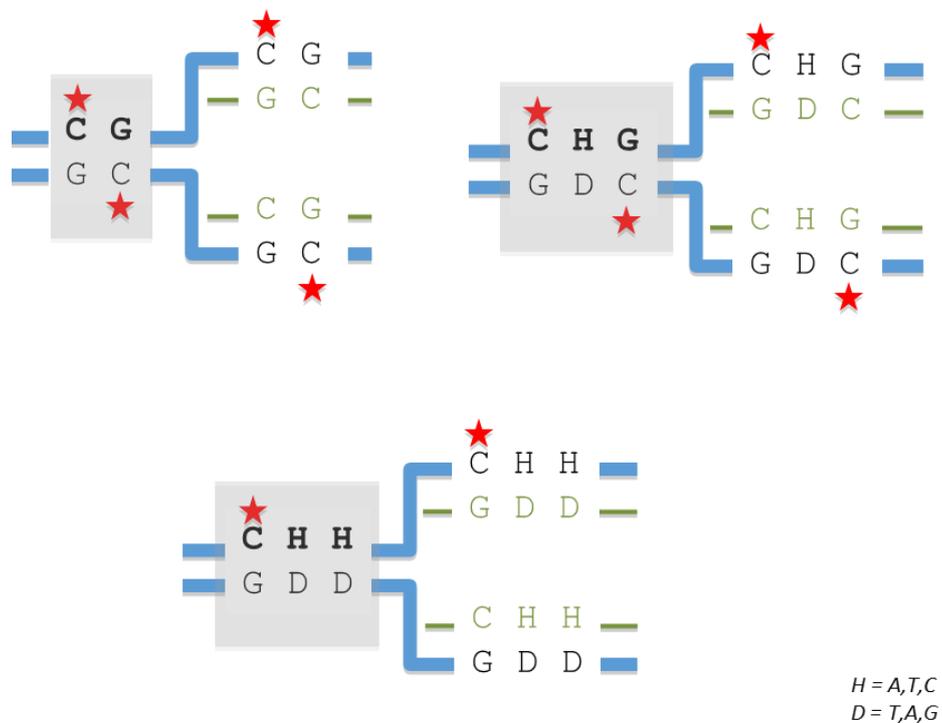


Figure I.11 : Schéma simplifié du maintien de la méthylation dans les trois contextes

A gauche : ADN avant réplication ; à droite : ADN après réplication. L'étoile indique que la cytosine est méthylée. Pour les contextes symétriques (CG et CHG), la cytosine méthylée du brin parent (en bleu) peut servir de matrice pour ajouter un groupement méthyl sur le brin néosynthétisé (en vert). Pour le contexte CHH, non symétrique, un des brins parents ne présente pas de cytosine méthylée (en bas) et l'autre possède une cytosine méthylée, mais pas de cytosine à proximité pour recopier le signal.

En absence de maintenance de méthylation d'ADN ou d'activité *de novo*, le signal est perdu de façon passive à travers les réplifications de l'ADN (Furner & Matzke, 2011, Penterman *et al.*, 2007). En plus de cette déméthylation passive, un processus de déméthylation active existe. Il est assuré par l'action des glycosylases de l'ADN (Figure I.12). Par exemple, ROS1, DML2 et DML3 agissent tout au long du

développement pour maintenir l'absence de méthylation au niveau des gènes contenant des éléments transposables (Furner & Matzke, 2011, Law & Jacobsen, 2010, Penterman *et al.*, 2007).

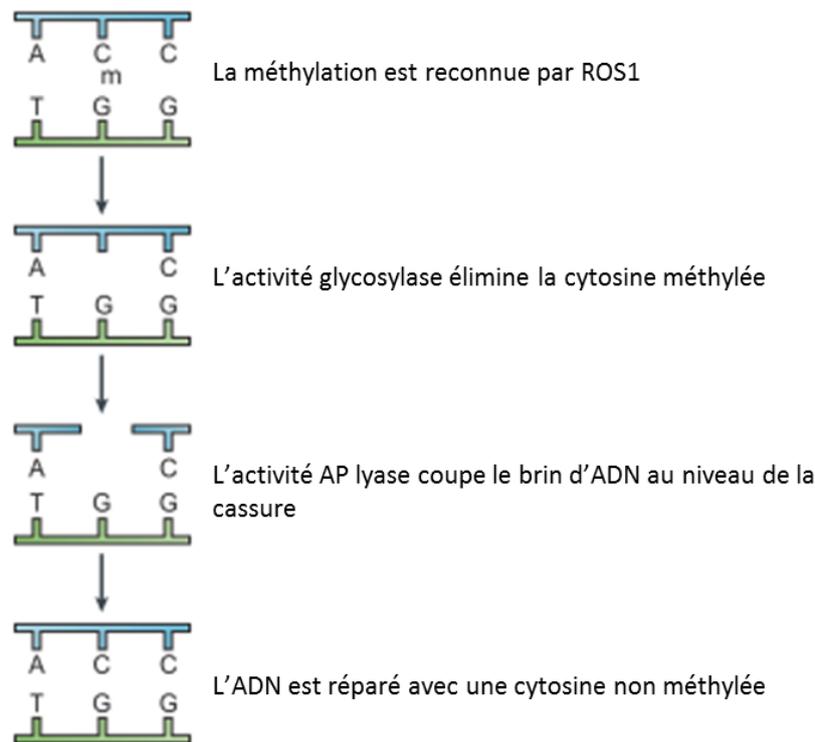


Figure I.12 : Perte de la méthylation de l'ADN par voie active (d'après Chan *et al.* (2005))

Ainsi, la mise en place de la méthylation est commune aux trois contextes, mais la maintenance diffère d'un contexte à l'autre (Figure I.11). L'analyse des niveaux de méthylation des cytosines de chaque contexte montre un niveau élevé pour le contexte CG, moyen pour CHG et plus faible pour le contexte CHH (respectivement >80%, 20–60%, <20%) Ceci est vrai pour *A. thaliana*, (Cokus *et al.*, 2008, Lister *et al.*, 2008) mais a été observé chez la plupart des espèces végétales qui ont été analysées (Niederhuth *et al.*, 2016) (Figure I.13). Bien que la méthylation de l'ADN semble guidée par une mécanistique permettant sa maintenance après la réplication de l'ADN, le niveau de méthylation à 100% est rarement observée même dans le contexte CG (Quadrana & Colot, 2016).

Cette différence est liée à une différence de fidélité des mécanismes de maintenance de la méthylation après la réplication de l'ADN, la maintenance de méthylation des cytosines en contexte CG étant la plus fidèle (Gehring & Henikoff, 2007). Des homologues aux acteurs de la voie RdDM et aux methyltransférases sont retrouvés chez les plantes cultivées comme le riz (Ahmad *et al.*, 2014, Wu *et al.*, 2010, Hu *et al.*, 2014), la cacahuète (Wang *et al.*, 2016a), et le fraisier (Gu *et al.*, 2016). Les voies décrites chez la plante modèle *A. thaliana* ne sont cependant pas complètement universelles. Par exemple, la protéine CMT3 est absente chez certains angiospermes comme *Eutrema salsugineum* (Bewick *et al.*, 2016).

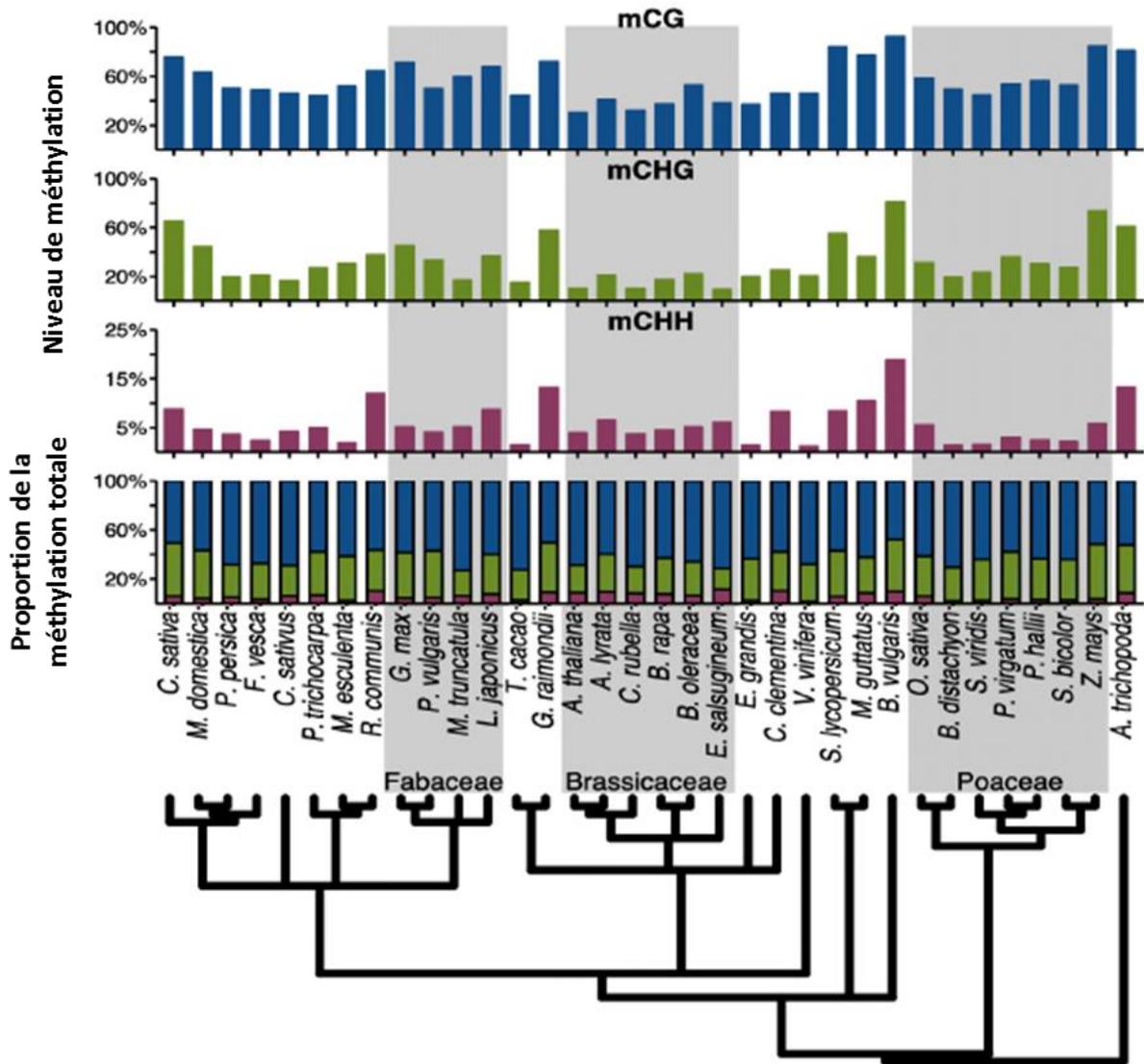


Figure I.13 : Taux de méthylation de l'ADN à l'échelle du génome chez différentes espèces d'angiospermes dans le contexte CG, CHG et CHH (Niederhuth *et al.*, 2016).

3. Profil de méthylation des différentes composantes génomiques

3.1. Méthylation des gènes

Pour la majorité des gènes, le profil type d'une méthylation génique est caractérisé par l'existence d'une méthylation au niveau des régions transcrites ou corps du gène (GbM pour « Gene body Methylation ») uniquement en contexte CG, avec une déplétion au niveau du TSS et TTS (Figure I.14A). Les gènes présentant un profil GbM sont plus larges, contiennent plus d'exons et évoluent plus lentement que les gènes non méthylés, et semblent avoir des fonctions plus importantes que les gènes non méthylés (Takuno & Gaut, 2012).

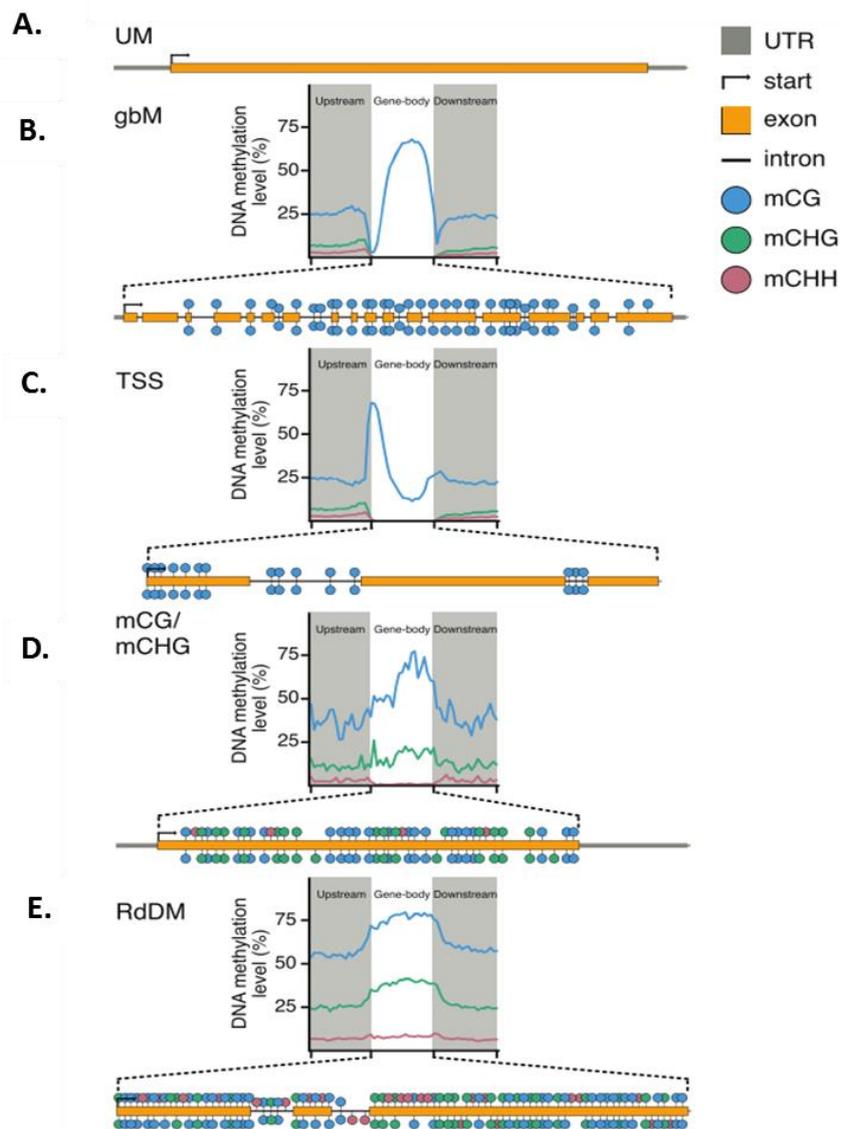


Figure I.14 : Profils de méthylation de l'ADN au niveau des gènes (d'après Bewick and Schmitz (2017))

A. Gène n'ayant aucune méthylation dans aucun contexte (UM : UMethylated). **B.** Gène ayant une méthylation uniquement en contexte CG, dans la région transcrite et absente au niveau du TSS et du TTS (gbM : gene body Methylation). **C.** Présence de méthylation CG uniquement au niveau du TSS. **D.** Présence de méthylation en contextes CG et CHG au niveau de la région transcrite. **E.** Présence de méthylation CHH guidée par la voie RdDM, suggérant la présence de la méthylation en CG et CHG.

Du fait du rôle primordial de la régulation des gènes dans l'établissement du phénotype, la contribution de la GbM au phénotype est un axe majeur de recherche. La fonction de la GbM reste, néanmoins, encore énigmatique et plusieurs hypothèses ont été proposées. Au niveau transcriptionnel, le profil observé pourrait être lié à une inhibition de l'initiation de la transcription PolIII dans le milieu du gène, où le niveau de méthylation est élevé. Par contre, la méthylation GbM ne semble pas avoir d'effet sur l'élongation de la transcription. Elle pourrait aussi augmenter la fidélité de l'épissage alternatif en marquant la position des exons, ou en excluant la méthylation non CG des sites accepteurs de l'épissage (Lev Maor *et al.*, 2015, Wang *et al.*, 2016b). La méthylation de l'ADN et la présence du variant d'histone H2A.Z sont anti-corrélés, mais le lien de cause à effet reste à éclaircir. La méthylation de

l'ADN pourrait éviter le dépôt de ce variant, évitant ainsi que des promoteurs internes au gène deviennent compétents. A l'inverse, le variant H2A.Z pourrait éviter un positionnement de la méthylation de l'ADN au niveau du promoteur, évitant ainsi l'extinction de l'expression du gène. Cependant, la comparaison de gènes avec et sans GbM n'a pas révélé de différence en terme d'épissage, de niveau d'expression, ou de quantité du variant H2A.Z, et ces hypothèses pourraient donc être remises en question (Coleman-Derr & Zilberman, 2012, Zilberman *et al.*, 2007).

D'autres études de variation naturelle de la méthylation de l'ADN chez *A. thaliana* et chez le maïs ont montré une corrélation faible mais significative entre les variations de méthylation et les variations d'expression des gènes (Eichten *et al.*, 2013, Li *et al.*, 2015, Schmitz *et al.*, 2013b). Mais ce lien reste souvent corrélatif et non explicatif, et plusieurs scénarios pourraient être possibles selon les locus. Notamment, la méthylation pourrait participer au contrôle de la régulation des gènes, ou au contraire être une conséquence de la transcription (Seymour & Becker, 2017). Ainsi, en absence de preuves évidentes sur le rôle de la GbM sur l'expression, elle pourrait être considérée comme un produit dérivé neutre de la transcription, sans pour autant avoir de conséquences fonctionnelles. D'un autre côté, chez des accessions de *A. thaliana*, la présence de GbM est corrélée avec la latitude, ce qui suggère un rôle adaptatif (Dubin *et al.*, 2015). De façon contradictoire, la GbM est très conservée (présente dans 47 espèces d'angiospermes sur 49 étudiées (Bewick *et al.*, 2016), ce qui suggère un rôle fonctionnel, mais son absence chez *Eutrema salsugineum* et *Conringia planisiliqua* ne semble pas avoir de conséquence négative sur la survie des individus (Bewick *et al.*, 2016). La GbM pourrait donc avoir un rôle fonctionnel, mais pas obligatoire. Par exemple, elle pourrait être un réservoir de variation épigénomique, qui pourrait être avantageux dans certaines conditions. Par exemple, un même gène peut être présent sous forme non méthylée, méthylée par GbM, ou méthylée du fait de l'insertion d'un élément transposable (Kawakatsu *et al.*, 2016).

La GbM est au cœur des discussions et le mécanisme de sa mise en place n'est pas encore complètement élucidé. La présence de méthylation uniquement en contexte CG est liée à l'action de la méthyltransférase MET1. Chez les mutants *met1* d'arabidopsis, la GbM est perdue sur l'ensemble du génome. Cependant, après retour en conditions MET1 sauvages, la méthylation de type GbM est très lente (plus de 8 générations), alors que l'accumulation de méthylation en contexte CG est plus rapide sur les séquences répétées, notamment éléments transposables (Stroud *et al.*, 2013). Le gène INCREASED IN BONSAI METHYLATION 1 (IBM1) semble aussi y jouer un rôle majeur. Dans les mutants *ibm1*, une accumulation de la marque H3K9me2 et une augmentation de la méthylation en contexte CHG est observée aux locus présentant un profil GbM en conditions sauvages. A partir de ces observations, un modèle a été proposé selon lequel de façon occasionnelle, IBM1 serait incapable de déméthyliser ou de décrocher la marque H3K9me2 au niveau des régions transcrites. Dans ce cas, CMT3 serait recrutée et induirait ainsi la méthylation en contexte CHG, qui serait ensuite renforcée par la boucle d'auto-renforcement médiée par KYP/SUVH4/5/6, H3K9me2 et CMT3. Par un mécanisme encore inconnu, le contexte CG serait également méthylé dans ces régions, et la méthylation à ces sites serait ensuite maintenue par l'activité de MET1, assurant sa transmission fidèle à travers les générations (Bewick *et al.*, 2016, Bewick & Schmitz, 2017, Niederhuth & Schmitz, 2017, Niederhuth *et al.*, 2016, Stroud *et al.*, 2013, Wendte & Schmitz, 2017) (Figure I.15A). De façon alternative, un transcrit aberrant pourrait être occasionnellement formé à partir de promoteurs cryptiques, induisant la production de petits ARN par la voie RdDM et la méthylation dans les trois contextes. En l'absence du maintien de la transcription, seuls les sites CG seraient ensuite maintenus à travers les générations. (Tran *et al.*, 2005, Zilberman *et al.*, 2007) (Figure I.15B).

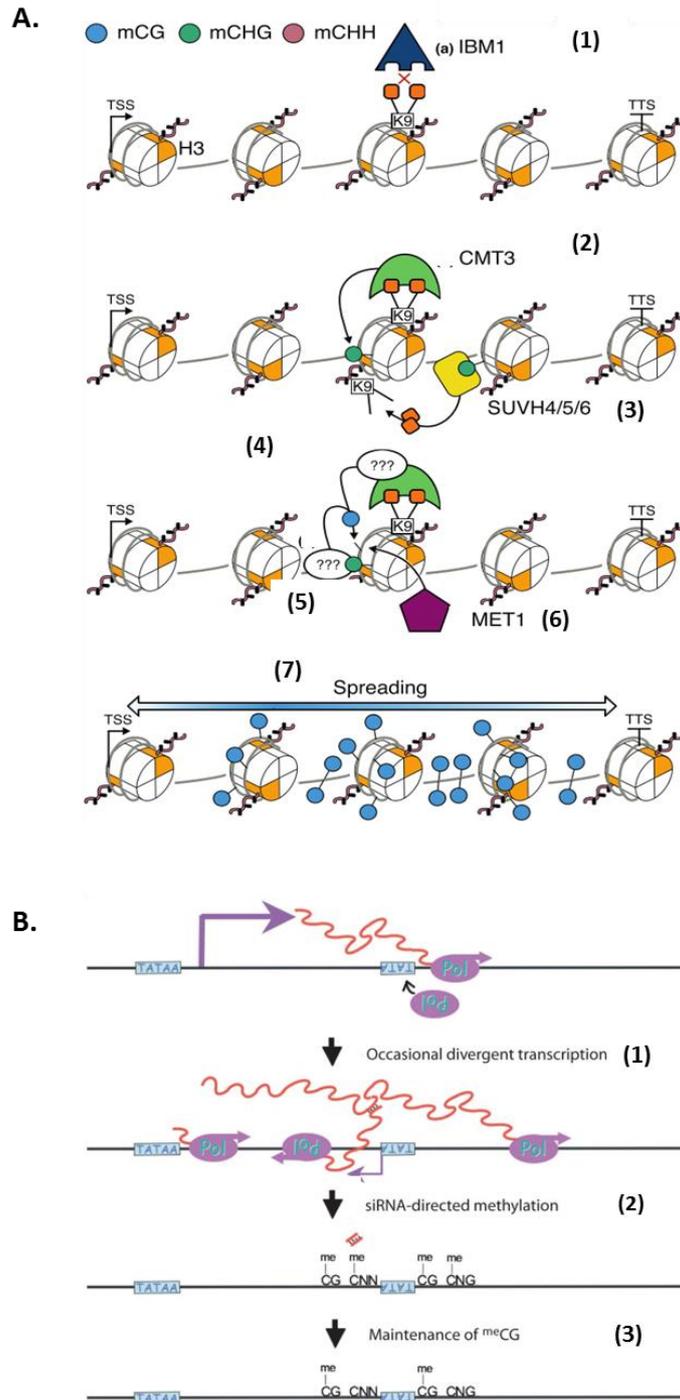


Figure I.15 : Modèles de mécanisme permettant d'expliquer le profil de méthylation gbM

Modèle proposé par Bewick and Schmitz (2017) faisant appel à l'action d'IBM1. Dans ce modèle, plusieurs étapes se succèdent : **(1)** IBM1 ne réussit pas à décrocher H3K9me2, **(2)** Recrutement de CMT3 et méthylation CHG, **(3)** la méthylation en contexte CHG recrute les protéines KYP/SUVH4/5/6 et H3K9me2, **(4)** renforcement du niveau de la méthylation, **(5)** méthylation des sites CG (par un mécanisme encore inconnu), **(6)** maintien de la méthylation CG par l'action de MET1, **(7)** la méthylation CG s'étale le long du corps du gène. **B.** Modèle de Tran *et al.* (2005) faisant appel à la transcription à partir de promoteurs cryptiques. Ce modèle fait appel à plusieurs étapes : **(1)** activation d'un promoteur cryptique, induisant une transcription divergente **(2)** l'ARN naissant est clivé en siARN qui peuvent induire de la méthylation de l'ADN **(3)** sans l'activité continue des siARN, la méthylation non-CG est perdue et seule la méthylation CG est maintenue. Les transcrits naissants sont représentés en rouge. TATA: promoteur. TATA: promoteur cryptique.

En plus de ce cas général de méthylation GbM, quatre autres types de profil de méthylation à l'intérieur du corps du gène ont été observés (Bewick & Schmitz, 2017). Ces quatre cas sont catégorisés en fonction du contexte de cytosine affecté et de sa localisation (promoteur, TSS), indiquant ainsi des voies de régulation différentes, et de leur impact sur la transcription: (i) les gènes non méthylés, montrent aucune méthylation dans les trois contextes CG, CHG et CHH, (ii) les gènes ayant une méthylation CG uniquement au niveau du TSS, conduisant à la répression de l'expression du gène (Niederhuth *et al.*, 2016, Niederhuth & Schmitz, 2017), (iii) les gènes présentant dans la région transcrite une méthylation en CG et CHG et une déplétion en méthylation CHH, et exprimés à un niveau plus bas par rapport aux gènes avec gbM (Niederhuth *et al.*, 2016), (iv) les gènes enrichis en méthylation CHH (et possiblement aussi CG et CHG), *via* la voie RdDM, qui ont une répression d'expression dans toute la plante à l'exception du pollen et de la graine (Niederhuth & Schmitz, 2017) (Figure I.14).

Enfin, la méthylation au niveau des régions régulatrices du gène peut avoir lieu dans les 3 contextes CG, CHG et CHH et peut affecter l'expression. De façon générale, la méthylation de l'ADN au niveau du promoteur diminue l'expression des gènes (Zilberman *et al.*, 2007, Roessler *et al.*, 2016), vraisemblablement en diminuant l'affinité des facteurs de transcription à la région régulatrice. Par exemple, chez *Arabidopsis*, la majorité (72%) des facteurs de transcription montrent une faible affinité quand la séquence cible est méthylée, 4% s'accrochent lorsqu'il y a présence de méthylation et 24% ne révèlent aucune différence (O'Malley *et al.*, 2016). Chez le riz, à l'inverse, la méthylation au niveau de la région terminale de transcription (TTR) réprime plus l'expression du gène que lorsqu'elle est présente au niveau du promoteur (Li *et al.*, 2012). D'autre part, la présence de méthylation en contexte CHH à environ 500pb en amont du TSS, est corrélée avec le niveau d'expression du gène (Gent *et al.*, 2013, West *et al.*, 2014, Lu *et al.*, 2015). Cette méthylation, médiée par la voie RdDM, pourrait avoir un rôle organisationnel en délimitant une zone d'hétérochromatine transcriptionnellement active mais méthylée, qui ferait la jonction entre l'hétérochromatine non transcrite, enrichie en H3K9me2, et l'euchromatine transcriptionnellement active.

3.2. Méthylation des éléments transposables

Les éléments transposables (ET) sont les constituants majoritaires des génomes de plantes. Ils sont présents dans les génomes sous forme de séquences répétées, ce qui illustre leur capacité à se dupliquer (revu dans Vitte *et al.* (2014)). La protection de l'intégrité d'un génome dépend de l'inhibition des ET, afin de minimiser les effets des mutations que pourraient avoir de nouvelles insertions. Le contrôle de la prolifération des éléments transposables est médié par la condensation de la chromatine, qui limite leur transcription. Plusieurs mécanismes ont ainsi été développés, afin d'assurer une grande efficacité. Ces différents mécanismes sont particulièrement efficaces car malgré l'abondance des ET dans les génomes de plantes, les échappements au système sont rares. Les mécanismes en jeu font intervenir les ARN interférents (RNAi), la modification des histones mais surtout la méthylation de l'ADN, qui est un acteur majeur de la répression transcriptionnelle des ET (Underwood *et al.*, 2017, Lisch, 2009, Slotkin & Martienssen, 2007).

Les éléments transposables sont essentiellement localisés dans les régions hétérochromatiques, et même participent à leur définition. Dans les régions denses en éléments transposables, à distance des gènes, le niveau de méthylation est élevé dans les trois contextes, et est lié à la machinerie de

maintenance des trois contextes. Les marques H3K27me2 et H2K9me2 sont également enrichies dans ces régions. Dans les régions moins denses en éléments transposables, et plus proches des gènes, la méthylation des éléments transposables est assurée par la voie RdDM. Chez *Arabidopsis thaliana*, elle cible essentiellement les éléments de petite taille et les extrémités des éléments de grande taille (Stroud *et al.*, 2013, Zemach *et al.*, 2013).

Dans certains cas, la régulation chromatinienne des éléments transposable peut déborder en dehors des régions cibles, ce qui modifie l'état chromatinien dans les régions adjacentes. C'est ce que l'on appelle le « spreading ». Ce phénomène a été notamment décrit chez *Arabidopsis* (Ahmed *et al.*, 2011) et le maïs (Eichten *et al.*, 2012) et est localisé à faible distance (jusqu'à 1kb) des séquences d'éléments transposables. L'analyse de la méthylation autour de 1543 insertions d'ET polymorphes entre 140 accessions d'*A. thaliana* a révélé 457 cas de spreading à faible distance (<300pb) et 243 cas de spreading à plus grande distance, pouvant aller jusqu'à 3.5kb du site d'insertion (Quadrona *et al.*, 2016a) (Figure I.16). La comparaison de différentes classes d'éléments transposables chez le maïs a révélé une différence entre types d'éléments. Par exemple, alors que les LINE ne montrent pas de trace de « spreading » de l'hétérochromatine, les rétrotransposons à LTR de type *gypsy* (RLG) montrent un débordement à la fois de la méthylation de l'ADN et de la marque H3K9me2, et les rétrotransposons de type *copla* (RLC), débordent uniquement la marque H3K9me2 (Eichten *et al.*, 2012).

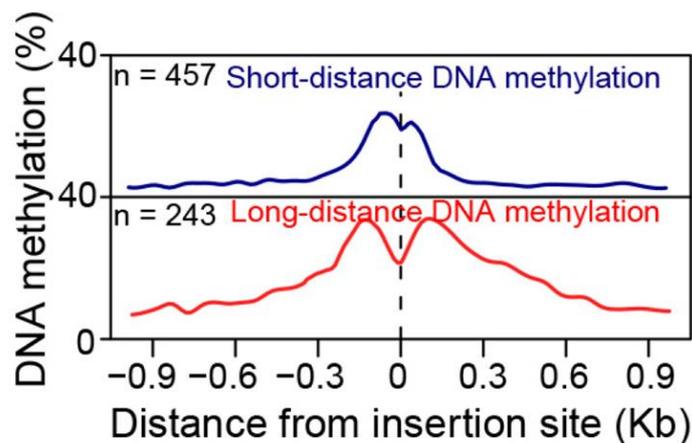


Figure I.16 : Méthylation de l'ADN dans les régions flanquantes des sites d'insertion des ET (d'après Quadrona *et al.* (2016a))

En haut (bleu) : profil de méthylation observé pour 457 cas, et présentant un « spreading » à courte distance (<300bp). En bas (rouge) : profil de méthylation observé pour 243 insertions, et présentant un « spreading » à plus longue distance (jusqu'à 3.5kb, non représenté ici)

Le contrôle de la méthylation des ET peut induire des modifications d'expression des gènes adjacents. La Figure I.17 présente les cas les plus emblématiques. Chez la tomate, l'augmentation de la méthylation de l'ET localisé en amont du gène *Cnr* induit une mauvaise maturation du fruit (Figure I.17). Chez *Arabidopsis*, l'hyperméthylation (induite par la mutation *ddm1*) du LINE localisé en aval du gène BNS est responsable du phénotype *Bonsai* (Saze & Kakutani, 2007). Chez le palmier à huile (*Elaeis guineensis*), l'hypométhylation (induite par culture *in vitro*) du LINE *Karma* localisé dans l'intron du gène DEFICIENS engendre des fruits couverts et parthénocarpiques (Ong-Abdullah *et al.*, 2015). Le « spreading » de la méthylation de l'élément *Gyno-hAT* localisé au niveau du promoteur du gène *CmWIP1* induit un changement de sexe chez le melon (*Cucumis melo*) (Martin *et al.*, 2009).

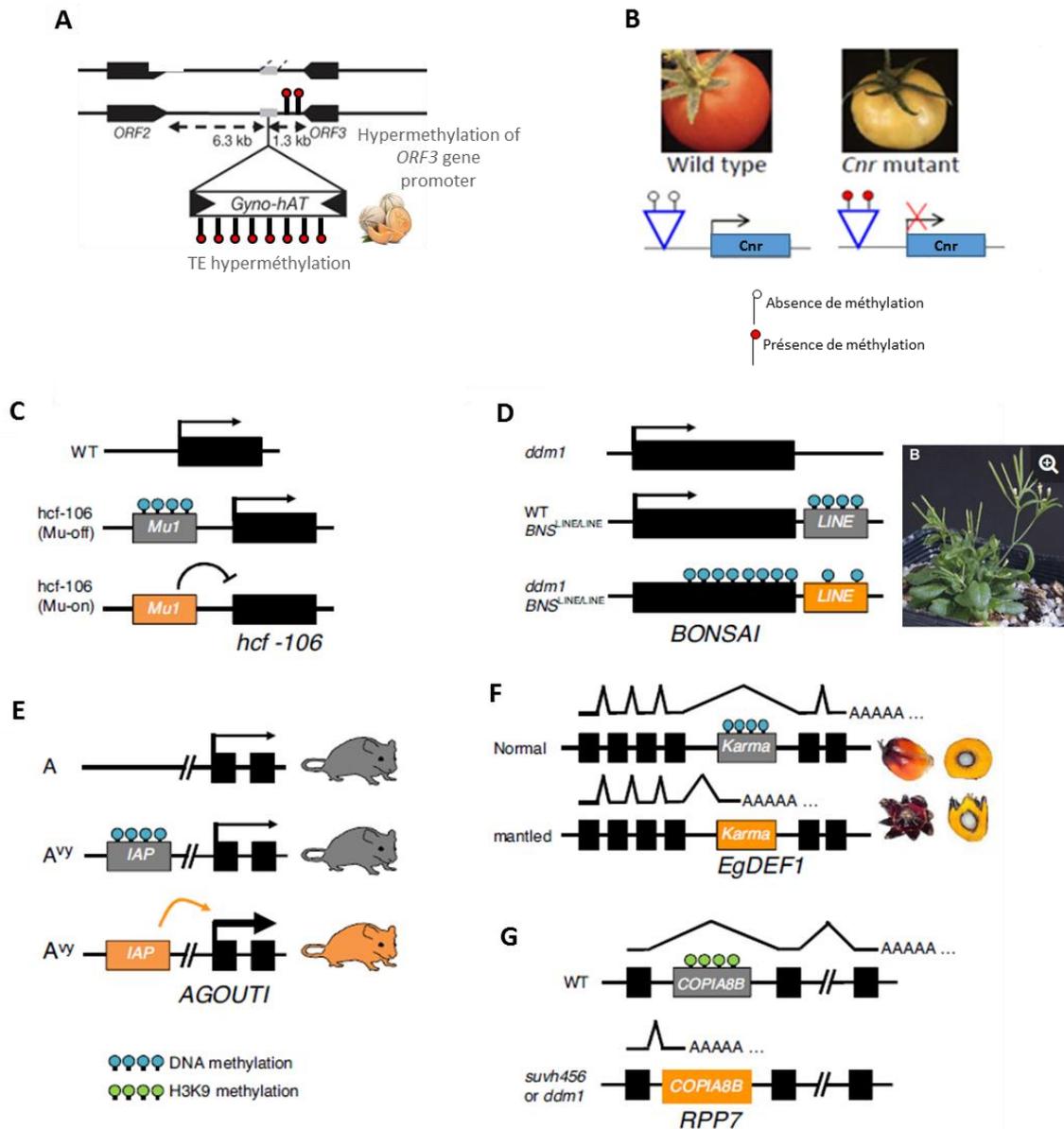


Figure I.17 : Cas de méthylation des éléments transposables à côté des gènes résultant dans une variation phénotypique

A. Le « spreading » de l'hyperméthylation à côté du transposon *Gyno-hAT* engendre le déterminisme sexuel chez le melon (*Adapté de Martin et al. (2009)*). **B** Hyperméthylation du locus *Cnr* induisant une déficience de maturation chez la tomate (*Adapté de Manning et al. (2006)*). **C à G** (revu dans Hosaka and Kakutani (2018)) **C.** L'insertion du transposon *Mu* dans le promoteur du gène *hcf-106* (*High chlorophyll fluorescence 106*) est associée à une variabilité dans l'expression du gène. Elle dépend de l'activité des copies et de leurs niveaux de méthylation (Martienssen & Baron, 1994). **D.** En contexte mutant pour *DDM1*, le « spreading » de la méthylation de l'ADN du *LINE* en aval du gène *BNS* chez *Arabidopsis* engendre un phénotype nain (Saze & Kakutani, 2007). **E.** La présence d'un élément transposable *IAP* (Intracisternal A-particle) en amont du locus *Agouti viable yellow* (*Avy*) modifie la couleur du pelage de la souris. Quand le transposon est faiblement méthylé, la souris a un pelage de couleur jaune, elle est obèse et présente un diabète (Duhl *et al.*, 1994). **F.** L'hypométhylation de l'élément *Karma* induit la formation de fruits couverts et parthénocarpiques (Ong-Abdullah *et al.*, 2015). **G.** La perte de H3K9me2 (induite dans le triple mutant *suvh4-suvh5-suvh6* ou dans le mutant *ddm1*) au niveau de l'élément *Copia* situé dans l'intron du gène *RPP7* induit une terminaison prématurée du transcrit (Tsuchiya & Eulgem, 2013).

L'impact de la méthylation de l'ET sur l'expression du gène varie suivant la localisation de l'insertion de l'ET (intron, exon) (Le *et al.*, 2015). L'insertion d'un ET en amont d'un gène suivie de sa répression transcriptionnelle chromatinienne peut avoir plusieurs conséquences : le « spreading » de l'état chromatinien de l'ET pourrait conduire à l'extinction transcriptionnelle du gène adjacent, et une perte de la méthylation (totale ou partielle) de l'élément pourrait révéler un nouveau promoteur dans la séquence de l'ET, ou révéler de nouveaux éléments régulateurs (Figure I.18) (Barkan & Martienssen, 1991, Springer *et al.*, 2016, Hirsch & Springer, 2017, Lisch, 2013).

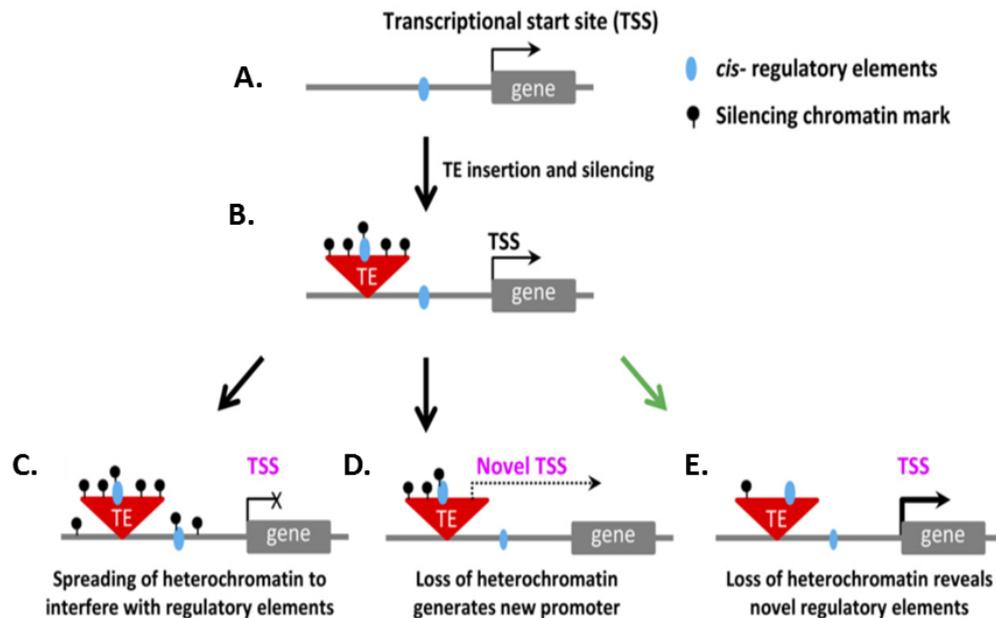


Figure I.18 : Différent scénarios suite à une nouvelle insertion de ET (Springer *et al.*, 2016)

A. En l'absence d'élément transposable, la transcription du gène est régulée par une séquence régulatrice localisée en amont du TSS du gène (oval bleu). **B.** Insertion d'un élément transposable en amont du site régulateur, régulée par l'ajout de méthylation (cercles noirs) **C.** La méthylation de l'ADN (ou autre marque de la chromatine) s'étale dans les régions à côté de l'ET (« spreading ») et empêche ainsi l'accessibilité de l'élément régulateur en *cis*. **D.** perte des modifications de la chromatine dans les bords de l'ET induisant la création d'un nouveau TSS issu de la séquence de l'ET **E.** perte des modifications de la chromatine le long de l'ET révélant un nouveau site d'accrochage pour l'élément régulateur en *cis*. La suppression des marques de la chromatine dans l'ET peut se faire de manière stochastique ou peut avoir lieu suite au « spreading » de l'euchromatine après activité importante de l'expression du gène à côté de l'ET.

3.3. Méthylation des centromères

La chromatine du centromère est caractérisée par la présence d'un variant d'histone spécifique, CENH3, qui est un marqueur de l'identité centromérique (Talbert *et al.*, 2002, Zhong *et al.*, 2002). Chez le maïs, les séquences centromériques *CentC* et *CRM* interagissent avec les nucléosomes portant CENH3. Le niveau de méthylation est important pour déterminer l'identité du centromère, en participant à la régulation de la chromatine centromérique (Choo, 2001, Schueler & Sullivan, 2006). Chez *A. thaliana*, le riz et le maïs, les séquences centromériques actives sont associées à un état

d'hypométhylation de l'ADN, en particulier en comparaison avec les régions péri-centromériques adjacentes (Koo *et al.*, 2011, Yan *et al.*, 2010, Zhang *et al.*, 2008). De manière contrastée, les mêmes répétitions sont hyperméthylées lorsqu'elles sont localisées dans les régions péri-centromériques. Ainsi, les mêmes séquences peuvent être hypo ou hyperméthylées selon leur localisation et leur association à CENH3. Une hypothèse serait que la méthylation de l'ADN joue un rôle important dans la démarcation de la chromatine au niveau du centromère par rapport aux régions flanquantes et que l'hétérochromatine jouerait un rôle de barrière pour empêcher le « spreading » de la marque CENH3 (Zhang *et al.*, 2008). L'analyse des séquences sous-jacentes a révélé que malgré leur forte identité de séquences, les répétitions associées ou non à CENH3 ont des profils particuliers de composition en cytosines CG et CHG. Chez *A. thaliana*, l'hypométhylation est associée à un niveau réduit de la marque H3K9me2 (Zhang *et al.*, 2008).

Le mécanisme qui sous-tend la localisation spécifique de la marque CENH3 au niveau du centromère reste imprécis. Woo *et al.* (2007) montrent que la protéine VIM1 est indispensable pour la méthylation de l'ADN au niveau du centromère chez *Arabidopsis* et que la méthylation centromérique et péri-centromérique est régulée par deux mécanismes différents mais qui restent inconnus. Des transcrits issus des séquences répétées centromériques ont été observés dans plusieurs espèces. Ces transcrits sont essentiels pour la régulation du chargement du variant CENH3 dans les régions centromériques (Chan & Wong, 2012) et pour l'assemblage du complexe des kinétochores (Scott, 2013). Chez le maïs, des transcrits des répétitions CentC et CRM2 ont été détectés (Du *et al.*, 2010, Topp *et al.*, 2004). Une hypothèse serait que la présence du variant CENH3 est liée à la transcription au niveau du centromère. La voie RdDM ne joue pas un rôle dans l'assemblage de l'hétérochromatine péri-centromérique. La méthylation de l'ADN ainsi que la méthylation de H3K9 de ces régions sont indépendantes de l'activité des siRNA et sont dues à l'activité de DDM1, de MET1 et de CMT3 et CMT2 (Matzke & Mosher, 2014).

3.4. Méthylation le long des chromosomes

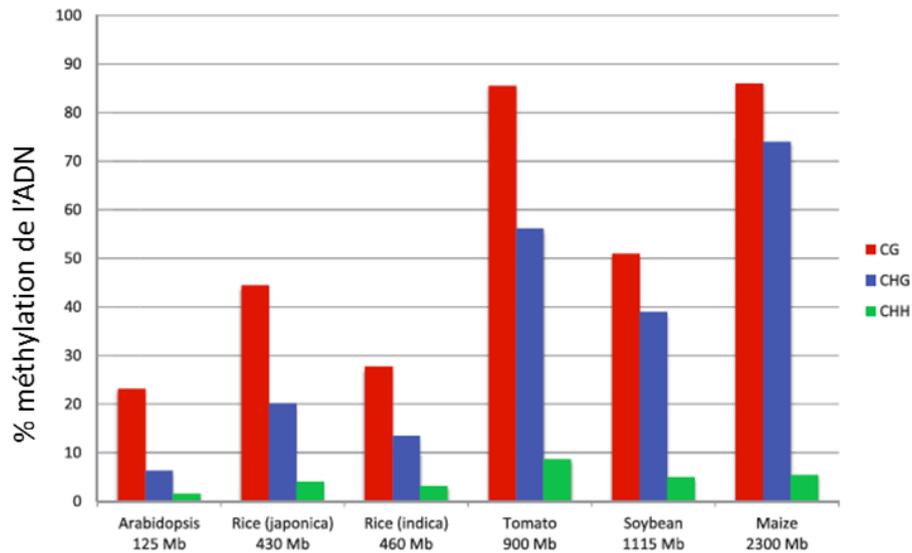
Le niveau global de méthylation d'un génome varie en fonction de sa taille et du nombre et de la nature des gènes et d'éléments transposables (Niederhuth & Schmitz, 2017, Mirouze & Vitte, 2014) (Figure I.19A). En moyenne, une augmentation d'un coefficient de 10 dans la taille du génome implique une augmentation de 10% du niveau global de méthylation des cytosines (Alonso *et al.*, 2015, Vidalis *et al.*, 2016).

Les variations du niveau de méthylation de l'ADN correspondent à la fois à une variation du pourcentage des sites méthylés et de leur niveau de méthylation (Niederhuth *et al.*, 2016, Takuno *et al.*, 2016). Par exemple, le génome du maïs (2.3Gb) est très méthylé en comparaison à d'autres espèces dont le génome est moins riche en éléments transposables, et ceci est vrai pour les trois contextes CG, CHG et CHH, et en particulier en comparaison avec *A. thaliana* (125Mb) (Mirouze & Vitte, 2014) (Figure I.19A). Chez les angiospermes, le niveau de méthylation global pour les trois contextes CG, CHG et CHH est corrélée avec la taille du génome. Il est de l'ordre de 5.3% chez *A. thaliana* et 39.2% chez *Narcissus* (Bewick *et al.*, 2016, Niederhuth *et al.*, 2016) (Figure I.19B).

Le profil de méthylation le long des chromosomes dépend très largement de la quantité et de la dispersion des éléments transposables. Par exemple, chez les espèces à petit génome comme *A. thaliana*, la méthylation est essentiellement concentrée dans les régions péri-centromériques riches en éléments transposables. Le reste du génome est assez peu méthylé, du fait de la faible densité en

éléments transposables dans les régions géniques. A l'inverse, chez les espèces à plus gros génome riches en ET comme le riz (40%) et le maïs (85%), les chromosomes sont très méthylés sur toute leur longueur (Mirouze & Vitte, 2014) (Figure I.20).

A.



B.

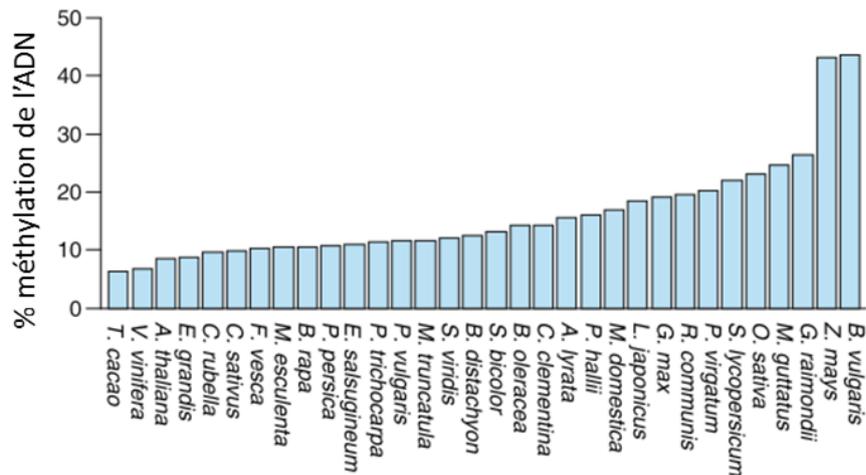


Figure I.19 : Variation du niveau de méthylation de l'ADN entre espèces d'angiospermes

A. Variation du niveau de méthylation par contexte de cytosine chez 6 espèces d'angiospermes présentant des tailles de génome contrastées. Les trois contextes sont représentés, et montrent des différences entre espèces (d'après Mirouze and Vitte (2014)).

B. Distribution du niveau de méthylation global des cytosines avec la taille du génome chez 32 espèces d'angiospermes, illustrant le lien fort entre ces deux paramètres (d'après Vidalis *et al.* (2016)).

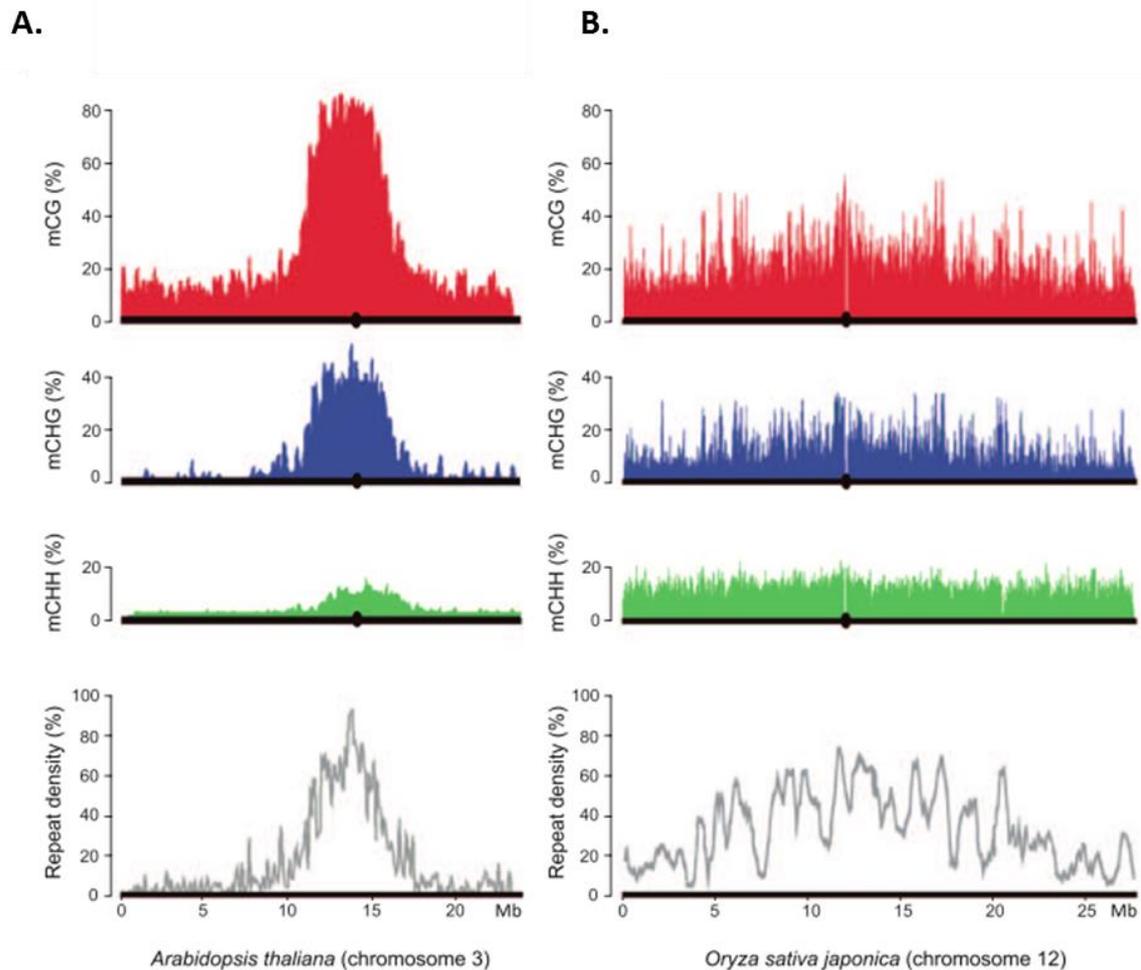


Figure 1.20 : Distribution de la méthylation de l'ADN dans les trois contextes de cytosine en fonction de la distribution des éléments transposable (d'après Mirouze and Vitte (2014))

A. Chez *A.thaliana* (chromosome 3), la distribution de la méthylation de l'ADN est localisée au niveau des régions péri-centromérique riches en ET.

B. chez le riz *Oryza sativa japonica* (chromosome 12), la densité en ET est importante tout le long des chromosomes, ce qui induit un niveau de méthylation élevé tout le long du chromosome.

Bien que les ET soient, de par leur nombre et leur richesse en cytosines méthylées, les principaux déterminants du méthylome, les gènes contribuent également à la distribution de la méthylation de l'ADN le long du génome, en particulier pour le contexte CG qui est abondamment méthylé dans les régions transcrites d'un grand nombre de gènes (GbM) (Bewick & Schmitz, 2017, Takuno & Gaut, 2013). Les îlots CHH, localisés à proximité des gènes, montrent des niveaux de méthylation CHH assez élevés (15%) (Gent *et al.*, 2013). Cependant, leur taille réduite (moins de 500pb), ne leur permet pas de jouer un rôle important dans le profil de méthylation à l'échelle chromosomique, même dans les régions riches en gènes.

Comme exposé plus haut, les différentes composantes génomiques telles que les gènes et les ET sont ciblées de manière spécifique par des mécanismes différents. Les différences en densité de ces composantes le long des chromosomes induit ainsi une différence en terme de méthylation des trois

contextes. Springer and Schmitz (2017) classent les régions génomiques en cinq classes de méthylation, selon que aucun, un, deux ou trois des contextes sont méthylés. Ils distinguent également une classe de régions de niveau de méthylation intermédiaire. La proportion de ces cinq types de méthylation varie le long du chromosome et entre espèces, en lien avec la quantité d'éléments transposables (Figure I.21).

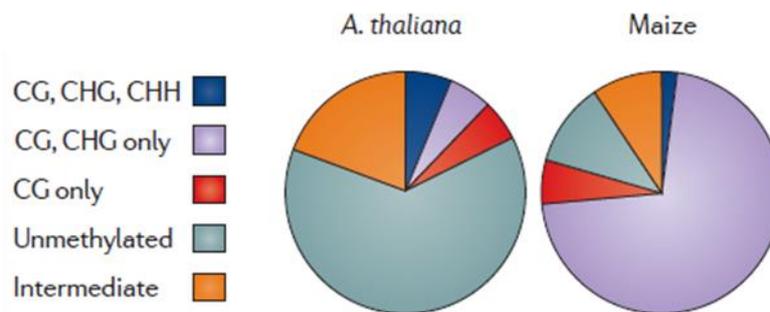


Figure I.21 : Distribution des 5 types de méthylation proposée par Springer and Schmitz (2017) chez Arabidopsis et chez le maïs

Les proportions sont celles du nombre de fenêtres de 100pb annotées comme appartenant à chaque catégorie. Les proportions sont très différentes entre les deux génomes, avec peu de régions non méthylées chez le maïs, et peu de régions à méthylation de niveau intermédiaire.

Enfin, la distribution de la méthylation de l'ADN dans les différents contextes dépend aussi de la machinerie disponible dans l'espèce en question. Par exemple, la GbM n'est pas retrouvée dans toutes les espèces, et pourrait être spécifique des angiospermes (Bewick & Schmitz, 2017). L'absence de CMT2 chez le maïs pourrait expliquer la plus grande importance de la RdDM pour la maintenance de la méthylation en contexte CHH.

Ainsi, le profil du méthylome en CG, CHG et CHH d'une espèce, et même celui d'un individu, dépend à la fois de la composition en gène et ET, de leur organisation relative, et de la composition des voies de régulation de la méthylation de l'ADN. Cet ensemble détermine la façon spécifique qu'a le méthylome de chaque individu à changer en réponse aux voies développementales, et à l'environnement (Mirouze & Vitte, 2014).

III. Origines de la variation du méthylome

1. Origine génétique

Comme nous l'avons vu plus haut, le méthylome participe à réaliser une lecture sélective de l'information génomique, en assurant une régulation parallèle de séquences qui doivent rester éteintes (les ET) et de séquences qui doivent être exprimées au bon moment au cours du développement (les gènes). Cette lecture est primordiale au pour la survie et la « fitness » des individus, et est étroitement contrôlée, par la mise en place de plusieurs voies mécanistiques

différentes. Malgré ce contrôle étroit, des variations peuvent apparaître, générant ainsi une variation locale du niveau de méthylation, sans modification de la séquence nucléotidique. Si cette variation est maintenue à travers les générations, il y a alors apparition d'un épiallèle, c'est-à-dire un locus qui diffère entre deux individus par son état chromatinien (notamment la méthylation de l'ADN) sans changement dans la séquence nucléotidique, et qui peut se transmettre à travers les générations (Hofmeister *et al.*, 2017). Richards (2006) a proposé de classer les épiallèles en trois classes théoriques, selon le niveau de dépendance de la méthylation de l'ADN à la séquence génétique locale (Figure I.22A). Les épiallèles « purs » apparaissent indépendamment de la variation génétique suite à des événements stochastiques comme par exemple des erreurs dans le maintien de la méthylation après réplication de l'ADN. Les épiallèles « facilités » apparaissent avec une plus grande fréquence en présence d'un élément génétique, mais d'une manière probabiliste. Les épiallèles « obligatoires » sont complètement dépendants de la variation génétique et leur présence est conditionnée par la présence dans le génome de séquences particulières (Richards, 2006).

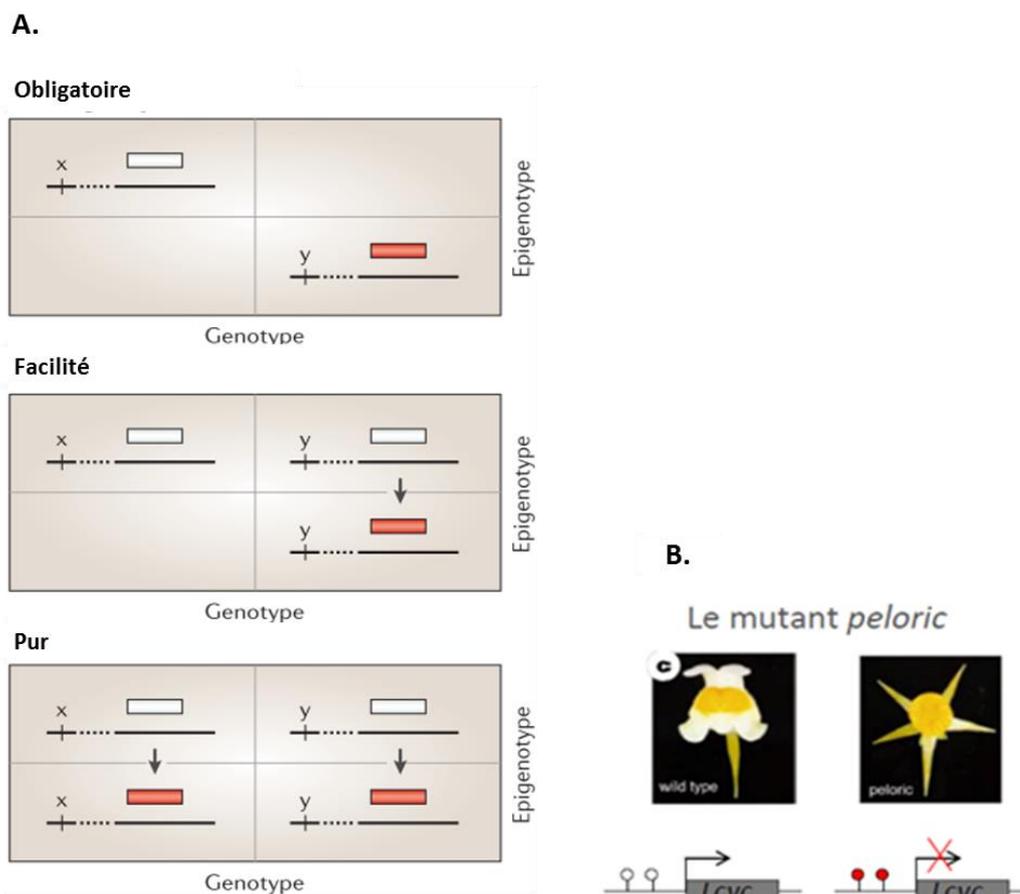


Figure I.22 : Classification d'épiallèles (d'après Richards (2006))

A. Trois classes d'épiallèles suivant le lien entre le génome et l'épigénome. L'axe horizontal distingue 2 génotypes avec les 2 allèles x et y. Les rectangles désignent une région génomique. Les rectangles rouges désignent l'installation d'un épiallèle.

B. Exemple d'un épiallèle pur (au moins en *cis*). La méthylation en amont du gène *Lcyc* engendre un phénotype pélorique chez la linéaire. Le changement de méthylation (rouge : méthylé, blanc, non méthylé) est corrélé au phénotype, mais aucun polymorphisme génétique n'est détecté dans la région (d'après Cubas *et al.* (1999)).

L'exemple le plus connu d'épiallèle est sans doute celui du locus *Lcyc* chez la linaria (*Linaria vulgaris*), où l'hyperméthylation en amont du gène affecte la symétrie radiale de la fleur, engendrant un phénotype pélorique (Figure I.22B) (Cubas *et al.*, 1999). Ce cas est considéré comme un épiallèle « pur », car aucune variation génétique n'a pu être décrite en association. La variation observée au niveau du méthylome peut aussi être la conséquence d'une variabilité dans la transmission de la méthylation. Par exemple, l'analyse des profils de méthylation des gènes méthylés dans l'écotype Columbia et non méthylés dans l'écotype *Landsberg* en F1 (16 individus) puis en F2 (80 individus) montre que la méthylation peut être perdue en F1 (1/16), puis héritée en F2. La méthylation des gènes est ainsi héritable mais de façon instable, ce qui pourrait expliquer le polymorphisme observé entre les accessions d'arabidopsis (Vaughn *et al.*, 2007). Cependant, ce type d'épiallèle n'est pas le plus fréquent. Par exemple, l'analyse de 80 RIL (Recombinant Inbred Line) de soja (*Glycine max*) montre que la majorité (90%) des DMR identifiées co-ségrègent avec des variations de la séquence d'ADN (avec une majorité pour un contrôle en *cis*) suivant les proportions de Mendel (Schmitz *et al.*, 2013a).

De façon assez fréquente, la variation de la méthylation de l'ADN entre individus est liée à l'insertion de l'ET ou au « spreading » du signal de méthylation associé à cet élément sur les régions adjacentes à l'insertion. Selon la stabilité de la méthylation associée, ces cas peuvent correspondre à des cas d'épiallèle facilité ou obligatoire. Un exemple d'épiallèle facilité est celui du locus *Agouti*, où l'expression du gène (et par suite le phénotype) dépend de l'état de méthylation de l'élément transposable (IAP) localisé en amont du gène (Morgan *et al.*, 1999) (Figure I.17E). C'est un cas d'épiallèle facilité, car son apparition est facilitée par la présence de la séquence IAP, mais pas complètement associée. Des épiallèles liés aux ET ont également été observés au sein des populations, ou d'ensemble d'accessions. Par exemple, la comparaison de 13 accessions américaines d'arabidopsis ayant divergé d'un ancêtre commun il y a environ une centaine d'années et présentant de faibles variations de séquences a permis de détecter 3 200 DMR avec des niveaux de méthylation élevés. Ces régions sont essentiellement localisées dans des régions intergéniques, sont héritables de façon stable, et sont associées à des variations de séquence (Hagmann *et al.*, 2015). L'analyse du niveau de méthylation d'insertions de l'ET polymorphes entre 140 accessions d'*A. thaliana* a révélé le lien fort entre présence d'un ET et présence de méthylation dans les régions adjacentes (Quadrana *et al.*, 2016a) : 50% des cas sont méthylés presque exclusivement dans les accessions portant l'insertion. Néanmoins, dans 40% des cas, la séquence adjacente est non méthylée dans toutes les accessions, incluant celles présentant l'insertion (Figure I.23). Ceci suggère que certains ET seraient incapables de générer du « spreading ». L'origine de cette différence reste cependant à élucider. De façon similaire, Stuart *et al.* (2016) décrit que l'effet de l'insertion sur la méthylation locale est dépendant de la distance, les DMR localisées à moins de 1kb ayant une plus forte corrélation avec la présence de l'ET que celles localisées plus loin. Enfin, l'observation d'une méthylation plus élevée (pour les trois contextes de cytosine) dans les ET non partagés entre accessions induit un niveau global de méthylation plus élevé dans les accessions ayant eu des activités de transposition récentes (Quadrana *et al.*, 2016b, Stuart *et al.*, 2016).

Des cas d'épiallèles obligatoires plus complexes ont été proposés (Eichten *et al.*, 2013). Ils impliquent des facteurs en *cis* et en *trans* pour l'origine de la variation méthylation de l'ADN dans le locus en question. Enfin, la paramutation peut définir un autre type d'épiallèle, où un allèle peut influencer l'état de la méthylation de l'ADN d'un autre allèle quand ils sont présents ensemble dans un hétérozygote.

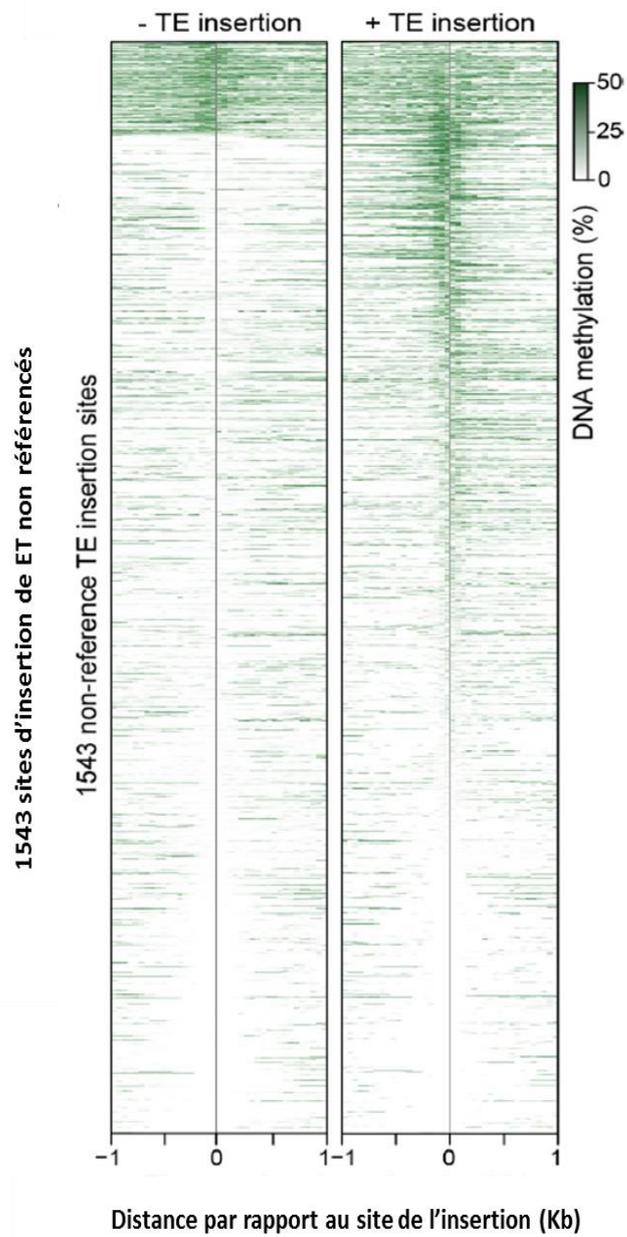


Figure I.23 : Niveau de méthylation de l'ADN dans les régions localisées 50pb en amont et en aval de 1543 sites d'insertion pour des accessions avec présence (+) ou absence de l'ET (Quadrona *et al.*, 2016a)

2. Régulation développementale

Comme je l'ai présenté plus haut, un des rôles de l'épigénome est de contrôler le bon déroulement du développement, en exprimant les gènes au moment approprié. D'ailleurs, c'est dans le but de définir une théorie permettant de comprendre comment la croissance et la différenciation embryonnaire peuvent générer des tissus différents à partir d'une même cellule indifférenciée et portant une seule information génétique que Waddington a défini le terme d'épigénèse, qui a donné ensuite son nom à l'épigénétique (Waddington, 1942a, Waddington, 1942b). Ce terme a par la suite connu différentes définitions, faisant référence aux changements de l'activité des gènes qui sont mitotiquement et/ou méiotiquement héréditaires et qui ne reposent pas sur une modification de la séquence de l'ADN (Holliday, 1987, Holliday, 1994). Avec la description des modifications de la chromatine, l'épigénétique a ensuite été définie par Bird comme l'étude des changements de l'activité des gènes transmis à travers les divisions cellulaires suite aux changements des états de la chromatine, en l'absence des conditions qui lui ont données naissance (Bird, 2007). En effet, ces modifications permettent une lecture programmée du génome au cours du développement et la transmission mitotique de l'état de l'expression des gènes. Par exemple, les gènes du groupe *polycomb* (*PcG*) agissent sur la méthylation des histones et permettent le contrôle de l'expression des gènes lors du développement. Ils contribuent à préserver cet état d'expression en constituant une mémoire moléculaire, et ainsi de maintenir une information mitotiquement héréditable (Orlando, 2003, Eichten *et al.*, 2014).

La comparaison de la méthylation de l'ADN entre différents organes chez les plantes révèle l'existence de régions différentiellement méthylées ce qui montre le rôle de cette marque chromatinienne dans le développement. Ceci a été observé chez *A. thaliana*, en comparant les cotylédons et les organes végétatifs et reproductifs (Ruiz-Garcia *et al.*, 2005). L'initiation de la floraison et la longueur des racines sont des processus développementaux liés à des différents profils au niveau de la méthylation de l'ADN chez *Arabidopsis* (Cortijo *et al.*, 2014). Chez le maïs, la comparaison entre trois tissus montre que le niveau de méthylation le plus bas est trouvé chez l'épi en comparaison avec deux types de feuilles, suggérant l'existence d'un rôle différent de la méthylation au niveau des organes végétatifs et reproducteurs (Lu *et al.*, 2008). Chez le pommier (*Malus domestica* Borkh.), le niveau de méthylation du contexte CHH est plus élevé dans l'ADN du fruit que dans celui des feuilles, suggérant l'activité de la machinerie RdDM dans cet organe. Cette différence de méthylation a lieu à proximité de 249 gènes annotés comme régulateurs de la floraison et du développement du fruit (Daccord *et al.*, 2017), illustrant le rôle de cette voie de régulation de la méthylation dans le développement du fruit. Des variations de méthylation entre organes sont également observées chez le coton (*Gossypium*) (Osabe *et al.*, 2014), *Brachypodium distachyon* (Roessler *et al.*, 2016), le sorgho (*Sorghum bicolor*) (Turco *et al.*, 2017) et la fraise (*Fragaria vesca*) (Gu *et al.*, 2016). La méthylation de l'ADN varie aussi en fonction de la maturité des tissus. En effet, les organes matures (feuilles matures, graines, fruits) sont souvent plus méthylés que les organes immatures (endospermes, apex de la tige, jeunes feuilles, tiges et racines des plantules) (Teyssier *et al.*, 2008, Zemach *et al.*, 2010).

Au sein d'un même organe, des différences de méthylation sont également observées au cours du temps. Chez la tomate, la maturation du fruit est associée à une déméthylation de l'ADN au niveau du site d'accrochage du facteur de transcription (RIN) au niveau du promoteur de plusieurs gènes responsables de la maturité du fruit (Zhong *et al.*, 2013). Chez la pomme, la présence de changement de méthylation des promoteurs de trois gènes contrôlant le nombre de cellules dans le fruit (et ainsi

sa taille) suggère également, le rôle de la méthylation de l'ADN dans le développement du fruit (Daccord *et al.*, 2017).

A l'inverse de ce qui est observé chez les mammifères, la méthylation de l'ADN est moins reprogrammée chez les plantes. Chez *A. thaliana*, on observe une déméthylation des séquences d'ET des cellules centrales du sac embryonnaire et de l'albumen, qui affecte essentiellement les cytosines en contexte CG (Quadrana & Colot, 2016). Du côté paternel, une déméthylation est observée dans le noyau végétatif des grains de pollen, mais celle-ci affecte moins les ET. Au sein des cellules produisant directement le zygote, on observe une absence de méthylation en CHH dans les spermatozoïdes. Il semble que le profil de méthylation de ces cellules soit également maintenu chez le zygote et l'embryon, vraisemblablement par l'action des méthyltransférases MET1 et CMT3 et indépendamment de l'action de la voie RdDM. Ainsi, les plantes présentent une absence de reprogrammation au niveau des cytosines CG et CHG. Au cours du développement de l'embryon, la méthylation CHH se met progressivement en place par la voie RdDM (Calarco *et al.*, 2012, Ibarra *et al.*, 2012, Jullien *et al.*, 2012, Kawashima & Berger, 2014). Enfin, le méristème présente une activité RdDM spécifique qui renforce le contrôle des ET pendant le passage entre la phase végétative et la phase reproductive (Baubec *et al.*, 2014).

3. Réponse du méthylome aux conditions environnementales

3.1 Réponse sous contrainte

3.1.1. Analyses ciblées sur des locus choisis

Les variations de l'épigénome constituent une cible de plus en plus étudiée pour comprendre les mécanismes de la réponse des plantes aux variations de l'environnement. En effet, la chromatine possède une structure flexible qui varie en condition de stress (Santos *et al.*, 2017).

Plusieurs études se sont intéressées à l'étude de la méthylation de l'ADN au niveau de locus particuliers dans le génome, notamment les gènes de réponse aux contraintes abiotiques, afin d'étudier le lien entre les variations de la méthylation et l'expression de ces gènes, en conditions expérimentales. Ces études ont montré que des contraintes abiotiques variées induisent des modifications de méthylation de ce type de gènes dans différentes espèces. Chez la tomate (*Solanum lycopersicum*), le déficit hydrique est accompagné par une déméthylation d'environ 100 cytosines du gène *Asr1*. Ce changement est accompagné par une diminution de la marque H3K27me3 et une induction de l'expression du gène (González *et al.*, 2011). Le stress salin chez le riz (*Oryza sativa*) induit une déméthylation des éléments transposables ainsi qu'une déméthylation et une hyperméthylation des gènes de réponses au stress abiotiques et des gènes impliqués dans la modification de la chromatine (Karan *et al.*, 2012). Chez le soja (*Glycine max*) le stress salin induit la déméthylation de quatre gènes codant des facteurs de transcription, et est associée à une triméthylation de H3K4 et une acétylation de H3K9 dans les promoteurs (Song *et al.*, 2012). D'autres études ont cherché le lien entre les variations de méthylation et la sensibilité à la contrainte appliquée. La comparaison de la méthylation entre deux génotypes de blé (*Triticum aestivum* L.) de sensibilité différente montre une différence de méthylation au niveau du gène TaGAPC1 (Cytosolic glyceraldehyde-3-phosphate dehydrogenase) suite à un déficit hydrique. Les stress salin et osmotique induisent une déméthylation CG et CHG au niveau du

promoteur chez le génotype tolérant à la sécheresse en comparaison au génotype sensible où il y a augmentation au niveau de la méthylation CHG et CHH (Fei *et al.*, 2017).

L'influence de l'environnement sur le méthylome est aussi appréhendée en analysant les niveaux d'expression des gènes de régulation de l'épigénome. Par exemple, la température élevée impacte les protéines impliquées dans la voie RdDM tels que DRM2 et les ARN polymérases (NRPD1 et NRPE1) chez *Arabidopsis* (Naydenov *et al.*, 2015a). Chez la cacahuète, le froid et les températures élevées influencent aussi les niveaux d'expression des méthyltransférases, tout comme le déficit hydrique chez le maïs (Wang *et al.*, 2016a, Qian *et al.*, 2014). Chez le riz (*Oryza sativa*), les stress salin et osmotique modifient les niveaux d'expression des méthyltransférases, notamment celle de *OsMET1-1* qui se retrouve réduite dans les racines. L'expression de *OsDNMT2* est augmentée en présence de stress salin et réduite en présence de contraintes osmotiques. (Ahmad *et al.*, 2014, Sharma *et al.*, 2009).

3.1.2. Analyses à l'échelle du génome

Au contraire des analyses citées auparavant et qui ciblent des régions particulières du génome, d'autres études se sont intéressées aux variations du méthylome à l'échelle du génome. Chez *A. thaliana*, la majorité des variations de la méthylation de l'ADN détectées le long du génome suite à un déficit en phosphate sont localisées dans les séquences d'ET, suggérant que la méthylation intervient pour renforcer le contrôle et le silencing des régions génomiques hautement méthylées. La méthylation au niveau des transposons corrèle avec des changements dans l'expression des gènes d'à proximité. Les variations de méthylation de l'ADN sont présentes au niveau des régions codantes des gènes (gbM) (Yong-Villalobos *et al.*, 2015). Le stress osmotique chez *A. thaliana* induit des changements de méthylation essentiellement dans/à côté des éléments transposables (Wibowo *et al.*, 2016). D'un autre côté, l'application de traitement au plomb (Pb) chez le maïs à des intervalles de temps successifs (12, 24 et 48h) montre que la méthylation de l'ADN est la plus élevée dans les régions intergéniques suivies par les introns et les promoteurs et affecte essentiellement le contexte CG. Dans les cas où la variation de méthylation est localisée dans les promoteurs, le niveau d'expression des gènes est négativement corrélé avec la présence de méthylation (Ding *et al.*, 2014). Chez *Leymus chinensis*, les températures élevées et les carences en azote induisent des hyper- et des hypo-méthylations, qui sont le plus souvent localisées dans des régions riches en éléments transposables (Yu *et al.*, 2013). A l'inverse, chez le maïs, l'analyse de l'impact du froid, du chaud et du rayonnement UV montre que ces contraintes induisent des changements de méthylation au niveau des feuilles, et que les DMRs correspondantes sont enrichies au niveau des gènes. Cependant, une variation est observée entre différentes plantes du lot témoin, suggérant que le niveau de méthylation varie d'une plante à l'autre même en conditions standard. Cette variation est plus marquée au sein des plantes ayant été exposées aux contraintes abiotiques (Eichten & Springer, 2015).

L'ensemble de ces études montre que la méthylation des ET est souvent affectée par les contraintes appliquées. Ces cas peuvent être des hyper- ou des hypo-méthylation, suggérant que certains ET pourraient être contrôlés de façon plus efficace en réponse au stress alors que d'autres pourraient ainsi échapper au contrôle chromatinien de leur expression. Dans le cas de certaines contraintes abiotiques, la déméthylation entraîne la transposition de certains ET. Par exemple, chez *Antirrhinum majus*, le stress froid induit l'hypométhylation de l'ADN et par suite la transposition de l'élément *Tam-3* (Hashida *et al.*, 2006). La méthylation de l'ADN, et même la régulation chromatinienne n'est cependant pas le seul intervenant : l'activation de plusieurs éléments transposables chez *A. thaliana*

suite à un stress thermique prolongé est accompagnée par une perte du nucléosome et une décondensation de l'hétérochromatine, mais qui a lieu sans perte de méthylation d'ADN et avec uniquement des changements mineurs dans les modifications d'histone (Pecinka *et al.*, 2010). La mutation dans les voies de régulation de la méthylation, combinée à une contrainte environnementale, est parfois nécessaire à la transposition. C'est le cas du rétrotransposon de type *Copia ONSEN* chez *A. thaliana*, dont la transcription et les formes extra-chromosomiques sont observées suite à un stress chaud, mais qui ne produit de nouvelles copies que quand la voie des petits ARN est affectée (Ito *et al.*, 2011). Le promoteur d'*ONSEN* ne contient cependant pas de sites CG ou CHG, et la réduction en CHH n'est pas suffisante pour sa prolifération. Celle-ci pourrait être liée à la présence de régions régulatrices de réponse à la température au sein de la séquence de l'élément (Cavrak *et al.*, 2014). Enfin, *ONSEN* est actif dans des mutants de la polymérase II d'*A. thaliana*, ce qui indique que la déméthylation engendrée par une déficience dans l'activité de la PolII est suffisante pour son activation (Thieme *et al.*, 2017).

3.1.3. Réponse en association avec d'autres marques de la chromatine

Si la méthylation de l'ADN est modifiée par les contraintes environnementales, cette modification est souvent associée à la modification d'autres marques de la chromatine. Par exemple, chez le maïs, l'expression du gène de réponse au froid *ZmICE1* (inducer of CBF/DREB expression 1) est modulée par une hyperacétylation et une déméthylation qui ont lieu dans des régions du promoteur, du premier exon et des domaines de liaison au facteur ICE (Hu *et al.*, 2011). D'autre part, une diminution de l'acétylation de la marque H3K9 est observée dans les régions géniques et intergéniques après traitement au froid (4°C, chaque 6h pendant 24h) chez des plantules de maïs. D'autres études montrent l'existence de variations au niveau d'autres marques de la chromatine, sans que la méthylation de l'ADN soit affectée (Zhao *et al.*, 2014a, Zhao *et al.*, 2014b).

Ainsi, il est parfois difficile de savoir si la méthylation de l'ADN est directement affectée par la contrainte elle-même, ou si sa modification est un effet secondaire, lié à la modification d'autres marques (Pikaard & Mittelsten Scheid, 2014, Probst & Mittelsten Scheid, 2015). D'autre part, il n'est pas encore bien clair si les variations de l'épigénome sont la cause ou la conséquence de la réponse de la plante aux conditions de l'environnement. L'augmentation de l'expression des gènes de réponse au stress pourrait générer des transcrits de type PolII, qui pourrait induire la mise en place de la méthylation dans les régions adjacentes, et ainsi augmenter la méthylation des éléments transposables localisés à côté de ces gènes. Ce mécanisme a été proposé comme une réponse à la déficience en phosphate chez le riz (Secco *et al.*, 2015) (Figure I.24). Cependant, il reste encore assez hypothétique.

Comprendre l'interaction entre l'environnement et l'épigénome reste une tâche complexe. Certaines limites sont dues à la difficulté de mettre en place des approches expérimentales bien contrôlées. De plus, dans la nature les plantes sont simultanément confrontées à différents types de contraintes résultant dans des réponses additives ou antagonistes, et leur analyse combinée reste difficile à appréhender en conditions contrôlées (Meyer, 2015, Probst & Mittelsten Scheid, 2015, Atkinson & Urwin, 2012). Enfin, les métabolites pourraient influencer l'acquisition ou la perte des marques de la chromatine (Vriet *et al.*, 2015) (Figure I.25). Du fait de la plasticité de la méthylation, l'analyse de la stabilité de la réponse entre individus est importante. Quand plusieurs individus ont été analysés, une variabilité importante a été notée entre individus, et il semble que les profils de méthylation soient

plus instables suite à un stress qu'en conditions « standard » (Eichten & Springer, 2015). Cette instabilité pose notamment la question de la détection des signaux induits, et du nombre d'individus à analyser pour détecter un signal et analyser sa stabilité.

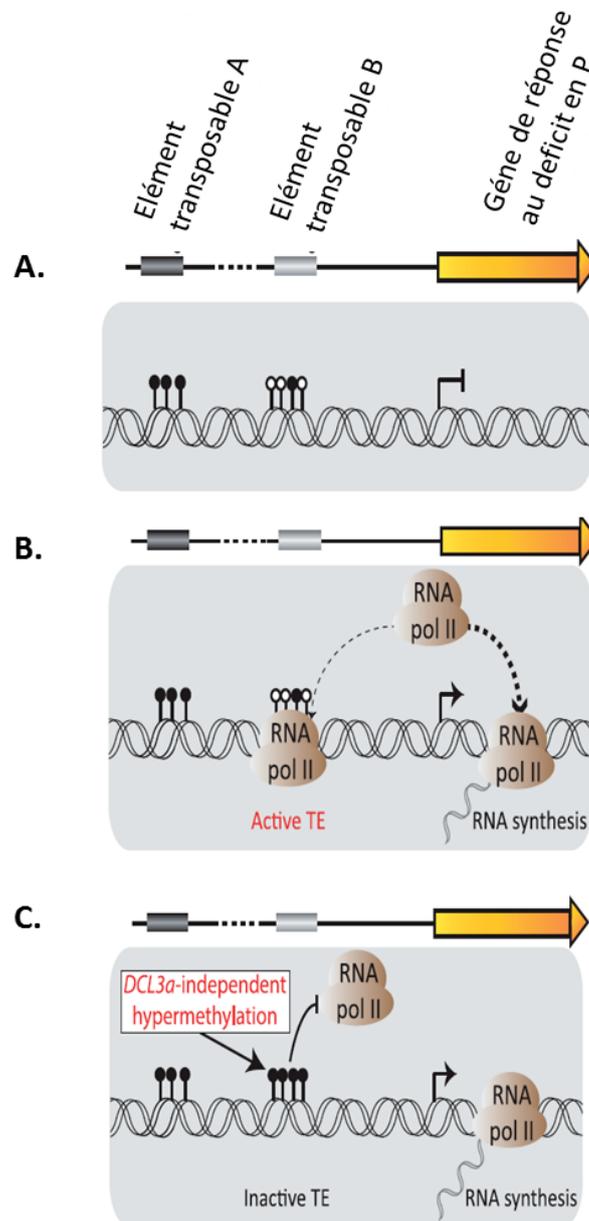


Figure I.24 : Interaction entre l'expression des gènes et la méthylation d'ADN comme réponse au déficit en phosphate (P) (Secco *et al.*, 2015)

A. En absence de déficit en phosphate, l'élément transposable (ET) B (proche du gène de réponse au déficit en phosphate) est moyennement méthylée, et l'ET A (localisé plus loin) est très méthylé. **B.** En cas de déficit en phosphate à court terme, la transcription du gène induit une activité élevée de Pol II qui peut également induire la transcription de l'ET B. **C.** En cas de déficit en phosphate prolongé, l'augmentation de la méthylation permet d'inactiver l'activité de Pol II sur l'ET B.

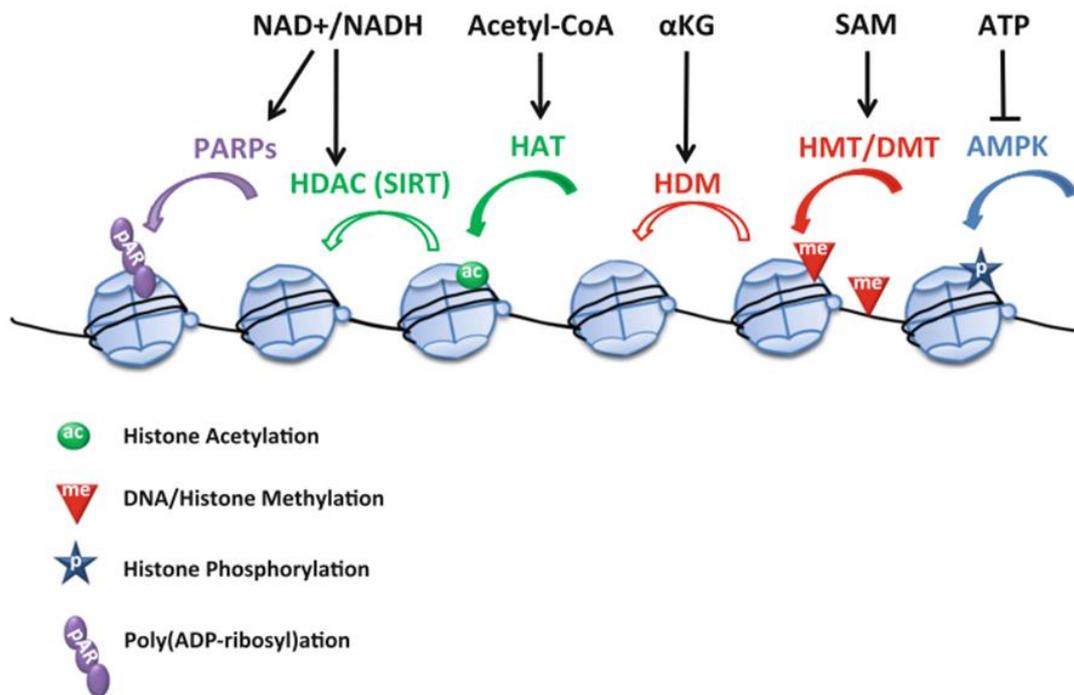


Figure I.25 : Interaction entre le métabolisme et la chromatine (Vriet *et al.*, 2015)

Dans ce modèle, l'acquisition ou la perte des marques de la chromatine est influencée par les métabolites. Le ratio NAD⁺/NADH régule l'activité de la déacétylase HDAC et de la polymérase (PARPs), deux enzymes impliquées dans la réparation de l'ADN, la réplication, la compaction de la chromatine et la transcription des gènes. αKG est un cofacteur qui agit sur HDM pour la déméthylation de l'histone. L'Acetyl-CoA régule HAT qui induit l'acétylation de l'histone. SAM est un substrat pour les histones et les méthyltransférases. La concentration faible de l'ATP active la kinase AMPK responsable de la phosphorylation de l'histone.

Sous des conditions des stress, il y a réduction de la production d'énergie par la plante. Le métabolite ATP diminue, induisant l'augmentation de AMPK et celle du ratio NAD⁺/NADH, induisant l'activité de SIRT. Il se produit aussi une modification du cycle des citrates, ce qui fait appel à acetyl-CoA et αKG, affectant ainsi l'acétylation et la déméthylation des histones.

SIRT: sirtuin NAD⁺-dependent histone deacetylases, PARPs: poly(ADP-ribose) polymerases, αKG: α-ketoglutarate, HDM: Histone lysine Demethylase, SAM: S-adenosylmethionine, HMT: Histone methyltransferase, DMT: DNA methyltransferase.

3.2. Maintien et transmission de la réponse induite

Si l'environnement peut induire une variation de la méthylation de l'ADN chez les plantes, une question majeure pour l'adaptation est son maintien après retour à des conditions standard, et sa transmission à travers la mitose. Ces étapes sont indispensables à la détection de variations épigénétiques (Roudier *et al.*, 2009).

Dans certains cas, la méthylation est réversible une fois que les conditions redeviennent « normales ». Par exemple, des perforations réalisées sur des feuilles de maïs montrent une diminution de la teneur en 5-méthylcytosine mais cet état est réversible, avec un retour à l'état initial quelques heures après arrêt du traitement. Cependant, la déméthylation transitoire observée dans le premier exon du gène *ZmCPK11* (Calcium-Dependent Protein kinase) codant une protéine impliquée dans les voies de

signalisation des stress chez le maïs, est associée à une augmentation dans la transcription de *ZmCPK11* quelques heures après (Lewandowska-Gnatowska *et al.*, 2014). Ceci suggère que des changements même transitoires de la méthylation pourraient avoir un effet à plus long terme. A l'inverse, certains changements de méthylation induits par des contraintes abiotiques peuvent se maintenir durant la vie de la plante même après arrêt de la contrainte. Par exemple, Secco *et al.* (2015) montrent que les changements de méthylation apparus suite à une déficience en phosphate (une hyperméthylation des séquences d'ET proches des gènes) persiste à travers la mitose après l'arrêt du traitement.

Une deuxième question est celle de la transmission sur une ou plusieurs générations des changements induits. Comme je l'ai présenté plus haut, contrairement aux mammifères où la régulation du méthylome dans les gamètes, le zygote et l'embryon est essentielle pour reprogrammer le développement des cellules (Seisenberger *et al.*, 2013), les plantes connaissent moins de reprogrammation au niveau de leur méthylome, et celui-ci est plutôt renforcé que reprogrammé durant la reproduction sexuée (revu dans (Kawashima & Berger, 2014, Quadrana & Colot, 2016)). De plus, chez les plantes, la lignée germinale apparaît assez tard dans la vie de la plante et à partir des cellules somatiques différenciées (Baulcombe & Dean, 2014). De ce fait, les cellules à l'origine de la lignée germinale ont un profil de méthylation qui est la somme de toutes les modifications accumulées au long de la vie de la plante, en particulier en réponse à des contraintes environnementales.

Afin de tester la stabilité transgénérationnelle de la méthylation, des populations dédiées ont été construites, les EpiRIL, pour « Epigenetic Recombinant Inbred Lines ». Les parents de ces lignées sont génétiquement identiques sauf à un locus, où l'un des deux parents porte un allèle mutant pour un gène impliqué dans une voie de régulation de la méthylation : *met1* (Reinders *et al.*, 2009) ou *ddm1* (Johannes *et al.*, 2009). Après croisement avec le parent sauvage et sélection des individus F2 homozygotes sauvages pour ce gène, les plantes sont autofécondées sur six générations pour obtenir les EpiRIL. Les EpiRIL montrent une héritabilité de la méthylation parentale de la méthylation de l'ADN jusqu'à la huitième génération (Johannes *et al.*, 2009), ainsi que des gains progressifs de méthylation à certains locus (Teixeira & Colot, 2009).

Cependant, dans le cas des mutants, la variation de méthylome observée affecte toute la plante. Dans le cas des contraintes abiotiques, une partie seulement de la plante peut être affectée, ce qui change vraisemblablement les chances de transmission à la descendance. Notamment, il est clair que le type de lignée cellulaire affectée peut, en partie, expliquer la transmission ou non des variations du méthylome dans la descendance. Par exemple, dans le cas de la réponse du riz au phosphate, les changements observés affectent essentiellement le niveau de méthylation des racines, un organe qui ne participe pas à la lignée cellulaire produisant les tissus reproducteurs (Secco *et al.*, 2015). De surcroît, la mémoire des modifications induites par le stress peut être transmise à la génération suivante, mais être perdue par la suite si la contrainte n'est pas réappliquée. C'est le cas de l'exposition de la plante *A. thaliana* à des contraintes osmotiques (Wibowo *et al.*, 2016). Dans ces cas, la transmission est, de plus, différente entre mâle et femelle : la mémoire est transmise par la lignée maternelle, mais dans la lignée germinale paternelle la méthylation de l'ADN est reprogrammée de façon active par l'activité de l'ADN glycosylase (Wibowo *et al.*, 2016).

Pour ces raisons, la méthylation de l'ADN pourrait être transmise à la descendance, et ainsi participer à l'adaptation des populations. Afin de tester le rôle des épiallèles dans la variation phénotypique, des QTL ont été analysés à partir des populations d'epiRIL. Ces analyses ont montré une variation phénotypique héritable pour des traits complexes comme la hauteur de la plante, la longueur des

racines et la précocité de floraison (Johannes *et al.*, 2009, Cortijo *et al.*, 2014), illustrant ainsi la contribution des épiallèles à la variation phénotypique.

L'efficacité de cette transmission n'est cependant pas totale, ce qui pourrait induire une variabilité assez élevée dans la descendance (Pecinka & Mittelsten Scheid, 2012) et permettrait une héritabilité qualifiée de « douce » (Boyko & Kovalchuk, 2011). Cette héritabilité semble être plus élevée quand les variations de méthylome induites sont associées à des variations génétiques, comme l'insertion d'éléments transposables (Rey *et al.*, 2016). Par exemple, la déficience en azote chez le riz induit une hyperméthylation au niveau de deux locus d'éléments transposables. Ces altérations sont détectées jusqu'à la troisième génération en absence d'induction de la même contrainte (Kou *et al.*, 2011).

Enfin, une partie de la variabilité de méthylation induite pourrait être due à la réponse de gènes impliqués dans le contrôle de la méthylation. Les changements observés seraient alors un effet secondaire de l'impact de la contrainte, et non un effet direct sur plusieurs locus. Ainsi, la culture de 150 accessions suédoises d'*Arabidopsis* dans deux conditions contrôlées de température (10 ou 16 °C) a montré que l'augmentation de la température induit une augmentation de 14% de la méthylation CHH, essentiellement dans les séquences d'ET. Des analyses GWAS utilisant les profils de méthylation comme phénotype ont permis d'identifier une variation allélique au niveau du locus CMT2 qui corrèle avec les conditions climatiques naturelles des accessions (Dubin *et al.*, 2015).

En conclusion, le méthylome est variable sous influence des contraintes abiotiques de l'environnement, et peut affecter différentes régions génomiques, en particulier des ET. Le lien entre méthylation et expression génique reste corrélatif et ne permet pas de comprendre la relation de cause à effet entre ces variations. Les variations de la méthylation peuvent sous-tendre une partie de la plasticité phénotypique et de l'adaptation, et la transmission entre générations semble être d'autant plus forte que l'épiallèle est facilité par la présence d'une séquence génomique particulière. Ces séquences sont souvent des éléments transposables, et l'interaction insertion/méthylation au site d'insertion pourrait permettre à la fois une transmission forte (car facilitée) tout en conférant une flexibilité du signal (« spreading » plus ou moins étendu par exemple).

IV. Le maïs, un modèle d'étude pour analyser l'impact de l'environnement sur le méthylome

1. Organisation du génome

Le maïs a un génome de 2.3Gb. Il est considéré comme une espèce à « gros génome » mais cette taille est en fait légèrement en dessous de la médiane des tailles de génome des angiospermes (Figure I.26), et fait de cette espèce une espèce « moyenne » en terme de taille de génome. Le séquençage du génome de la lignée Américaine B73 (Schnable *et al.*, 2009) a révélé un total de 32,000 gènes répartis sur les 10 chromosomes (Figure I.27). La dernière version du génome en contient 39,324 (Jiao *et al.*, 2017). Le génome de B73 contient environ 85% d'éléments transposables appartenant à différentes classes. Les rétrotransposons à LTR sont largement majoritaires (60% du génome de la version 4), suivis des héliçons (3,6%). Les transposons à ADN, les LINE et les SINE ne représentent qu'un faible pourcentage du génome (Jiao *et al.*, 2017). Ces nombres sont vraisemblablement sous-évalués. Par exemple, ils estiment la contribution des LINE à 0,01% du génome, alors qu'une analyse approfondie des rétrotransposons a estimé la part génomique des LINE à 1% sur la version 2 (Baucom *et al.*, 2009).

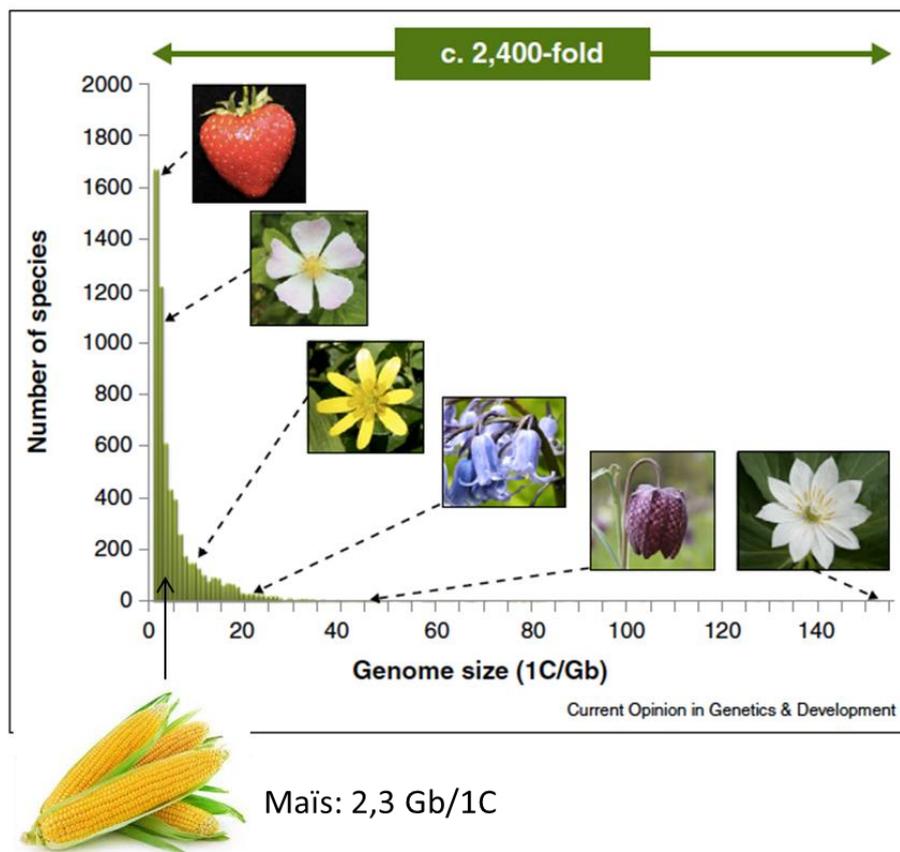


Figure I.26: Distribution de la taille du génome de 7542 angiospermes (adapté de Dodsworth *et al.* (2015))

Les tailles sont en Gb pour 1C, et proviennent de la base de données *Plant DNA C-values database* (www.data.kew.org/cvalues). Les images montrent des plantes représentatives de différentes tailles de génome. De gauche à droite : *Fragaria ananassa* 1C/0.60 Gb; *Rosa canina* 1C/1.39 Gb; *Ranunculus ficaria* 1C/9.12 Gb; *Hyacinthoides non-scripta* 1C/20.73 Gb; *Fritillaria meleagris* 1C/46.26 Gb; *Paris japonica* 1C/148.8 Gb. La position de la taille du génome du maïs (2,3Gb) illustre que celle-ci correspond à une taille moyenne pour les angiospermes.

Les séquences répétées du génome sont distribuées tout le long des chromosomes. Elles sont essentiellement localisées dans les régions péri-centromériques et moins abondantes dans les extrémités chromosomiques riches en gènes (Figure I.27). Cette distribution est largement façonnée par la densité en rétrotransposons à LTR (Figure I.28). Les deux grandes superfamilles de rétrotransposons à LTR (RLG pour « Retrotransposon LTR *Gypsy* », et RLC pour « Retrotransposon LTR *Copia* ») ont des distributions chromosomiques différentes : la superfamille RLG est concentrée dans les régions de l'hétérochromatine péri-centromérique, alors que les éléments de type RLC montrent une préférence pour s'accumuler dans les régions euchromatiniennes des extrémités des chromosomes (Baucom *et al.*, 2009). Les LINE et SINE sont enrichis dans les régions géniques.

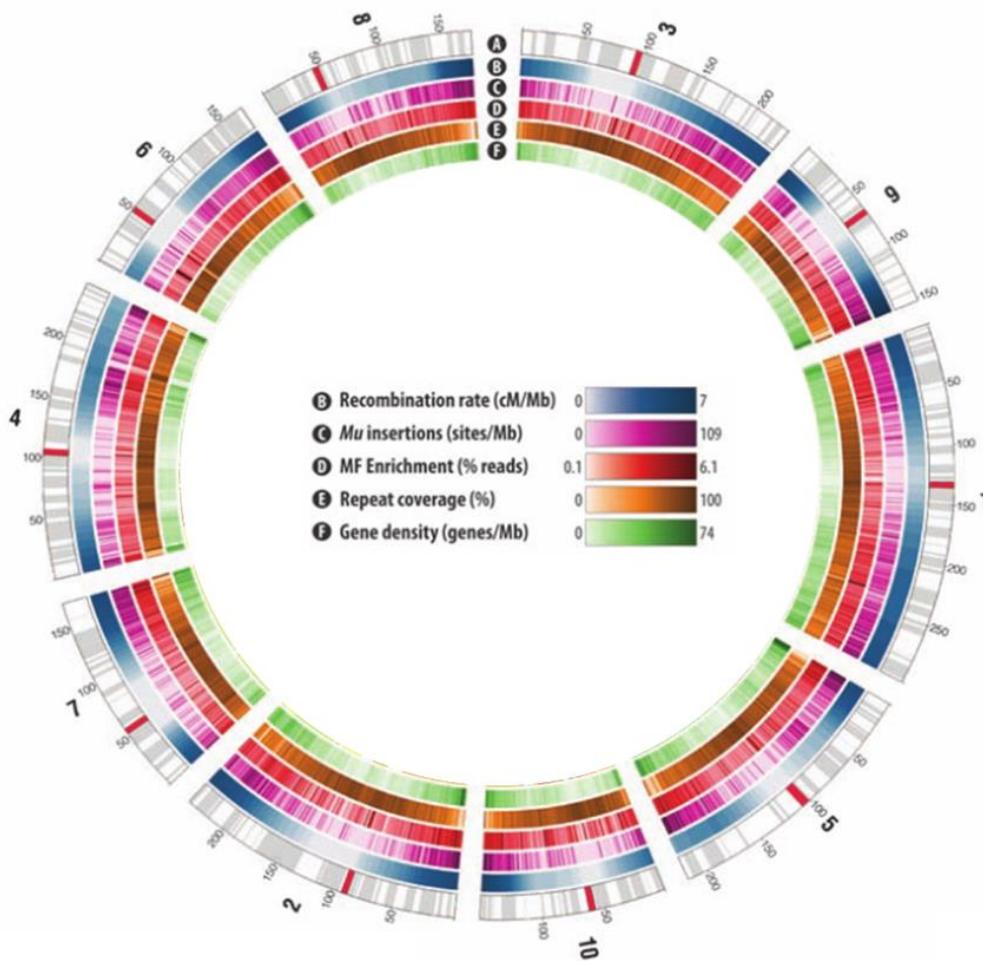


Figure I.27 : Génome de référence du maïs (B73 RefGen_v1) (adapté de Schnable *et al.* (2009))
 Les cercles montrent la structure du génome en 10 chromosomes. **A.** Carte génétique Assemblage de contigs en gris et en blanc. En rouge le centromère. **B.** Carte de déséquilibre de liaison fondée sur 6363 marqueurs physiques et génétiques. **C.** Sites d'insertions de l'élément *Mu*. **D.** Pourcentage des reads alignés (MF:Methyl-filtration). **E.** Pourcentage de couverture en éléments transposables (Repeat Masker). **F.** Densité en gènes à partir du comptage du nombre de gènes sur des fenêtres glissantes de 1Mb avec un pas de 200Kb.

L'analyse fine des copies de rétrotransposons à LTR a permis d'identifier un total de 406 familles ont qui constituent 1,7 Gb (75%) de l'ADN nucléaire total (Baucom *et al.*, 2009). Ces familles ont un nombre très variable de copies. Les 20 familles les plus abondantes constituent environ 70% du génome ; ils appartiennent en majorité à la superfamille *gypsy* (13/20). A l'inverse, beaucoup de familles ne possèdent qu'un petit nombre de copies (Figure I.29). Les familles possédant le plus grand nombre de copies sont *Huck* (3341), *ji* (4093), *Cinful-zeon* (9844), et *Opie* (3530) (Baucom *et al.*, 2009). La densité des différentes familles varie beaucoup le long des chromosomes : les familles à faible nombre de copies (<10) sont localisées dans les régions enrichies en gènes (dans ou à côté des gènes), alors que les familles avec un nombre de copies élevé (100 millions) sont préférentiellement localisées dans d'autres rétrotransposons LTR. Les familles ayant entre 10 et 100 copies sont localisées dans les deux espaces de façon équilibrée (Baucom *et al.*, 2009) (Figure I.30).

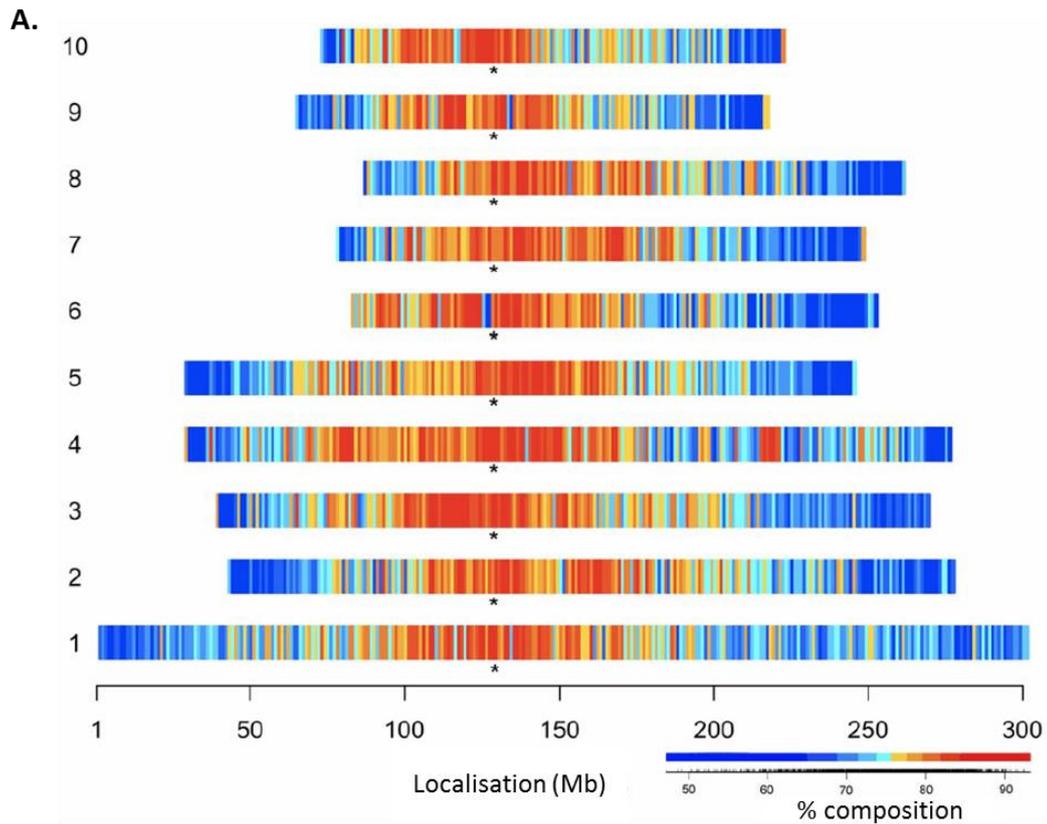


Figure I.28 : Densité en séquences de rétrotransposons à LTR dans le génome du maïs (Baucom *et al.*, 2009)

Distribution de la densité en séquences de rétrotransposon à LTR le long des 10 chromosomes du maïs. Les couleurs froides indiquent une densité faible, les couleurs chaudes une densité forte. La distribution montre un enrichissement dans les régions péri-centromériques.

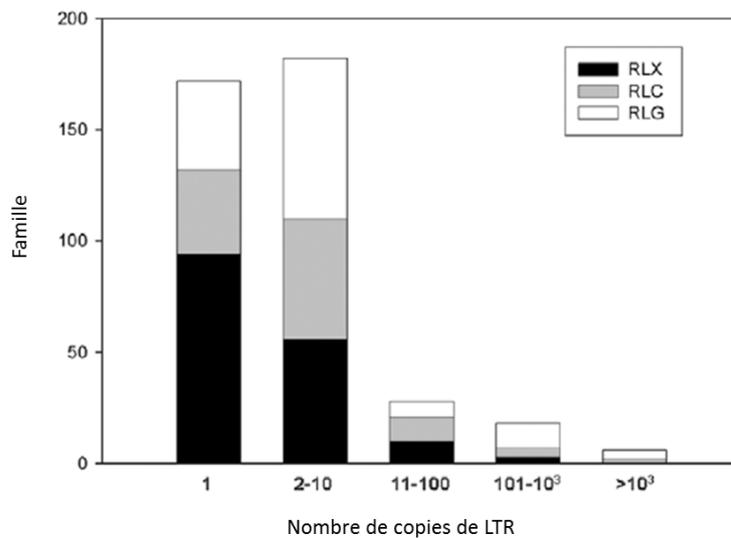


Figure I.29 : Distribution du nombre de copies par famille de rétrotransposons à LTR dans le génome du maïs (Baucom *et al.*, 2009)

La grande majorité des familles comporte moins de 10 éléments. Les éléments présentant le plus grand nombre de copy sont de la superfamille *gypsy* (RLG, en blanc).

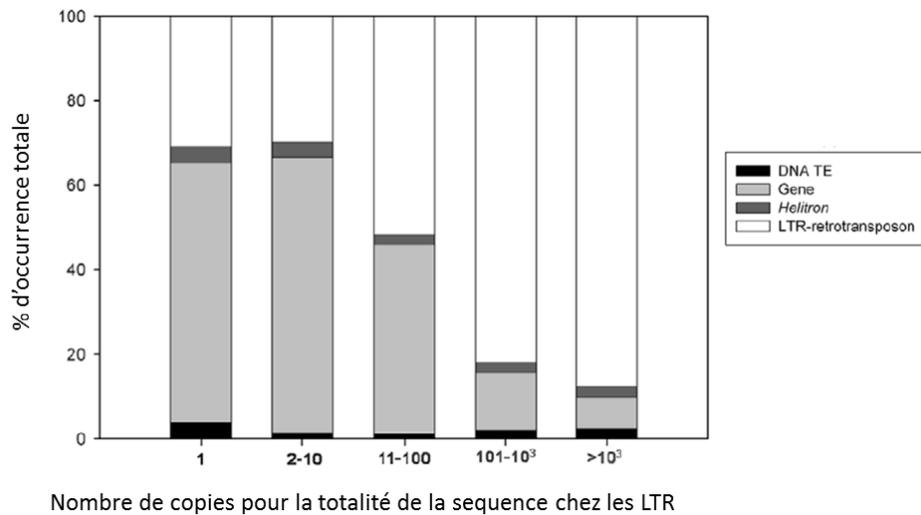


Figure I.30 : Localisation des différentes familles de rétrotransposon à LTR en fonction du nombre de copie (Baucom *et al.*, 2009)

Les familles les moins abondantes (<10 copies) sont majoritairement localisées dans les gènes (en gris). A l'inverse, les familles présentant un grand nombre de copies (>100) sont essentiellement localisées dans d'autres rétrotransposons à LTR.

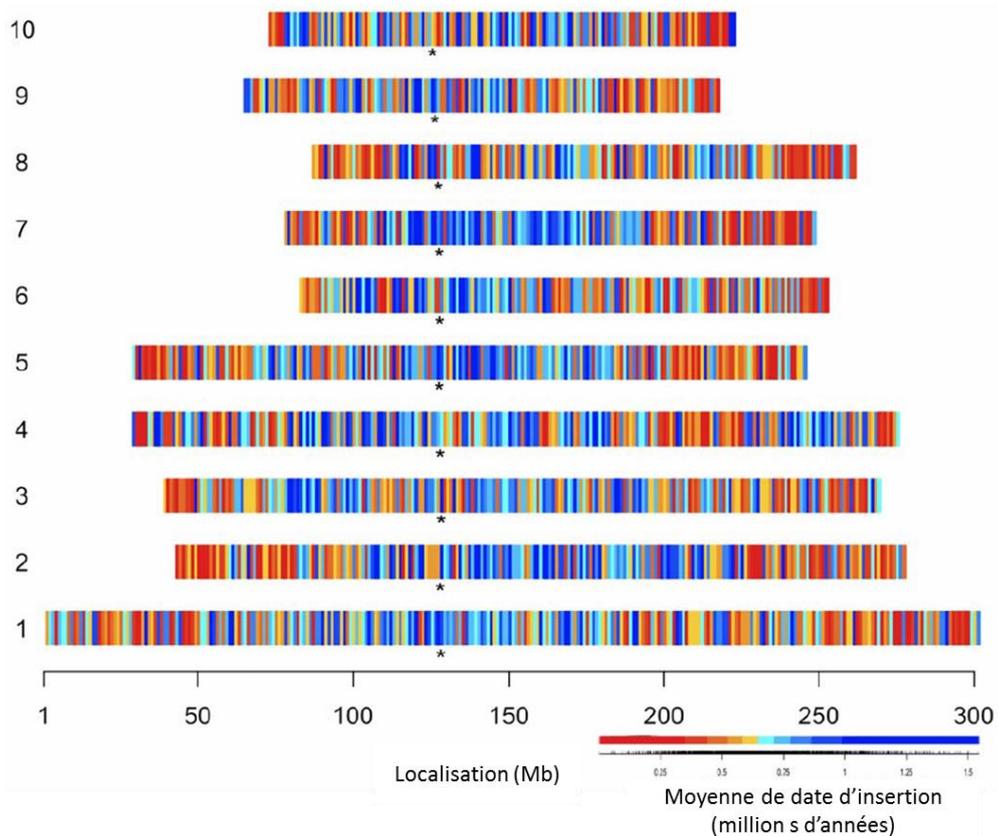


Figure I.31 : Densité en rétrotransposons à LTR en fonction de leur date d'insertion (Baucom *et al.*, 2009)

Les couleurs chaudes indiquent des régions portant des éléments jeunes, les couleurs froides des éléments anciens. Les copies anciennes sont enrichies dans les régions péri-centromériques.

Les copies de rétrotransposons à LTR sont également organisées le long des chromosomes en fonction de leur date d'insertion : les insertions récentes (< 0,25 millions d'années) sont enrichies dans les régions de l'euchromatine, alors que les insertions plus anciennes sont plutôt enrichies dans les régions péricentromériques (Baucom *et al.*, 2009) (Figure I.31). L'insertion des éléments transposables a aussi un impact sur la taille des introns, qui sont plus longs chez le maïs comparé aux orthologues chez le riz et le sorgho (Schnable *et al.*, 2009).

2. Organisation du méthylome

Ainsi, la structure génomique d'un chromosome de maïs révèle une succession de régions riches en ET imbriquées les uns dans les autres, et des régions moins denses où les copies sont isolées en dehors des gènes ; ou dans les gènes. Les gènes sont très associés à des séquences d'ET : 50% des gènes ont un ET en amont, 45% en aval, et 33% en chevauchement (Anderson, Maize Genetics Conference 2018). Cette structure complexe induit une organisation particulière de l'épigénome, et en particulier du méthylome. En effet, l'euchromatine et l'hétérochromatine sont très imbriquées. La portion hétérochromatique est majoritairement définie par H3K9me2, H3k27me2 et la méthylation CHG. Entre cette hétérochromatine inactive et l'euchromatine des gènes, on trouve des régions d'hétérochromatine activement transcrite, régulée par RdDM. Ces régions présentent une forte méthylation en CHH, et des « îlots CHH » ont été décrits en amont et en aval des gènes (Gent *et al.*, 2013, Gent *et al.*, 2014). Cette zone permet de contrôler l'expression de ET, tout en laissant la chromatine sous forme activement transcrite et décondensée (dépourvue de H3K9me2) à proximité des gènes (Gent *et al.*, 2013, Gent *et al.*, 2014, Springer *et al.*, 2016, West *et al.*, 2014) (Figure I.32). Cette organisation est particulièrement différente de celle d'*A. thaliana*, à petit génome, où l'hétérochromatine est restreinte au niveau des régions péricentromériques et télométriennes, et nécessite de façon moins cruciale la présence d'une hétérochromatine active à proximité des gènes (Feng & Jacobsen, 2011, Springer *et al.*, 2016) (Figure I.32).

3. Spécificité des voies de régulation de la méthylation

Le maïs est un ancien allotétraploïde. L'évènement de polyploïdie a eu lieu il y a entre 5 et 12 millions d'années (Swigonova *et al.*, 2004). Le retour à l'état diploïde est associé à des épisodes de remaniements chromosomiques incluant d'importantes fusions et inversions (Wei *et al.*, 2007), et les analyses de syntenie avec le génome du riz et du sorgho indiquent que 25% des gènes dupliqués sont encore dans le génome sous forme de paire. Cette duplication du génome a eu un impact sur les voies de régulation de la méthylation : MET1 est codé par les deux gènes ZmMET1a et ZmMET1b, CMT3 par ZmMET2a et ZmMET2b et DRM par trois gènes, ZmMET3a, ZmMET3b et ZmMET3c. ZmMET4 est équivalent à DNMT2 et n'est pas dupliqué dans B73. La duplication des copies a permis une spécialisation de ces protéines pour différents types d'éléments transposables (Qian *et al.*, 2014, Li *et al.*, 2014b). Par exemple, la duplication de CMT3 en ZmMET2 et ZmMET5 a permis une spécialisation de ZmMET5 dans la reconnaissance des LINE (Figure I.33). Enfin, aucun orthologue au gène CMT2 de maïs n'a pu être détecté dans le génome de B73. Ceci pourrait expliquer un rôle plus important de la RdDM dans cette espèce.

4. Variabilité intra-spécifique du génome et du méthylome

En plus de la structure complexe du génome et de l'épigénome, le maïs est également caractérisé par une forte variation génomique entre individus, due à la présence/absence de renflements chromatinien (« knobs ») et de l'insertion différentielle d'éléments transposables (Wang & Dooner, 2006) (Figure I.34). Cette variabilité est à l'origine d'une partie importante des variations épigénomiques entre génotypes de maïs. La comparaison des méthylomes des génotypes Américains Mo17 et B73 a permis de détecter 690 DMR dont 71 sont localisées dans des régions géniques identiques entre les deux génotypes. Ce type de DMR suggère l'existence d'épiallèles purs (Eichten *et al.*, 2011). La transmission de ces variations du méthylome sont, dans la plupart des cas, associées avec des variations génomique avec un contrôle en *cis*, en *trans* ou de type paramutation (Eichten *et al.*, 2013, Regulski *et al.*, 2013, Li *et al.*, 2014a). La comparaison des méthylomes de cinq lignées américaines (B73, Mo17, Oh43, CML322 et Tx303) a révélé des niveaux de méthylation similaires à 99%. La comparaison des génotypes deux à deux a révélé un nombre de DMR entre 5000 et 20000. La majorité de ces variations est liée à l'insertion différentielle d'éléments transposables (Li *et al.*, 2015).

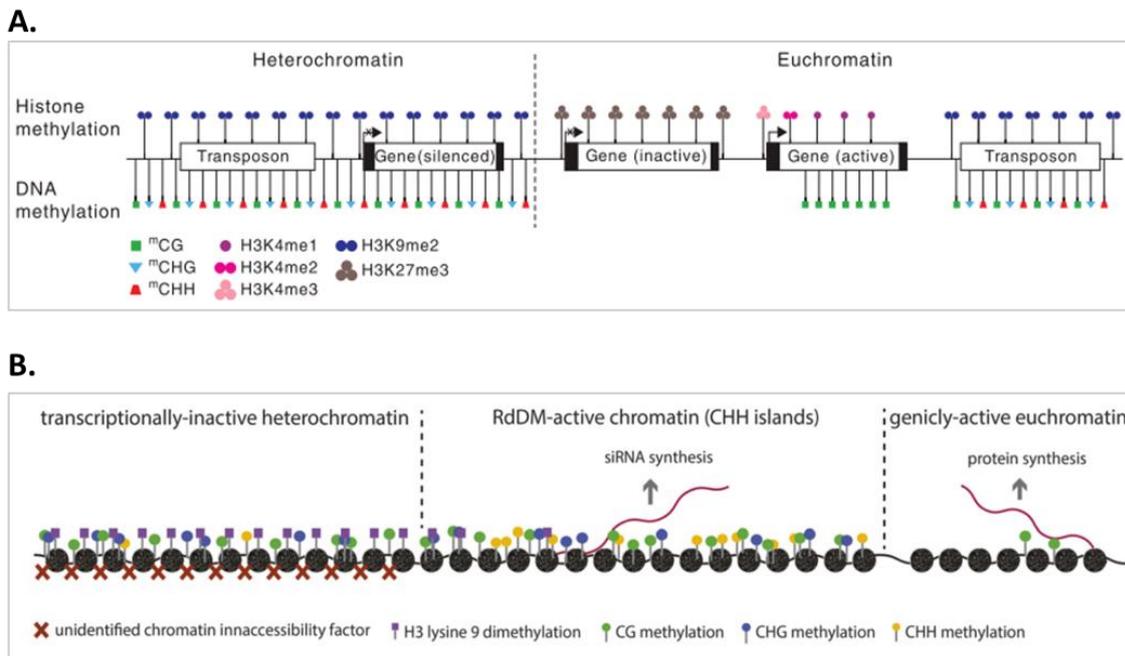


Figure I.32 : Comparaison de l'organisation chromatinienne entre deux espèces ayant une organisation génomique différente

A. Chez la plante à petit génome *Arabidopsis thaliana*, deux domaines majeurs de la chromatine sont observés. A gauche, une région d'hétérochromatine péri-centromérique liée à la présence d'ET et de gènes transcriptionnellement silencieux, caractérisée par des niveaux élevés de H3K9me2 et de méthylation de l'ADN dans les 3 contextes de cytosines (d'après Feng and Jacobsen (2011)). A droite, une région d'euchromatine. Les ET présents de l'euchromatine contiennent aussi un enrichissement en H3K9me2 et une forte méthylation de l'ADN. Les gènes inactifs sont enrichis en H3K27me3. Les gènes actifs sont enrichis en H3K4 et méthylés en contexte CG (gbM). **B.** Chez le maïs, une espèce à génome plus complexe, trois domaines peuvent être caractérisés. A gauche, un domaine d'hétérochromatine où la transcription est inactive (absence de méthylation en contexte CHH suggérant l'absence d'activité de la voie RdDM). Au milieu, un domaine d'hétérochromatine transcrite, riche en CHH méthylés, et dont le contrôle est assuré par la voie RdDM. A droite, un domaine d'euchromatine, où les gènes sont activement transcrits (d'après Gent *et al.* (2014)).

A.

Gene Allele	Maize Gene Name	<i>Arabidopsis</i> Ortholog	Gene Function
<i>zmet2-m1</i>	<i>Zmet2/Dmt102</i>	<i>CMT3</i>	Methyltransferase
<i>zmet2-m2</i>	<i>Zmet2/Dmt102</i>	<i>CMT3</i>	Methyltransferase
<i>zmet5-m1</i>	<i>Zmet5/Dmt105</i>	<i>CMT3</i>	Methyltransferase
<i>zmet5-m2</i>	<i>Zmet5/Dmt105</i>	<i>CMT3</i>	Methyltransferase
<i>zmet7-m1</i>	<i>Zmet7/Dmt107</i>	<i>DRM1/2</i>	Methyltransferase
<i>zmet7-m2</i>	<i>Zmet7/Dmt107</i>	<i>DRM1/2</i>	Methyltransferase
<i>zmet7-T03</i>	<i>Zmet7/Dmt107</i>	<i>DRM1/2</i>	Methyltransferase
<i>chr101-m1</i>	<i>Chr101</i>	<i>DDM1</i>	Chromatin remodeler
<i>chr101-m3</i>	<i>Chr101</i>	<i>DDM1</i>	Chromatin remodeler
<i>chr106-m1</i>	<i>Chr106</i>	<i>DDM1</i>	Chromatin remodeler
<i>chr106-T11</i>	<i>Chr106</i>	<i>DDM1</i>	Chromatin remodeler
<i>vim102-1</i>	<i>Vim102</i>	<i>VIM1</i>	Methylcytosine binding protein
<i>mop1</i>	<i>Mop1</i>	<i>RDR2</i>	RNA-dependent RNA polymerase
<i>mop2</i>	<i>Mop2</i>	<i>NRPD2/E2</i>	RNA polymerase D2/E2
<i>mop3</i>	<i>Mop3</i>	<i>NRPD1</i>	RNA polymerase D1
<i>tgr1</i>	<i>Tgr1</i>	Unknown	Unknown
<i>tgr9</i>	<i>Tgr9</i>	Unknown	Unknown

B.

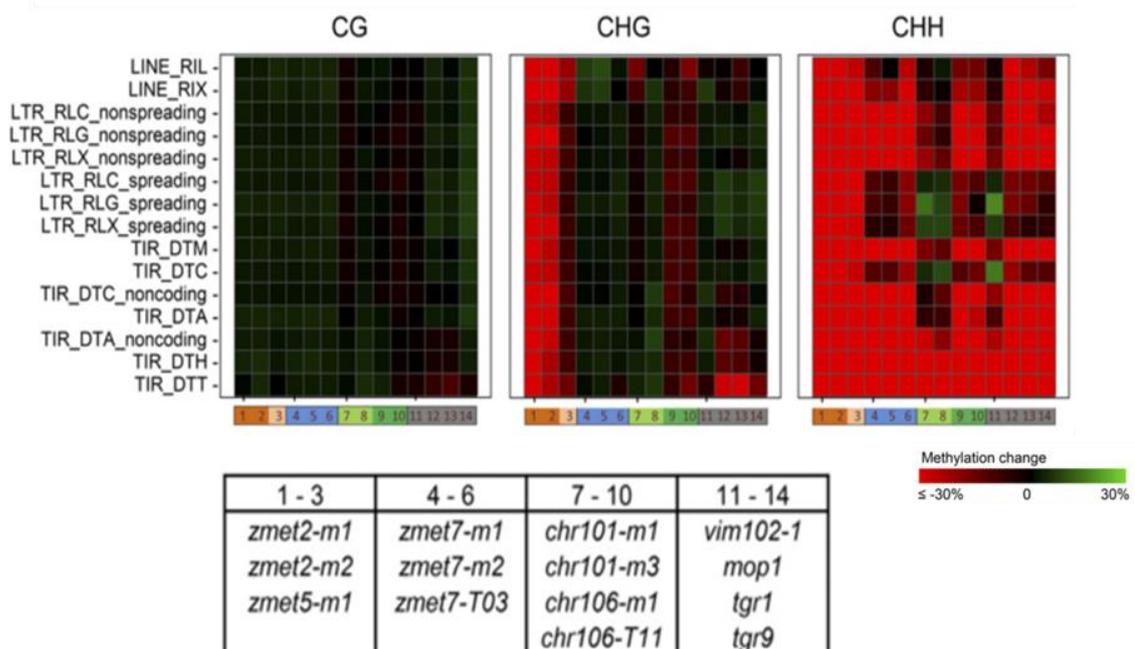


Figure I.33 : Présentation des méthyltransférases homologues à *Arabidopsis* chez le maïs (d'après Li *et al.* (2014b))

A : Mutants disponibles chez le maïs pour plusieurs méthyltransférases. Les noms des gènes sont présentés en premier, suivi du nom du mutant. La fonction CMT3 est assurée par deux gènes, *Zmet2* et *Zmet5*. La fonction DDM1 par deux gènes, *Chr101* et *Chr106*.

B : Changement du niveau de méthylation dans les 3 contextes CG, CHG et CHH pour différentes familles d'ET (à l'intérieur de l'ET). Le tableau décrit les mutants utilisés pour cette étude. Les lignes et les colonnes des heatmaps représentent, respectivement, les familles des ET et les mutants. Le rouge indique une perte de méthylation et le vert indique un gain de méthylation.

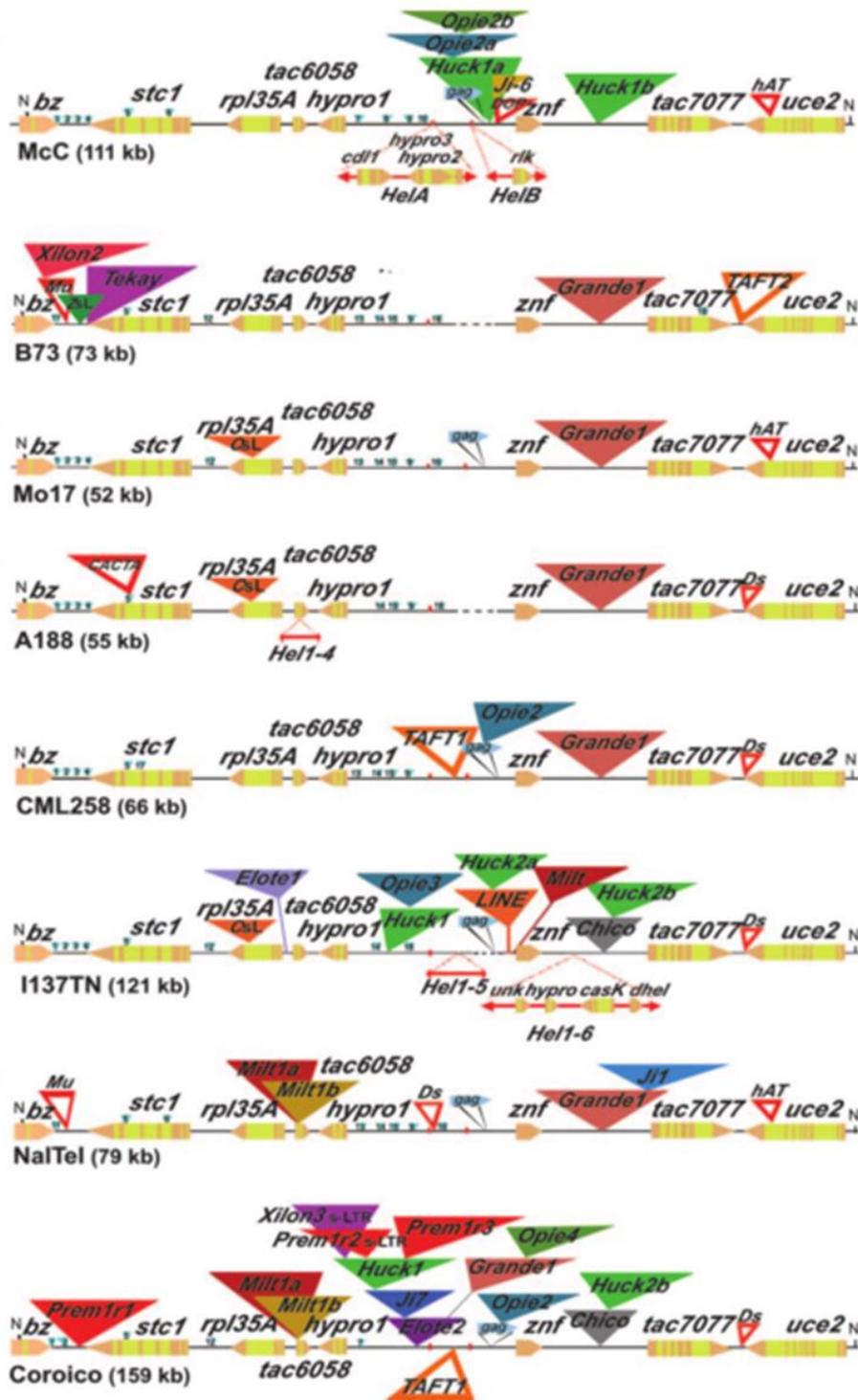


Figure I.34 : Variation de la région génomique Bronze entre 8 génotypes de maïs (Wang & Dooner, 2006)

Le nom des génotypes est indiqué à gauche. Les triangles représentent des ET. Les boîtes jaunes et orange des gènes (range : exons, jaune, intron). La comparaison des séquences deux à deux montre un pourcentage de séquences communes qui varie de 25 à 84% selon les génotypes comparés.

5. Impact des contraintes abiotiques sur la méthylation chez le maïs

Chez le maïs, plusieurs travaux ont été réalisés pour étudier l'impact de l'environnement sur la variation de l'épigénome et la régulation de l'expression des gènes. Par exemple, le traitement froid provoque une activation transitoire de la transcription des répétitions 180bp et TR-1 des « knobs ». L'expression des « knobs » est due à une hyperacétylation en H3K9, une diminution de la méthylation de la marque H3K9me2 et une déméthylation de l'ADN, induisant ainsi un remodelage du nucléosome au sein de ces régions répétées en tandem (Hu *et al.*, 2012). Une déméthylation de l'ADN dans le transposon *Ac/Ds* dans les racines du maïs est observée après un traitement froid (Steward *et al.*, 2000). D'un autre côté, l'application de traitement au plomb à des intervalles de temps successives (12, 24 et 48h) induit une élévation de la méthylation de l'ADN (essentiellement en contexte CG) dans les régions intergéniques, suivies par les introns et les promoteurs. Quand la méthylation est localisée dans les promoteurs, le niveau d'expression des gènes est négativement corrélé avec la méthylation de l'ADN (Ding *et al.*, 2014). Enfin, des analyses de la quantité de méthylation par HPLC montrent une déméthylation globale du génome dans des racines de plantules de maïs âgées de 13 jours sur les quelles des « pulses » de froid ont été appliqués pendant 6 jours. La déméthylation de l'ADN continue à diminuer même après retour des plantes dans des conditions de température de 23°C (Steward *et al.*, 2002). Enfin, l'analyse des méthylomes de plusieurs individus après exposition au froid, aux rayonnements UV ou à une température élevée a révélé une variabilité importante entre individus, et les auteurs suggèrent que les contraintes appliquées augmentent la variabilité observée entre individus (Eichten & Springer, 2015).

6. Le maïs : qualités et contraintes

Pour ces raisons, le maïs est un bon modèle pour l'étude du lien entre le génome, le méthylome et l'environnement. Cependant, du fait de la grande variabilité génomique existant entre génotypes, l'existence d'un seul génome de référence chez le maïs (B73) rend complexe l'alignement des séquences issues d'autres génotypes. Beaucoup de séquences générées sont ainsi éliminées, et l'intégralité des génomes des autres génotypes ne peut pas être analysée. De plus, la position des éléments transposables étant variable d'un génotype à l'autre, l'annotation des éléments transposables de B73 ne peut pas être utilisée pour analyser la proximité d'une région méthylée d'un autre génotype.

V. Analyse du méthylome : défis et techniques dédiées

L'ADN extrait pour l'analyse de la méthylation provient, dans la majorité des cas, d'un tissu, constitué d'une population de plusieurs cellules, parfois même de plusieurs types cellulaires. Le niveau de méthylation d'une cytosine d'un échantillon dépend donc à la fois du signal porté par chaque cellule, et du nombre de cellules mélangées dans l'échantillon. Pour contourner ce problème, des techniques se basant sur l'analyse d'une cellule isolée sont en cours de développement (Ji *et al.*, 2015, Yong *et al.*, 2016b). Mais la plupart des analyses réalisées à ce jour est fondée sur l'analyse d'un signal moyen issu d'un échantillon de plusieurs cellules. Alors que chaque cellule a un signal qualitatif (méthylé ou non méthylé) pour chaque cytosine, l'analyse d'un ensemble de cellule rend le signal quantitatif

(pourcentage de cellules ayant un signal de méthylation à une cytosine donnée par rapport à l'ensemble des cellules (Figure I.35). Cette variabilité cellulaire pourrait contribuer à l'observation d'une stochasticité entre individus. Enfin, les signaux de méthylation ayant le plus fort impact phénotypique consistent en une succession de cytosines méthylées. De ce fait, le signal analysé est la somme horizontale des signaux quantitatifs de chaque cytosine de la région.

Les premières investigations de la méthylation de l'ADN ont été réalisées en utilisant des techniques de type digestion-amplification PCR. Ces méthodes s'appuient sur des enzymes de restriction sensibles à la méthylation de l'ADN et correspondent à une méthode AFLP modifiée qui utilise des enzymes isochizomères comme par exemple *MspI* et *HpaII* qui reconnaissent le même site de restriction (CCGG) mais présentent une sensibilité différente à la méthylation de l'ADN (Reyna-Lopez *et al.*, 1997). Cependant, ces techniques restent limitées car elles permettent une révélation de type présence/absence mais non pas une analyse quantitative du signal à analyser.

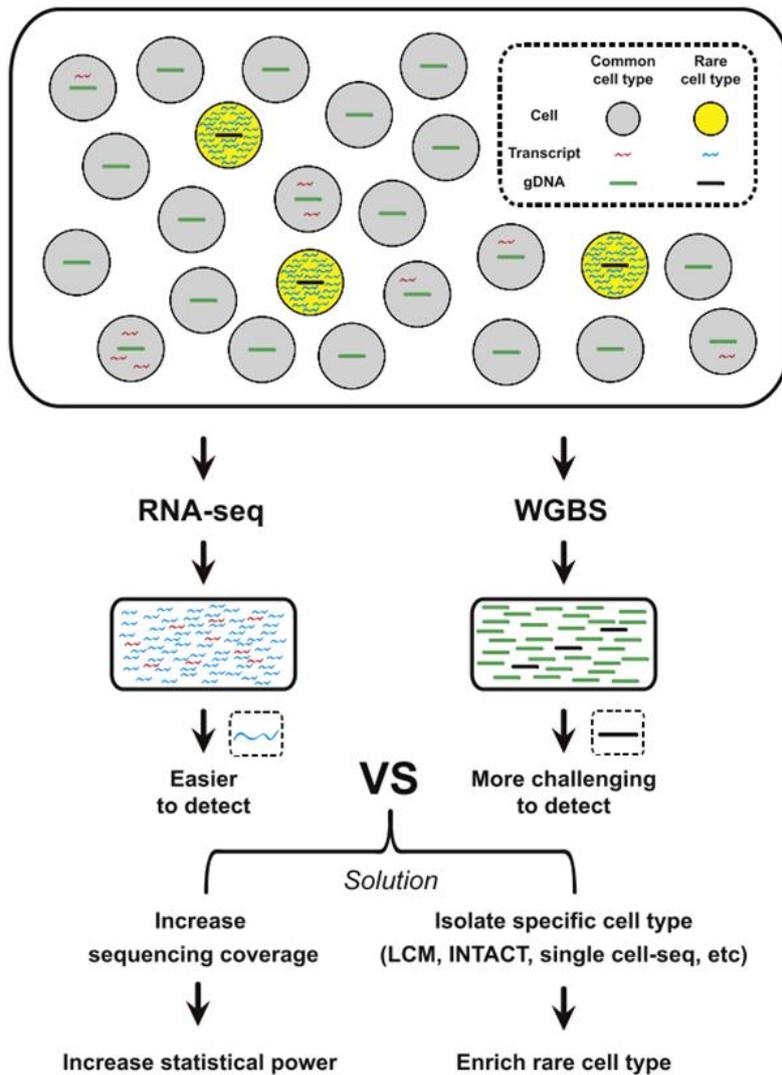


Figure 1.35 : Différence dans la détection des gènes différentiellement exprimés et des régions différentiellement méthylées (DMR) (Ji *et al.*, 2015)

Dans le cas de l'analyse des transcrits, c'est l'abondance des transcrits qui est analysée. Même si celle-ci dépend d'un petit nombre de cellules, l'amplification du signal permet son identification. Dans le cas de l'analyse de la méthylation, chaque cellule porte un état différent. De ce fait, la quantité de méthylation estimée dépend du nombre de cellules portant l'événement de méthylation. Si ce nombre de cellule est faible, le signal est difficile à détecter. L'analyse d'un seul type cellulaire spécifique ou par cellule (single cell) peut faciliter l'étude du méthylome en dehors de l'augmentation de la profondeur de séquençage. INTACT: isolation of nuclei tagged for specific cell types, LCM: laser capture microdissection, WGBS: whole-genome bisulfite sequencing.

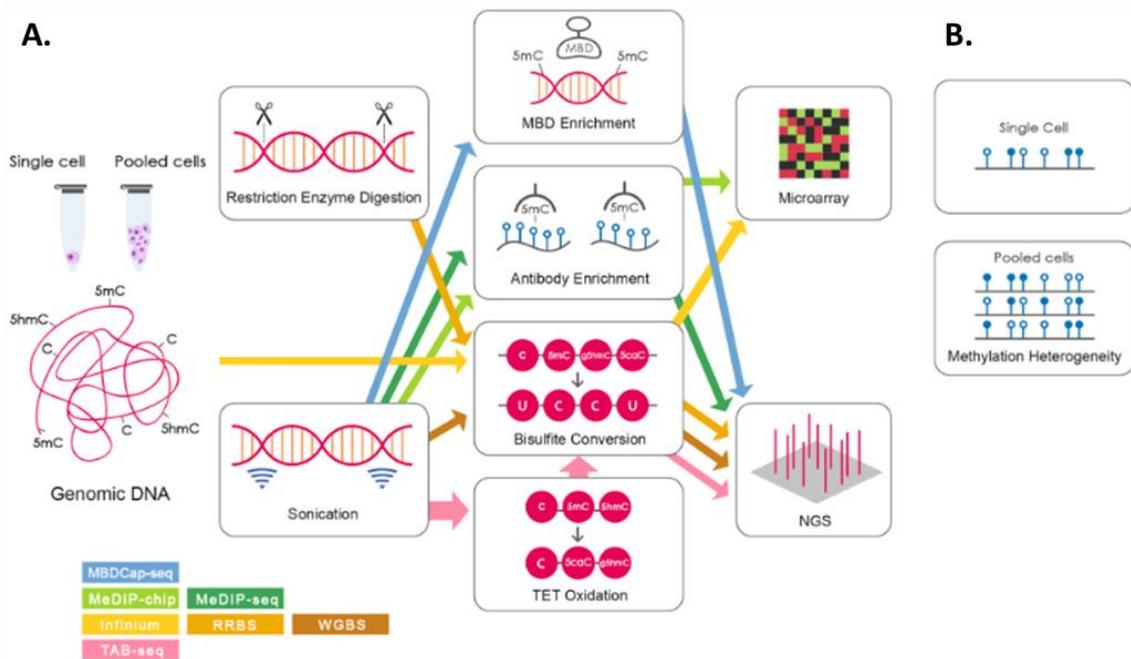


Figure I.36 : Méthodes communément utilisées pour l'analyse du méthylome (Yong *et al.*, 2016a)

A. La procédure peut impliquer (i) une fragmentation de l'ADN génomique par des enzymes de restriction ou une sonication, (ii) un enrichissement en régions méthylées (par immunoprécipitation par un anticorps reconnaissant la 5-méthylcytosine ou par affinité avec une protéine reconnaissant la méthylation - MBD : methyl-binding domain), (iii) une conversion au bisulfite de sodium ou (iv) une oxydation par les protéines TET (Ten-Eleven Translocation, pour analyse des cytosines hydroxyméthylées) avant d'être hybridé sur puce ou séquencé. **B.** L'isolation de l'ADN à partir d'une seule cellule permet de révéler l'hétérogénéité dans la méthylation de l'ADN observée dans une population de cellules. Les points colorés en bleu représentent les cytosines méthylées (5mC) et les autres points montrent les cytosines non méthylées.

La détection de ce signal quantitatif peut se faire selon plusieurs méthodes (Figure I.36). La quantité de l'ADN, la résolution, la couverture des régions et les analyses bio-informatiques varient en fonction de la technique utilisée (Yong *et al.*, 2016b). Deux grands types peuvent être distingués :

Le premier consiste en une analyse globale de la méthylation de fragments d'ADN génomique. Ces fragments sont récupérés soit après fractionnement par une enzyme sensible à la méthylation comme *McrBC* (Irizarry *et al.*, 2008) ou après immunoprécipitation par des anticorps reconnaissant spécifiquement la 5-méthylcytosine. Ces fragments peuvent ensuite soit être hybridés sur une puce contenant des régions particulières du génome que l'on souhaite analyser (« Methylated DNA Immuno-Precipitation on chip », ou MeDIP on chip) (Weng *et al.*, 2009), soit séquencée et alignée sur un génome de référence (MeDIP-Seq) (Zhao *et al.*, 2014c). Dans ces techniques, le niveau de précision du signal est limité par la taille du fragment généré, qui est en général d'environ 400pb. Elles ne permettent pas l'analyse fine de chaque cytosine du génome, et ne permettent donc pas d'analyser séparément les différents contextes de cytosines, ce qui est regrettable pour analyser des méthylomes de plante.

Le deuxième type de méthode est fondé sur la détection quantitative du niveau de la méthylation pour chaque cytosine après conversion de l'ADN par l'agent chimique, le bisulfite de sodium (Figure I.36) : suite à une réaction de déamination, les cytosines non méthylées sont converties en Uracil (U). Les cytosines méthylées ne sont pas affectées du fait de l'encombrement stérique dû au groupement méthyl. Cette conversion est suivie d'une amplification par PCR, au cours de laquelle l'Uracil (U) généré est converti en thymine (T) (Frommer *et al.*, 1992). Le produit peut être séquencé pour analyser une région particulière. L'amplification par PCR peut également être intégrée à la construction de banques de type Illumina, et séquencé de façon intensive. On parle alors de « bisulfite-sequencing » (BS-seq) ou de « whole genome bisulfite sequencing » (WGBS). Ce type de méthode est particulièrement intéressant pour l'analyse du méthylome de plantes, du fait de la différence de régulation des trois contextes de cytosines (CG, CHG et CHH). D'ailleurs, le premier méthylome résolu à la cytosine a été celui d'*A. thaliana* (Cokus *et al.*, 2008, Lister *et al.*, 2008). Toutes les cytosines du génome peuvent théoriquement être analysées. Cependant, la détection de la quantité de méthylation par cytosine dépend de l'étape d'alignement des lectures générées sur le génome de référence. La profondeur de séquençage employée détermine le niveau de la précision du signal. L'étape d'alignement requiert des méthodes spécifiques du fait que la conversion génère du polymorphisme C/T, et que les ADN essentiellement à la non complémentarité entre les deux brins Watson et Crick suite à la conversion bisulfite (Bock, 2012). De ce fait, de façon encore plus importante que pour l'analyse des SNP, l'utilisation d'un seul génome de référence est problématique. Le logiciel Bismark (Krueger & Andrews, 2011) est largement utilisé pour l'alignement des lectures de BS-seq. Il est fondé sur l'alignement, par Bowtie2, des lectures sur deux génomes représentant chacun le brin Watson et le brin Crick après conversion. Or, Bowtie2 nécessite de choisir, en amont de l'alignement, le nombre de bases différentes tolérées par alignement. Dans le cas d'alignements sur un génome de référence distant, ceci pose des problèmes dans les régions moyennement méthylées. Ce problème est mieux pris en compte par le logiciel Novoalign (<http://www.novocraft.com/documentation/novoalign-2/novoalign-user-guide/bisulphite-treated-reads/novomethyl-analyzing-methylation-status/>), qui ne nécessite pas de choix *a priori*, mais qui teste un grand nombre de positions et filtre *a posteriori* sur les scores de qualité.

Avant séquençage, les fragments amplifiés après conversion au bisulfite de sodium peuvent également être utilisés pour une hybridation en solution avec des sondes choisies, représentant un ensemble de régions d'intérêt. Seules les régions ainsi « capturées » sont par la suite séquencées, ce qui permet d'obtenir une couverture importante, et donc un niveau élevé de précision du signal de méthylation de chaque cytosine de ces régions en employant de fortes couvertures de séquençage. Le détail des différentes étapes de cette technologie est présenté dans la Figure I.37.

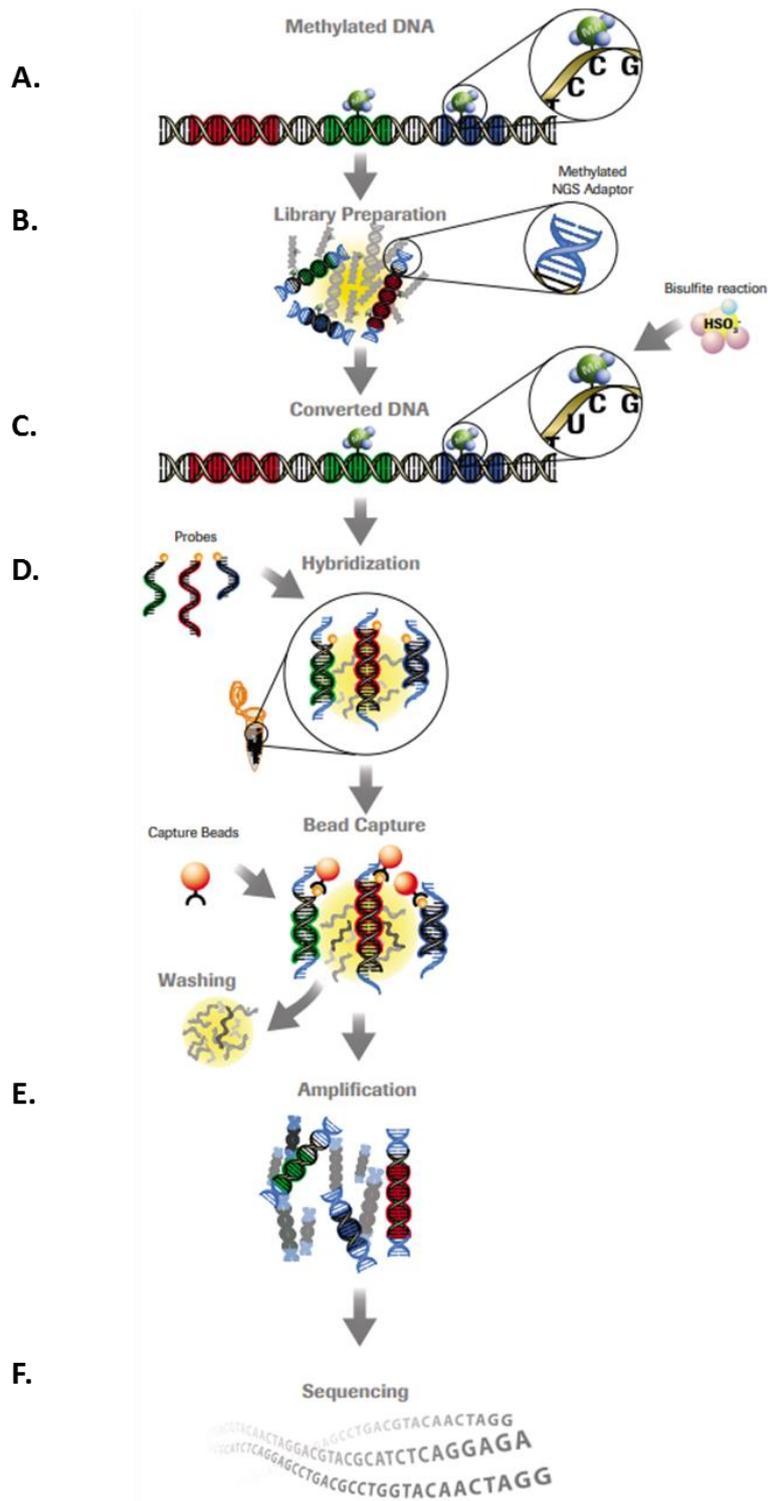


Figure I.37 : Etapes de la méthode bisulfite se basant sur la capture des séquences ciblées (SeqCapEpi)¹

A. Disposition de régions d'ADN ciblées. **B.** Préparation d'une librairie en employant des adaptateurs méthylés. **C.** Conversion de l'ADN au bisulfite de sodium et amplification par PCR. **D.** Hybridation en solution avec des sondes issues des séquences cibles, étape dite de « capture ». **E.** Amplification des fragments d'ADN capturés. **F.** Séquençage haut débit.

¹ www.roche.de/res/content/10527/seqcap_gesamtbrochure.pdf

VI. Présentation des objectifs de la thèse

Les études réalisées au cours de ma thèse s'inscrivent dans la thématique de recherche de l'équipe DyGAP (Dynamique du génome et adaptation des plantes cultivées) dont l'objectif principal est de caractériser les variations génomiques (SNP, variation structurale) et épigénomiques impliquées dans l'évolution et la régulation de l'expression des génomes. Au sein de cette équipe, Clémentine Vitte développe une thématique de recherche visant à comprendre l'interrelation entre la structure du génome et son expression, et comment cette relation peut être modifiée par des contraintes environnementales. Un des axes de recherche est en particulier d'analyser l'interaction entre variation génomique et variation du méthylome entre lignées de maïs, en conditions standard et en condition de stress abiotique (froid et déficit hydrique). Cet axe de recherche est développé au sein du workpackage3 (WP3) du programme d'investissement d'avenir *Amaizing* (www.amaizing.fr), et repose sur l'analyse de données haut débit de type BS-seq de forte profondeur. Il est réalisé en étroite collaboration avec les autres tâches du workpackage, qui visent à caractériser (i) la variation génomique entre lignées de maïs et son impact sur les réseaux d'expression, et (ii) les bases moléculaires de l'adaptation du maïs aux conditions Européennes. Cette partie du programme du WP3 est réalisée en collaboration avec Catherine Giauffret et en lien étroit avec l'INRA de Mons, dont l'expertise scientifique a été primordiale pour le choix du matériel végétal, des conditions de culture et du suivi des organes de mon travail de thèse.

Dans ce cadre, mon travail de thèse a bénéficié de l'expertise scientifique et technique développée au sein de l'équipe DyGAP et du programme *Amaizing*, en particulier de l'INRA de Mons. Plus largement, j'ai bénéficié de l'expertise de nombreux collègues intéressés par l'épigénétique, dans le cadre du réseau pluridisciplinaire « Epigénétique en Ecologie et Evolution » (RTP 3E).

Ma problématique de thèse est articulée suivant 3 principaux axes de recherche qui ont pour but de :

(i) Caractériser la réponse du méthylome au froid chez un génotype de maïs sensible (B73), après un traitement froid prolongé, et juste à la sortie de ce traitement. En particulier, cette partie vise à estimer l'étendue des modifications de la méthylation de l'ADN induites par le froid et leur orientation (hyper- ou hypo-méthylation), à déterminer si elles affectent le génome de façon globale ou de façon ciblée, et à estimer les mécanismes pouvant être à l'origine de ces modifications. Pour répondre à ces questions, j'ai caractérisé les variations de méthylation de l'ADN des feuilles entre plantes ayant poussé en conditions contrôle et en conditions de froid à partir de données BS-seq. En parallèle aux analyses du méthylome, des données de mRNA-seq ont été exploitées sur les mêmes plantes, afin d'obtenir une caractérisation moléculaire de leur état physiologique et de l'expression des gènes impliqués dans les voies de régulation de la méthylation de l'ADN. Ce chapitre est présenté sous forme d'un article scientifique en cours de rédaction, qui sera soumis au journal *Genome biology* au printemps 2018.

(ii) Caractériser l'étendue de la maintenance des variations détectées à la sortie du froid au cours du développement de la plante et dans la descendance. En particulier, l'analyse au cours du développement avait pour but de déterminer si le niveau de maintenance est différent pour des feuilles à différents stades de développement au moment de l'exposition au froid. Cette analyse est fondée sur des données de SeqCapEpi. Ces données ont été réceptionnées à la fin de ma thèse, ce qui n'a pas permis d'inclure leur analyse dans le manuscrit. Le chapitre 2 présente les résultats des analyses qui m'ont permis de choisir le matériel végétal à analyser et les régions génomiques à suivre, et présente les analyses à réaliser.

(iii) Déterminer si la méthylation de l'ADN est une composante de la tolérance au froid chez le maïs. Pour cela, j'ai cherché à évaluer si la réponse du méthylome au froid varie entre génotypes sensibles et tolérants au froid. Cette partie reste limitée par la disponibilité d'un seul génome de référence, celui de B73. Le chapitre 3 présente les résultats préliminaires que j'ai obtenus à partir de l'alignement des séquences BS-seq des trois génotypes sur le génome de B73, et discute des limites de cette approche et des moyens à mettre en œuvre pour les contourner.

Le développement de ces trois axes a nécessité la mise en place d'un dispositif expérimental dédié, qui est présenté en préambule à ces trois chapitres.

Un dernier chapitre est consacré à une discussion générale intégrant les différents résultats présentés dans les trois premiers chapitres, ainsi qu'aux perspectives qui en découlent.

Deux annexes présentent (i) le détail des dispositifs expérimentaux (Annexe 2) et (ii) la mise au point d'une méthode de détection de régions différentiellement méthylées induites par un stress abiotique chez le maïs (Annexe 3).

Préambule.

Vue d'ensemble de l'expérimentation mise en place pour répondre aux grands axes de la thèse

Pour répondre aux questions de ma thèse, il a fallu mettre en place un dispositif approprié qui permette à la fois de soumettre les plantes à une contrainte forte, suivre leur développement, et récolter des grains sur les plantes matures afin de pouvoir produire la génération suivante.

1. Choix des génotypes et des conditions de culture

Dans le cadre des analyses génétiques réalisées par l'équipe de Catherine Giauffret, la comparaison du développement de lignées de maïs après 45 jours à une température de 14°C jour/10°C nuit en chambre de culture a permis de mettre en évidence une variabilité phénotypique de réponse au froid : selon les génotypes, les plantes meurent, ou présentent à la fois un retard de croissance et un jaunissement foliaire, ou seulement un retard de croissance en se basant sur le suivi de l'émission foliaire et la mesure de la biomasse (Figure P.1). Parmi ces génotypes, nous avons cherché trois génotypes à la fois contrastés pour leur réponse au froid, mais capables de survivre au traitement appliqué. Nous avons également cherché à représenter des génotypes d'intérêt agronomique, et pour lesquels une séquence génomique était disponible ou en cours de réalisation. L'ensemble de ces critères nous a conduit à choisir les lignées B73 (très sensible), F2 (sensible) et F331 (tolérante).

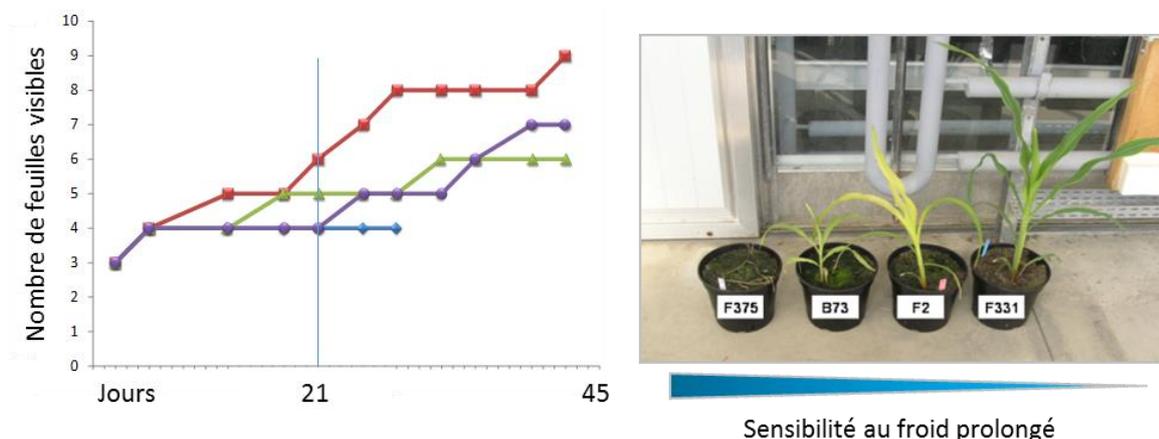


Figure P.1 : Réponse contrastée du froid prolongé sur le maïs

En réponse à un froid prolongé, les différents génotypes de maïs montrent une variabilité phénotypique. Certaines plantes meurent au stade 4 feuilles visibles (courbe en bleu), elles sont représentées par le génotype F375. D'autres sont chlorotiques et montrent un retard de développement (courbes verte et violette), elles sont représentées par les génotypes B73 et F2. Enfin, certaines ne montrent qu'un retard de développement; elles sont représentées par le génotype F331 (courbe rouge). Adapté de Catherine Giauffret.

La lignée B73 est une lignée Américaine du groupe « Stiff Stalk ». C'est la lignée de référence Américaine, pour laquelle de nombreux outils de génétique (carte génétique, carte physique) (Lee et

al., 2002b), génomique (première séquence génomique assemblée) (Schnable *et al.*, 2009) et mutagénèse (populations de TILLING) (Till *et al.*, 2004) ont été développés.

La lignée F2 est une lignée française. Elle appartient au groupe des cornées Européens (« European Flints »). Elle a été développée à partir de la population Lacaune, cultivée dans le sud-ouest de la France. Dans les années 1950, c'était la lignée à meilleur rendement en valeur hybride en croisement avec du matériel Denté Américain. Depuis, elle a été très utilisée dans de nombreux programmes d'amélioration variétale. Cette lignée a été choisie comme lignée modèle en génomique dans le cadre du projet *Amaizing*, et sa séquence génomique a fait l'objet d'un assemblage Illumina tôt dans le projet, conduisant à la détection de régions de F2 absentes de B73 et à la construction d'un pangénome B73/F2 contenant la séquence génomique de B73 ainsi que les contigs spécifiques de F2 (Darracq *et al.*, 2018).

La lignée F331 est une lignée issue de la population 86, une population précoce à grain jaune sélectionnée par le CIMMYT, composée à 55% de matériel tropical d'altitude, à 25% de matériel tempéré et à 20% de matériel tropical/subtropical (Eagles & Lothrop, 1994). Elle a été fixée génétiquement au CIMMYT puis à l'INRA d'Estrées-Mons par l'équipe de Catherine Giauffret. Sa séquence génomique fait l'objet d'un assemblage dans le cadre du programme *Amaizing*, mais qui n'a été disponible qu'à la fin de ma thèse (mars 2018).

Afin d'appliquer une contrainte forte tout en permettant la croissance et la reproduction des plantes traitées au froid après retour en conditions normales, nous avons choisi d'adapter les conditions de culture froide contrôlées utilisées précédemment par Catherine Giauffret (chambre climatisée) en diminuant la durée du traitement à 21 jours au lieu de 45 jours. Dans les deux analyses, les plantes ont poussé en conditions standard (serre) du semis jusqu'à la levée, afin d'éviter la mort prématurée des génotypes sensibles.

2. Mise en place du dispositif expérimental

Pour répondre aux grands axes de ma thèse, nous souhaitons mettre en évidence l'effet du froid sur le méthylome à la fois (i) juste après exposition au froid, (ii) un mois après la fin de l'exposition au froid et retour en conditions standard, et (iii) dans la descendance (génération G1).

L'installation de l'expérimentation a été réalisée au sein du site INRA d'Estrées-Mons (Picardie) qui possède une chambre climatisée permettant de contrôler à la fois le froid et la lumière. L'expérimentation de la génération G0 a été réalisée de fin mars 2015 à aout 2015. Elle a été séparée en trois grandes étapes (voir détail Annexe 2) :

- 1) Croissance en parallèle des lots « stressé » et « non stressé »
- 2) Mise en commun de deux lots en conditions standard et culture jusqu'à floraison
- 3) Autofécondations et croisements, puis récolte des grains

L'expérimentation de la génération G1 a été réalisée de début mars 2017 à juillet 2017. Elle a concerné uniquement les génotypes B73 et F2 car le génotype F331 n'avait pas donné de grains. Elle a consisté en deux grandes étapes :

- 1) Croissance des plantes en conditions standard
- 2) Autofécondations et récolte des grains

L'équipe de l'INRA de Mons n'avait encore jamais fait pousser un lot « stressé » et un lot « non stressé » de façon synchronisée en conditions contrôlées. La mise en place de cette expérimentation a donc demandé une réflexion importante en amont pour faire face à certaines contraintes :

- le fait de vouloir analyser le profil de méthylation au cours du développement et après retour dans des conditions « standard » des couples de plantes « stressée/non stressée » a impliqué de cultiver en parallèle les deux lots. Comme il n'y avait qu'une chambre climatisée, un lot a été cultivé en chambre climatisée (jour : 16h à 14°C ; nuit : 8h à 10°C), l'autre lot en serre (jour : 16h à 24°C ; nuit : 8h à 18°C). Afin de limiter les différences entre ces deux environnements, nous avons adapté la durée de photopériode de la chambre climatisée à celles de la serre.

- le stress froid prolongé en début de vie induit un retard de développement. Pour apparier des plantes au même stade de développement, il a fallu estimer le décalage attendu, et faire pousser les deux lots de façon décalée. A partir d'une date de semis estimée, nous avons semé trois fois le lot « non stressé », à trois dates différentes. A l'arrêt du traitement froid et au moment de la mise en commun des deux lots, des paramètres de croissance (nombre de feuilles dégainées, nombre de feuilles visibles, longueur de la dernière feuille dégainée) ont été mesurés. Ces paramètres ont permis de déterminer celui des trois lots « non stressé » ayant un état développemental le plus proche de celui des plantes du lot « stressé ».

- du fait des contraintes de place dans la chambre climatisée, les plantes ont été semées dans des pots de 4,5L, puis transférées dans des pots de 18L lors de la mise en commun en serre pour permettre leur croissance jusqu'au stade grain. Afin de limiter l'impact du transfert sur les racines, un filet a été disposé au fond de chaque pot de 4,5L, et les mottes ont été transférées sans dommage aux racines (Annexe 2, figure 1). Le même système de pot et de filet a été utilisé pour le lot « non stressé » cultivé dans la serre afin que les conditions soient comparables.

3. Prélèvements

Pour l'analyse de la génération G0, deux grandes séries de prélèvement foliaire ont été organisées :

Après 21 jours de traitement froid, nous avons prélevé 3 plantes de chaque lot et de chaque génotype pour étudier les variations de la méthylation de l'ADN juste à la sortie du froid. Cette série de 18 échantillons (3 échantillons x 3 génotypes x 2 conditions) a été utilisée pour des analyses de BS-seq (dernière feuille visible) et mRNA-seq (avant dernière feuille visible).

Après un mois de culture des deux lots en commun en conditions standard, nous avons prélevé des feuilles qui étaient à différents stades de développement pendant l'exposition au froid. Les prélèvements ont été réalisés pour les trois génotypes, et la série d'échantillons de B73 a été utilisée pour des analyses de SeqCapEpi à partir des DMR issues de l'analyse des échantillons de cette lignée juste après exposition au froid.

Pour l'analyse de la génération G1, deux grandes séries de prélèvement foliaire ont été organisées :

- 1) Prélèvement de plantules entières
- 2) Prélèvement de plusieurs feuilles au cours du développement

Les prélèvements de la génération G1 ont été réalisés pour les génotypes B73 et F2, et la série d'échantillons de B73 a été utilisée pour des analyses de SeqCapEpi à partir des DMR issues de l'analyse des échantillons de cette lignée juste après exposition au froid.

L'ensemble du dispositif et des prélèvements est présenté dans la figure P.2, en parallèle des questions auxquelles chaque partie permet de répondre et des numéros de chapitres dans lesquels les résultats correspondant sont présentés. Cependant, pour faciliter la lecture du manuscrit, le détail du dispositif expérimental et de chaque étape de la culture sont décrits dans l'annexe 2. Le détail des prélèvements est présenté dans chaque chapitre. Pendant cette étape expérimentale, j'ai contribué à la réflexion du dispositif expérimental à installer et j'ai participé aux différentes étapes des deux cultures (préparation de grains, installation, rempotage, dissection et prélèvement des feuilles, croisement), en bénéficiant de l'aide technique de l'équipe GCIE à INRA d'Estrées-Mons.

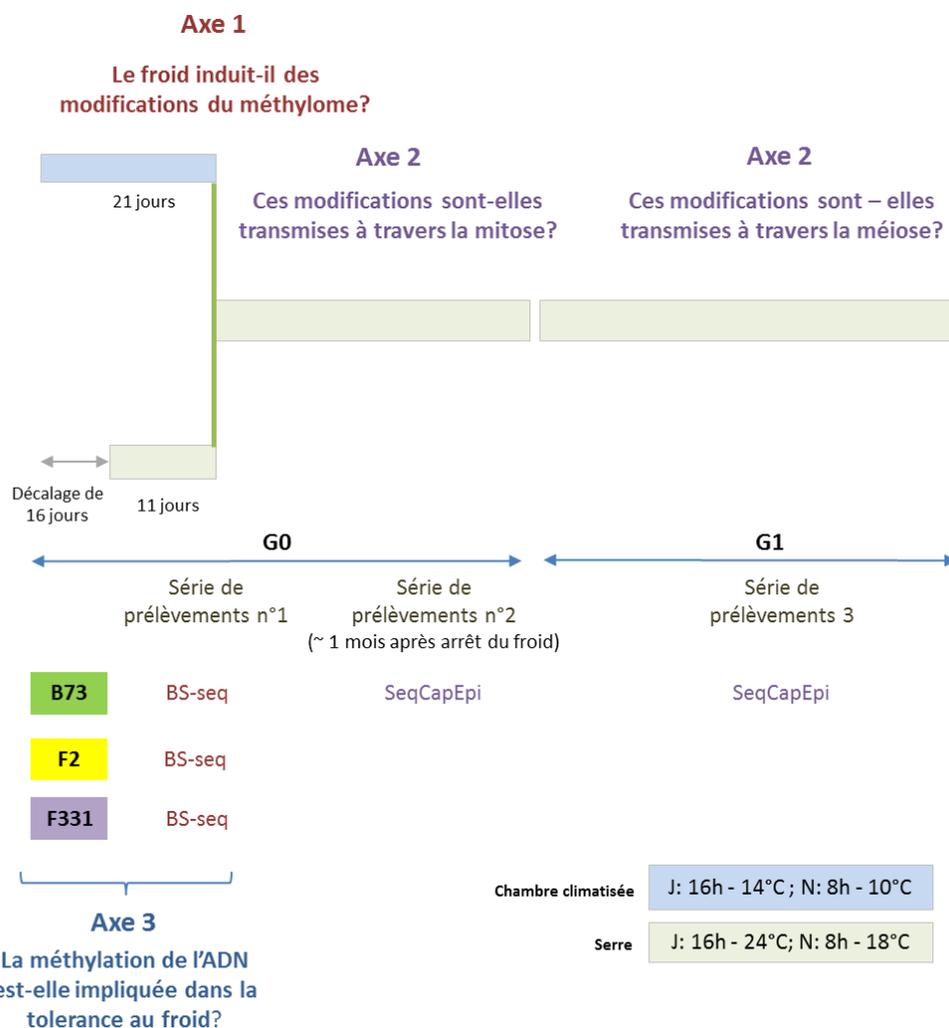


Figure P.2 : Présentation générale des questions, expérimentations et prélèvements

G: Génération. J: jour. N: nuit.

Chapitre II.

Le froid prolongé induit des variations de la méthylation de l'ADN chez une lignée de maïs sensible au froid

1. Contexte et objectifs

La méthylation de l'ADN est une marque de la chromatine impliquée dans la régulation de l'expression des gènes et des éléments transposables. Plusieurs études ont montré que les contraintes abiotiques induisent des modifications de la méthylation de l'ADN chez les plantes cultivées.

A ce jour, une seule étude a testé la réponse du méthylome de maïs à des contraintes abiotiques (Eichten & Springer, 2015). Cette étude a permis de démontrer que des contraintes abiotiques comme le froid, la radiation des UV et les hautes températures sont capables de modifier la méthylation de l'ADN à plusieurs locus. Cependant, peu de régions ont pu être caractérisées comme différentiellement méthylées (DMR) entre plantes « test » et plantes « contrôle », et l'article conclut essentiellement à une variation stochastique de la méthylation chez l'ensemble des individus, qui est accrue chez les individus « test ». Ceci est lié à plusieurs facteurs. D'abord, le choix de la contrainte appliquée (pour le froid 4°C pendant 4h), du moment d'observation (feuille drapeau, longtemps après exposition aux contraintes), de l'organe analysé (feuille drapeau, qui s'est développée après l'exposition aux contraintes) qui ont vraisemblablement eu un impact fort sur les résultats obtenus. De plus, cette étude a été réalisée en utilisant une méthode fondée sur l'immunoprécipitation des séquences portant des cytosines méthylées suivie d'une hybridation sur une puce contenant des régions ciblées (« MeDIP on chip »), ce qui a entraîné deux limites majeures : (i) la restriction des régions génomiques analysables à la partie peu répétée (« low copy ») du génome, (ii) l'absence d'analyse du niveau de méthylation pour les différents contextes (CG, CHG et CHH).

Afin d'intégrer les deux limites de cette étude et de mieux caractériser la réponse du méthylome de maïs au froid, nous avons choisi une approche différente. D'abord, nous avons choisi d'exposer les plantes de maïs à un froid prolongé au début de leur cycle de développement, un traitement qui est proche des conditions du maïs en champs, et qui permet de réaliser plusieurs cycles de divisions cellulaires pendant l'exposition au froid, un paramètre sans doute important pour acquérir un signal de méthylation assez fort pour être observé. Ensuite, nous avons choisi d'utiliser la technique de BS-seq, avec de fortes profondeurs de séquençage, ce qui permet à la fois d'étudier les contextes séparément et d'estimer assez précisément les niveaux de méthylation. Enfin, en parallèle des analyses de méthylation de l'ADN, nous avons souhaité déterminer l'impact du froid appliqué sur le transcriptome des mêmes plantes, ce qui permet à la fois d'avoir une première caractérisation de l'impact physiologique du froid appliqué, d'estimer les changements de transcription des gènes impliqués dans la régulation de la méthylation de l'ADN, et comparer les modifications de méthylome et de transcriptome. Pour ce faire, nous avons appliqué un froid de 16h jour à 14°C; 8h nuit à 10°C pendant 21 jours, et nous avons comparé 3 plantes « froid » et 3 plantes « contrôle ». Pour des raisons de quantité de matériel, l'analyse du méthylome (BS-seq) a été réalisée à partir de la dernière feuille

visible et l'analyse du transcriptome (mRNA-seq) a été réalisée à partir de l'avant dernière feuille visible (voir Annexe 2).

Notre étude est la première analyse de données de BS-seq chez le maïs avec trois répliques. Du fait de la stochasticité observée dans Eichten et Springer (2015), nous avons choisi de développer deux types d'approche : (i) une analyse exploratoire du méthylome complet en analysant la différence de méthylation entre les signaux moyens des individus « froid » et « contrôle » sur les 3 contextes de cytosines le long du génome. Cette analyse est réalisée sans fixation de seuils de variations de méthylation *a priori*. et permet d'analyser les profils de méthylation le long des chromosomes, en comparaison avec le contenu (gène, éléments transposables) de ceux-ci afin de déterminer le type de régions génomiques affectées; (ii) une détection de régions statistiquement différenciellement méthylées (DMR) entre les deux lots de plantes. Cette analyse a été réalisée avec des tailles de fenêtre de petite taille (200pb, 500pb et 1kb) avec une différence de méthylation de 10% (voir Annexe 3), dans le but de détecter des régions avec un potentiel plus élevé d'impact sur le phénotype (expression des gènes notamment). Ces DMR ont été catégorisées en s'appuyant sur la caractérisation des variations de chaque contexte de cytosines et sur l'annotation des régions ciblées, afin de proposer un modèle mécanistique d'apparition de ces epialléles. Enfin, la relation entre méthylome et transcriptome est appréhendée sous deux formes : (i) le lien entre présence d'une DMR à proximité d'un gène et le niveau de transcription de celui-ci, et (ii) le lien entre profils de méthylation observés et la modification transcriptionnelle des gènes intervenant dans les voies de régulation de la méthylation de l'ADN.

L'ensemble des résultats de cette étude est présenté ci-après sous forme d'un article en cours de préparation pour le journal « Genome Biology ».

2. Achour *et al.* en préparation pour soumission dans *Genome Biology* : Prolonged cold in maize triggers genome-wide hypermethylation and local strong methylation variation

Prolonged cold in maize triggers genome-wide hypermethylation and local strong methylation variation

Zeineb ACHOUR¹, Johann Joets¹, Martine Leguilloux¹, H  l  ne SELLIER², Jean-Philippe PICHON³, Magalie LEVEUGLE³, Herve DUBORJAL³, Jos   Caius^{4,5}, V  ronique Brunaud^{4,5}, Christine Paysant-Le Roux^{4,5}, Tristan Mary-Huard^{1,6}, Catherine GIAUFFRET⁷, and Cl  mentine VITTE¹

¹INRA, CNRS, Universit   Paris–Sud, AgroParisTech/UMR G  n  tique Quantitative et Evolution (GQE) - Le Moulon, Gif-sur-Yvette, France ; ²INRA UE Grandes Cultures Innovation Environnement (GCIE), Estr  es-Mons, France ; ³Biogemma, Chappes, France ; ⁴CNRS, INRA, Universit   Paris-Sud, Universit   d'Evry, Universit   Paris-Saclay / Institute of Plant Science (IPS2) Orsay, France ; ⁵CNRS, INRA, Universit   Paris-Diderot, Sorbonne Paris-Cit  /Institute of Plant Science (IPS2), Orsay, France; ⁶MIA, INRA, AgroParisTech, Universit   Paris-Saclay, Paris, France ; ⁷INRA UR Agrossources et Impacts environnementaux (AgroImpact), Estr  es-Mons, P  ronne, France.

Abstract

Understanding how plants respond to environmental constraint is a major task to both better understand local adaptation and improve agriculture. Early and prolonged cold is a major abiotic constraint for maize, and sensitive genotypes show phenotypic damage such as delay in development and chlorosis. To better characterize maize response to cold, we analyzed the impact of prolonged cold on transcriptome (through mRNA-seq) and methylome (through whole genome bisulfite seq) of the sensitive line B73. We show that DNA methylation is modified genome-wide, mainly increasing levels in the three cytosine contexts (CG, CHG and CHH) all over chromosomes. Methylation changes do not affect all contexts in a similar manner across chromosomes: CG and CHG changes are observed in transposable element (TE)-dense regions, while CHH changes are observed close to genes. Modifications however mainly target TE sequences, highlighting the parallel activation of several machineries to tame these sequences through an increase in DNA methylation, in their different genomic locations. Besides these global changes, we also observe high local changes at centromeric and knob locations, as well as at 750 loci detected as Differentially Methylated Regions (DMRs). These DMRs show various types of locations as compared to genes and TEs, and associate with particular changes in the different contexts, suggesting activation of feature-targeted machineries. This study highlights the complex interplay between genome and methylome in response to cold stress in a complex TE-rich genome such as this of maize.

Keywords

Cold, abiotic stress, methylome, transcriptome, DMR, transposable element, maize

Introduction

Plants, as sessile organisms, must cope with constantly changing environments. When unfavorable, environmental constraints can limit normal growth and development, through stress induction at the cellular level. Stressful environmental conditions include biotic constraints (herbivore attacks, pathogen infection) and abiotic constraints such as environmental factors (heat, cold, drought) or soil content (lack of nutrient, excess of salt or heavy metals). Because they affect plant growth and development, stressful constraints impact plant fitness, as well as crop yield. Understanding how plants respond to environmental cues is therefore a major task to both better understand local adaptation and improve agriculture.

DNA methylation is a chromatin mark involved in silencing of transposable elements, maintenance of genome stability and regulation of gene expression (Eichten *et al.*, 2014, Jones, 2012, Law & Jacobsen, 2010, Niederhuth & Schmitz, 2017, Reinders *et al.*, 2009). Analysis of DNA methylation patterns of several loci, in particular genes, has shown that DNA methylation is challenged by abiotic stresses such as cold, water deficit, osmotic pressure, metal concentration, nitrogen deficiency in several species including crops (Choi & Sano, 2007, Labra *et al.*, 2002, Song *et al.*, 2012, Verhoeven *et al.*, 2010). But global genomic response of DNA methylation to abiotic stress remains poorly characterized. When whole genome data have been generated, studies have focused on the detection of Differentially Methylated Regions (DMRs) between the “stressed” and “unstressed” lots, that is, small regions with statistically significant variation. For instance, phosphate deficit in rice roots was shown to induce 175 DMRs, 84% of which were hypermethylated in the “stressed” plants (Secco *et al.*, 2015). Interestingly, these were mostly of the CHH context, and essentially located in transposons close to phosphate-induced genes. These changes were shown to be transmitted through mitosis but not to the progeny. In maize, flag leaves methylome analysis after cold, heat and UV treatments were revealed a total of 389 DMRs, which were enriched near genes (Eichten & Springer, 2015). In *Arabidopsis thaliana*, Ganguly *et al.* (2017) detected 40 drought-induced DMRs, with no transmission to the progeny. In maize, whole genome HPLC-based analysis revealed a global cold-induced demethylation change in roots (Steward *et al.*, 2002) and specific analysis of knob regions revealed an increase in expression accompanied by local chromatin remodeling, increase in H3K9ac and decrease in DNA methylation (Hu *et al.*, 2012). While it is now clear that DNA methylation is affected by abiotic stresses, and in particular by cold in maize, further studies are required to decipher whether induced changes are regulated at the chromosome

scale, or target specific regions of the genome, such as stress responsive genes or particular transposable elements. In addition, the stresses applied have been mainly cold-pulses of a few days, and the effect of early prolonged cold on the methylome remains to be elucidated. This kind of stress has a double interest for detecting an impact at the plant level: (i) long exposure is particularly interesting as plant cells divide under the abiotic constraint, which could induce higher changes due to accumulation of a stronger signal over DNA replication, (ii) exposure early in development affects the meristem that will give rise to all the cells of the organism.

Here, we investigate the impact of prolonged cold on maize methylome and transcriptome. Transcriptome shows that plants did sense cold treatment as a stress, and reveal that expression of genes regulating DNA methylation is modified. Methylome is also modified by cold, mainly by increasing methylation at the three contexts. CG and CHG increase in TE-dense regions, while CHH increases in gene-rich regions, showing that the cold-induced methylation increase involves multiple different pathways, thus pointing to a crucial role of global genome hypermethylation in response to cold. High short and local hyper- and hypo-methylation are also observed, and are enriched in different features and cytosine context, thus highlighting cold-induced recruitment of specific pathways at particular local sequence content.

Results

B73 is sensitive to prolonged cold

To analyze maize transcriptome and methylome response to cold, we generated two sets of plants. In the “stressed” set, plants were subjected at the emergence stage to a long (21 days) but moderate (16h day at 14°C; 8h night at 10°C) cold treatment. Conditions of the “stressed” set were chosen to mimic as much as possible cold conditions that could be experienced in a maize field. In the “unstressed” set, plants were grown in standard greenhouse conditions (16h day at 24°C; 8h night at 18°C) (Additional file 1: Figure S1, see “Materials and Methods”). Plants from the “stressed” set are both delayed in development and show chlorotic leaves, thus showing that B73 phenotype is altered by prolonged cold treatment (Additional file 1: Figure S1).

To better characterize the phenotypic effect observed, we first compared transcriptome from the two plant sets. For this, high coverage strand specific 2x150 bases paired-end mRNA-seq data were generated on three plants of each set, using samples collected from the one before last visible leaf. 196,873,798 pairs of reads were generated, with an average of 32,812,299 read

pairs per sample. Pairs of reads were mapped to B73 reference genome sequence (RefGene v2) completed with gene sequences from other inbred lines. For each sample, over 93% of pairs were unambiguously mapped, and used for subsequent count analysis. Statistical analysis with a Benjamini-Hochberg (FDR<0.05) revealed a total of 6958 differentially expressed genes between “stressed” and “unstressed” plants. Among these, 3021 and 3937 are up- and down-regulated in stressed plants, respectively. These two categories of genes were separately analyzed for Gene Ontology (GO) enrichment. Downregulated genes show an enrichment in photosynthesis (40/117 genes, pvalue= $5.14e^{-11}$), lipid biosynthesis (70/313, pvalue= $6.88e^{-09}$), oxydo-reduction (297/1929 genes, pvalue= $6.16e^{-09}$), transmembrane transporter (163/988 genes, pvalue= $3.95e^{-07}$), cellulose biosynthesis (18/50 genes, pvalue= $4.56e^{-06}$), and glutamine biosynthetase (6/7 genes, pvalue= $1.34e^{-05}$), while upregulated genes are enriched in RNA processing (71/297 genes, pvalue= $4.97e^{-12}$), ribosome biogenesis (23/68, pvalue= $5.43e^{-11}$), protein refolding (17/36 genes, pvalue= $3.53e^{-11}$), DNA replication (25/102 genes, pvalue= $1.55e^{-08}$), DNA repair (40/188 genes, pvalue= $5.34e^{-07}$), in particular mismatch repair (9/16 genes, pvalue= $2.36e^{-07}$). These enrichments explain the visible phenotype, and show that the cold treatment to which plants were exposed induced a physiological stress at the cellular level.

Cold induces a genome-wide increase in DNA methylation

We then investigated changes in methylome. For this, high coverage whole genome Bisulfite-sequencing (BS-seq) was performed on the same set of plants, using the last visible leaves as samples. A total of 1,452,302,823 read pairs (2x100 bases) was generated, with 226,087,680 to 260,391,650 read pairs per sample, thus leading to a theoretical coverage depth of 21 to 24 X.

Read pairs were mapped onto B73/F2 pangenome and cytosines from unambiguously mapped reads were extracted together with corresponding methylation levels (see “Materials and Methods”). About 68% of reads were uniquely mapped, most of which (98%) being mapped as pairs. No obvious difference was observed among all samples or between conditions (Additional file 2: Table S1). Average methylation levels were 0.93%, 0.86% and 0.04% for CG, CHG and CHH in the chloroplast sequence, leading to a global bisulfite conversion efficiency above 99.7%. Based on cytosine coverage per strand (Additional file 1: Figure S2), these cytosines were then filtered to remove too low (below 4x) and too high (over 25x) covered bases to keep only cytosines with precise enough methylation level and to remove regions with possible remaining spurious mapping. These are stringent criteria, but we believe this is necessary to precisely estimate methylation levels and compare them among samples. With

these criteria, a total of 247 303 357 cytosines could be analyzed, among which 54 444 473 CGs, 44 451 905 CHGs and 148 406 979 CHHs (Additional file 2: Table S2).

To get an overview of DNA methylation levels, we first analyzed genome-wide methylation levels. As expected, average methylation levels for the 6 samples were high (82%) for CGs, medium-high (70%) for CHGs and low (1%) for CHHs. Interestingly, methylation levels were a few percent higher in stressed than unstressed plants, overall and for each context separately. This is particularly marked for CHGs (68% vs. 71%) (Additional file 2: Table S2). To further characterize this difference, we then compared the ratio of methylated/all cytosines between the two conditions. Based on distribution of methylation level per cytosine and previous publications (Gent *et al.*, 2013, Niederhuth *et al.*, 2016), we considered a cytosine to be methylated when methylation level was over 80% for CG, over 70% for CHG and over 5% for CHH. Again, difference was detected between samples of the two conditions, with stressed samples showing a higher number of methylated cytosines for all three contexts, and was significant for CHG and CHH (Student's test, pvalue <0.05, Additional file 2: Table S2).

Cold-induced methylation increase varies across chromosomes

To investigate whether the methylation increase observed targets specific chromosome regions, we analyzed average DNA methylation difference between stressed and unstressed plants across the 10 maize chromosomes in the three contexts CG, CHG and CHH using 100kb bins (Additional file 1: Figure S3, S4, S5, S6). As shown in Fig. 1a for the chromosome 10, an increase was observed throughout the chromosome, but was higher in some regions. Methylation difference increased in the middle of chromosome for CG and CHG and was higher in chromosome arms for CHH. As observed at the whole genome scale, difference was the highest for CHG (1 to 4%), high for CG (0 to 2%) and low for CHH (0 to 0.8%) (Fig. 1a). To investigate whether these patterns could be explained by enrichment in particular genomic features, we then compared these profiles to bin coverage of TEs and genes along the chromosome. Interestingly, CG and CHG profiles followed TE coverage, while CHH profile followed gene coverage (Fig. 1a, b, Additional file1: Figure S3, S4, S5 S6). We also noted regions with very high CHG increase (>5% over several bins of 100kb) at centromeres and at particular positions of chromosomes 3, 6, 7 and 10 (Fig. 1a). Extraction of these regions using segmentation and subsequent analysis revealed the presence of knob repeats and enrichment in CENH3 (Fig. 1c, Additional file 1: Figure S7, S8).

Because these regions show a lower CHG methylation level as compared to average genome in unstressed plants (75 vs 80%), we tested whether the increase observed is due to this particular genomic content, or whether it is due to lower original CHG methylation levels. Analysis of methylation change for all genomic regions with methylation levels below 75% in unstressed plants revealed a behavior similar to this of the rest of the genome, thus showing that the CHG response is specific to centromere and knob regions (Fig. 1d).

Cold-induced increase in DNA methylation varies upon genomic features

To better characterize the pattern observed, we then investigated whether the increase in methylation observed at the 100kb scale could be attributed to changes in particular genomic features. In particular, we wanted to test whether CG and CHG increase in TE rich regions are related to a change in TE sequences, and whether CHH increase in gene rich regions is related to a change in genic TEs, or affect any TE located close to genes. To investigate this, we separated the genome features into four categories: “gene only”, “TE only”, “gene and TE” and “no gene no TE” based on overlap or absence of overlap of gene and TE annotation. While genes show almost no variation in any of the three contexts, TE sequences outside of genes show the highest variation. TEs within genes, and regions outside of genes or TEs show an intermediate pattern, most regions being invariable and some showing an increase (Fig. 2a, Additional file 2: Table S3). As expected, methylation levels were always higher for CGs than for CHGs, and were low for CHHs. Nevertheless, the increase observed was higher for CHG, for all types of features (Fig. 2, Additional file 2: Table S3).

Analysis of TE classes separately revealed different patterns. For transposons, an increase in CHG and CHH was observed independently of their location relative to genes, with the exception of DTA where no CHH variation was observed outside of genes. For CHG, this increase is higher outside genes and for CHH it is higher inside genes (Additional file 2: Table S4). For LINEs, both RIX and RIL vary in CHG and CHH when outside of genes. When inside of genes, RIX does not show any variation in neither in CHG nor in CHH, while RIL varies only in CHG (Student’s test, pvalue <0.05). LTR retrotransposon families were mainly located outside of genes. Most of them (95%, 87% and 92% for RLC, RLG and RLX families, respectively) showed an increase in all context, but particularly in CHG.

In a second step, we wanted to investigate the behavior of the so-called CHH islands (*Gent et al.*, 2013). Indeed, while these regions are too short to affect much the 100kb pattern observed, their analysis could help better characterize the mechanisms involved. Analysis of 500pb

upstream gene TSS revealed typical CHH island patterns, which showed a 1% level increase in stressed plants as compared to unstressed ones. A similar pattern was observed in the 500bp downstream genes, but was less marked. Interestingly, genes with no expression had both lower CHH levels (2%) and lower CHH increase (0.5%) than other genes, and the highest peak was observed for genes with no differential expression and genes with expression increase (6% in unstressed plants and 2% increase in stressed plants). However, CHH levels did not correlate with expression levels, as both up- and down-regulated genes showed a CHH increase. The CHH peak observed for down-expressed genes was less sharp than this of up-expressed genes (Fig. 3, Additional file 2: Table S5).

Cold induces strong DNA increase or decrease at short local scale

While whole chromosome analyses on average values are helpful to investigate cold-induced large scale methylation remodeling, it is not adapted to investigate the impact of DNA methylation on particular regions, or possible impact on phenotype. To do so, we studied whether cold could induce strong (>10%) changes in DNA methylation. For this, we characterized Differentially Methylated Regions (DMRs) between “stressed” and “unstressed” plants using information on the three replicates per condition. Cytosines from the CG and CHG contexts were analyzed separately and DMRs were detected using the methylkit R package (Akalin et al., 2012) with windows of 200, 500 and 1000kb (See “Materials and Methods”). DMRs were then filtered to retain only those containing at least 5 Differentially Methylated Cytosines (DMCs) each with a 10% methylation difference. Considering the low levels of CHH methylation, DMRs were not analyzed for the CHH context. Rather, CHH levels were, further, used to categorize CG and CHG DMRs. With these criteria, we detected a total of 750 DMRs, among which 354 were hypermethylated in stressed plants, and 396 were hypomethylated in stressed plants (Table 1). Size of CG DMRs ranges from 200 to 3900 bp, contain 5 to 154 cytosines and 7 methylated cytosines. CHG DMRs are larger (300 to 5700 bp), have more cytosines (5 to 242) and more DMCs (5 to 80) DMCs. For both contexts, DMCs represent about 20% of total cytosines, and methylation differences range from 10% to 75%. Hyper- and hypomethylated DMRs do not show striking differences in any of these characteristics (Table1, Additional file1: Figure S9).

Local DMRs differ both in context-specific methylation levels and genomic location

To get insights into possible underlying mechanisms, we then classified these DMRs based on their variation in either CG (“CG only”), CHG (“CHG only”) or variation in both CG and CHG

("CG plus CHG") using a 2-component Gaussian mixture model (see "Materials and methods", Fig. 4, Additional file 1: Figure S10). This classification allowed classifying 91 % of all DMRs. Among classified cases, the majority of DMRs (52%) varies both in CG and CHG levels, while 31 % vary only in CHG and a smallest fraction (7%) varies only in CG (Table 2). Comparison of hyper- and hypo-methylated cases revealed similar amounts of the different classes. However, "CG only" cases showed higher variation when hypermethylated (from 10% to 30% vs 14% to 19% for hypomethylated cases) (Fig. 4, Additional file 1: Figure S10). Interestingly, CG variation was higher for "CG plus CHG" than for "CG only" cases (4 times for hypermethylated cases, 2.5 times for hypomethylated cases), suggesting different underlying mechanisms (Fig. 4b, Additional file 1: Figure S10). In contrast, no obvious difference is observed for CHG variation between "CHG only" and "CG plus CHG" cases (Fig. 4, Additional file 1: Figure S10). Interestingly, when the DMR shows variation in both CG and CHG, variation levels of the two contexts are correlated, suggesting a concerted mechanism (Fig. 4, Additional file 1: Figure S10). "CG plus CHG" cases are also the only ones to show a variation in CHH, although not proportional (Fig. 5c, Additional file 1: Figure S11c). To analyze methylation signal conservation among individuals, we then performed a chi square test on the number of methylated and unmethylated reads for each DMR. CG stability was analyzed for "CG only" and "CG plus CHG" DMRs, and CHG stability was analyzed for "CHG only" and "CG plus CHG" DMRs. This analysis revealed that about 26% of the cases (25% for CG and 27% for CHG) were stable in unstressed plants, while 22% were stable in stressed plants (23% for CG and 22% for CHG). These two categories did not much overlap (2.8% of stable cases in both conditions for CG, 5% for CHG), the largest number of cases being instability in both conditions (51% for CG and 49% for CHG). No difference was observed between hyper- and hypomethylated cases.

To further characterize the molecular origin of these DMRs, we then investigated methylation levels in stressed and unstressed plants. As shown in Fig. 5a (Additional file 1: Figure S11a), "CG only" DMRs show a wide range of CG methylation levels in unstressed plants (from 0 to 68% and from 16 to 74% for hyper- and hypo-methylated cases, respectively) and are free of methylation in CHG and CHH contexts. "CHG only" DMRs also occur from a large range of original methylation levels (0 to 72% and 17 to 91% CHG methylation in unstressed plants for hyper- and hypomethylated cases) and an absence of CHH methylation but, in contrast to "CG only" cases, they also show a high (>90%) and invariable level of CG methylation (Fig. 5b, Additional file 1: Figure S11b). "CG plus CHG" DMRs are characterized by an even wider

range of original methylation level in both CG and CHG as compared to the two other classes (0 to 94% and 25 to 100% for CG and 0 to 72% and 0 to 91% in hyper- and hypo-methylated cases, respectively). They also harbor CHH methylation (up to 24%), which is absent from the two other classes. For all classes of DMRs, hypomethylated cases have 10% to 20% higher methylation levels in unstressed plants than hypermethylated cases (Fig. 5c, Additional file 1: Figure S11c).

To get complementary insights into the molecular origin of methylation variation, we then analyzed the distance of DMRs to genes and to TEs (see “Materials and methods”, Additional file 1: Figure S12). Distance analysis to TEs shows that all three categories of DMRs are enriched upstream and downstream of TEs but depleted inside TE sequences. In addition, “CG only” DMRs are also enriched away from TEs (Fig. 6a). Distance analysis to genes show that “CG only” and “CHG only” DMRs are enriched inside genes, while “CG plus CHG” DMRs are enriched upstream and downstream of genes. These patterns are similar for hyper- and hypomethylated cases (Fig. 6b). “CG only” and “CHG only” DMRs being enriched within genes, we then investigated whether they could be related to genic TEs. Indeed, for hypermethylation cases, “CG only” DMRs are enriched upstream and downstream of TEs. In contrast, hypomethylation cases are depleted upstream of TEs. For both hypermethylated and hypomethylated categories, “CG only” DMRs exist in genes independently of TE, but as often as expected by chance. “CHG only” DMRs are enriched around and inside TEs, and depleted away from TEs (Fig. 6c). Analysis of DMR location across chromosomes did not reveal any particular clustering, suggesting that they occur in an independent manner (Additional file 1: Figure S13).

Genes involved in DNA methylation regulation vary in expression levels between stressed and unstressed plants

To further investigate origin of the methylome response observed, we specifically analyzed the expression pattern of a series of 74 genes described in the literature as involved in DNA methylation regulation. A total of 27 genes (36%) showed significant difference between stressed and unstressed plants, among which 5 were down-expressed and 22 were up-expressed in stressed plants (Table 3). Interestingly, the 22 up-expressed cases contained 7 genes involved in the RdDM pathway: PolIV and PolIV components (NRP(B/D/E)6a, NRP(D/E)7)), PolIV/PolV interacting factors (IWR1/DMS4/RDM4, SHH2c and MOP1), or AGO proteins (AGO104 and AGO105). Genes involved in CG and CHG maintenance such as MET1b, MET6 and MET8, and KYP and SUVH5/6 also showed an expression increase in stressed plants, as

well as DDM1. On the other hand, while the gene coding the largest PolIII unit (NRPB1) was increased in expression in stressed plants, genes coding the subunits NRPB2a, NRPB5a and NRPB9a showed significant expression decrease.

DMRs located inside of genes are related with an increase in expression, and correspond to CHG only cases located nearby TEs

To investigate the possible link between methylome and transcriptome variation, we analyzed the transcription change of the 377 genes with a DMR located inside or in the near vicinity (<2kb). Among the 192 genes with an hypermethylated DMR inside or nearby, only 40 showed a significant change in expression, most of which corresponded to cases with the DMR located inside (31 cases) and downstream of gene (8 cases). Most of these (20/31 and 6/8, respectively) corresponded to an increase in expression. Among the 185 genes with a hypomethylated DMR inside or nearby, only 50 showed a significant change in expression, most of which corresponded to cases with the DMR located inside (31 cases) and downstream of gene (13 cases). Genes with DMR located inside were enriched in increase in expression (21/31 cases) while genes with DMR located downstream of genes showed enrichment in down expression (9/13 cases). Interestingly, genes with a DMR located upstream were rarely affected in expression (2 cases for hypermethylated DMRs, and 6 cases for hypomethylated DMRs). Analysis of the 41 DMRs (20 hypermethylated cases and 21 hypomethylated cases) located within genes with a transcriptional change were mostly of the CHG type (35 cases) and close (<500bp) of a TE (39 cases). This is true for both hyper- and hypomethylated cases.

Discussion

In this study, we use whole genome bisulfite-seq and mRNA-seq methods to decipher transcriptome and methylome variations between stressed and unstressed seedling leaves at the end of a 21-days cold treatment. Transcriptome analysis confirms previous observation of a repression of genes involved in photosynthesis (photosynthesis, glutamine biosynthetase) and cell wall (cellulose synthase, lipids, transmembrane transporter) (Nguyen *et al.*, 2009, Sobkowiak *et al.*, 2014). However, we did not observe modified expression of anthocyan biosynthesis or circadian clock as observed in Sobkowiak *et al.* (2014). In addition, we show that the cold treatment applied was strong enough to destabilize proper basal cellular functions such as DNA replication, RNA processing and protein folding. In particular, the observation of

an increase of DNA mismatch repair function highlights that the conditions used induced mutagenesis, and thus that the plants were subjected to a strong physiological stress.

This stress induced also hypermethylation at the chromosome scale. This hypermethylation is low (few percent) but consistent all across chromosomes, and affects the three cytosine contexts. Interestingly, while the three contexts are modified, they do not affect the same genomic regions: while CG and CHG are both hypermethylated in regions where TEs are enriched, and are modified in TEs inside and outside of genes, CHH methylation difference is higher in gene rich regions, and affects TEs outside of genes as well as CHH islands located around genes. This suggests that the hypermethylation induced by cold involves the parallel set-up of several mechanisms. These mechanisms all target TE sequences, but depends on the type of heterochromatin involved. In the TE-rich heterochromatin, which is typically devoid of RdDM and rich in H3K9me3 (Gent *et al.*, 2014), the increase in CG and CHG methylation is likely induced by an increase in methylation maintenance. This is in accordance with the observed transcriptional increase in MET1b which mediates CG maintenance, and for KYP and SUVH5/6 which are involved in the self-reinforcing loop of CHG maintenance (Matzke & Mosher, 2014, Zemach *et al.*, 2013). In regions located around genes, where chromatin is mediated by RdDM to silence TEs while keeping chromatin in a relaxed form for possible transcription of the genes when needed in development, increase in CHH methylation is likely explained by an increase in the RdDM pathway. This is in accordance with the expression increase observed for PolIV/PolV interacting proteins such as NRP(B/D/E)6a, NRP(D/E)7, and for PolIV-associated proteins such as IWR1/DMS4/RDM4, SHH2 and MOP1. An increase in DDM1 expression is also detected in our dataset, and may participate in the hypermethylation observed at TE sequences.

In addition to these chromosome-scale changes, we also observed particularly high increase of CHG methylation at CENH3-binding centromeres and knob regions located on chromosomes 3, 6, 7 and 10. As well as TE-dense regions, centromere and knob regions are enriched in H3K9me2 and H3K27me2 (Gent *et al.*, 2011, Regulski *et al.*, 2013, Shi & Dawe, 2006). So the mechanism involved may also be related to CG and CHG maintenance. The particularly high (10%) increase observed in these regions seems however to be specific of this particular type of sequences, thus suggesting a possible role of CENH3 itself.

Expression change of genes involved in DNA methylation pathways has been reported for other abiotic constraints. For example, heat upregulate RdDM genes such DRM2, ARN PolIV (NRPD1) and PolIV (NRPE1) in *Arabidopsis* (Naydenov *et al.*, 2015b), osmotic stress reduces

the expression of OsMET1-1 in rice roots (Ahmad *et al.*, 2014, Sharma *et al.*, 2009) and cold, heat and drought stress increase methyltransferases genes transcription in peanut (Wang *et al.*, 2016a). In maize, four of the expressed methyltransferases (ZmMET1a, ZmMET1b, ZmMET2a and ZmMET2b) showed decreased expression levels in leaves after drought and salt stress (Qian *et al.*, 2014). Here, among 74 maize genes involved in the regulation of DNA methylation, we observed a total of 27 genes (36%) with differential expression between stressed and unstressed plants, among which only 5 are downregulated in stressed plants. This shows that in our experiment, stress induced an intensive activation of these pathways, which led to a global hypermethylation. This hypermethylation is mainly observed in TE sequences, which suggests an increase in TE silencing. Observation of expression increase of genes involved in DNA repair suggests that the stress experienced was strong. If so, it may also have induced mobilization of TEs. Interestingly, two of the 5 genes with significant decrease in expression in stressed plants are PolIII-related genes (NRPB9a and NRPB5a). In *A.thaliana*, polII mutants lead to mobilization of some TEs (Zheng *et al.*, 2009). Although this is still speculative, the stress-induced decrease in the PolIII function we observe in our plants could also lead to mobilization of some TEs. This could explain the need to tighten TE silencing in stressed plants, and may explain the overall increase in methylation at TE regions. However, we also observe an increase of expression for the large NRPB1 PolII subunit, so further analyses will be needed to investigate this point.

With our data, we also observe a total of 750 short regions with differential methylation over 10%. This is higher than the 175 DMRs observed in rice roots after phosphate starvation (Secco *et al.*, 2015), or than the 40 DMRs associated with drought stress in *A.thaliana* (Ganguly *et al.*, 2017). Again, this could be due to the severity of the stress, or to the particular nature of the maize genome. Observation of many (half of the total cases) hypomethylated cases was unexpected considering the overall hypermethylation observed in our stressed plants, and suggest a different mechanism. While not emphasized by the authors, such hypomethylated cases have also been observed in *Arabidopsis* and rice. In the rice case, hypermethylated regions dominated over hypomethylated regions (147 vs. 28, respectively, Secco *et al.*, 2015). But in *A.thaliana*, hypomethylated cases were as abundant as hypermethylated cases for the CG and CHG DMRs detected (9 and 8 cases, respectively, Ganguly *et al.*, 2017).

Comparison of methylation levels among individuals for all the DMRs revealed a large number of unstable cases. While previous studies have reported higher instability in stressed plants as compared with unstressed ones (Eichten & Springer, 2015), we did not observe this. Rather,

instability was almost similar in each of the two conditions and represented three quarter of the total cases. The large variability observed in each condition is mainly explained by the particular behavior in one plant of each set. Origin of this remains to be fully elucidated. In the stressed lot, the outlier plant had an intermediate phenotype between stressed and unstressed, which was visible at the whole plant level, as well as the transcriptome level. So we anticipate that this plant was not as affected by cold as the others. This difference could originate from cold variability in the room. Pots were rotated to avoid a local effect, but an interaction between developmental stage and position cannot be excluded. The observation of an outlier plant in the unstressed lot is less clear, and a possible maternal effect cannot be ruled out, as seeds did not derive from a single plant. Despite the presence of these outliers, we could detect statistically significant cases of DMR with repeatable pattern in each of the plant lots (6 for CG and 13 for CHG), as well as many cases (50 and 64 for CG and CHG, respectively) showing stability in the stressed lot. This highlights that the same regions are repeatedly targeted through the cold treatment. As for landscapes variation, these DMRs seem to originate from different mechanisms, likely including gene-body methylation (some “CG only” cases), the action of IBM and CG and CHG maintenance in nearby genic TEs (some “CG only” and “CHG only” cases), and RdDM in non-genic TEs (“CG plus CHG” cases) (Bewick & Schmitz, 2017, Wendte & Schmitz, 2017). Link between presence of these DMRs and transcription of the genes is limited, but several cases highlight the link between transcriptional change and TE-derived CHG methylation. Overall, TEs are involved in many cases of the methylation changes observed, both at the global and local scale. This suggests that cold modifies the methylome in a way that increases silencing of TEs, likely to reduce the impact of possible cold-induced TE mobilization.

Conclusion

Our study shows that prolonged cold induces moderate global hypermethylation in the maize genome, as well as stronger changes at the local scale, including hypermethylation of centromere and knobs, and hypermethylation and hypomethylation in genes. These changes involve different underlying mechanisms, thus highlighting the need for the genome to increase DNA methylation at any cost. While the methylome changes seem to have a limited impact on phenotype, the global increase in methylation observed, if pervasive over generations, could promote future C to T transition rates. This suggests that strong abiotic stresses may impact genome evolution rather than (or together with) physiological response.

Materials and Methods

Plant material and growth conditions

Seeds of maize inbred line B73 were provided by INRA, Saint-Martin-de-Hinx, France. Plants were grown in greenhouse (16h day at 24°C; 8h night at 18°C) for 3 days until emergence stage (1-2 visible leaves). Then, plants for the “stressed” set were placed in an air conditioned room (16h day at 14°C; 8h night at 10°C) for a prolonged cold treatment of 21 days. To compare developmentally synchronized plants between “stressed” and “unstressed” plants, the “unstressed” set was sown with a shift of 16 days from the “stressed” set. At the end of cold treatment, for both “stressed” and “unstressed” sets, the last and one before last visible leaves were harvested independently from 3 distinct seedlings (biological replicates) (Additional file 1: Figure S1), directly frozen in liquid nitrogen and stored at -80°C for further analysis. Last visible leaves and one before last visible leaves were used for transcriptome and methylome studies, respectively.

mRNA extraction, RNA-seq library preparation and RNA sequencing

mRNAs were extracted from the one before last visible leaf with Trizol (Invitrogen ref.15596018) and β -mercaptoethanol (SIGMA ref. M3148-25ML) reagents. Supernatant was recovered and RNA purified using Qiagen RNeasy Plant Mini kit (ref. 74904) following manufacturer’s instructions. Then, a Qiagen RNase-free DNase set (ref.79254) was applied to remove the residual DNA. Library construction and sequencing were generated by the IPS2-POPS platform with Illumina NexSeq500 sequencer using barcoded adaptors and 18 samples per lane. Briefly, mRNA were polyA selected, fragmented to 260 bases and libraries were performed using the TruSeq stranded mRNA kit (Illumina®, California, U.S.A.) and sequenced in paired-end (PE) with 150 bases read length.

mRNA_seq reads quality trimming and mapping

mRNA reads were trimmed for library adapters and for Phred Quality Score Qscore >20, read length >30 bases using Trimmomatic version 0.36 (Bolger *et al.*, 2014) . Ribosome sequences were removed using sortMeRNA version 2.1 (Kopylova *et al.*, 2012). Paired-end reads were then aligned onto the reference genome sequence B73 RefGenv2, completed with genomic regions of the F2 inbred line (Darracq *et al.*, 2018), and transcriptome sequences from Hirsh *et al.* (Hirsch *et al.*, 2014) . STAR version 2.5.2b was used for mapping, with options --outSAMprimaryFlag AllBestScore --outFilterMultimapScoreRange 0 to keep the bests results.

Only unambiguously mapped reads were used, and multihits were removed. With this procedure, over 80% of the reads were associated to a gene.

Detection of differentially expressed genes

Differential analysis followed the procedure described in Rigai *et al.* (2016). Briefly, genes with less than 1 read after a counts per million (CPM) normalization in at least one half of the samples were discarded. Library size was normalized using the trimmed mean of M-value (TMM) method and count distribution was modeled with a negative binomial generalized linear model where the environment factor (“stressed” or “unstressed” plants) was taken into account. Dispersion was estimated by the edgeR method (Version 1.12.0, McCarthy 2012) in the statistical software ‘R’ (Version 2.15.0, R Development Core Team (2005)). Expression differences were compared between stressed and unstressed plants using likelihood ratio test and p-values were adjusted by the Benjamini-Hochberg procedure to control False Discovery Rate (FDR). A gene was declared differentially expressed if its adjusted p-value < 0.05. FPKMs-like counts were calculated for visual analysis only, and were obtained by dividing normalized counts by gene length.

DNA extraction, Sodium bisulfite treated library preparation and DNA sequencing

DNAs were extracted from the last visible leaf following manufacturer’s protocol using Macherey Nagel midi extraction kit. For each sample, 5 µg of genomic DNA was fragmented by acoustic shearing to a target value of 250 bp using a Covaris S220 (Covaris, Woburn, MA), according to the manufacturer recommendations. Fragmented DNA was purified using the QIAQuick PCR purification kit (Qiagen, Valencia, CA). Then, DNA fragments underwent an NGS library preparation procedure consisting in end repair and methylated adaptor ligation using the TruSeq DNA Sample Preparation kit (Illumina, San Diego, CA). Adapter-ligated fragments ranging from 200 to 300 nt were isolated by agarose gel electrophoresis (Bio-Rad, Hercules, CA) and purified using the MinElute Gel Extraction Kit (Qiagen, Valencia, CA). Quality of the resulting NGS libraries was assessed by capillary electrophoresis using a Bioanalyzer DNA High-Sensitivity Chip (Agilent, Santa Clara, CA) and quantified using the Quant-It PicoGreen (Life Technologies, Carlsbad, CA). 500 ng of each native library were treated by sodium bisulfite using the EpiTect Bisulfite kit (Qiagen, Valencia, CA) and purified according to manufacturer’s instructions, followed by an additional round of purification using the MinElute PCR Purification Kit with a final resuspension volume of 15 µl (Qiagen, Valencia, CA). DNA fragments carrying both adapters after sodium bisulfite conversion were enriched

by 8 cycles of PCR using the Pfu Turbo Cx Hotstart DNA Polymerase (Agilent, Santa Clara, CA). Primer dimers were subsequently removed using two rounds of purification using AMPure beads (Beckman Coulter, Indianapolis, IN). Quality and quantity of this amplified library was assessed as precedently. For each sample NGS data were generated by Illumina paired-end sequencing (2x100 bases) on a HiSeq2000 machine. Sequence quality was evaluated using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>).

BS-seq reads mapping and extraction of methylation levels

Read pairs were mapped onto a pangenome sequence containing the B73 (RefGenv2) genome sequence and additional genomic regions from the F2 genome assembly (Darracq *et al.*, 2018). Fasta file from this pangenome sequence was indexed with the novoindex tool (<http://www.novocraft.com/documentation/novoindex/>) using the bisulfite option, thus creating two indexes: a “CT” reference where all the Cs of the forward (“Watson”) strand are converted into Ts, and a “GA” reference where all the Gs of the forward strand are converted into As. Bisulfite treated reads were then aligned onto these two indexed references using Novoalign v3.04.06 using plant adapted stringent parameters. Alignment resulting files were split according to the mapping strand (CT or GA) and converted into bam compressed files using samtools 1.2 (Li *et al.*, 2009). Only uniquely mapped reads were kept for further analysis. Samtools flagstat (<http://www.htslib.org/doc/samtools.html>) was used to assess the quantity of paired, proper paired or single reads among these uniquely mapped reads (Additional file 2: Table S1). Methylation level per cytosine was estimated by Novomethyl following the Novocraft protocol (<http://www.novocraft.com/documentation/novoalign-2/novoalign-user-guide/bisulphite-treated-reads/novomethyl-analyzing-methylation-status/>). An in-house perl script was then used to exclude heterozygous sites and low quality consensus calling, and to format them into methylKit (Akalın *et al.*, 2012) compatible files for further analyses. Methylation files from the 6 samples were split by cytosine context, and cytosines with 4x to 25x coverage depth in all 6 samples were extracted using a custom perl script.

Profiling methylation difference between stressed and unstressed plants

Average methylation levels per cytosine were calculated for the 3 “stressed” and for the 3 “unstressed” samples independently, and the difference between “stressed” and “unstressed” samples was computed. These values were used for variation analysis at the chromosome and feature scales. For chromosomes landscapes, windows were based on chromosome bins of 100 kb, or on genomic features. TE and gene annotation was based on B73 refGenv2.

For statistical comparison of methylation levels between features, cytosines overlapping gene only, gene plus TE, TE only and “other” were extracted for each context and each sample using *Bedtools intersect* (Quinlan & Hall, 2010). All cytosines and methylated cytosines were extracted for each feature type. Cytosines are considered as methylated when the methylation level >80% for CG, >70% for CHG and >5% for CHH. Based on these counts, Student’s test (pvalue < 0.05) was performed in order to compare “stressed” and “unstressed” sets.

For targeted analysis of TEs, these were categorized into superfamilies for LINEs (RIL, RIX) and transposons (DTC, DTM, DTA, DTH, DTT) and into families for LTR retrotransposons. Analysis of CHH methylation patterns upstream and downstream (500pb regions) of genes was done after classification of genes into “no expression”, “expressed with no difference”, “up-expressed in stressed samples” and “down-expressed in stressed samples” based on m-RNAseq results.

Detection of Differentially Methylated Regions (DMRs) between “stressed” and “unstressed” plants

Differentially Methylated Regions (DMRs) and Differentially Methylated Cytosines (DMCs) were detected between “stressed” and “unstressed” plants’ for CG and CHG cytosines, separately with the methylKit R package (Akalin *et al.*, 2012). Only cytosines with [4,25X] coverage per strand in each of the 6 replicates were analyzed (Additional file 2: Figure S2, Additional file 2: Table S2). A q-value below 0.01 and a methylation difference over 10% were required. For DMRs, sliding windows of 200, 500 and 1000 bp with a step size of 100 bp were used. Number of DMCs per DMR was estimated by comparing DMRs and DMCs positions using *Bedtools intersectBed* (Quinlan & Hall, 2010) with the option -c, using the cytosine context used for original detection. Only DMRs with at least 5 DMCs were kept for further analysis as in (Secco *et al.*, 2015). Finally, DMRs detected in different window size experiments were merged together using a custom Python script: when the same amount of DMC was found in the three window sets, DMR with the shortest size was kept. Otherwise, DMR with the largest number of DMC was kept. Methylation level of these regions was then estimated for each sample with roimethstat (<http://smithlabresearch.org/software/methpipe>) using the methylation call files used for original DMR detection.

DMR classification by context

DMR classification into “CG only”, “CHG only” and “CG plus CHG” categories was performed by a 2-component Gaussian mixture model using the mixture R package. The

Gaussian covariance structure was selected based on Bayesian Information Criterion (BIC). Each DMR was assigned to a cluster when its posterior probability was higher than 90%. It was considered unclassified otherwise. For “CG plus CHG” cases, redundancy between cases originally found by CG difference and originally found by CHG difference were merged using a 50% overlap threshold using Bedtools *intersectBed* (Quinlan & Hall, 2010), keeping the largest boundaries. DNA methylation levels were re-estimated for these new locations using roimethstat (<http://smithlabresearch.org/software/methpipe>).

DMR annotation

DMR location was compared to this of genes and transposons from B73 refgenv2 using Bedtools *intersectBed* (Quinlan & Hall, 2010). DMRs were categorized as “in”, “upstream”, “downstream” or “away” based on distance, strand and relative start and end positions (Additional file1: Figure S12). A distance superior to 500pb or 2kb was required for a DMR to be considered as “away” from transposons and genes, respectively. Enrichment of DMRs in the different categories was based on relative abundance as compared to a random set of regions. For each DMR, a random region with same size and from same chromosome was retrieved from the B73 genome using *shuffleBed* (V.2.7.3) (Quinlan & Hall, 2010) and subsequently checked to contain at least 5 cytosines from the [4x-25x] dataset used for original detection using custom Python script. Thus, random regions for “CG only” DMRs were checked to carry at least 5 CGs, these for “CHG only” DMRs to carry at least 5 CHGs, and these for “CG plus CHG” DMRs to carry at least 5 CGs and 5 CHGs.

Acknowledgements

We thank INRA Saint Martin de Hinx for seed stocks. The POPS platform benefits from the support of the LabEx Saclay Plant Sciences-SPS (ANR-10-LABX-0040-SPS).

Funding

This work was supported by the Investment for the Future ANR-10-BTBR-01-01 Amazing program. Salary of Z. Achour was supported by a scholarship from the Tunisian government.

Additional files

Additional file 1: Supplementary Figures S1–S13.

Additional file 2: Supplementary Tables S1–S5.

References

- Ahmad, F., X. Huang, H. X. Lan, T. Huma, Y. M. Bao, J. Huang and H. S. Zhang, 2014: Comprehensive gene expression analysis of the DNA (cytosine-5) methyltransferase family in rice (*Oryza sativa* L.). *Genet. Mol. Res*, **13**, 5159-5172.
- Akalin, A., M. Kormaksson, S. Li, F. E. Garrett-Bakelman, M. E. Figueroa, A. Melnick and C. E. Mason, 2012: methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*, **13**, R87.
- Bewick, A. J. and R. J. Schmitz, 2017: Gene body DNA methylation in plants. *Curr Opin Plant Biol*, **36**, 103-110.
- Bolger, A. M., M. Lohse and B. Usadel, 2014: Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.
- Choi, C. S. and H. Sano, 2007: Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. *Mol Genet Genomics*, **277**, 589-600.
- Darracq, A., C. Vitte, S. Nicolas, J. Duarte, J.-P. Pichon, T. Mary-Huard, C. Chevalier, A. Bérard, M.-C. Le Paslier and P. Rogowsky, 2018: Sequence analysis of European maize inbred line F2 provides new insights into molecular and chromosomal characteristics of presence/absence variants. *BMC Genomics*, **19**, 119.
- Eichten, S. R., R. J. Schmitz and N. M. Springer, 2014: Epigenetics: Beyond Chromatin Modifications and Complex Genetic Regulation. *Plant Physiol*, **165**, 933-947.
- Eichten, S. R. and N. M. Springer, 2015: Minimal evidence for consistent changes in maize DNA methylation patterns following environmental stress. *Front Plant Sci*, **6**, 308.
- Ganguly, D. R., P. A. Crisp, S. R. Eichten and B. J. Pogson, 2017: The Arabidopsis DNA Methylome Is Stable under Transgenerational Drought Stress. *Plant Physiol*, **175**, 1893-1912.
- Gent, J. I., Y. Dong, J. Jiang and R. K. Dawe, 2011: Strong epigenetic similarity between maize centromeric and pericentromeric regions at the level of small RNAs, DNA methylation and H3 chromatin modifications. *Nucleic acids research*, **40**, 1550-1560.
- Gent, J. I., N. A. Ellis, L. Guo, A. E. Harkess, Y. Yao, X. Zhang and R. K. Dawe, 2013: CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. *Genome Res*, **23**, 628-637.
- Gent, J. I., T. F. Madzima, R. Bader, M. R. Kent, X. Zhang, M. Stam, K. M. McGinnis and R. K. Dawe, 2014: Accessible DNA and relative depletion of H3K9me2 at maize loci undergoing RNA-directed DNA methylation. *Plant Cell*, **26**, 4903-4917.
- Hirsch, C. N., J. M. Foerster, J. M. Johnson, R. S. Sekhon, G. Muttoni, B. Vaillancourt, F. Peñagaricano, E. Lindquist, M. A. Pedraza and K. Barry, 2014: Insights into the maize pan-genome and pan-transcriptome. *The Plant Cell*, **26**, 121-135.
- Hu, Y., L. Zhang, S. He, M. Huang, J. Tan, L. Zhao, S. Yan, H. Li, K. Zhou, Y. Liang and L. Li, 2012: Cold stress selectively unsilences tandem repeats in heterochromatin associated with accumulation of H3K9ac. *Plant Cell Environ*, **35**, 2130-2142.
- Jones, P. A., 2012: Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*, **13**, 484-492.
- Kopylova, E., L. Noé and H. Touzet, 2012: SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, **28**, 3211-3217.
- Labra, M., A. Ghiani, S. Citterio, S. Sgorbati, F. Sala, C. Vannini, M. Ruffini-Castiglione and M. Bracale, 2002: Analysis of cytosine methylation pattern in response to water deficit in pea root tips. *Plant Biology*, **4**, 694-699.

- Law, J. A. and S. E. Jacobsen, 2010: Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews. Genetics*, **11**, 204.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing, 2009: The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.
- Matzke, M. A. and R. A. Moshier, 2014: RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics*, **15**, 394-408.
- Naydenov, M., V. Baev, E. Apostolova, N. Gospodinova, G. Sablok, M. Gozmanova and G. Yahubyan, 2015: High-temperature effect on genes engaged in DNA methylation and affected by DNA methylation in Arabidopsis. *Plant Physiology and Biochemistry*, **87**, 102-108.
- Nguyen, H. T., J. Leipner, P. Stamp and O. Guerra-Peraza, 2009: Low temperature stress in maize (*Zea mays* L.) induces genes involved in photosynthesis and signal transduction as studied by suppression subtractive hybridization. *Plant Physiol Biochem*, **47**, 116-122.
- Niederhuth, C. E., A. J. Bewick, L. Ji, M. S. Alabady, K. Do Kim, Q. Li, N. A. Rohr, A. Rambani, J. M. Burke and J. A. Udall, 2016: Widespread natural variation of DNA methylation within angiosperms. *Genome biology*, **17**, 194.
- Niederhuth, C. E. and R. J. Schmitz, 2017: Putting DNA methylation in context: from genomes to gene expression in plants. *Biochim Biophys Acta*, **1860**, 149-156.
- Qian, Y., Y. Xi, B. Cheng and S. Zhu, 2014: Genome-wide identification and expression profiling of DNA methyltransferase gene family in maize. *Plant Cell Rep*, **33**, 1661-1672.
- Quinlan, A. R. and I. M. Hall, 2010: BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841-842.
- Regulski, M., Z. Lu, J. Kendall, M. T. Donoghue, J. Reinders, V. Llaca, S. Deschamps, A. Smith, D. Levy, W. R. McCombie, S. Tingey, A. Rafalski, J. Hicks, D. Ware and R. A. Martienssen, 2013: The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res*, **23**, 1651-1662.
- Reinders, J., B. B. Wulff, M. Mirouze, A. Marí-Ordóñez, M. Dapp, W. Rozhon, E. Bucher, G. Theiler and J. Paszkowski, 2009: Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes & Development*, **23**, 939-950.
- Rigaill, G., S. Balzergue, V. Brunaud, E. Blondet, A. Rau, O. Rogier, J. Caius, C. Maugis-Rabusseau, L. Soubigou-Taconnat and S. Aubourg, 2016: Synthetic data sets for the identification of key ingredients for RNA-seq differential analysis. *Briefings in bioinformatics*, **19**, 65-76.
- Secco, D., C. Wang, H. Shou, M. D. Schultz, S. Chiarenza, L. Nussaume, J. R. Ecker, J. Whelan and R. Lister, 2015: Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *Elife*, **4**.
- Sharma, R., R. K. Mohan Singh, G. Malik, P. Deveshwar, A. K. Tyagi, S. Kapoor and M. Kapoor, 2009: Rice cytosine DNA methyltransferases - gene expression profiling during reproductive development and abiotic stress. *FEBS J*, **276**, 6301-6311.
- Shi, J. and R. K. Dawe, 2006: Partitioning of the maize epigenome by the number of methyl groups on histone H3 lysines 9 and 27. *Genetics*, **173**, 1571-1583.
- Sobkowiak, A., M. Jonczyk, E. Jarochowska, P. Biecek, J. Trzcinska-Danielewicz, J. Leipner, J. Fronk and P. Sowinski, 2014: Genome-wide transcriptomic analysis of response to low temperature reveals candidate genes determining divergent cold-sensitivity of maize inbred lines. *Plant Mol Biol*, **85**, 317-331.

- Song, Y., D. Ji, S. Li, P. Wang, Q. Li and F. Xiang, 2012: The dynamic changes of DNA methylation and histone modifications of salt responsive transcription factor genes in soybean. *PLoS One*, **7**, e41274.
- Steward, N., M. Ito, Y. Yamaguchi, N. Koizumi and H. Sano, 2002: Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *J Biol Chem*, **277**, 37741-37746.
- Verhoeven, K. J., J. J. Jansen, P. J. van Dijk and A. Biere, 2010: Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol*, **185**, 1108-1118.
- Wang, P., C. Gao, X. Bian, S. Zhao, C. Zhao, H. Xia, H. Song, L. Hou, S. Wan and X. Wang, 2016: Genome-Wide Identification and Comparative Analysis of Cytosine-5 DNA Methyltransferase and Demethylase Families in Wild and Cultivated Peanut. *Front Plant Sci*, **7**, 7.
- Wendte, J. M. and R. J. Schmitz, 2017: Specifications of targeting heterochromatin modifications in plants. *Molecular plant*.
- Zemach, A., M. Y. Kim, P.-H. Hsieh, D. Coleman-Derr, L. Eshed-Williams, K. Thao, S. L. Harmer and D. Zilberman, 2013: The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*, **153**, 193-205.
- Zheng, B., Z. Wang, S. Li, B. Yu, J.-Y. Liu and X. Chen, 2009: Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. *Genes & development*, **23**, 2850-2860.

Tables

Table 1 Characteristics of DMRs detected in each cytosine context

Type context	Hypermethylated		Hypomethylated	
	CG	CHG	CG	CHG
DMR number	192	162	228	168
Median DMR size (bp)	400	800	400	700
Median methylation difference (%)	29	35	25	30
Median cytosine number per DMR	21	34	30	39
Median DMC number	6	7	7	7

Table 2 Amount of different DMR types as defined by methylation differences observed in CG and CHG contexts

Type	Hypermethylated	Hypomethylated
CG only	30	18
CHG only	100	120
CG plus CHG	180	183
Unclassified	12	56

Table 3 Genes involved in DNA methylation regulation which show differential expression between stressed and unstressed plants.

Genes are ranked based on Log2FoldChange between the two conditions. Unstressed: Average normalized count values for unstressed plants. Stressed: Average normalized count values for stressed plants. Log2FoldChange are positive when the gene is more expressed in stressed than unstressed plants. FDR: False discovery rate based on Benjamini-Hochberg correction.

Gene name	Gene ID	Unstressed	Stressed	Log2FoldChange	P-value	FDR
NRPB9a	GRMZM2G046061	574	40	-3,8	5,5 E-11	3,1E-09
MET2	GRMZM2G025592	200	27	-2,9	1,4E-05	1,3E-04
MET3	GRMZM2G092497	2894	567	-2,4	3,0E-05	2,5E-04
DCL101	GRMZM2G040762	1384	588	-1,2	2,9E-03	1,1E-02
NRPB5a	GRMZM2G476009	2198	1079	-1,0	8,9E-03	2,8E-02
NRPB1	GRMZM2G044306	429	1056	1,3	3,9E-03	1,4E-02
AGO105 (AGO4)	GRMZM2G589579	95	234	1,3	7,3E-03	2,3E-02
SUVH5/6	GRMZM2G054380	56	158	1,5	7,5E-04	3,6E-03
DCL105	GRMZM2G301405	101	299	1,6	2,4E-03	9,4E-03
NRP(B/D/E)6a	GRMZM2G013600	422	1250	1,6	7,6E-04	3,6E-03
MET6	GRMZM2G065599	62	184	1,6	4,7E-03	1,6E-02
MET1b	GRMZM2G334041	36	112	1,6	4,4E-04	2,3E-03
MET8	GRMZM2G334041	36	112	1,6	4,4E-04	2,3E-03
IWR1/DMS4/RDM4	GRMZM2G098603	310	1000	1,7	1,3E-04	8,5E-04
SHH2c	GRMZM5G843174	45	190	2,1	4,8E-04	2,5E-03
MOP1/RDR101	GRMZM2G042443	27	115	2,1	4,4E-05	3,3E-04
NRP(D/E)7	GRMZM2G040702	83	360	2,1	1,1E-04	7,4E-04
AGO104 (AGO9)	GRMZM2G141818	78	404	2,4	6,5E-09	1,9E-07
KYPb	GRMZM2G336909	79	413	2,4	2,9E-08	7,1E-07
DDM1a	GRMZM2G177165	2	18	2,8	5,5E-05	4,0E-04

Figures

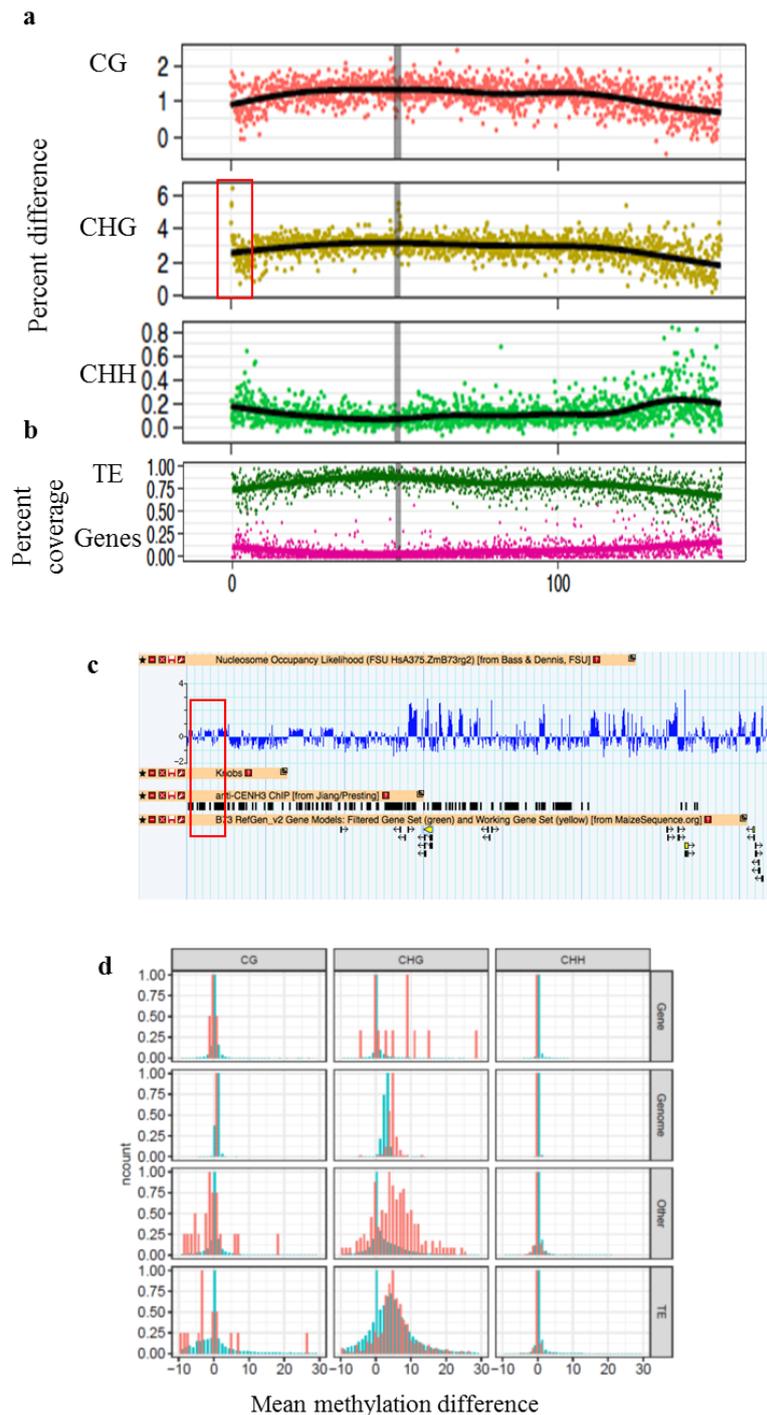


Fig. 1 Global DNA methylation analysis

a Chromosome 10 landscape showing average difference between methylation levels of stressed and unstressed plants over 100kb bins, for each of the three cytosine contexts (CG, CHG and CHH). Centromere is highlighted by a grey bar. Red square highlights the location of a region with particularly high difference. **b** Chromosome 10 landscape highlighting percent of coverage of each bin by TE (green, MTEC annotation), and genes (pink, Filtered gene set annotation). **c** Screenshot of the MaizeGDB website, showing the annotation of knob repeats and enrichment in CENH3 in the region detected with high variation at the tip of chromosome 10. **d** Difference of methylation observed in centromere regions (red) or other genomic regions with similar level of methylation (<75%) in unstressed plants (green).

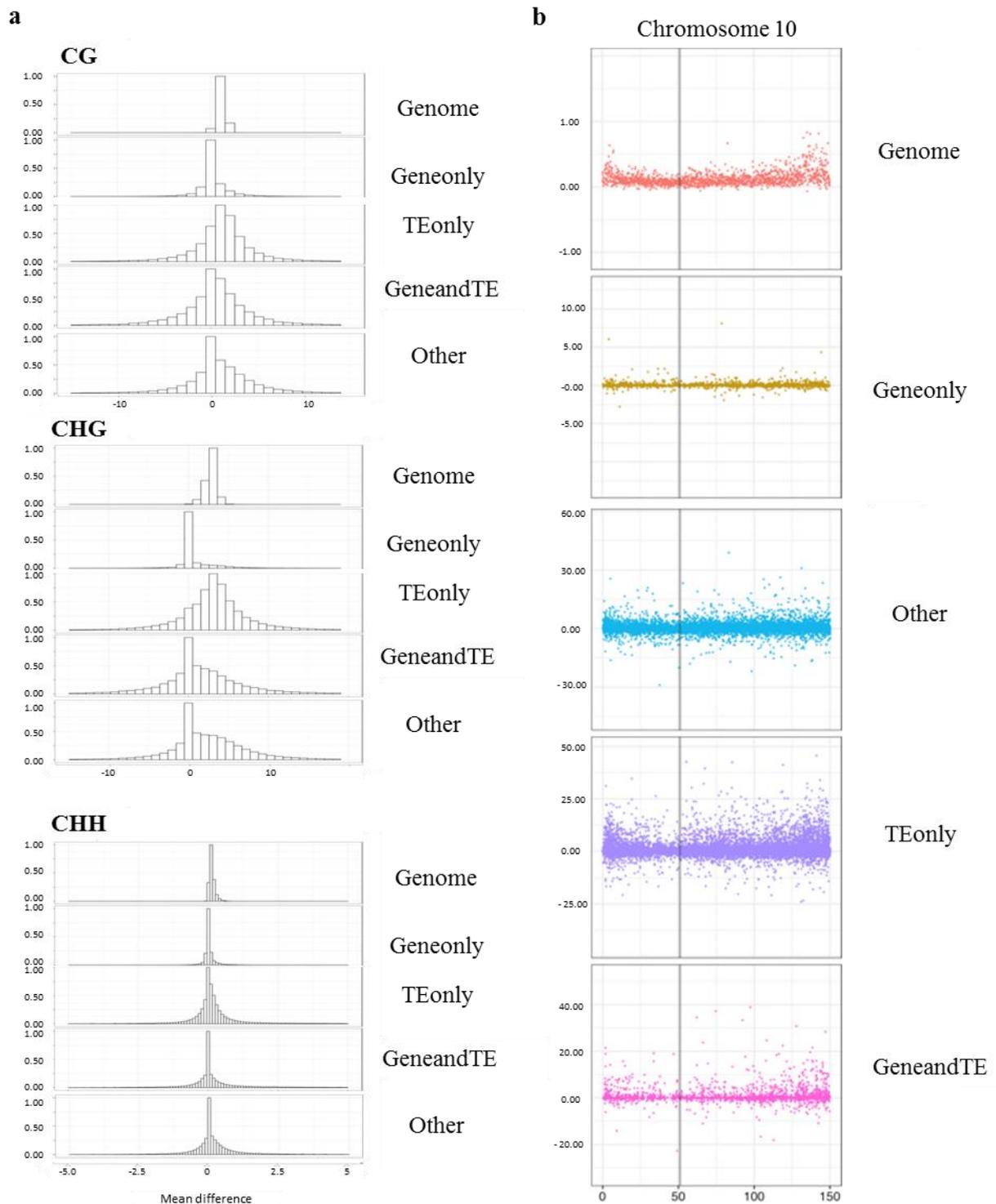


Fig. 2 Average methylation difference between stressed and unstressed plants

a Methylation difference distribution in the different genomic features for the three cytosine contexts CG, CHG and CHH. Genome: 100kb bins, Gene only: genes with TE bases removed, TE only: TEs located outside of genes, TEgene: TEs located inside of genes. Other: Sequence stretches separating genes and TEs. **b** Methylation difference chromosome landscape for the same genomic features for the CHH context, highlighting that the CHH increase observed on chromosome tips is due to changes in TEs whether in genes or not, but does not affect gene sequence or regions with no gene and no TE (other). Chromosome 10 is shown.

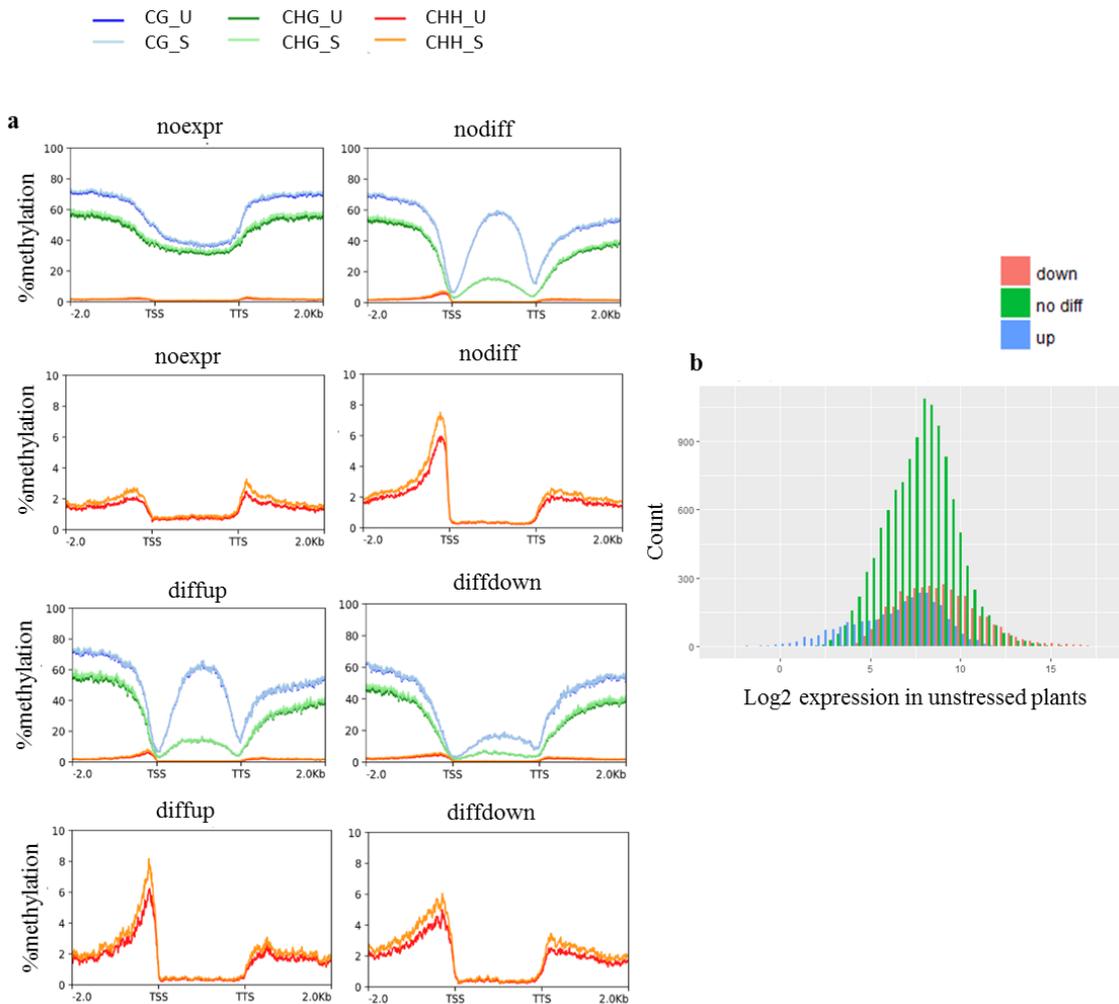


Fig. 3 Profile of average methylation values for stressed and unstressed plants in and around genes, in relation with expression

a Methylation profiles for CG (blue), CHG (green) and CHH (red) for stressed (light colors) and unstressed (dark colors) plants. noexpr: Gene with no expression in both conditions (11,542 genes), nodiff: genes with expression but not differential between the two conditions (11,841 genes), diffup: genes with increased differential expression in stressed plants (3,021 genes), diffdown: genes with decrease differential expression in stressed plants (3937 genes). U: unstressed. S: Stressed. **b** Distribution of expression levels in unstressed plants, showing that down-expressed genes have higher expression in unstressed plants than up-expressed genes, thus highlighting that CHH levels do not follow expression levels in unstressed plants.

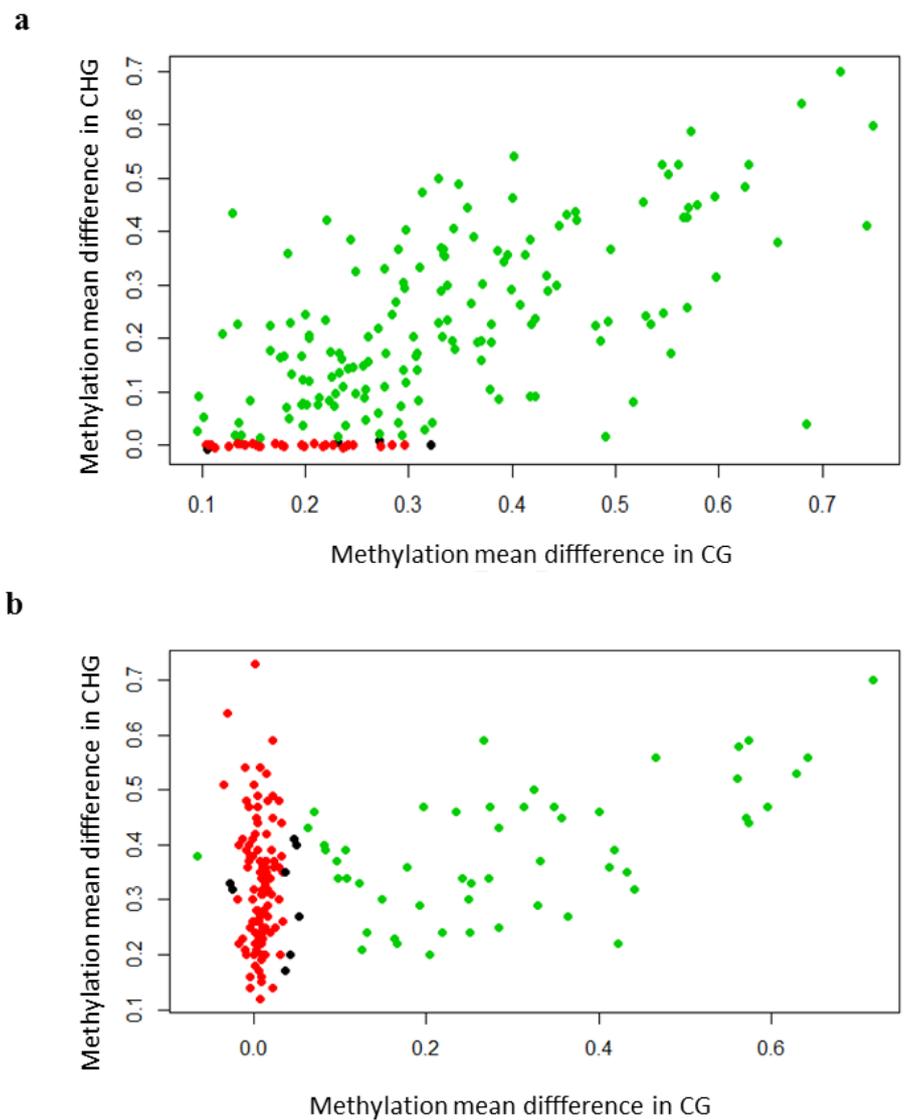
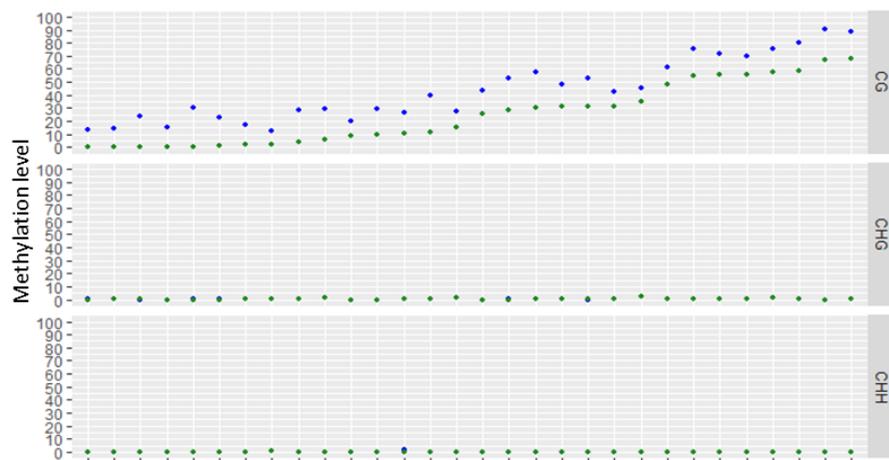


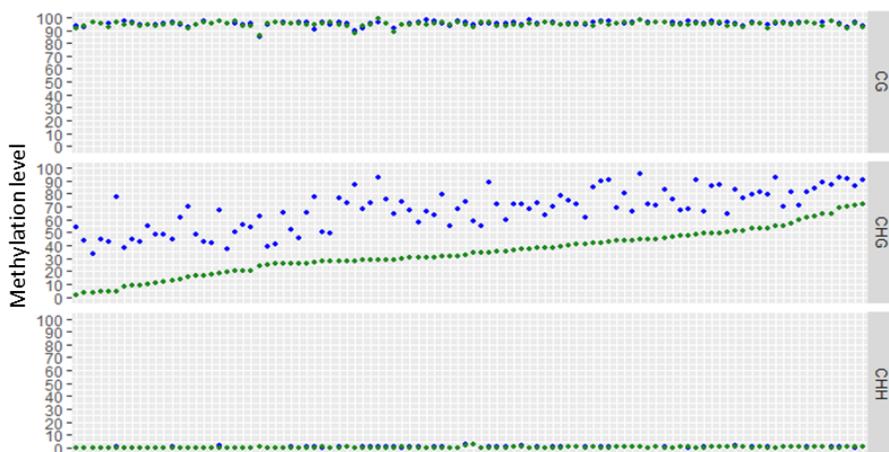
Fig.4 Comparison of CG and CHG methylation difference mean between stressed and unstressed plants for each hypermethylated DMR

For DMRs detected using CG cytosines (**a**) and using CHG cytosines (**b**). Each dot represents one DMR, and shows corresponding average methylation difference in CG context (x-axis) and in CHG context (y-axis). Green dots highlight DMR classified as variable in both contexts. Red dots represent « CG only » (a) or CHG only (b) cases. Black dots show unclassified cases.

a



b



c

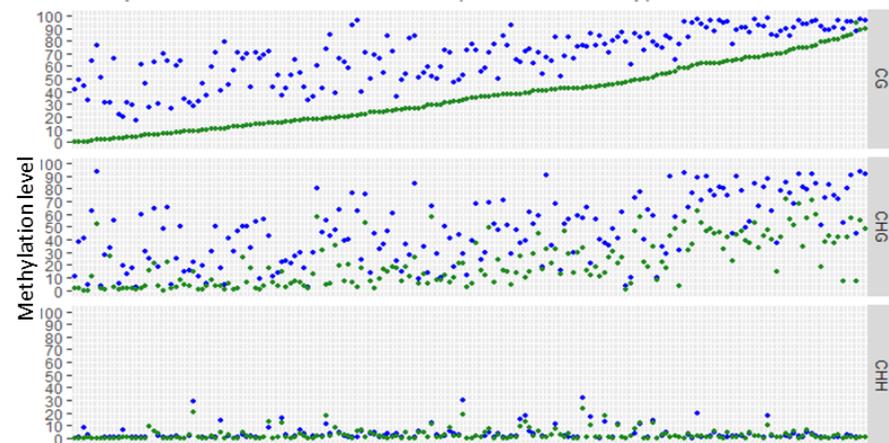


Fig. 5 Methylation levels in unstressed and stressed samples in hypermethylated DMRs
a « CG only » DMRs, **b** « CHG only » DMRs, **c** « CG plus CHG » DMRs. Top: levels for CG methylation. Middle: levels for CHG methylation. Bottom: levels for CHH methylation. Each DMR (x-axis) is represented by two dots, showing average methylation levels in unstressed (green) and stressed (blue) plants.

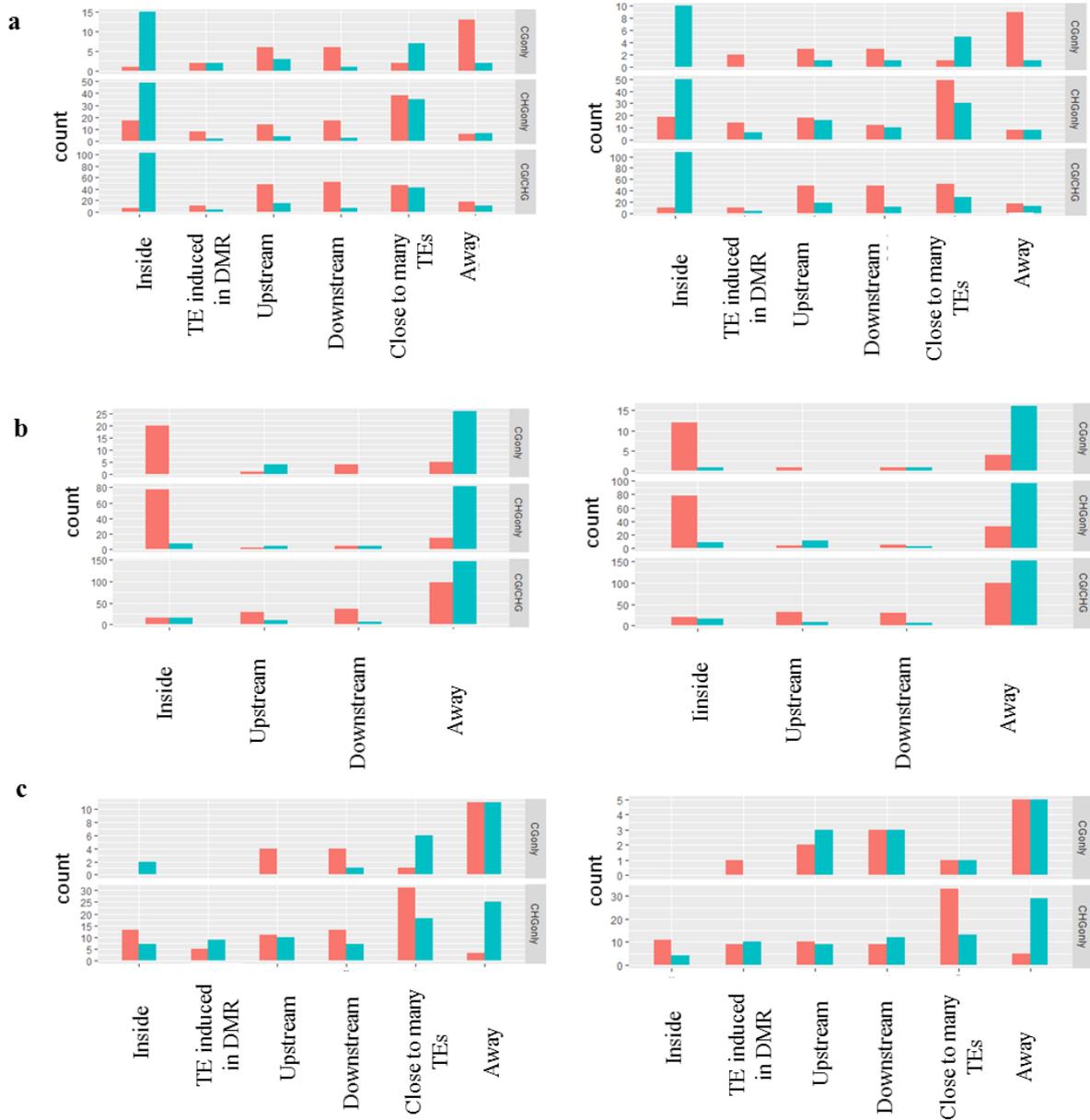


Fig. 6 Annotation of hypermethylated (left) and hypomethylated (right) DMRs
a Distance comparison with TE annotation. **b** Distance comparison with gene annotation. **c** For cases of DMRs located inside genes, distance comparison with TE annotation. Pink: DMRs. Green: random regions with same size and same chromosome as DMRs.

Additional file 1

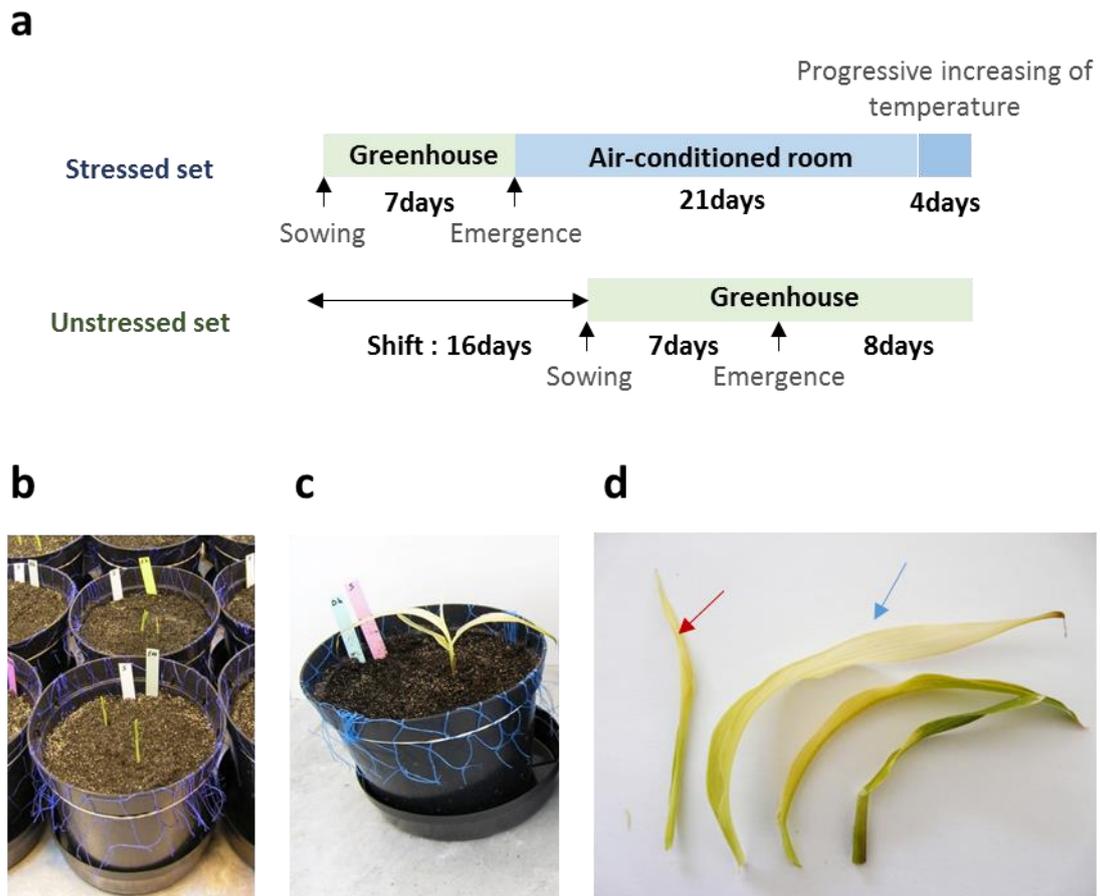
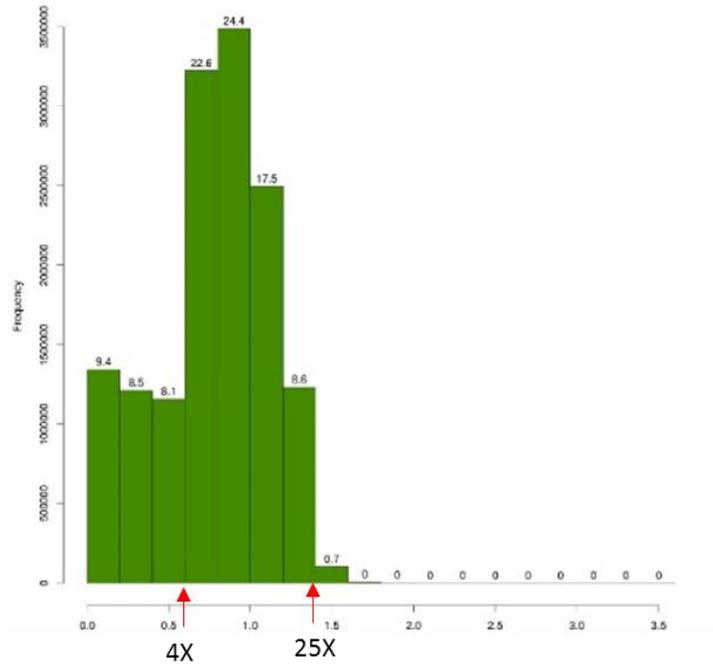


Figure S1

a. Experimental setting. At emergence stage, plants were placed for 21 days either in cold conditions (air conditioned room. 16h day at 14°C ; 8h night at 10°C) or in standard conditions (greenhouse. 16h day at 24°C; 8h night at 18°C) leading to two sets of plants referred to as “stressed” and “unstressed”. **b.** Plants at emergence stage. **c and d.** Plant phenotype after cold exposure showing growth delay and chlorosis. The first leaf (elongated before plants were placed in cold) is almost not affected while all other leaves are chlorotic highlighting the effect of cold on developing leaves only. The last visible leaf (red arrow) of 3 individual plants of each condition was sampled for methylome analysis, the one before visible leaf (blue arrow) was sampled for transcriptome analysis

a



b

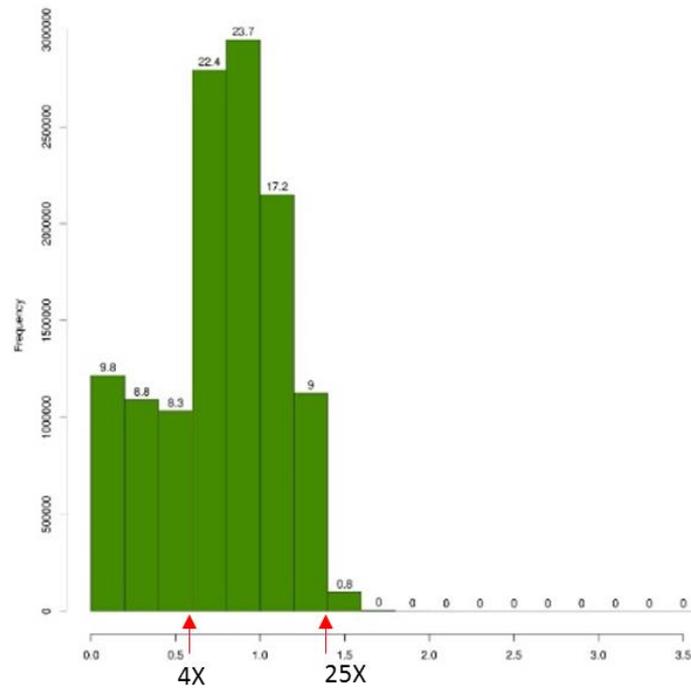


Figure S2
Distribution of cytosine coverage for CG (**a**) and CHG (**b**) contexts

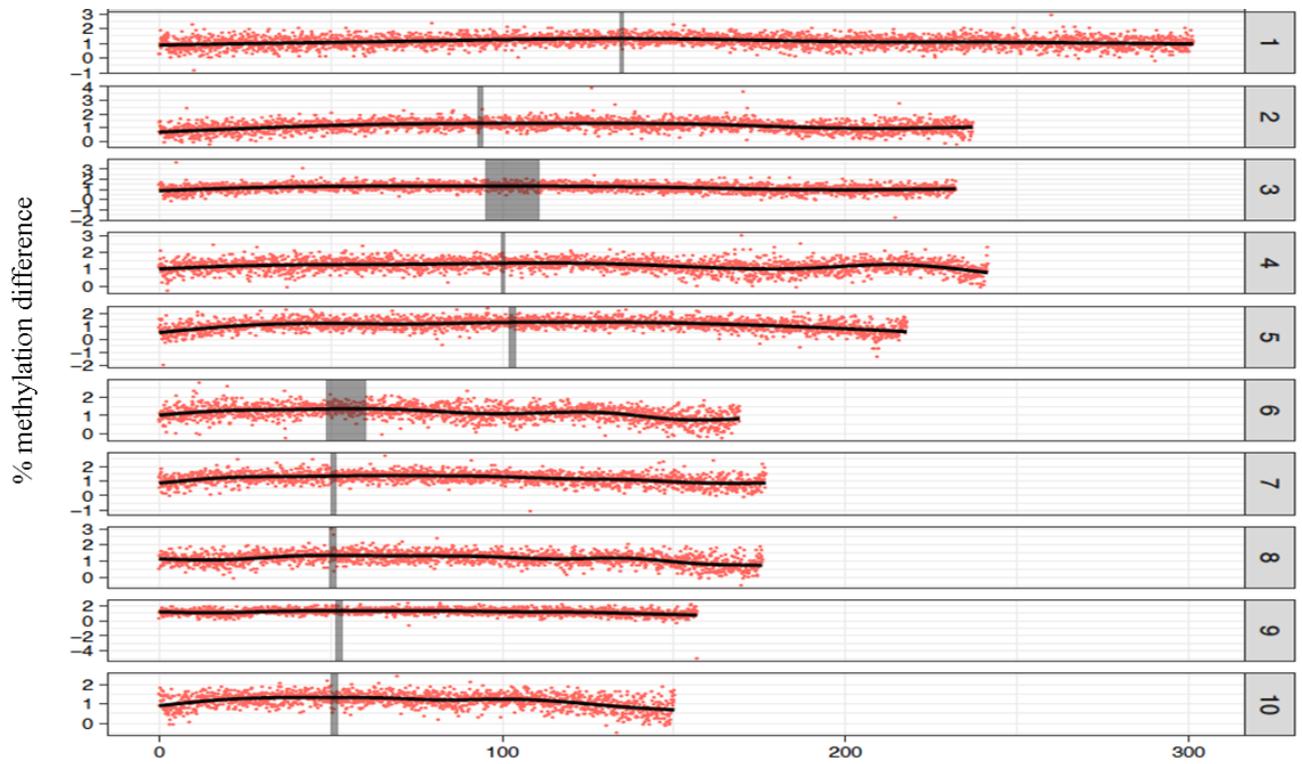


Figure S3
 Mean of methylation difference between stressed and unstressed plants in CG context along the 10 chromosome of maize over 100kb bins

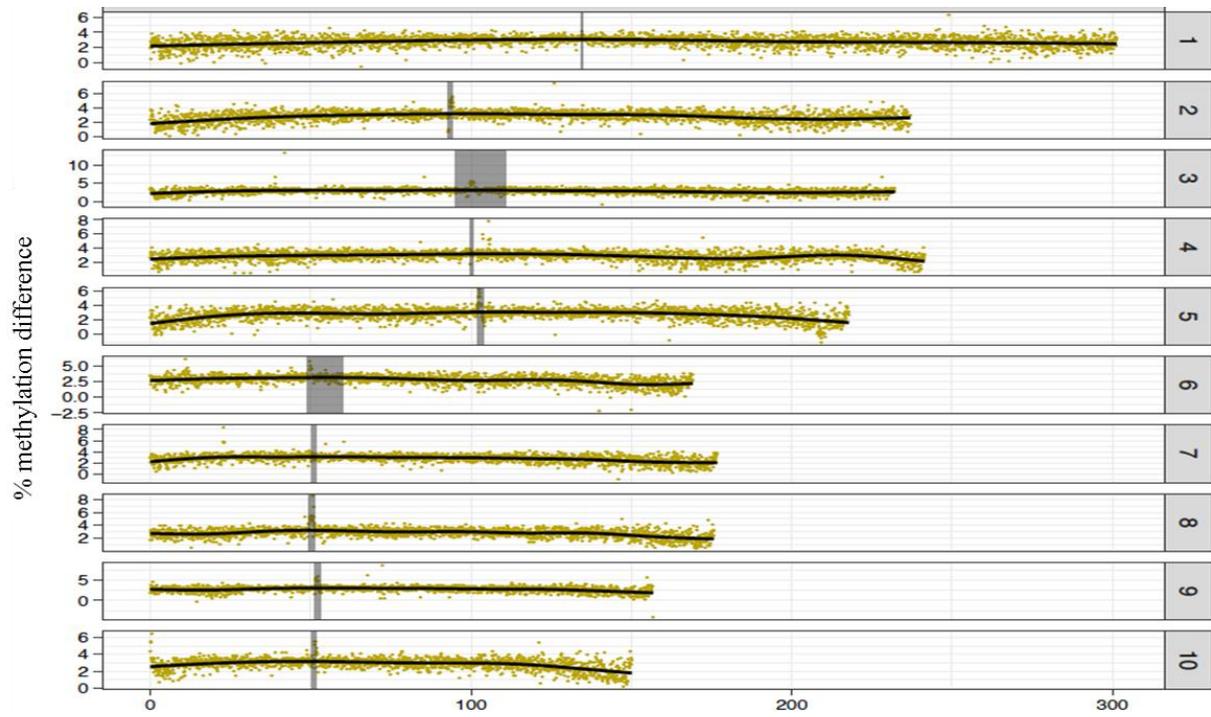


Figure S4
 Mean of methylation difference between stressed and unstressed plants in CHG context along the 10 chromosome of maize over 100kb bins

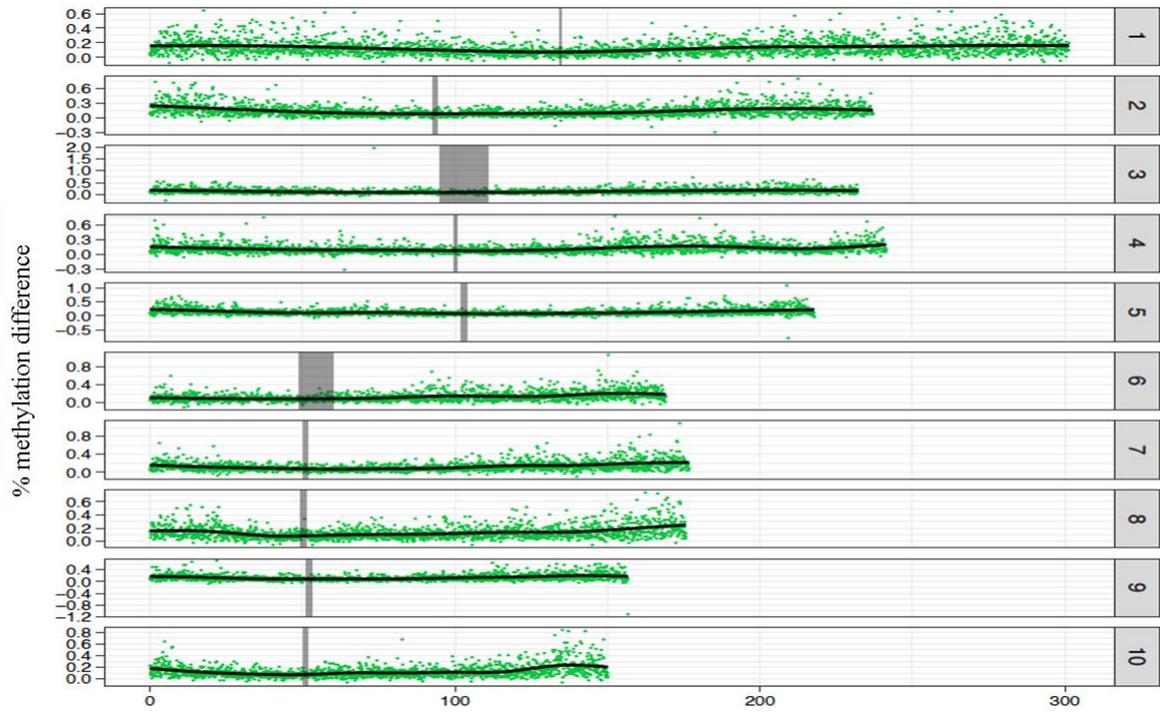


Figure S5
 Mean of methylation difference between stressed and unstressed plants in CHH context along the 10 chromosome of maize over 100kb bins

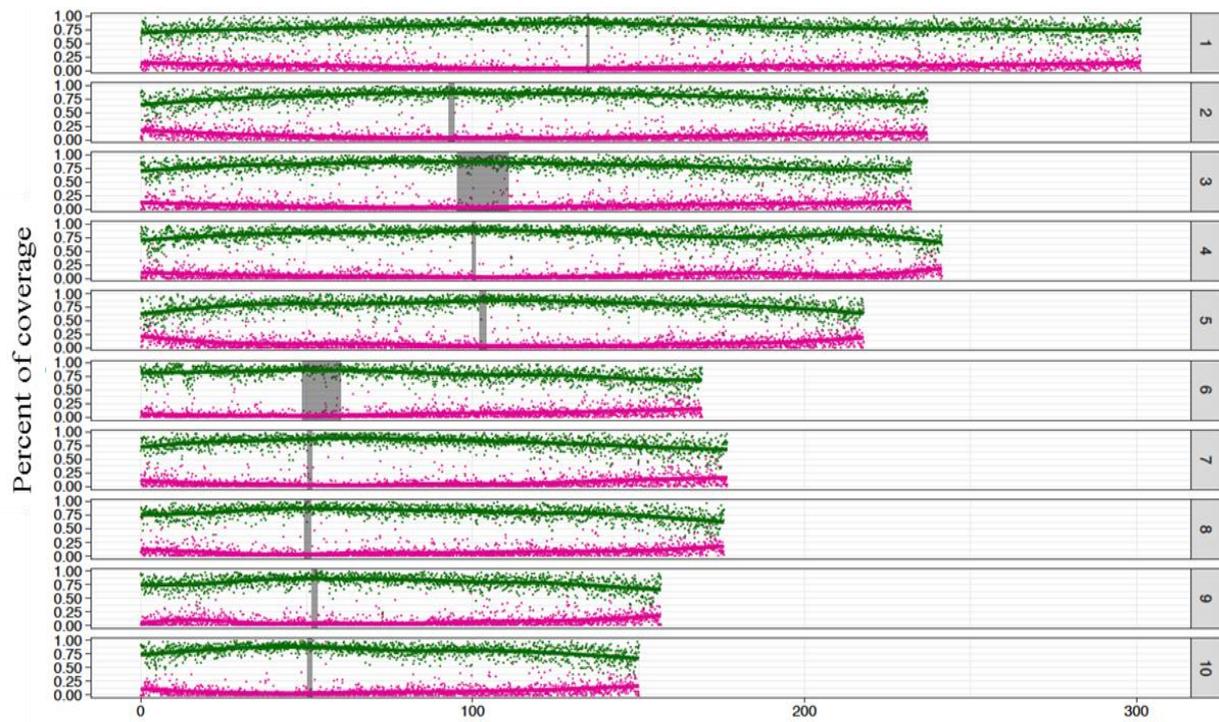


Figure S6
 Percent of coverage of each bin by TE (green, MTEC annotation), and genes (pink, Filtered gene set annotation) in the 10 chromosomes of maize

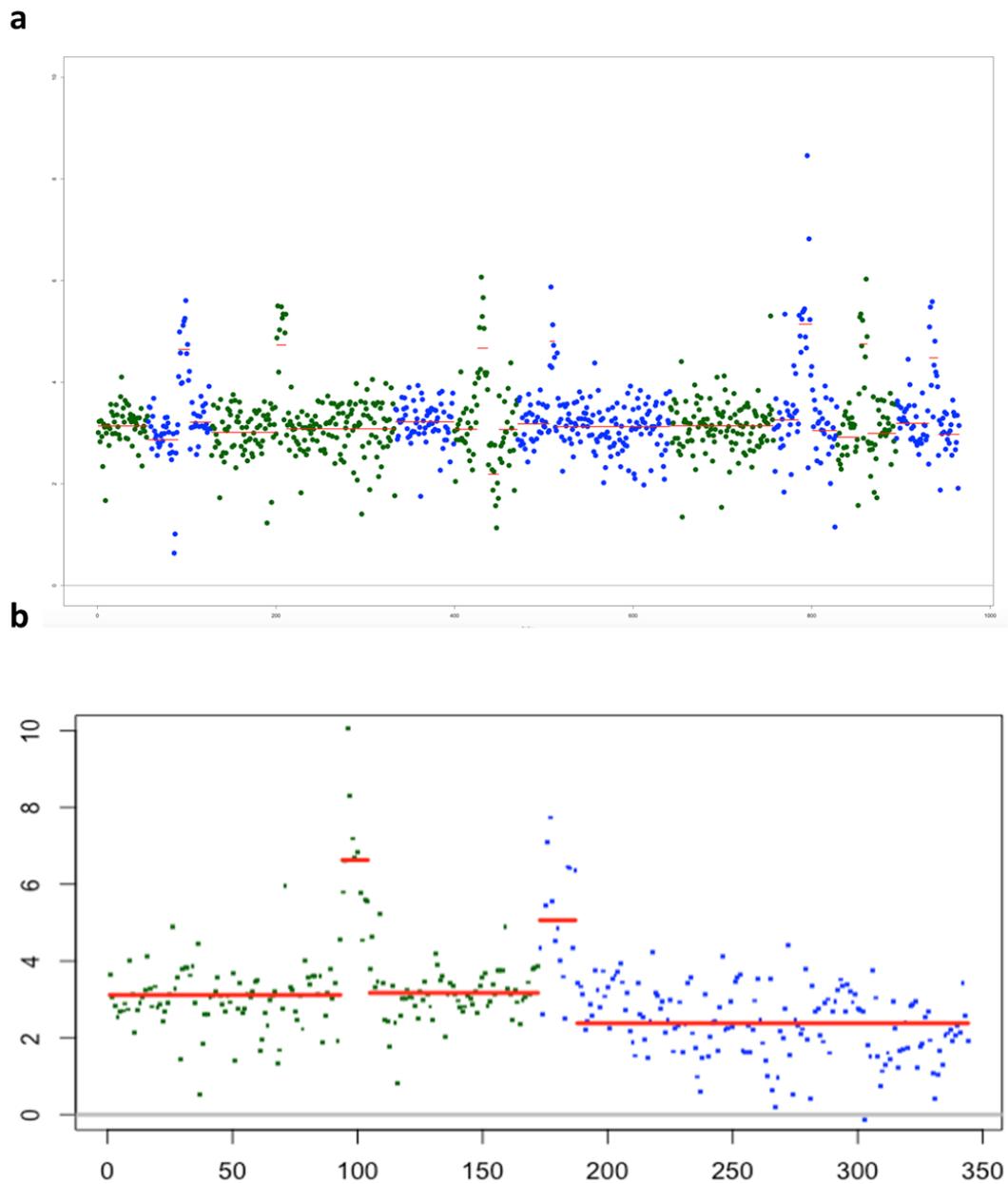


Figure S7

Highlight of CHG methylation level

a CHG methylation peaks in centromeres for each chromosome. Green and blue colors distinguish each chromosome. **b** CHG peaks in specific regions of chromosome 7 and 10. Colors show the succession of analyzed regions using segmentation method.

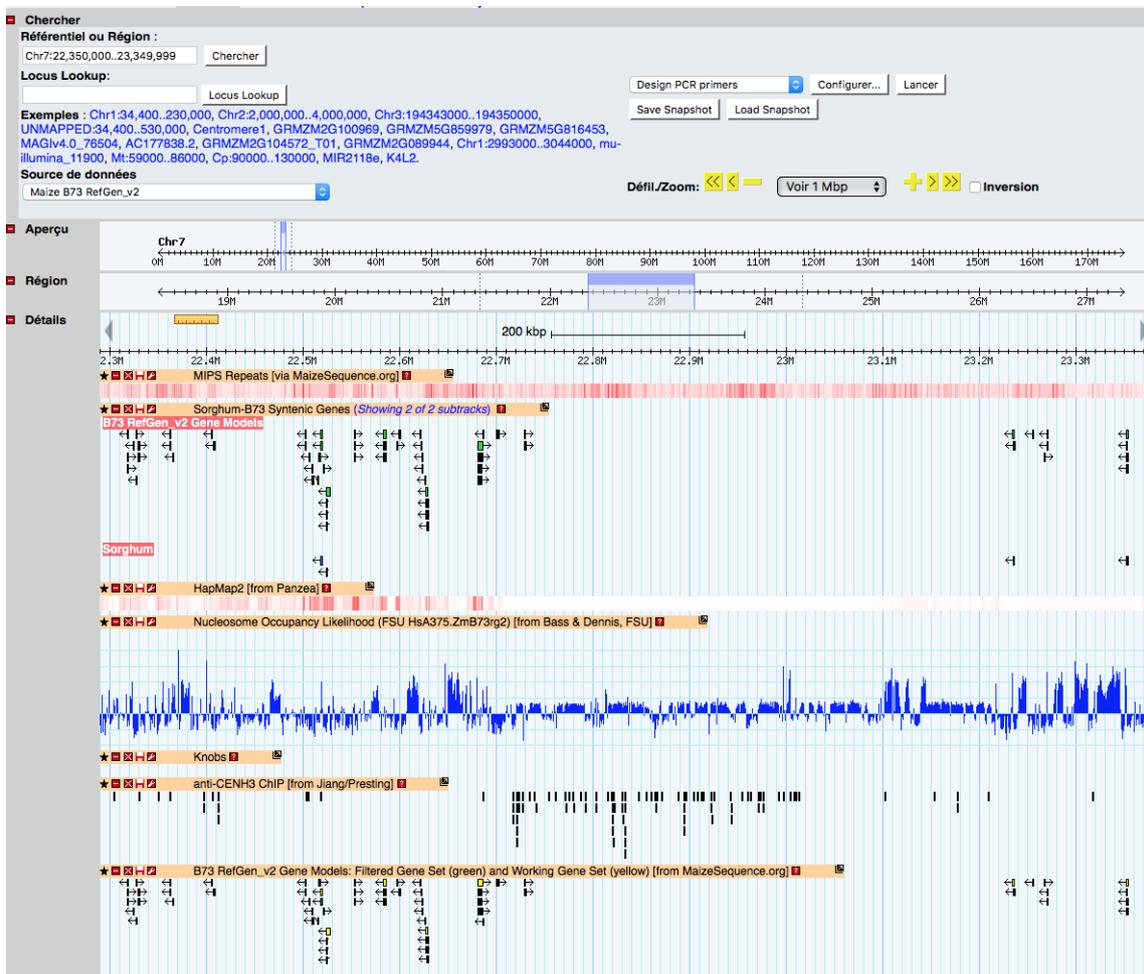


Figure S8

Screenshot of the MaizeGDB website, showing the annotation of knob repeats and enrichment in CENH3 in the region detected with high variation at the tip of chromosome 7.

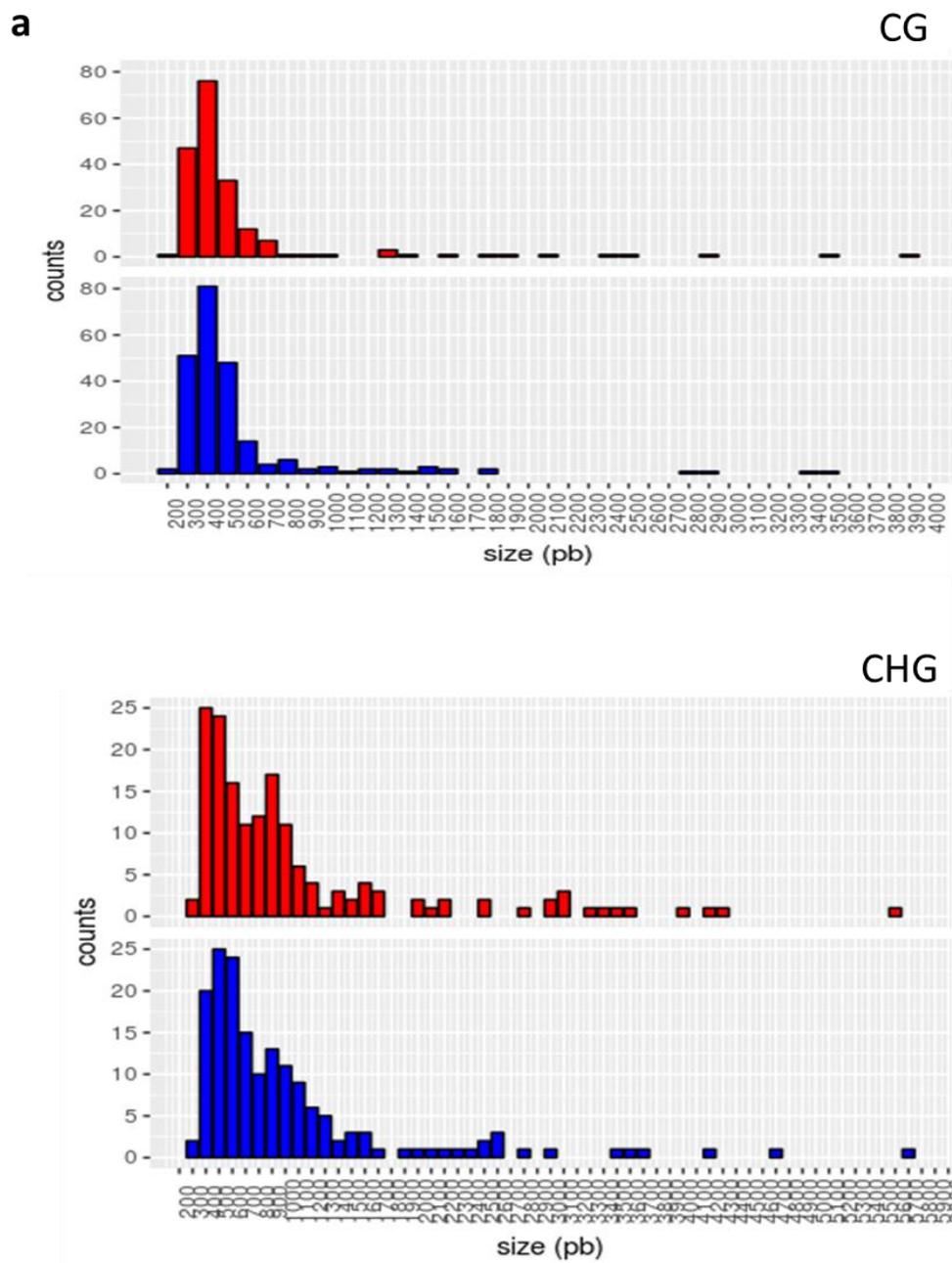


Figure S9

Distribution of DMRs characteristics in hyper (red) and hypomethylation (blue) cases. **a** DMR size. **b** DMR methylation difference between stressed and unstressed plants. **c** Number of cytosines in DMR. **d** Number of DMCs in DMR

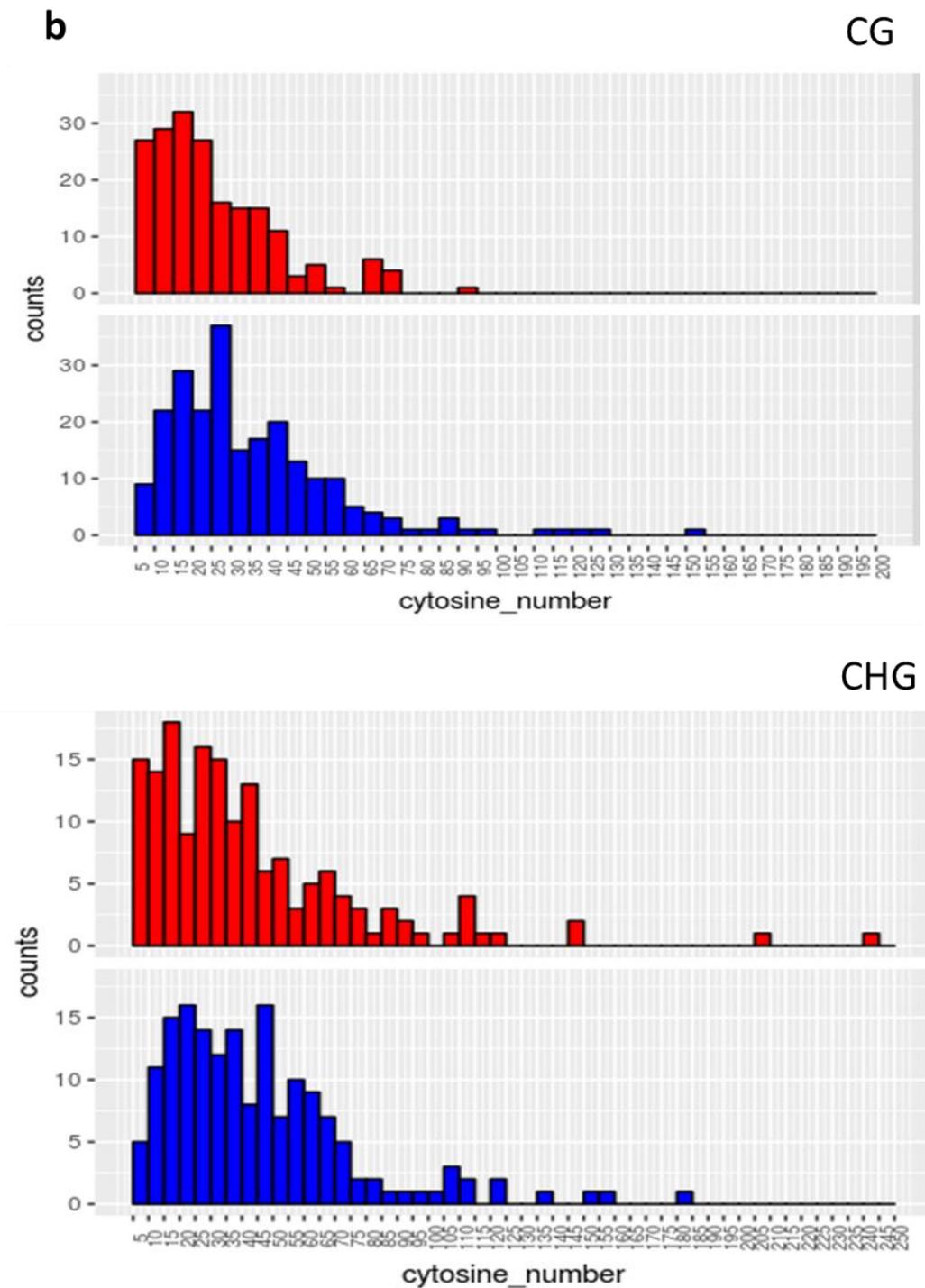


Figure S9 (continued)

Distribution of DMRs characteristics in hyper (red) and hypomethylation (blue) cases. **a** DMR size. **b** DMR methylation difference between stressed and unstressed plants. **c** Number of cytosines in DMR. **d** Number of DMCs in DMR

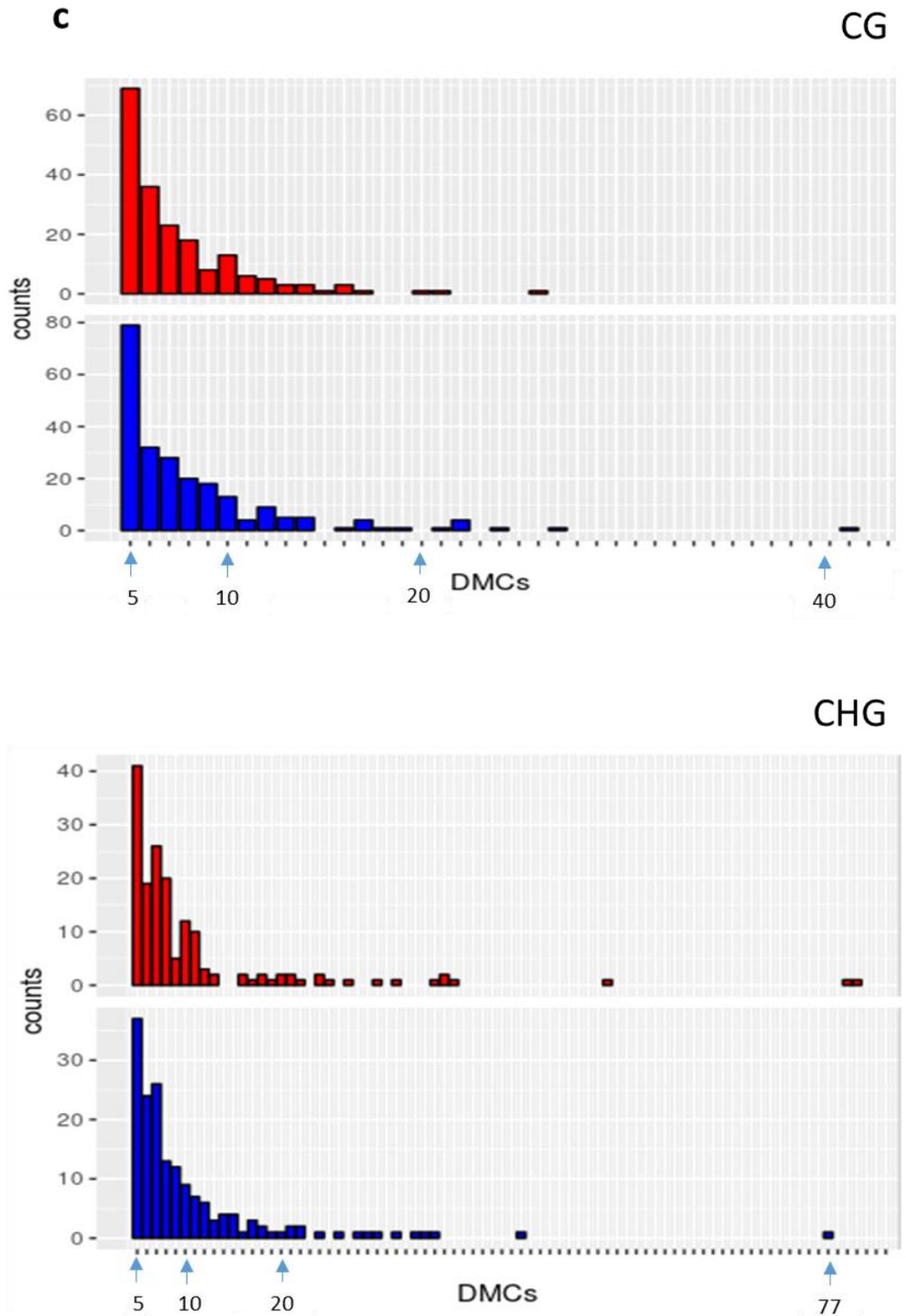


Figure S9 (continued)

Distribution of DMRs characteristics in hyper (red) and hypomethylation (blue) cases. **a** DMR size. **b** DMR methylation difference between stressed and unstressed plants. **c** Number of cytosines in DMR. **d** Number of DMCs in DMR

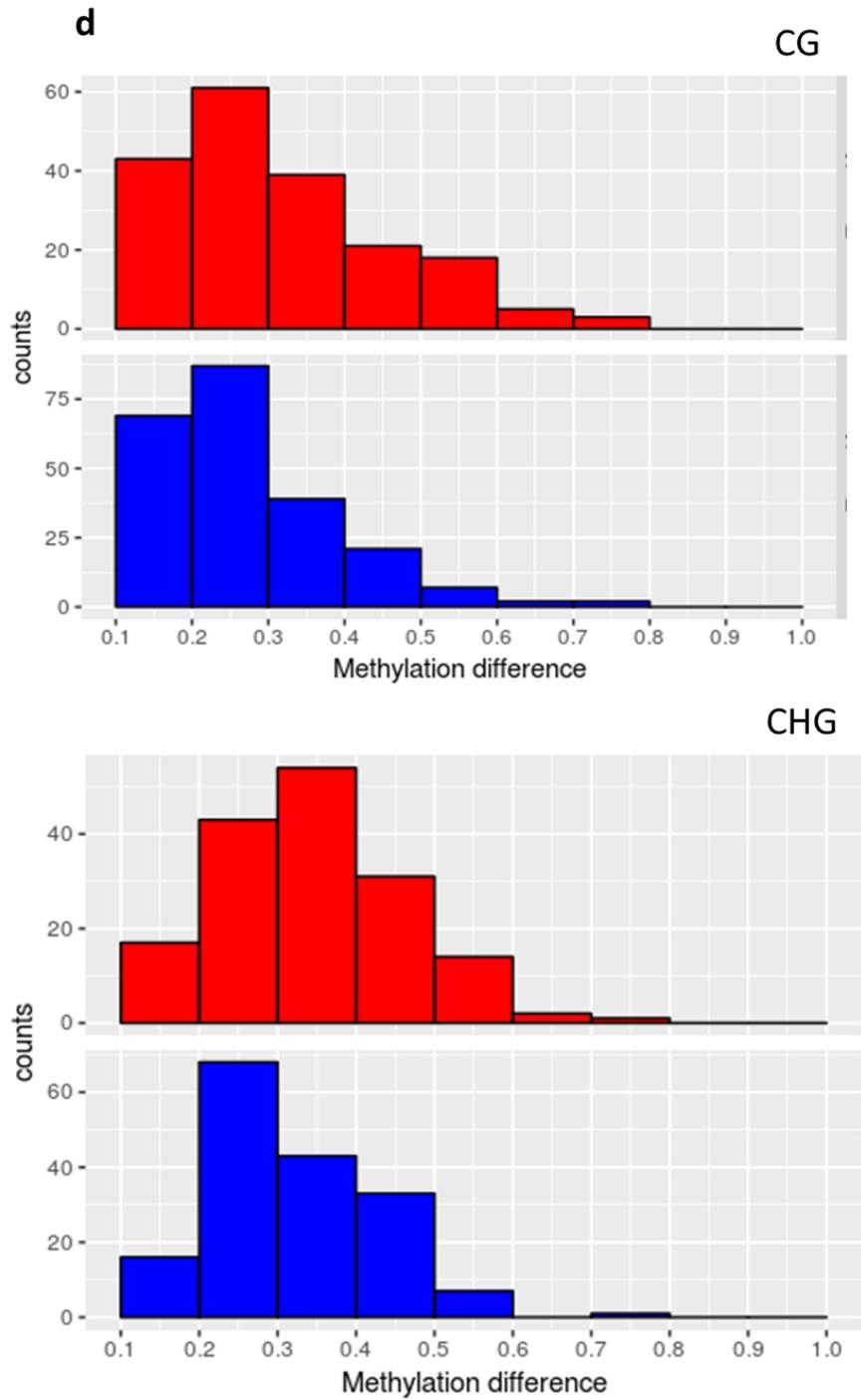


Figure S9 (continued)

Distribution of DMRs characteristics in hyper (red) and hypomethylation (blue) cases. **a** DMR size. **b** DMR methylation difference between stressed and unstressed plants. **c** Number of cytosines in DMR. **d** Number of DMCs in DMR

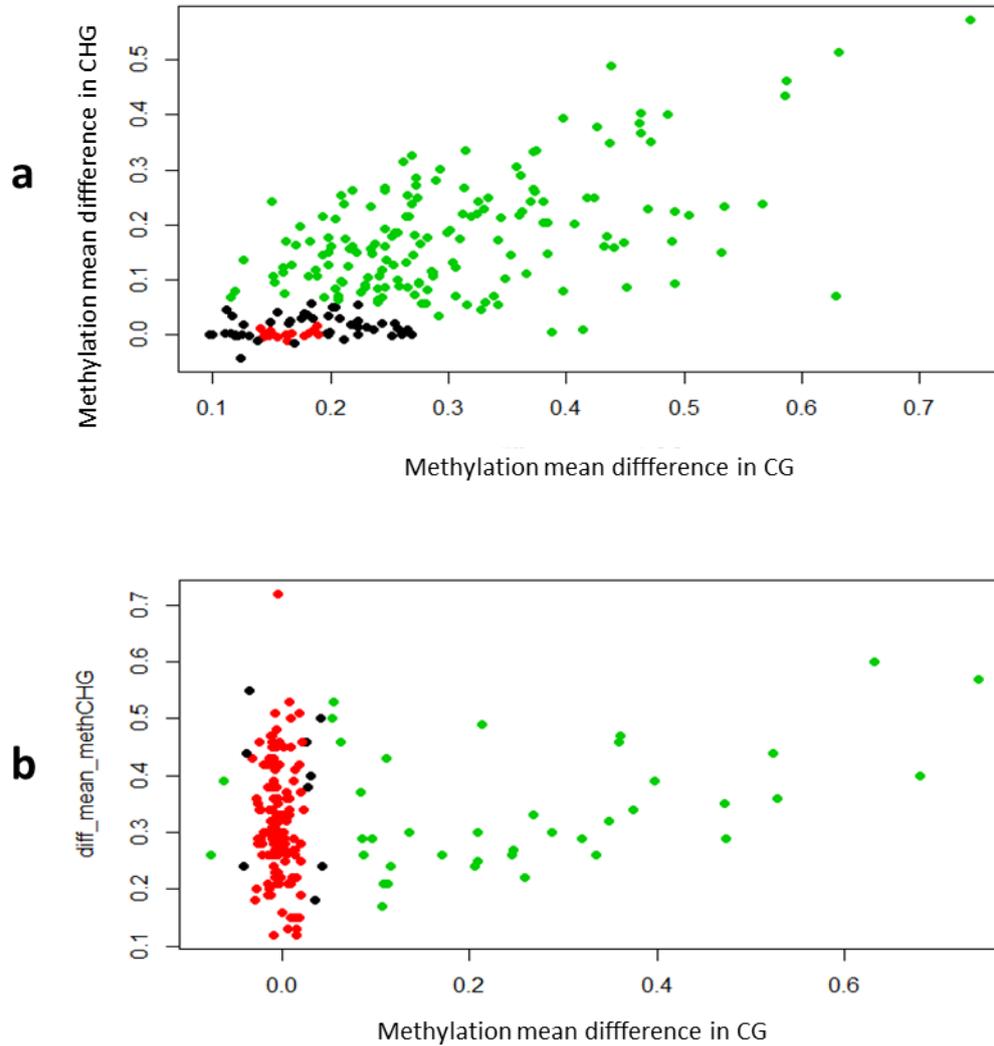


Figure S10

Comparison of CG and CHG methylation difference mean between stressed and unstressed plants for each hypomethylated DMR

For DMRs detected using CG cytosines (a) and using CHG cytosines (b). Each dot represents one DMR, and shows corresponding average methylation difference in CG context (x-axis) and in CHG context (y-axis). Green dots highlight DMR classified as variable in both contexts. Red dots represent « CG only » (a) or CHG only (b) cases. Black dots show unclassified cases.

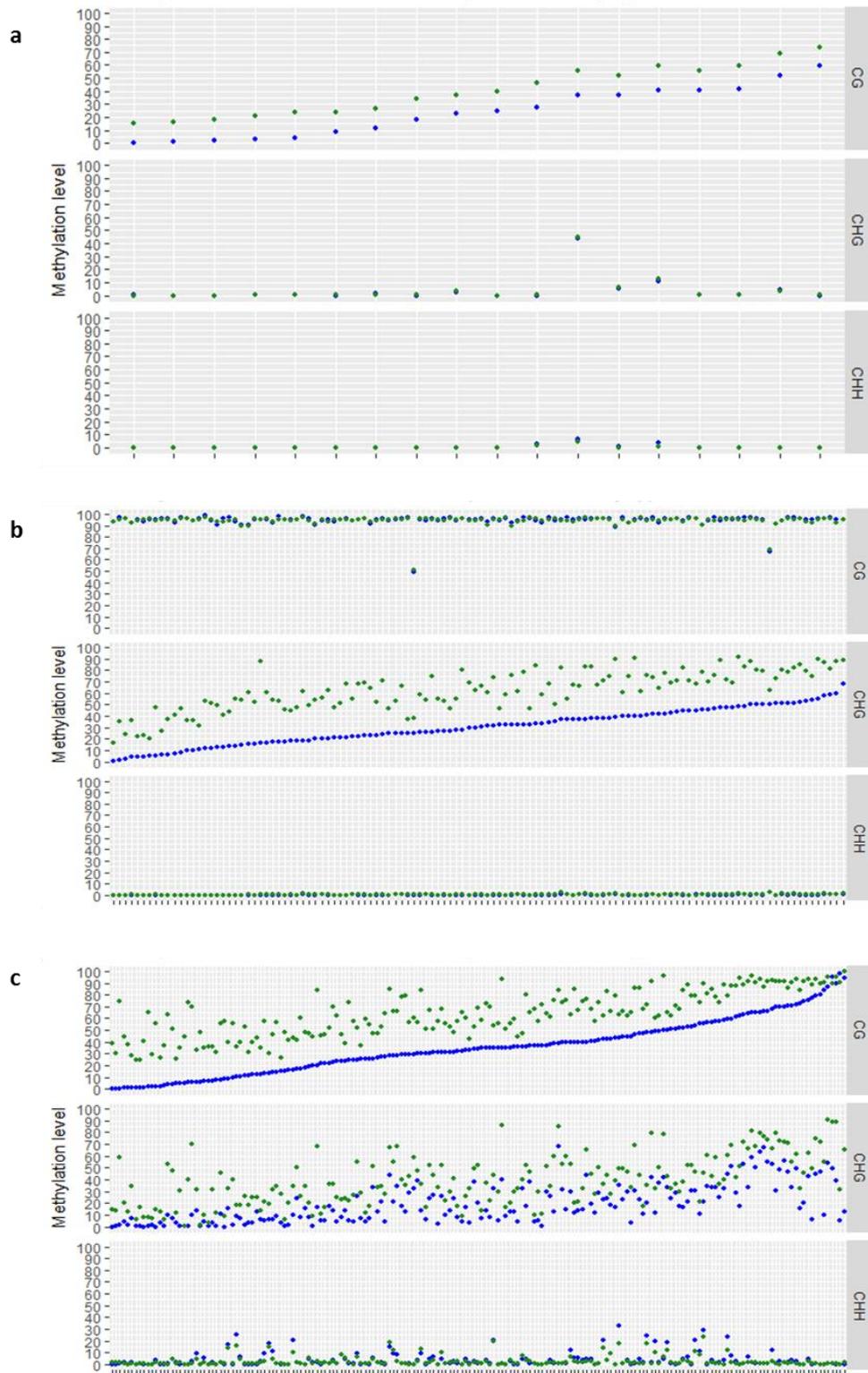


Figure S11

Methylation levels in unstressed and stressed samples in hypomethylated DMRs

a « CG only » DMRs, **b** « CHG only » DMRs, **c** « CG plus CHG » DMRs. Top: levels for CG methylation. Middle: levels for CHG methylation. Bottom: levels for CHH methylation. Each DMR (x-axis) is represented by two dots, showing average methylation levels in unstressed (green) and stressed (blue) plants

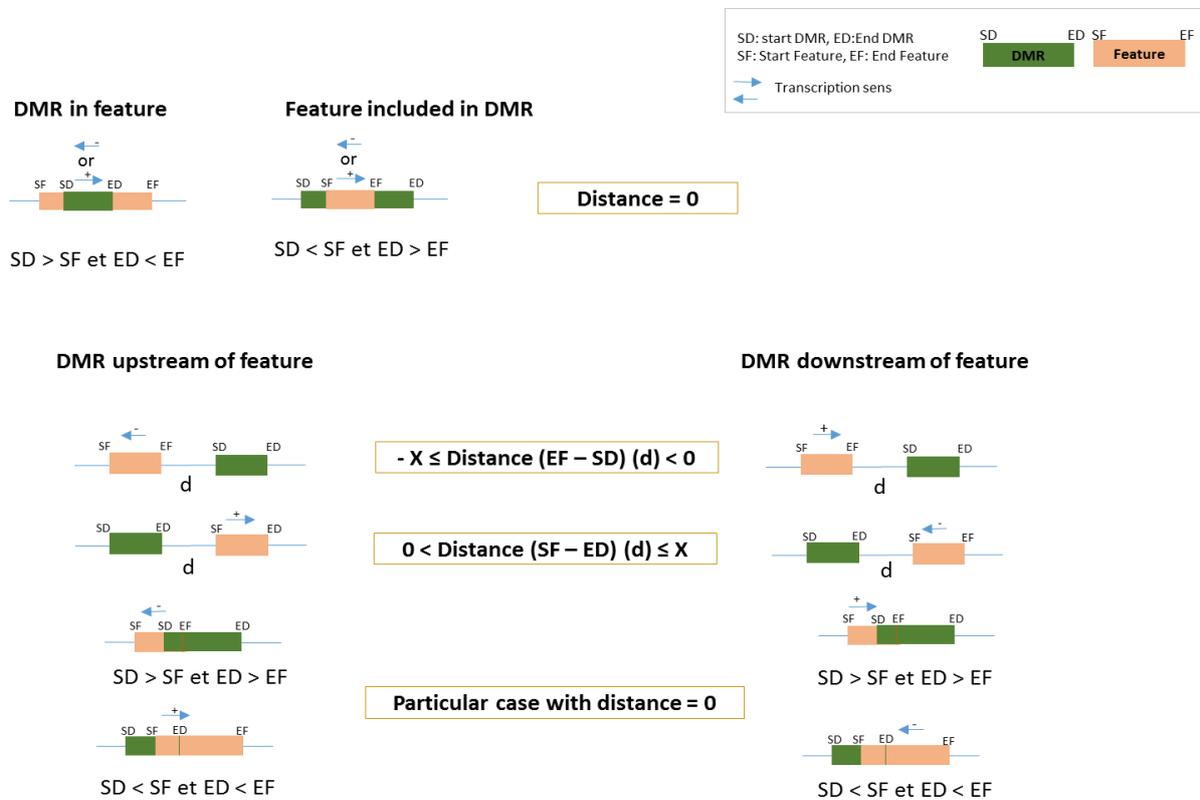


Figure S12

Diagram showing the different annotation of DMRs with features

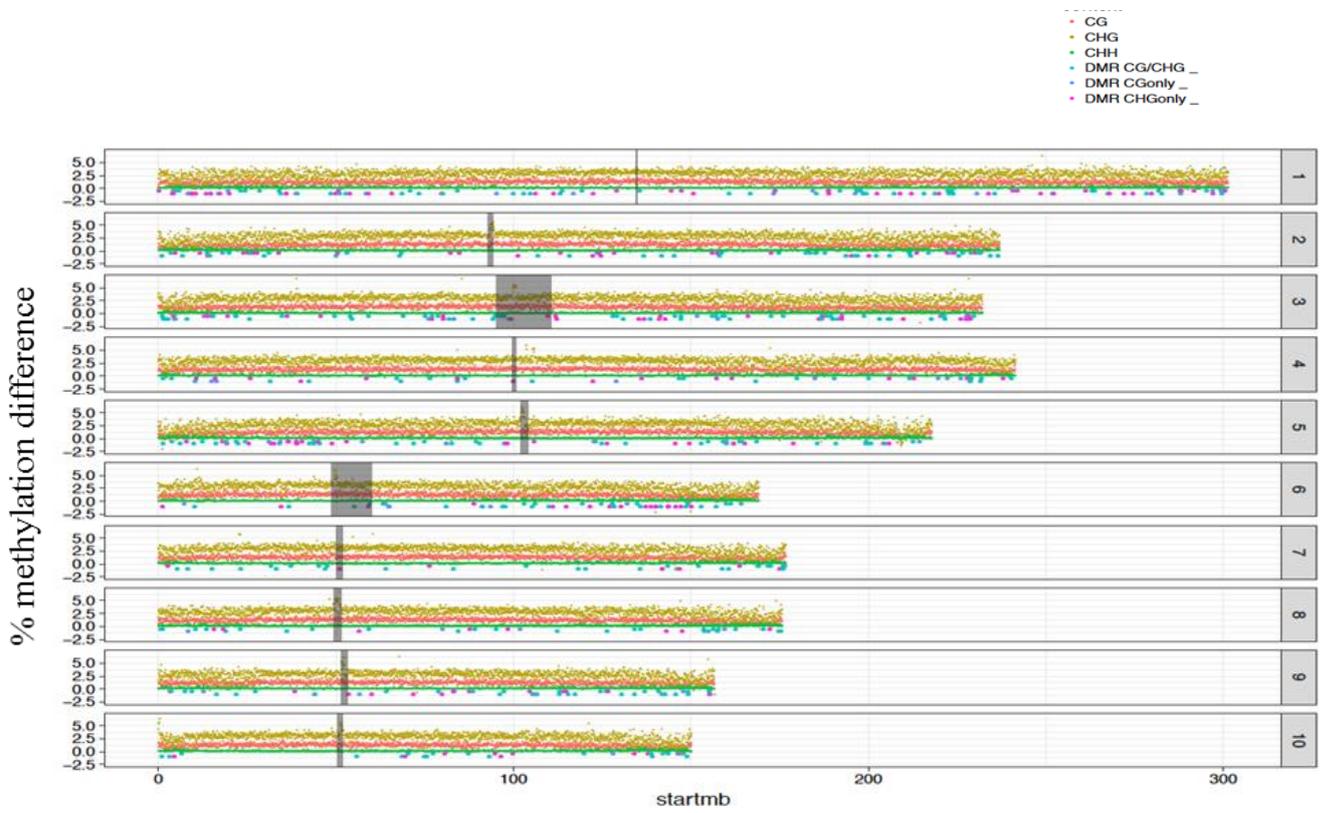


Figure S13
Localisation of detected DMR along chromosomes

Additional file 2

Table S1

Sequencing and mapping data for the 6 replicates of B73 (U:Unstressed. S:Stressed)

Sample	Number of generated reads	Sequenced Gb	Genome size	Theoretical coverage depth	Number of uniquely mapped reads	Percent of uniquely mapped reads	Reads mapped as pairs	Number of reads mapped as single reads	Percent of reads in pairs among uniquely mapped reads	Percent of reads in single reads among uniquely mapped reads
B73_U_rep1	497 596 575	49.76	2.15	23.1	331 326 243	66.6	325 307 528	6 018 715	98.2	1.8
B73_U_rep2	465 665 603	46.57	2.15	21.7	304 503 290	65.4	297 953 498	6 549 792	97.8	2.2
B73_U_rep3	508 161 305	50.82	2.15	23.6	341 923 410	67.3	336 300 618	5 622 792	98.4	1.6
B73_S_rep1	452 175 360	45.22	2.15	21.0	304 345 919	67.3	297 861 476	6 484 443	97.9	2.1
B73_S_rep3 (30S)	460 223 502	46.02	2.15	21.4	309 038 356	67.1	302 542 512	6 495 844	97.9	2.1
B73_S_rep3 (34S)	520 783 300	52.08	2.15	24.2	372 880 268	71.6	366 863 962	6 016 306	98.4	1.6

Table S2

Number of analyzed cytosines and percent of methylation level in each replicate and context

Sample	Number of cytosines with 4x and 25x and common to the 6 replicates in each genotype			Number of methylated cytosines			Percent of methylated cytosines among analyzed common cytosines between 4X-25x			Mean percent of methylation among analyzed common cytosines between 4x-25x		
	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH
B73_U_rep1				44 688 548	29 258 868	10 542 362	82.08	65.82	7.10	82.08	68.98	1.17
B73_U_rep2				43 018 217	27 957 895	9 489 133	79.01	62.89	6.39	80.93	67.00	0.98
B73_U_rep3				42 620 101	28 137 326	9 790 534	78.28	63.30	6.60	80.78	67.23	0.98
B73_S_rep1	54 444 473	44 451 905	148 406 979	44 700 947	30 349 943	10 682 882	82.10	68.28	7.20	82.79	71.01	1.20
B73_S_rep3 (30S)				44 919 371	30 479 139	10 673 699	82.50	68.57	7.19	82.94	71.26	1.19
B73_S_rep3 (34S)				44 012 376	29 482 710	10 621 782	80.84	66.32	7.16	81.50	69.35	1.09

Table S3

Number and percentage of methylation level in each replicate and context in cytosines overlapping with gene only (**a**), TE only (**b**), gene and TE (**c**), non gene no TE (other, **d**)

a

Sample	Number of cytosines in genes only between 4x-25x and common to the 6 replicates in each genotype			Number of methylated cytosines in genes only between 4x-25x and common to the 6 replicates in each genotype			Percent of methylated cytosines among all cytosines in genes only			Mean percent of methylation in genes only		
	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH
B73_U_rep1	4 771 011	4 294 487	10 282 322	951 664	441 389	210 285	19.95	10.28	2.05	20.50	11.05	0.33
B73_U_rep2				932 956	426 182	192 467	19.55	9.92	1.87	20.26	10.75	0.27
B73_U_rep3				926 950	425 325	194 374	19.43	9.90	1.89	20.17	10.69	0.26
B73_S_rep1				950 030	451 397	218 390	19.91	10.51	2.12	20.50	11.21	0.34
B73_S_rep3 (30S)				954 927	452 154	213 325	20.02	10.53	2.07	20.55	11.23	0.33
B73_S_rep3 (34S)				938 127	441 243	203 653	19.66	10.27	1.98	20.22	10.98	0.30

b

Sample	Number of cytosines in TEs only between 4x-25x and common to the 6 replicates in each genotype			Number of methylated cytosines in TEs only between 4x-25x and common to the 6 replicates in each genotype			Percent of methylated cytosines among all cytosines in TEs only			Mean percent of methylation in TE only		
	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH
B73_U_rep1	40 114 900	32 710 356	113 611 825	37 380 572	25 079 801	8 627 238	93.18	76.67	7.59	93.06	80.21	1.23
B73_U_rep2				35 934 906	23 949 295	7 752 961	89.58	73.22	6.82	91.73	77.90	1.03
B73_U_rep3				35 591 631	24 116 874	7 990 824	88.72	73.73	7.03	91.57	78.21	1.04
B73_S_rep1				37 388 134	26 032 137	8 667 720	93.20	79.58	7.63	93.90	82.63	1.27
B73_S_rep3 (30S)				37 573 934	26 145 342	8 694 044	93.67	79.93	7.65	94.07	82.93	1.25
B73_S_rep3 (34S)				36 803 418	25 276 586	8 694 901	91.75	77.27	7.65	92.41	80.68	1.15

Table S3 (continued)

Number and percentage of methylation level in each replicate and context in cytosines overlapping with gene only (**a**), TE only (**b**), gene and TE (**c**), non gene no TE (other, **d**)

c

Sample	Number of cytosines in genes and TE between 4x-25x and common to the 6 replicates in each genotype			Number of methylated cytosines in genes and TE between 4x-25x and common to the 6 replicates in each genotype			Percent of methylated cytosines among all cytosines in genes and TE			Mean percent of methylation in genes and TE		
	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH
B73_U_rep1	1 086 888	882 682	2 867 186	1 008 042	583 365	239 763	92.75	66.09	8.36	92.15	69.42	1.62
B73_U_rep2				979 568	560 081	214 706	90.13	63.45	7.49	90.98	67.41	1.34
B73_U_rep3				974 127	559 286	216 980	89.63	63.36	7.57	90.92	67.35	1.30
B73_S_rep1				1 000 457	591 970	243 539	92.05	67.06	8.49	92.41	70.44	1.73
B73_S_rep3 (30S)				1 004 403	593 601	239 506	92.41	67.25	8.35	92.61	70.64	1.68
B73_S_rep3 (34S)				993 459	575 673	234 902	91.40	65.22	8.19	91.41	68.79	1.50

d

Sample	Number of cytosines in other features ("no genes no TEs") between 4x-25x and common to the 6 replicates in each genotype			Number of methylated cytosines in other features ("no genes no TEs") between 4x-25x and common to the 6 replicates in each genotype			Percent of methylated cytosines among all cytosines in other features ("no genes no TEs")			Mean percent of methylation in other features ("no genes no TEs")		
	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH
B73_U_rep1	8 471 674	6 564 380	21 645 646	5 348 270	3 154 313	1 465 076	63.13	48.05	6.77	63.49	50.84	1.18
B73_U_rep2				5 170 787	3 022 337	1 328 999	61.04	46.04	6.14	62.65	49.41	0.97
B73_U_rep3				5 127 393	3 035 841	1 388 356	60.52	46.25	6.41	62.49	49.48	0.95
B73_S_rep1				5 362 326	3 274 439	1 553 233	63.30	49.88	7.18	64.01	52.30	1.26
B73_S_rep3 (30S)				5 386 107	3 288 042	1 526 824	63.58	50.09	7.05	64.13	52.49	1.21
B73_S_rep3 (34S)				5 277 372	3 189 208	1 488 326	62.29	48.58	6.88	63.05	51.18	1.10

Table S4

Number and percentage of methylation level in each replicate and context in cytosines of each transposon superfamilies and LINE families when **(a)** TEonly, and **(b)** gene and TE

a

TE	Context	Number of cytosines per superfamily in TEonly	Number of methylated cytosines in TEonly						Percent of methylated cytosines among all cytosines per superfamily in TEonly		Methylation percentage in TEonly	
			Unstressed			Stressed			Mean for unstressed plants	Mean for stressed plants	Mean for unstressed plants	Mean for stressed plants
			rep1	rep2	rep3	rep1	rep3 (30S)	rep3 (34S)				
DTA	CG	512 460	459 751	443 566	441 938	455970	458 753	451 289	87.50	88.85	90.13	91.02
	CHG	327 912	191 490	178 003	180 474	202576	203 733	193 687	55.91	60.99	66.65	70.01
	CHH	1 072 678	231 121	201 378	205 412	248215	238 228	225 196	19.82	22.11	4.16	5.33
DTC	CG	1 221 824	1 126 284	1 086 205	1 082 373	1 131 051	1 135 776	1 113 182	89.89	92.21	91.63	92.96
	CHG	854 874	595 097	564 764	570 225	620 459	622 983	602 196	67.46	71.97	74.35	77.65
	CHH	3 212 913	373 603	326 549	342 212	393 290	384 344	371 798	10.81	11.93	1.86	2.23
DTH	CG	153 626	138 293	134 647	134 300	138 863	139 717	137 129	88.36	90.20	89.94	90.97
	CHG	131 403	100 339	96 615	96 876	103 205	103 698	101 549	74.54	78.25	77.32	80.33
	CHH	616 786	122 017	112 191	116 920	134 695	130 772	125 181	18.98	21.11	4.83	5.93
DTM	CG	429 801	393 228	380 214	379 649	392 814	394 876	388 276	89.43	91.20	90.86	91.99
	CHG	282 266	207 455	199 304	200 968	215 652	216 557	210 857	71.77	75.94	75.89	79.09
	CHH	1 084 248	150 153	13 4977	140 981	156 753	153 882	149 130	13.10	14.13	2.83	3.33
DTT	CG	16 871	14 322	13 989	13 766	14 341	14 436	14 222	83.13	84.96	86.74	87.64
	CHG	11 361	7 643	7 097	7 105	8 001	8 057	7 792	64.09	69.98	69.94	74.00
	CHH	49 394	21 049	19 798	20 623	22 938	22 226	21 646	41.48	45.09	14.58	18.00
RIL	CG	12 013	11 099	10 766	10 799	11 104	11 175	10 976	90.64	92.28	91.56	92.56
	CHG	16 792	12 266	11 826	11 871	12 535	12 573	12 268	71.39	74.19	72.96	75.43
	CHH	62 046	4 109	3 892	4 203	4 287	4 414	4 262	6.56	6.96	0.93	1.00
RIX	CG	97 392	87 748	85 587	85 299	88 589	89 216	87 052	88.52	90.65	89.97	91.10
	CHG	146 028	105 592	10 2376	10 2620	108 188	108 798	106 775	70.90	73.90	72.21	74.71
	CHH	569 143	41 133	3 9496	41 878	46 057	45 295	43814	7.17	7.92	1.13	1.26

Table S4 (continued)

Number and percentage of methylation level in each replicate and context in cytosines of each transposon superfamilies and LINE families when (a) TEonly. and (b) gene and TE

b

TE	Context	Number of cytosines per superfamily in gene and TE	Number of methylated cytosines in gene and TE						Percent of methylated cytosines among all cytosines per superfamily in gene and TE		Methylation percentage in gene and TE	
			Unstressed			Stressed			Mean for unstressed plants	Mean for stressed plants	Mean for unstressed plants	Mean for stressed plants
			rep1	rep2	rep3	rep1	rep3 (30S)	rep3 (34S)				
DTA	CG	64483	57965	56641	56516	57417	57575	56888	88.46	88.85	90.31	90.83
	CHG	37173	14880	13989	14093	15526	15584	14762	38.52	41.13	47.69	49.52
	CHH	115565	18188	16092	16025	19266	18769	18129	14.51	16.20	2.99	3.84
DTC	CG	32713	27194	26308	26193	27041	27133	26756	81.21	82.46	83.68	84.41
	CHG	24359	12242	11562	11606	12681	12751	12310	48.46	51.65	56.27	58.82
	CHH	81728	14514	13290	13766	16751	16105	15201	16.95	19.60	3.79	4.86
DTH	CG	11445	9767	9599	9517	9708	9730	9609	84.12	84.60	85.96	86.38
	CHG	11138	5868	5676	5655	6103	6103	5931	51.47	54.28	56.44	58.79
	CHH	50012	11035	10243	10578	12759	12143	11576	21.23	24.31	5.62	7.13
DTM	CG	22760	18835	18372	18262	18673	18726	18577	81.24	81.98	83.22	83.69
	CHG	16539	9209	8870	8758	9495	9553	9200	54.09	56.93	59.35	61.41
	CHH	61062	13652	12441	12792	14417	14031	13550	21.23	22.93	6.26	7.49
DTT	CG	3563	2866	2784	2801	2833	2841	2827	79.06	79.53	83.54	83.84
	CHG	2069	675	626	593	717	705	678	30.51	33.83	38.52	41.28
	CHH	9668	3049	2809	3073	3705	3422	3345	30.79	36.11	9.22	11.85
RIL	CG	5577	5301	5176	5145	5273	5254	5240	93.37	94.24	93.21	93.97
	CHG	8868	4052	3916	3864	4272	4216	4056	44.47	47.15	47.36	49.99
	CHH	29958	1138	1019	983	1069	1113	1036	3.49	3.58	0.48	0.48
RIX	CG	47600	44110	43318	43366	44138	44391	43746	91.59	92.63	91.71	92.42
	CHG	79788	32660	31306	30977	32969	32894	32005	39.66	40.89	42.05	43.19
	CHH	271401	9903	9065	9271	9745	9891	9512	3.47	3.58	0.49	0.51

Table S5

Number and percentage of CHH methylation level in each replicate in 500bp upstream of genes (a) and downstream of genes (b)

No.exp genes: genes with no expression in both conditions. No.diff.genes: genes without significant change in expression between stressed and unstressed plants, Diff.up: genes with up differential expression in stressed plants, Diff-down: genes with down differential expression in stressed plants

a

Sample	Number of CHH				Number of methylated CHH				percent methylated CHH among all CHH				Mean percent of CHH methylation			
	No.expr. genes	No.diff. genes	Diff.Up	Diff_down	No.expr. genes	No.diff. genes	Diff.Up	Diff_down	No.expr. genes	No.diff. genes	Diff.Up	Diff_down	No.expr. genes	No.diff. genes	Diff.Up	Diff_down
B73_U_rep1	358 003	420 428	107 197	154 087	31 787	74 197	18 814	20 262	8.9	17.6	17.6	13.1	1.90	5.62	5.46	4.32
B73_U_rep2					28 010	68 009	17 236	18 590	7.8	16.2	16.1	12.1	1.57	4.70	4.52	3.56
B73_U_rep3					28 932	68 535	17 285	18 779	8.1	16.3	16.1	12.2	1.50	4.46	4.20	3.41
B73_S_rep1					33 810	80 169	20 583	22 162	9.4	19.1	19.2	14.4	2.19	6.49	6.24	5.00
B73_S_rep3 (30S)					33 110	76 656	19 612	21 214	9.2	18.2	18.3	13.8	2.11	6.13	5.91	4.74
B73_S_rep3 (34S)					31 989	73 098	18 831	20 258	8.9	17.4	17.6	13.1	1.83	5.34	5.18	4.13

b

Sample	Number of CHH				Number of methylated CHH				percent methylated CHH among all CHH				Mean percent of CHH methylation			
	No.expr. genes	No.diff. genes	Diff.Up	Diff_down	No.expr. genes	No.diff. genes	Diff.Up	Diff_down	No.expr. genes	No.diff. genes	Diff.Up	Diff_down	No.expr. genes	No.diff. genes	Diff.Up	Diff_down
B73_U_rep1	320 226	399 840	100 941	144 571	34 175	35 247	9 072	15 750	10.7	8.8	9.0	10.9	2.28	2.01	2.05	2.66
B73_U_rep2					30 099	30 784	7 854	14 148	9.4	7.7	7.8	9.8	1.80	1.57	1.59	2.14
B73_U_rep3					31 890	31 679	8 223	13 927	10.0	7.9	8.1	9.6	1.76	1.48	1.53	1.93
B73_S_rep1					38 064	39 027	10 587	17 876	11.9	9.8	10.5	12.4	2.68	2.31	2.41	3.20
B73_S_rep3 (30S)					36 821	36 468	9 662	16 574	11.5	9.1	9.6	11.5	2.52	2.13	2.19	2.94
B73_S_rep3 (34S)					34 967	33 880	8 833	15 610	10.9	8.5	8.8	10.8	2.21	1.78	1.83	2.47

3. Conclusion et perspectives

En conclusion, le froid prolongé induit une variation du transcriptome qui explique les caractères phénotypiques observés pour le génotype B73 à la sortie du froid (baisse de photosynthèse, retard de croissance). L'augmentation de l'expression des gènes impliqués dans les fonctions vitales de la cellule telles que le repliement des protéines et réparation de l'ADN montre en outre que les conditions de culture que nous avons utilisées ont été perçues comme un vrai stress au niveau cellulaire. Cette réponse s'accompagne d'une augmentation faible (quelques pourcents) de la méthylation le long des chromosomes, pour les 3 contextes de cytosines (CG, CHG et CHH) ainsi que des variations localisées mais fortes (>10%) détectées par l'analyse des DMR, en hyperméthylation et en hypométhylation.

A l'échelle chromosomique, la cellule met en place plusieurs mécanismes (maintenance des CG et CHG, RdDM) qui augmentent en parallèle la méthylation dans les différentes régions génomiques, mais ciblent toutes les séquences d'éléments transposables. Ceci suggère que la réponse au froid nécessite une augmentation du contrôle des éléments transposables, ce qui pourrait s'expliquer par la mobilisation possible de certains éléments transposables suite au stress ressenti. Ainsi, la cellule mettrait en place un mécanisme global d'augmentation de la méthylation, pour contrecarrer le mouvement de certains éléments.

La caractérisation des contextes des DMR ainsi que l'annotation de ces dernières suggère l'existence de plusieurs mécanismes de régulation du méthylome en réponse au froid à l'échelle locale (Figure II.1). Les cas de CG only, observés dans les gènes avec ou sans lien avec la présence d'un élément transposable pourraient s'expliquer par une modification de la méthylation "gene body" dont la voie n'est pas encore bien élucidée mais qui fait appel au gène MET1 pour le maintien et la propagation de la méthylation CG dans le corps du gène (revu dans (Bewick & Schmitz, 2017), voir aussi Introduction). Les DMR CHG only localisées dans les gènes à proximité des éléments transposables pourraient résulter d'une augmentation de la maintenance de la méthylation CHG par la boucle d'auto-renforcement CMT3/H3K9me2 qui pourrait être due à une insuffisance d'IBM1. A noter que dans ces régions la méthylation CG est déjà très élevée, ce qui ne permet pas de détecter une éventuelle augmentation et ne permet donc pas de savoir si la maintenance CG est également affectée. Les DMR CG/CHG, qui présentent aussi une variation du contexte CHH et sont localisés à côté de éléments transposables en dehors des gènes pourrait s'expliquer par l'activité RdDM.

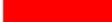
Régions	CGonly	CHGonly	CG/CHG
Génique sans ET			
ET génique			
Génique, à côté d'un ET		(En aval)	
Génique, à côté de plusieurs ET			
A côté de ET loin des gènes			

Figure II.1 : Résumé de l'annotation des DMR hyperméthylées (en rouge) et hypométhylées (en bleu)

Si les mécanismes proposés sont en accord avec les données bibliographiques, leur existence aux locus analysés reste hypothétique. L'analyse conjointe de la méthylation, des petits ARN et des marques d'histones, ou de la compaction de la chromatine (par exemple par MNase-seq ou ATAC-seq) permettrait également de mieux comprendre l'impact du froid sur les modifications de la chromatine et de faire le lien avec les modifications observées au niveau des DMR. Malheureusement, le matériel que nous avons prélevé a été presque entièrement utilisé, et de telles analyses ne pourraient être réalisées que sur d'autres feuilles des mêmes plantes.

La comparaison entre la réponse du transcriptome et du méthylome ne montre qu'une corrélation réduite. Cependant, quelques tendances sont observées : (i) les DMR localisées en amont des gènes ont peu de lien avec la transcription, (ii) les DMR localisées dans les gènes montrent au contraire un lien assez fort avec l'expression, et presque tous ces cas sont liés à une différence de méthylation en CHG, à proximité d'un élément transposable. Le lien de cause à effet entre transcription et méthylation reste cependant à analyser, par exemple en réalisant des cinétiques entre méthylome et transcriptome au cours du traitement au froid.

Dans le cadre de notre analyse, nous avons été très stringents en ne conservant que les cytosines analysables dans les 6 individus, ceci afin d'étudier la répétabilité du signal entre échantillons. L'analyse de la répétabilité a révélé que le signal est stable entre individus stressés pour un quart des cas, et stable entre individus non stressés dans un autre quart. A l'inverse de ce qui a été décrit dans d'autres études (Eichten & Springer, 2015), la stochasticité n'est donc pas plus forte après exposition à une contrainte abiotique. La variabilité observée est essentiellement due à la présence de deux individus « outliers », l'un dans le lot stressé (34S) et l'autre dans le lot non stressé (11NS) Ces « outliers » sont observables au niveau du méthylome, mais aussi du transcriptome. Ainsi, il semble que la variabilité observée soit plutôt due à une variabilité de réponse entre individus, et non à une variation locus-spécifique. La réponse différente de ces deux répétitions pourrait avoir une origine intrinsèque, liée à une provenance différente des graines (les graines utilisées proviennent d'un ensemble d'épis, et donc possiblement de plantes mères différentes) ou bien, malgré les efforts mis en place, à des variations locales des conditions de culture. Bien que nous ayons réalisé des rotations des plantes dans la chambre de culture, une interaction position dans la chambre climatisée/moment de développement, ou d'une différence faible de germination induisant des différences légères de stade de développement à l'entrée de la chambre climatisée ne sont pas à exclure. L'analyse des plantes ayant poussé dans les mêmes pots pourrait permettre de commencer à répondre à cette question (voir chapitre III). L'analyse d'un plus grand nombre de plantes pourrait permettre aussi de mieux caractériser l'ampleur de la variation entre individus.

Enfin, le plan expérimental compare un lot « stressé » qui a poussé en chambre climatisée et un lot « non stressé » qui a poussé en serre. Ainsi, les variations de méthylation détectées à la sortie du froid résultent d'un contraste entre ces deux conditions, dont un paramètre est le froid. Mais d'autres paramètres comme le rayonnement lumineux, l'évaporation, la teneur en CO₂ varient également entre les deux conditions et pourraient être responsables des variations de méthylation observées. Dans le cadre de cette expérience, nous avons choisi de faire pousser les deux lots en parallèle, car nous souhaitons mettre les plantes à pousser ensemble après le traitement froid, afin de suivre l'évolution des profils de méthylation au cours du développement, et dans les générations suivantes. Afin de valider que les variations observées proviennent bien du froid, il pourrait être envisagé de refaire une

culture du lot « non stressé » dans la chambre climatisée, avec des conditions de températures similaires à celles de la serre. Cependant, la chambre climatisée est très utilisée, et une telle culture n'est pas prioritaire. D'autre part, les analyses de méthylome sont coûteuses, et ces analyses ne sont pas prioritaires, notamment par rapport à l'analyse des petits ARN ou des marques d'histone sur les plantes de cette première expérimentation.

Chapitre III.

Dynamique de la variation de la méthylation de l'ADN au cours du développement et entre deux générations

I. Contexte et objectifs

Dans le chapitre 1, j'ai montré la réponse du méthylome suite à l'exposition des plantes au froid prolongé (21 jours) en début de développement. Afin de déterminer si ces modifications sont de nature épigénétique, une question majeure est ensuite de déterminer si les modifications induites sont maintenues une fois disparition du signal qui les a générées (le froid) et à travers les divisions cellulaires (mitose et/ou méiose) (Figure III.1).

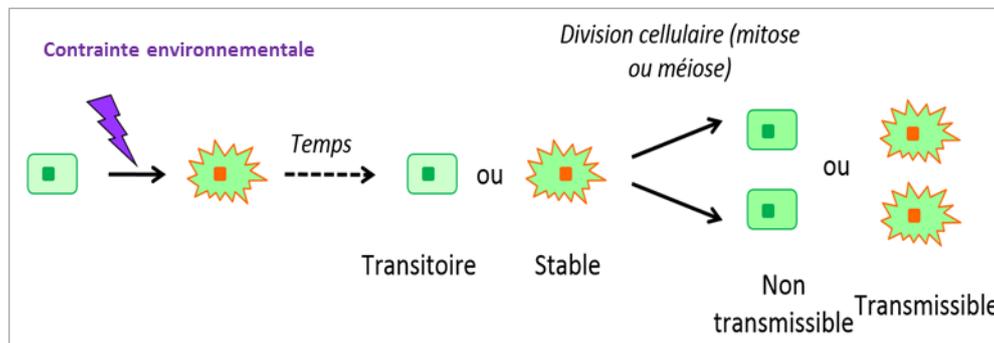


Figure III.1 : Schéma hypothétique de l'évolution de la variation d'une marque chromatinienne suite à l'application d'une contrainte abiotique (adapté de Roudier et *al.*, 2009)

Le schéma décrit la perception d'un signal environnemental par une cellule. Cette dernière subit une modification de la chromatine (méthylation de l'ADN par exemple) qui pourrait être transitoire (la cellule regagne rapidement son état initial) ou qui serait stable dans le temps. Cette marque peut ou non traverser la barrière des divisions cellulaires (mitose et méiose).

L'objectif de cette partie est de tester cette hypothèse. Pour cela, j'ai suivi la dynamique des variations détectées entre plantes « stressées » et « non stressées » (chapitre 1) au cours du développement, après retour des plantes dans des conditions « standard ». Les questions auxquelles j'ai cherché à répondre sont :

- (i) les DMR issues des différentes composantes génomiques (gènes, ET, etc.) et des différents contextes de cytosine (CG, CHG, CHH) sont-elles transmises avec la même efficacité ?
- (ii) l'efficacité du maintien dépend-elle du niveau de méthylation atteint à la sortie du froid ?
- (iii) l'efficacité du maintien est-elle variable entre individus ?
- (iv) Peut-on faire un lien entre le niveau de méthylation dans une feuille de rang donné et le stade de développement de cette feuille au moment de l'exposition au froid ?
- (v) Les modifications sont-elles transmises à la descendance ?
- (vi) Sont-elles transmises différemment par la voie maternelle et la voie paternelle ?

II. Choix du matériel végétal à prélever

1. Prélèvements dédiés à l'analyse de la maintenance des niveaux de méthylation au cours du développement

Un critère important dans le choix du matériel végétal à analyser était de pouvoir appréhender la question du lien entre stade de développement de la feuille au moment de l'exposition au froid, et l'efficacité du maintien du niveau de méthylation dans cette feuille plus tard dans le développement. Pour cela, nous souhaitions suivre les niveaux de méthylation de l'ADN sur des feuilles initiées avant, pendant et après l'exposition au froid, selon les définitions suivantes :

Feuille initiée avant l'exposition au froid : initiation avant le traitement froid (présence de petits primordia de feuille), divisions et début d'élongation cellulaire pendant le traitement froid.

Feuille initiée pendant l'exposition au froid : initiation pendant le traitement froid, divisions et élongation cellulaire pendant et après le traitement froid.

Feuilles initiées après l'exposition au froid : initiation, divisions et élongation cellulaire après le traitement froid.

En plus de ces prélèvements, nous souhaitions prélever la dernière « vraie » feuille produite avant la transition florale : la **Feuille drapeau-1** (la dernière feuille est en fait la feuille drapeau, mais c'est une feuille modifiée qui est physiologiquement assez différente des autres feuilles). Cette feuille correspond à une feuille initiée après l'exposition au froid, mais elle a la particularité d'être, au niveau développemental, la feuille la plus proche du méristème produisant les organes reproducteurs mâles. Elle sert donc comme témoin pour comparer les niveaux de méthylation observés dans la descendance. A noter que, chez le maïs, le nombre de feuilles varie d'un génotype à l'autre, et que de ce fait le nombre de divisions cellulaires pour atteindre la feuille drapeau (et donc la panicule mâle) varie d'un génotype à l'autre. Cette différence est intéressante car si le signal est perdu au cours des divisions cellulaires, les génotypes avec un plus grand nombre de rang de feuille pourraient montrer une moins bonne transmission à la descendance, indépendamment de la méiose.

Afin de déterminer le stade de développement de chacune des feuilles pendant l'exposition au froid, j'ai réalisé des dissections sur 3 plantules de chacun des 3 génotypes, au stade levée (avant entrée dans la chambre climatique) et à la sortie de la chambre climatique. Pour cela, j'ai réalisé la séparation manuelle des feuilles visibles, puis une dissection au scalpel sous loupe binoculaire jusqu'à observation de l'apex méristématique (Figure III.2). Ces comptages ont permis de déterminer les rangs des feuilles développées avant et après le traitement froid (Tableau III.1A), et ainsi de déterminer par déduction les feuilles qui étaient en cours de développement pendant le traitement froid (Tableau III.1B). Ces résultats montrent que les 3 génotypes initient le même nombre de feuilles avant exposition au froid, mais un nombre différent de feuilles pendant le traitement froid. Etant donné que le rythme de développement des 3 génotypes est identique avant le traitement les différences observées pendant le traitement sont liées à la variation de sensibilité au froid des génotypes choisis. Le nombre de feuilles différent initié après le traitement est lié aux différences de sensibilité mais aussi de précocité.

Le choix des feuilles prélevées pour chaque génotype est indiqué dans le tableau III.1. Ce choix est fondé à la fois sur les résultats des dissections (Tableau III.1A) et la nécessité de travailler sur les mêmes plantes pour les différents rangs de feuille, tout en évitant l'impact potentiel de coupures successives sur le méthylome. Nous avons donc réalisé l'ensemble des prélèvements à une même date pour

chaque plante. A cette date, la feuille initiée après le froid devait être ligulée et donc les feuilles 1 à 5 étaient en état sénescé. Ces prélèvements ont été réalisés à des dates variables pour chaque génotype, afin de prélever toutes les plantes de chaque génotype à un même stade de développement (B73 au stade 18 feuilles visibles, F2 au stade 15 feuilles visibles et F331 au stade 21 feuilles visibles). Les prélèvements de ces feuilles ont commencé environ un mois après l'arrêt du froid pour les génotypes B73 et F2 et un mois et demi pour le génotype F331. L'attente du développement de la feuille drapeau-1 n'aurait pas permis le prélèvement de la feuille 6. De ce fait, cette feuille a été prélevée après la première série de prélèvements.

De façon optimale, la comparaison avec les échantillons prélevés à la sortie du froid aurait dû nous conduire à prélever la feuille 4 pour l'analyse de la méthylation de l'ADN de la feuille initiée avant le froid. Mais cette feuille aurait été en état sénescé au moment où la feuille initiée après le froid aurait été au bon stade (*i.e* ligulée) pour être prélevée. De ce fait, nous avons choisi à la place de prélever la feuille 6. Pour la feuille initiée pendant le froid et la feuille initiée après le traitement froid, nous avons choisi la feuille n+1 du diagramme de dissection afin de prendre un rang de marge par rapport à la zone d'ambiguïté.

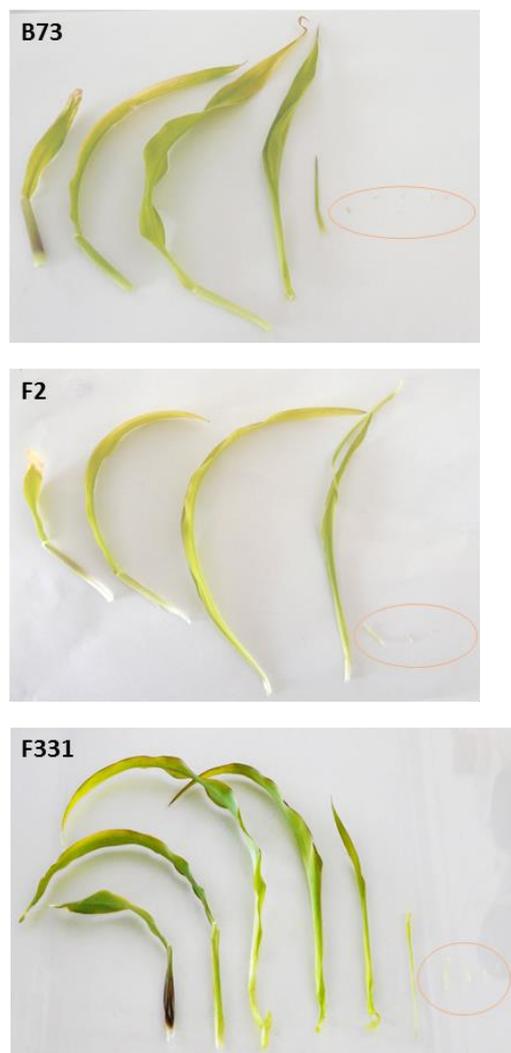


Figure III.2 : Exemple de plantule disséquée, illustrant le nombre de feuilles dénombré à la sortie de la chambre climatisée

Tableau III.1 : Rangs de feuilles prélevés sur les plantes adultes. **A.** Diagramme présentant le nombre de feuilles initiées avant, pendant et après le traitement froid pour les 3 génotypes B73, F2 et F331. La zone grisée correspond à une zone d'ambiguïté liée à une variation entre les répétitions au moment de la dissection. **B.** Feuilles choisies pour les analyses au cours du développement pour les 3 génotypes. * Le rang de feuille n'est pas estimé à partir des dissections mais correspond à la valeur moyenne observée sur l'ensemble des plantes

A.	Avant		Pendant		Après
B73	1 à 6	NA	7 à 10	11	> 12
F2	1 à 6	7	8 à 9	10 à 12	> 13
F331	1 à 6	7 à 9	10 à 13	14	> 15

B.

	Feuille initiée avant le froid	Feuille initiée pendant le froid	Feuille initiée après le froid	Feuille d-1*
B73	6	8	13	20
F2	6	9	14	14
F331	6	11	16	22

2. Prélèvements dédiés à l'analyse de la maintenance des niveaux de méthylation dans la descendance (génération G1)

Les plantes du lot « stressé » et « non stressé » prélevées au cours du développement ont été menées à floraison, puis des autofécondations et des croisements ont été réalisés, en privilégiant les croisements entre plantes du lot « non stressé » et plantes du lot « stressé » appariées dans le dispositif en serre (voir Annexe 2). Les autofécondations permettent de comparer la transmission du signal chez une plante « stressée » par rapport à une plante « non stressée ». Les croisements permettent de déterminer si l'origine maternelle ou paternelle des épiallèles générées par le froid a un effet sur la transmission.

En mars 2017, le semis de la génération G1 a été réalisé en conditions standard en serre (voir Annexe 2). Avec cette culture, nous souhaitons (i) avoir une estimation globale sur la transmission de la méthylation de l'ADN dans la descendance et (ii) suivre la dynamique de la transmission au cours du développement dans la descendance.

Ayant disposé deux graines par pot, la première plantule a été utilisée pour le prélèvement destructif au stade « plantule ». Elle a été prélevée au stade 3 feuilles visibles, pour être similaire au stade de la génération G0 utilisé pour la comparaison à la sortie du traitement froid (chapitre 1). La deuxième

plantule a été maintenue en pot jusqu'à la floraison pour les prélèvements non destructifs et la réalisation des autofécondations et des croisements. Sur ces plantes, deux séries de prélèvements ont été réalisés : la feuille 2 (au stade plantule), la feuille 8 et la feuille drapeau-1 sur plante mature (au stade 15 feuilles pour F2 et 18 pour B73). Le prélèvement au stade plantule permet d'estimer le niveau de transmission dans la descendance. La feuille 2 permet de valider la présence du signal au stade jeune sur la plante qui générera des descendants. La F8 permet de suivre le signal au cours du développement et la feuille d-1 est le témoin de méthylation (organe le plus proche du méristème en terme de développement) pour la génération d'après.

III. Suivi des variations de méthylation pour le génotype sensible B73

1. Sélection des régions génomiques à analyser

Pour permettre le suivi des variations de méthylation détectées à la sortie du froid (chapitre 1), nous avons mis en place une analyse ciblée par « Sequence Capture Epigenetics » (SeqCapEpi), une méthode permettant d'analyser de façon spécifique les niveaux de méthylation de régions choisies, sur un ensemble d'échantillons choisis. L'analyse SeqCapEpi a été réalisée en collaboration avec Biogemma, dans le cadre du programme d'investissement d'avenir *Amaizing*. J'ai réalisé la détection des DMR, puis nous avons interagit avec eux pour mettre au point le design de la capture, et ils ont réalisé la partie de biologie moléculaire (qui inclue la préparation des banques et la capture) et sous-traité le séquençage. Ils ont ensuite réalisé l'alignement des séquences sur le génome. Ces alignements ont été réceptionnés à la toute fin de ma thèse, ce qui ne m'a pas permis de les analyser avant la rédaction de ce manuscrit. Je présente ci-dessous les analyses auxquelles j'ai participé pour réaliser le design de la capture, et un début de contrôle qualité des résultats de séquençage.

Les DMR utilisées comme matériel de départ pour le design ont été détectées de façon similaire à celles présentées dans le chapitre 1. En particulier, elles ont été détectées avec les mêmes tailles de fenêtres, mais différent au niveau du nombre de DMC par DMR : 5 DMC pour les DMR du chapitre 1, et 2 DMC pour celles-ci (voir détails dans l'Annexe 3). Cette différence est liée au fait que des DMR représentées dans le design du SeqCapEpi ont, pour des raisons de calendriers liées à la collaboration avec Biogemma, dû être choisies avant d'avoir complètement stabilisé les critères de sélection des DMR finales étudiées dans le chapitre 1. Au total, 12784 DMRs ont été sélectionnées, représentant un total de 10,29 Mb. Le détail de ces DMR est présenté dans le tableau III.2. Les chiffres présentés comprennent, pour chaque région, un « padding » de 50 bases à chaque extrémité.

Tableau III.2 : Détails des DMRs détectées entre plantes « stressées » et « non stressées » à la sortie du traitement froid, et sélectionnées pour le suivi développemental et de la génération G1

Contexte de détection	DMRs hyperméthylées chez les plantes "stressées"		DMRs hypométhylées chez les plantes "stressées"	
	Nombre	Taille cumulée (Mb)	Nombre	Taille cumulée (Mb)
CG	1390	1,32	1608	1,48
CHG	509	0,73	505	0,68
CHH	4512	3,32	145	0,11
3 contextes ensemble	2348	1,52	1767	1,13
Total	8759	6,89	4025	3,40

Une contrainte majeure dans le design des régions à capturer est d'éliminer les régions répétées, qui conduisent à l'hybridation de nombreuses régions du génome, et risquent d'utiliser une grande partie de l'effort de séquençage. Malgré cette contrainte, nous souhaitons pouvoir analyser les régions proches d'éléments transposables. De ce fait, nous avons mis au point une stratégie spécifique. Traditionnellement, les séquences cibles sont éliminées suivant le pourcentage de recouvrement avec des séquences d'éléments transposables. L'analyse du nombre de régions restant après filtre de 100% N (régions à 100% couvertes par un élément transposable), à 50% N (régions couvertes à au moins 50% par des séquences d'élément transposable) montre une réduction importante du nombre de régions (Figure III.3) : elle passe de 10 Mb à 2,11Mb avec au moins 20% de N et de 10 Mb à 5 Mb quand la région est entièrement représentée par une séquence répétée (100% N).

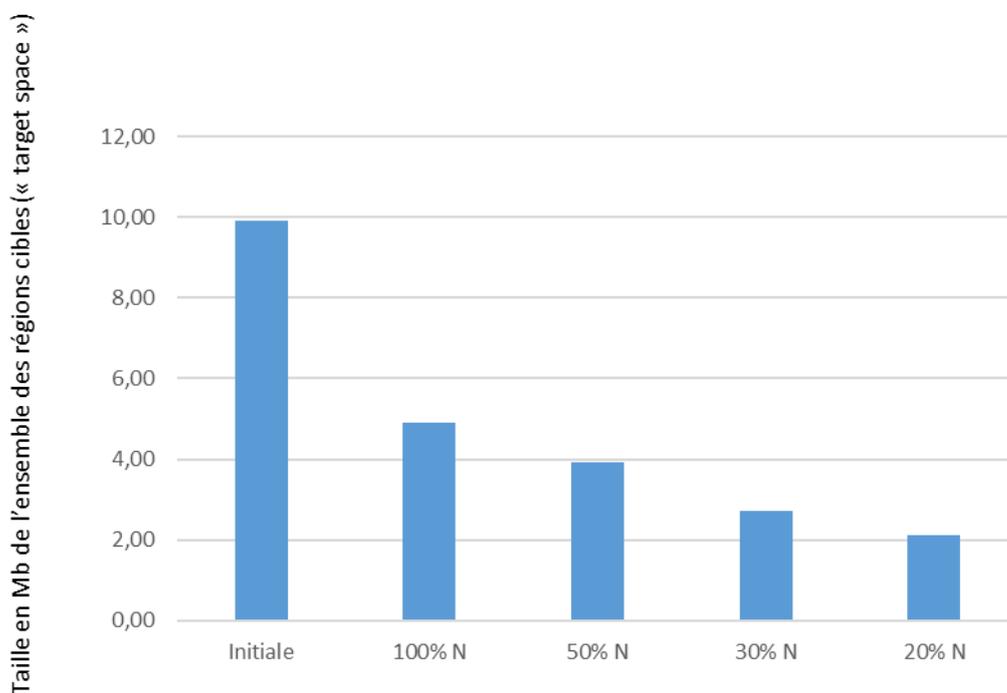


Figure III.3 : Evolution de la taille du « target space » en fonction du pourcentage des bases couvertes par des séquences d'éléments transposables (Biogemma)

%N: pourcentage de la région cible correspondant à des bases couvertes par une séquence d'élément transposable

Cette méthode nous a paru très stringente, et ne permet pas de conserver les séquences uniques adjacentes aux éléments transposables. Plutôt que d'éliminer les régions sur la base d'un pourcentage de couverture avec une séquence d'ET, nous avons analysé la taille des régions après élimination des bases couvertes par des séquences d'éléments transposables connus. Nous avons choisi de conserver toutes les régions contenant au moins 100 bases continues après élimination de ces bases. La figure III.4 illustre la diminution en taille des régions analysées après masquage et filtre à 100pb (Figure III.4). En particulier, la médiane évolue de 800pb à 475pb.

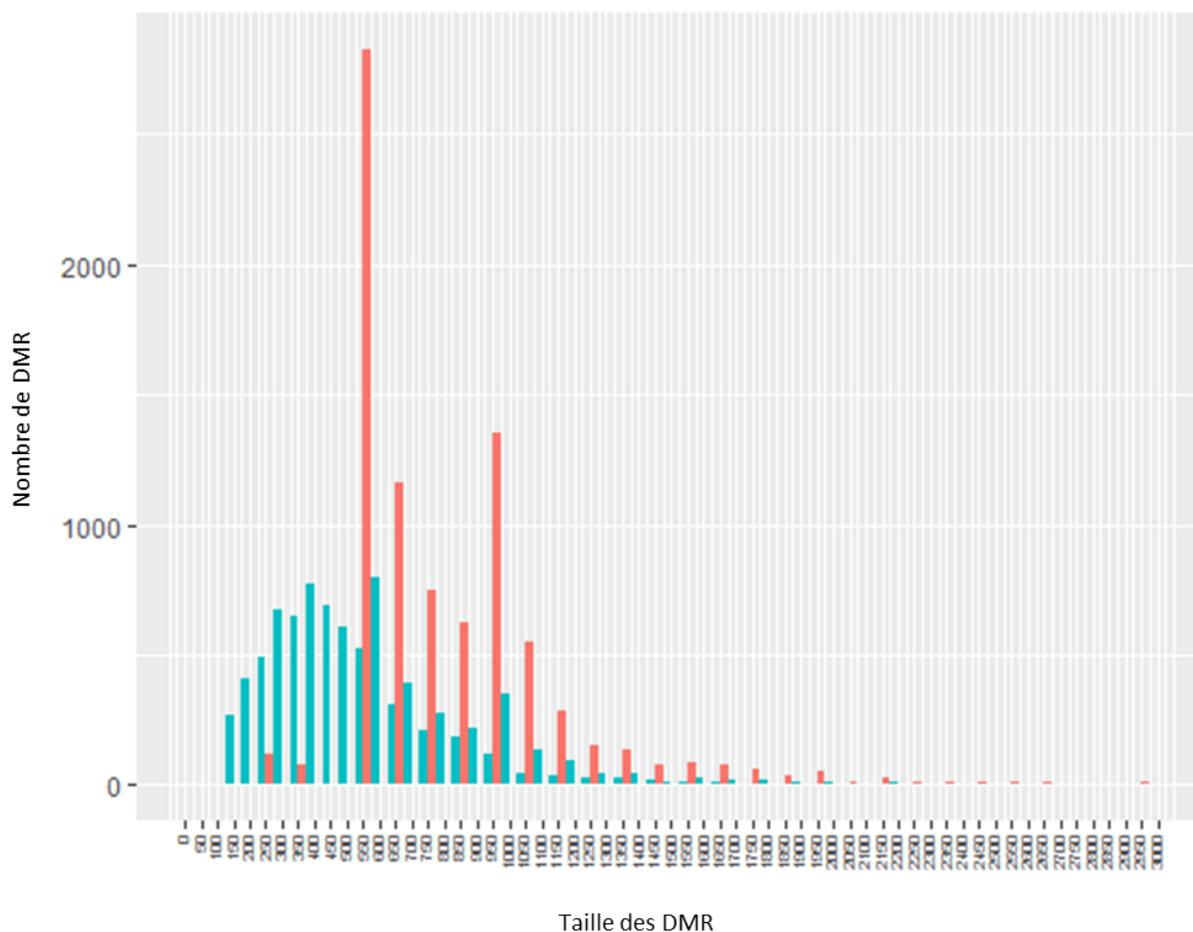


Figure III.4: Distribution de la taille des DMR avant (en rouge) et après (en vert) le premier masquage des séquences d'éléments transposables

Cette méthode nous a permis de conserver 8691 régions avec une taille cumulée de 4,79 Mb. A partir de ces régions, un deuxième filtre a été appliqué afin d'éliminer des séquences répétées non annotées (Roche, critères non communiqués). Ce deuxième filtre a éliminé un faible nombre de bases, ce qui n'a pas modifié le nombre de DMR analysées. Ces séquences ont ensuite été utilisées pour le design des sondes par Roche, puis seules les régions dont au moins 100 bases contiguës couvertes par des sondes ont été retenues. Ce filtre a éliminé un total de 121 séquences (99 non couvertes par des sondes, et 22 séquences couvertes par des sondes mais ayant une taille inférieure à 100 pb). Au final, un total de 8570 séquences présentant une taille de 4,62 Mb a pu être conservé pour l'analyse. La figure III.5 résume les différentes étapes qui ont été réalisées, des régions DMR de départ jusqu'aux régions sélectionnées pour la capture.

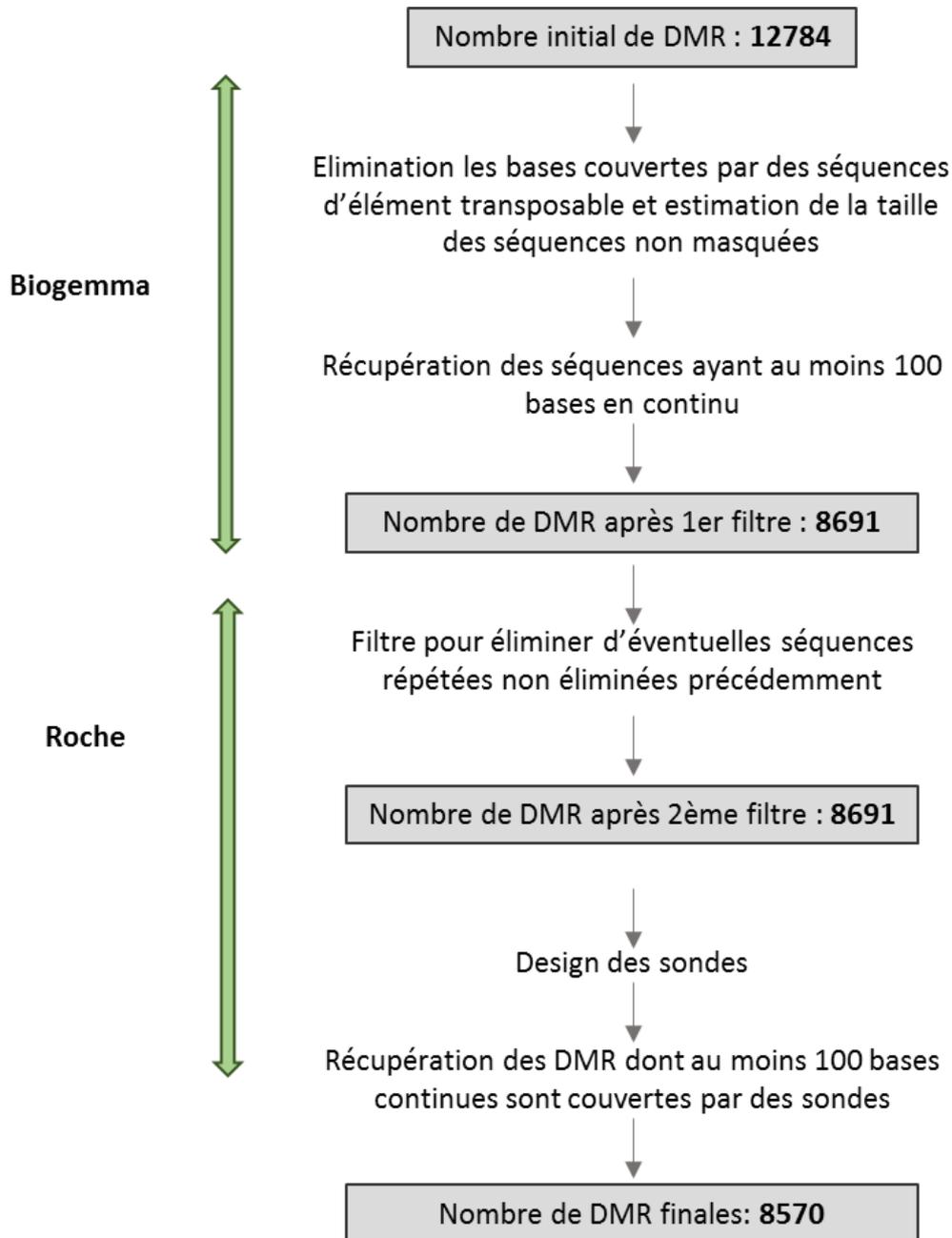


Figure III.5 : Etapes de sélection des DMR en amont de la capture

La figure III.6 montre que, en médiane, les régions choisies sont couvertes à 60 % par des sondes. Bien que cela puisse paraître peu, ceci est compensé par le fait que les cibles capturent des fragments d'environ 462pb, qui sont ensuite séquencés en « paired end » de 2x100 bases.

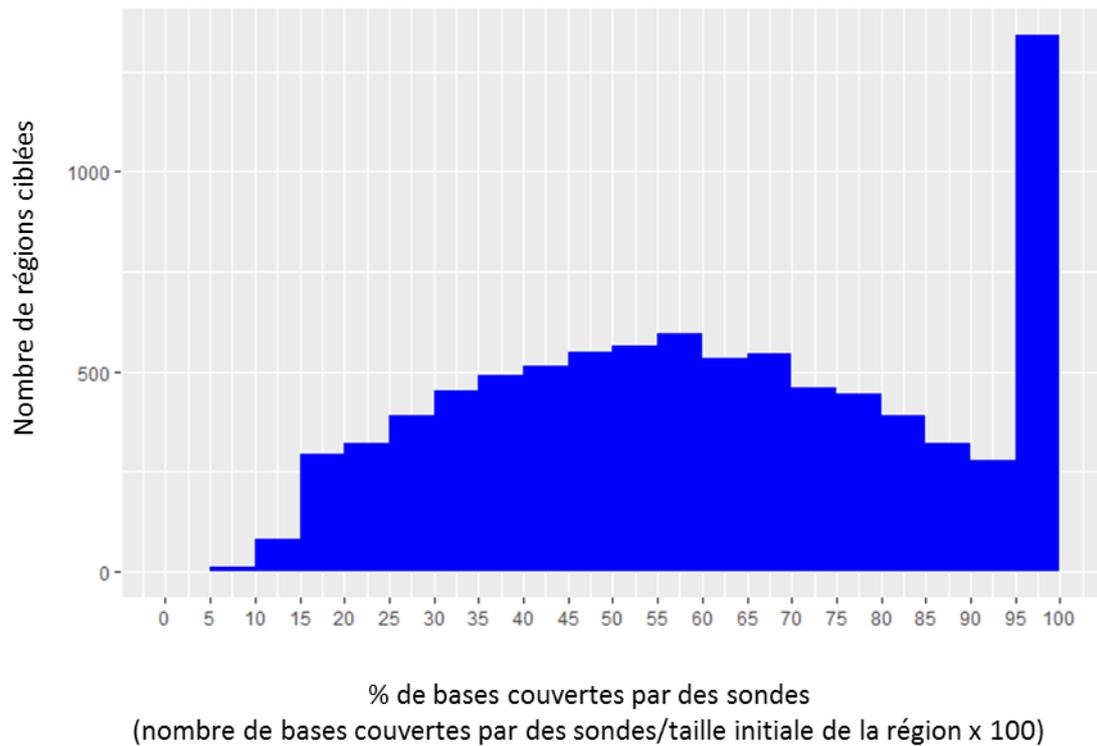


Figure III.6: Distribution du nombre des régions selon leur couverture en sondes

2. Choix des échantillons à analyser

Les régions ayant été dessinées à partir de la comparaison du lot « stressé » et « non stressé » à la sortie du froid pour le génotype B73, nous souhaitons analyser uniquement les échantillons relatifs à cette lignée. Pour des problèmes de coût, nous ne pouvions pas représenter tous les échantillons prélevés pour cette lignée (222 au total). Le choix des plantes à analyser a été fondé sur la réussite des autofécondations et des croisements effectués dans la génération 0, afin de pouvoir analyser les échantillons « développement » et « génération G1 » issus des mêmes plantes. Il a également été estimé pour permettre un séquençage de 80x par brin, en prenant en compte la taille du design, et le nombre d'échantillons que l'on peut mélanger par hybridation. Ainsi, l'analyse a ciblé un ensemble de 30 échantillons foliaires pour le suivi au cours du développement, et 42 pour le suivi dans la génération G1. Le nombre et le détail des plantes qui ont servi pour cette analyse sont décrits dans le tableau III.3.

Tableau III.3 : Feuilles analysées pour le suivi « développement » et « génération G1 » pour le génotype sensible B73

Pour l'analyse au cours du développement, les 4 feuilles correspondent aux feuilles initiées avant (F6), pendant (F8) et après le traitement froid (F13) et la feuille drapeau-1. Pour le suivi à la génération G1, le nombre d'échantillons correspond au nombre de grains issus du même épi. *Aucune plantule n'est prélevée en génération G1 du fait du faible nombre de grains. ** Plante utilisée uniquement pour le suivi au cours du développement.

Plante	Traitement	Type de croisement	Nombre de grains	Suivi développement	Suivi génération G1			
					Plantule	Feuille 2	Feuille 8	Feuille d-1
E1	froid	Autofécondation	2 grains	4	0*	2	2	2
H7	contrôle	Autofécondation	1 épi	4	3	3	3	3
F8	froid	femelle froid X mâle contrôle	17 grains	4	3	3	3	3
H6	contrôle			4				
P17	contrôle	femelle contrôle X mâle froid	1 épi	4	3	3	3	3
D1	froid			4				
P1**	contrôle	femelle froid X mâle contrôle	0 épi	3	0	0	0	0
C7**	froid			3				
Total				30	42			

3. Analyses à réaliser à partir des données

A réception des données, il est nécessaire de réaliser un contrôle qualité afin de déterminer si (i) la capture a bien fonctionné pour toutes les régions représentées et (ii) l'ensemble des échantillons a été séquencé à une profondeur similaire. La figure III.7 présente la distribution de la couverture globale (sur les 2 brins) des DMR pour deux échantillons. Elle montre une couverture médiane de 180X pour le premier échantillon et de 200X pour le deuxième. Le nombre de régions non couvertes ou à faible couverture (<80X) est assez faible, et l'analyse de ces régions révèle qu'elles sont systématiquement mal couvertes dans l'ensemble des échantillons. Il s'agit donc vraisemblablement d'un problème au niveau de la capture. A l'autre extrême, les fortes couvertures (>400X) montrent que malgré les différentes étapes de masquage, certaines régions répétées ont probablement été séquencées. Les deux individus analysés sont représentatifs de la forme de la distribution de couverture et de la médiane de l'ensemble des échantillons analysés (186X). La gamme de variation médiane de couverture est de 116X (minimum) à 302X (maximum), ce qui illustre une assez bonne homogénéité d'hybridation et de séquençage pour l'ensemble des échantillons. L'ensemble de ces résultats montre que les données issues du SeqCapEpi sont de bonne qualité.

Comme indiqué plus haut, les DMR du chapitre 1 sont incluses dans celles du SeqCapEpi, mais ne correspondent pas à l'intégralité de ces régions. Afin de suivre le niveau de méthylation des DMR détectées à la sortie du froid et catégorisées en fonction du contexte et de l'annotation (cf. chapitre1), une première étape va être d'extraire les niveaux de méthylation de chacune des DMR du chapitre 1, en appliquant le pipeline roimethstat (<http://smithlabresearch.org/software/methpipe>) aux fichiers de niveaux de méthylation des cytosines individuelles. Ceci permettra d'obtenir un diagramme de

l'évolution des niveaux de méthylation de chaque contexte pour chaque région au cours du développement, pour des plantes « stressées » et « non stressées » et leur descendance. Pour chaque DMR, les niveaux de méthylation de chaque contexte seront analysés, ainsi que la différence correspondante. Ces valeurs seront comparées aux valeurs obtenues à la sortie du traitement froid en G0 et entre elles, dans un ordre chronologique d'émergence des feuilles.

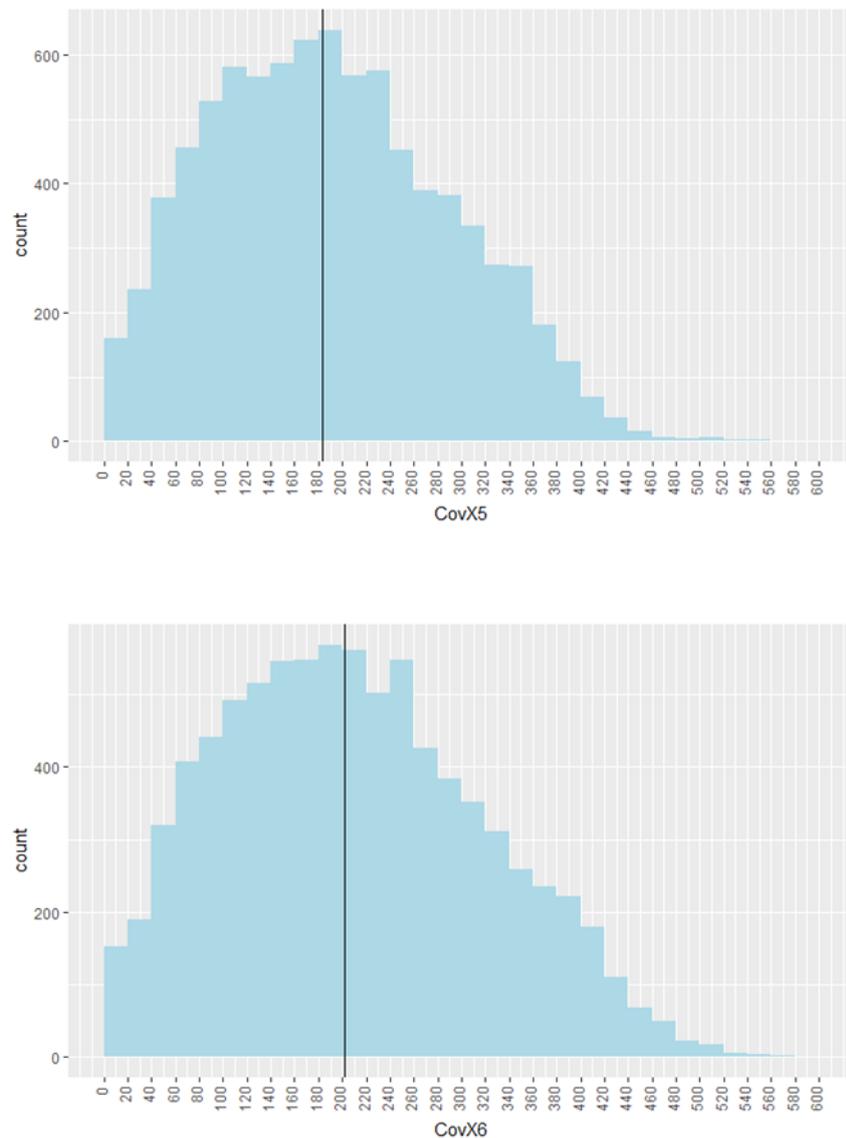


Figure III.7: Distribution de la couverture des DMR pour deux individus
Le trait indique la valeur médiane de couverture pour l'ensemble des régions

Les DMR détectées à la sortie du froid proviennent de plantules qui ont été sacrifiées au stade 4 feuilles visibles. De ce fait, ces plantes n'ont pas pu être incluses dans les analyses de suivi de méthylation au cours du développement, et ces analyses ont été réalisées sur des plantes différentes mais ayant poussé dans les mêmes conditions (dans la même culture). Une étape importante sera donc de vérifier si les niveaux de méthylation observés sont similaires sur ces plantes. S'ils ne le sont pas, cela pourra provenir (i) de la variabilité entre plantes et/ou (ii) de l'absence de maintien du signal après l'arrêt de l'exposition au froid. Dans le chapitre 1, j'ai présenté des DMR « stables » (de niveau de méthylation

similaire entre plantes « stressées »). Ces régions auront de plus grandes chances d'avoir été présentes sous forme de DMR au stade plantule dans les plantes analysées, et seront particulièrement intéressantes pour séparer les deux hypothèses.

L'analyse des niveaux de méthylation des feuilles F8, F13 et Fd-1 dans les individus stressés et non stressés permettra ensuite de tester différents cas d'évolution du signal au cours du développement. La figure III.8 présente sept cas théoriques qui pourraient être observés (1) maintien de la différence de méthylation entre individus « stressés » et « non stressés » au cours du développement, (2) perte du signal observé dans les individus « stressés » à la sortie du traitement froid, du fait du retour en conditions « standard » ou absence de signal du fait de la variabilité entre plantes, (3) perte progressive du signal observé dans les individus « stressés » à la sortie du traitement froid au cours du développement, (4) perte du signal dépendante de l'état développemental au moment de l'exposition au froid, (5) perte du signal induit, mais effet fort du développement sur les niveaux de méthylation, (6) signal aléatoire chez les individus stressés et non stressés. On peut aussi s'attendre à une augmentation du signal du fait du nombre plus élevé de divisions dans la feuille 6 dans le froid par rapport à la feuille 4. Un point important sera de déterminer si les DMR observées comme hyperméthylées et hypométhylées chez les individus « stressés » à la sortie du traitement froid ont les mêmes profils au cours du développement. On peut s'attendre à ce que cela ne soit pas le cas, car les mécanismes impliqués dans la reméthylation et la déméthylation sont très différents. En particulier, la déméthylation peut être passive, ce qui pourrait induire une perte progressive du signal.

Enfin, une question importante sera d'analyser la variabilité de la réponse entre individus. Comme nous l'avons vu au chapitre 1, il existe une variabilité assez forte entre individus, juste à la sortie du traitement froid. On peut s'attendre à ce qu'il y ait également une variabilité de maintenance du signal au cours du développement. Ainsi, pour chaque rang de feuille, les niveaux de méthylation observés chez les 4 individus (3 pour la feuille d-1) « stressés » et « non stressés » seront comparés afin d'estimer la répétabilité du signal observé.

L'analyse de la génération G1 sera réalisée de façon similaire, en analysant d'abord les niveaux de méthylation chez les plantules afin de déterminer si le signal a été transmis à travers la méiose et les premières phases de développement. Ces niveaux seront comparés à ceux de la feuille drapeau-1 de la génération G0, afin de vérifier que l'absence du signal provient bien du saut de génération et non d'une perte au cours du développement de la génération G0. La comparaison des profils pour les 3 plantes issues de chaque fécondation (autofécondation ou croisement) permettra de déterminer le niveau de variabilité de transmission du signal entre deux générations. La comparaison des autofécondations issues des plantes stressées avec celles issues des plantes non stressées permettra de déterminer si les niveaux de méthylation induits par le froid sont transmis au même niveau que celles préexistant dans le génome. Les résultats des croisements seront comparés entre eux afin de déterminer s'il existe une différence de transmission selon la voie (paternelle ou maternelle) apportant l'épiallèle induit par le froid (Figure III.8).

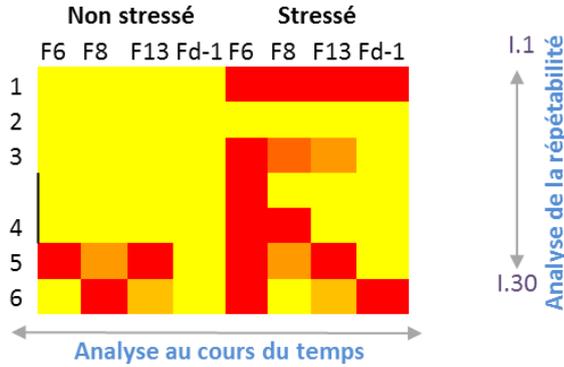
Ensuite, de même que pour la génération G0, le signal sera analysé au cours du développement. Il servira pour l'étude de la transmission du signal jusqu'à la feuille d-1 afin d'avoir un point de comparaison pour les individus de la génération G2 (les plantes de la génération G1 ont été autofécondées et les graines sont disponibles, voir Annexe2).

Sortie du Froid: détection d'un total de 631 DMR contexte spécifique et annotées par rapport aux gènes et aux ET

Typage de ces DMR dans les 3 contextes CG,CHG et CHH chez les 72 individus

Suivi en G0

DMR hyperméthylées chez les stressés

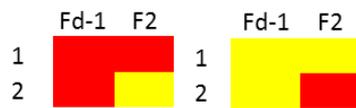


- (1) Maintien au cours du développement (2) perte dès retour en conditions standard (3) perte progressive au cours du développement (4) perte du signal en fonction de l'état du développement au moment de l'exposition au froid (5) effet du développement (6) signal aléatoire

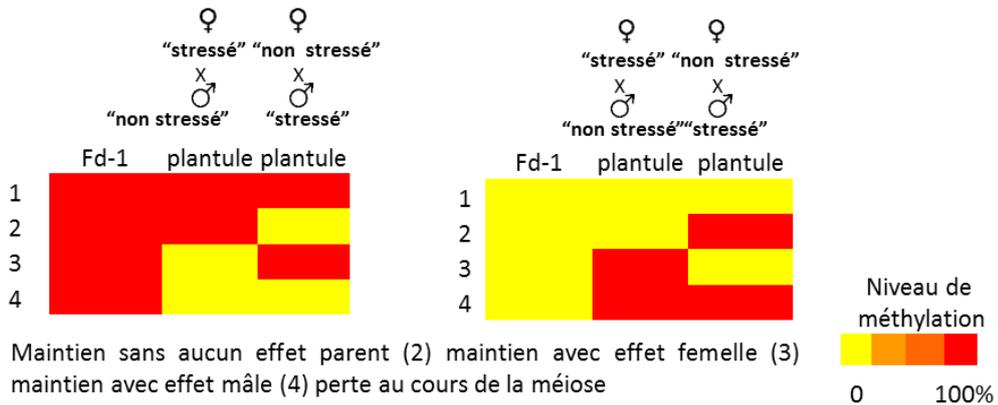
Suivi en G1

DMR hyperméthylées chez les stressés

⊗ Stressé ⊗ Non Stressé



- (1) Maintien (2) perte au cours de la méiose



- (1) Maintien sans aucun effet parent (2) maintien avec effet femelle (3) maintien avec effet mâle (4) perte au cours de la méiose

Figure III.8 : Schéma hypothétique de la dynamique de méthylation de l'ADN au cours du développement et dans la descendance pour le génotype B73

Les heatmpas illustrent la dynamique de la méthylation pour un seul individu au cours du temps

I : Individu, ⊗ : autofécondation

Comme indiqué précédemment, en plus des DMR du chapitre 1, le design comprend des régions différentiellement méthylées avec une différence de 10% entre les lots « stressé » et « non stressé », mais qui contiennent un nombre moins important de DMC. Ces DMR ont vraisemblablement un rôle biologique moins important. Mécanistiquement, il sera cependant intéressant de déterminer si elles sont maintenues de façon similaire à celles ayant un plus grand nombre de DMC. De même, pour les deux types de DMR, il sera intéressant de suivre le niveau de méthylation des DMC qu'elles contiennent, afin de déterminer si l'évolution du niveau de méthylation a lieu de façon concertée entre cytosines proches.

Une limite importante du SeqCapEpi est qu'il est fondé sur l'analyse de régions particulières. Comme les DMR analysées ont été obtenues à partir d'un lot de plantes qui n'a pas pu être analysé par la suite puisque les plantes ont été sacrifiées, et que nous avons observé une variation de réponse entre individus, il est possible que d'autres régions que celles représentées sur le design soient différentiellement méthylées entre les feuilles 6 du lot « stressé » et « non stressé ». Ces régions ne pourront pas être analysées par cette technique. Si trop peu de régions montrent un signal intéressant sur ces plantes, nous pourrions envisager de réaliser une analyse de BS-seq sur génome entier pour les échantillons de la feuille 6, afin de détecter des DMR entre feuilles 6 issues de plantes « stressées » et « non stressées », puis de redessiner une capture de SeqCapEpi pour les suivre sur les autres feuilles et dans la descendance.

Chapitre IV.

Comparaison de la réponse du méthylome au froid entre trois génotypes contrastés pour leur réponse phénotypique

I. Contexte et objectifs

Comme présenté en introduction, la recherche de régions génétiques impliquées dans la réponse au froid chez le maïs est délicate, notamment du fait de la présence probable de nombreux locus à effet faible. L'analyse transcriptomique de plantes ayant subi un traitement au froid pourrait permettre de proposer des gènes candidats. En parallèle, la meilleure compréhension de la réponse du méthylome au froid pour plusieurs génotypes pourrait permettre de déterminer si ces modifications pourraient avoir un rôle fonctionnel.

L'analyse de la croissance de lignées de maïs après un traitement de 21 jours à une température de 14°C jour/10°C nuit a permis de mettre en évidence une variabilité phénotypique de réponse au froid prolongé des lignées B73, F2 et F331. B73 a un phénotype très marqué (forte chlorose et retard de développement), F2 a un phénotype un peu moins affecté (chlorose et retard de développement) et F331 a un phénotype peu affecté (retard de développement). Ces mêmes phénotypes ont été observés dans une culture de 45 jours réalisée avant ma thèse (voir préambule présentant une vue d'ensemble de l'expérimentation). Les deux premières lignées sont considérées comme sensibles, et la dernière est considérée comme tolérante (Figure IV.1).



Sensibilité au froid prolongé

Figure IV.1 : Réponse phénotypique contrastée entre les 3 génotypes après 21 jours de traitement froid en chambre climatisée

A partir de ce matériel, nous avons analysé la réponse transcriptionnelle (par mRNA-seq) et du méthylome (par BS-seq) de chacune de ces lignées à la sortie du froid, afin de tester le lien entre ces deux réponses moléculaires et tenter d'expliquer le phénotype observé.

Les questions auxquelles je vise à répondre dans ce chapitre sont :

(i) est-ce que les trois génotypes présentent un nombre différent de gènes différentiellement exprimés ? Si oui, est-ce que les deux génotypes sensibles se distinguent du génotype tolérant ?

(ii) est-ce que les gènes différentiellement exprimés entre plantes « stressées » et « non stressées » dans chacun des génotypes sont enrichis dans des catégories fonctionnelles différentes ? Si oui, est-ce que les deux génotypes sensibles se distinguent du génotype tolérant ?

(iii) peut-on mettre en évidence des gènes clés de ces catégories fonctionnelles pour chacun des génotypes ?

(iv) Est-ce que la réponse du méthylome (différence de niveau de méthylation entre plantes « stressées » et « non stressées » à l'échelle chromosomique et nombre de DMR) est différente chez les trois génotypes ? Si oui, est-elle plus similaire entre les deux génotypes sensibles par rapport au génotype tolérant ?

(v) Les régions différentiellement méthylées sont-elles localisées au même endroit dans les trois génotypes ? Ou à côté du même type de séquences génomiques ? En particulier, sont-elles localisées à côté de gènes de fonctions différentes et pourraient expliquer la différence d'expression des gènes, voire le phénotype observé ?

Dans cette partie, je présente des analyses préliminaires sur les variations du méthylome des 3 génotypes. Le matériel végétal qui a servi pour l'extraction de l'ADN correspond aux feuilles prélevées à la sortie du froid (dernière feuille visible). Pour chacun des 3 génotypes, 3 plantes du lot « stressé » et trois plantes du lot « non stressé » sont analysées (voir préambule sur la description de l'expérimentation et annexe 2). Le séquençage a été réalisé en paires de séquences, et l'alignement des lectures a été réalisé sur le pangénome de B73/F2. Le pipeline utilisé pour la détection des DMR est celui utilisé pour l'analyse du génotype B73 présenté dans le chapitre 1 (voir aussi annexe 3 pour le détail méthodologique). Les analyses du transcriptome pour les 2 génotypes F2 et F331 sont en cours et n'ont pas pu être intégrés ici.

II. Résultats préliminaires et discussion

Le séquençage après traitement de l'ADN au bisulfite de sodium a permis de générer un total de 2 904 605 645, 2 760 644 289 et 3 096 753 166 de lectures, respectivement pour B73, F2 et F331. En moyenne, 49% et 47% des lectures ont pu être alignées de façon unique pour F2 et F331 (Tableau IV.1). Ce chiffre est inférieur à ce qui a été observé pour le génotype B73 (68% en moyenne).

Tableau IV.1 : Données de séquençage et de mapping pour les 3 génotypes B73, F2 et F331 (NS : Non Stressé et S : Stressé)

*estimée par rapport à la taille du pangéome B73/F2 qui est de 2.15Gb

Répétition	Taille de génome séquencé (Gb)	Couvertue théorique*	Pourcentage des reads mappés de façon unique	Pourcentage de reads en paire par rapport aux reads mappés de façon unique	Pourcentage de reads en single par rapport aux reas mappés de façon unique
B73_NS_rep1	49,8	23,1	66,6	98,2	1,8
B73_NS_rep2	46,6	21,7	65,4	97,8	2,2
B73_NS_rep3	50,8	23,6	67,3	98,4	1,6
B73_S_rep1	45,2	21,0	67,3	97,9	2,1
B73_S_rep3 (30S)	46,0	21,4	67,1	97,9	2,1
B73_S_rep3 (34S)	52,1	24,2	71,6	98,4	1,6
F2_NS_rep1	45,4	21,1	48,8	93,6	6,4
F2_NS_rep2	46,7	21,7	49,8	93,8	6,2
F2_NS_rep3	45,9	21,3	48,7	93,4	6,6
F2_S_rep1	47,3	22,0	49,2	93,8	6,2
F2_S_rep2	45,5	21,2	48,9	93,3	6,7
F2_S_rep3	45,3	21,1	50,0	94,0	6,0
F331_NS_rep1	50,7	23,6	45,6	92,9	7,1
F331_NS_rep2	55,0	25,6	48,1	93,7	6,3
F331_NS_rep3	51,4	23,9	47,5	93,7	6,3
F331_S_rep1	47,4	22,1	47,2	93,1	6,9
F331_S_rep2	50,9	23,7	46,9	93,2	6,8
F331_S_rep3	54,2	25,2	48,0	93,6	6,4

Pour chaque génotype, le nombre des cytosines couvertes entre 4 et 25X dans les 6 réplicas est présenté dans le tableau IV.2. Afin de caractériser les variations de méthylation globales pour chacun des 3 génomes, nous avons dans un premier temps analysé le pourcentage de méthylation moyen pour chaque génotype et chaque contexte, pour les individus « stressés » et les individus « non

stressés ». Nous avons ensuite estimé le nombre des cytosines méthylées par rapport au nombre total des cytosines pour chaque contexte (Tableau IV.2), une cytosine étant considérée comme méthylée si son niveau de méthylation est supérieur à 80% pour CG, supérieur à 70% pour CHG et supérieur à 5% pour CHH (Gent *et al.*, 2013, Niederhuth *et al.*, 2016).

Ces deux paramètres montrent que les plantes du lot « stressé » ont un niveau de méthylation global plus élevé que les plantes du lot « non stressé », pour les 3 génotypes (Tableau IV.2). Cependant, la différence entre les niveaux de méthylation moyens entre lot « stressé » et lot « non stressé » est plus élevée pour les deux génotypes sensibles (par exemple, 2,80% pour B73 et 3,66% pour F2 pour le contexte CHG) en comparaison avec F331 (1,13% pour le contexte CHG). Ceci est vrai pour les 3 contextes mais particulièrement marqué pour le contexte CHG (Tableau IV.2). Les niveaux de base de méthylation chez les individus « non stressés » montrent une variabilité entre les 3 génotypes : B73 a un niveau de méthylation plus fort que les deux autres génotypes (81% contre 75 et 76% pour CG, 68% contre 64 et 63% pour CHG) (Tableau IV.2).

L'analyse des niveaux de méthylation pour chaque génotype montre une variabilité entre répétitions d'un même traitement, essentiellement pour B73 (individus NS1 et S3) et pour F331 (individus NS1 et S3) (Tableau IV.2). Comme pour B73, cette différence pourrait être liée à une sensibilité différente des individus liée à un effet maternel (non maîtrisé ici), à une différence spatiale des conditions, ou à une différence légère de stade de développement au moment de l'entrée des plantes dans la chambre climatisée, ou au moment du prélèvement. Cependant, les variations sont ici analysées pour le génome complet, mais les régions analysables sont vraisemblablement différentes entre les trois génotypes, du fait de l'alignement des lectures sur le pangénome de B73/F2 et non sur chacun des génomes. Les différences observées pourraient donc provenir d'une différence de contenu des régions analysables (par exemple, moins d'éléments transposables analysables pour F2 et F331 par rapport à B73). Un moyen de déchiffrer cette variabilité serait de regarder la variation de couverture et de méthylation le long des chromosomes, en moyenne et par individu.

Tableau IV.2 : Nombre des cytosines analysées pour les 3 génotypes B73, F2 et F331 (NS : Non Stressé et S : Stressé)

Répétition	Nombre des cytosines couvertes entre 4 et 25X dans les 6 replicas			Nombre des cytosines méthylées			Pourcentage des cytosines méthylées par rapport aux nombre total des cytosines			Niveau de méthylation de l'ADN moyenné sur l'ensemble des cytosines du génome		
	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH	CG	CHG	CHH
B73_NS_rep1				44 688 548	29 258 868	10 542 362	82,08	65,82	7,10	82,08	68,98	1,17
B73_NS_rep2				43 018 217	27 957 895	9 489 133	79,01	62,89	6,39	80,93	67,00	0,98
B73_NS_rep3				42 620 101	28 137 326	9 790 534	78,28	63,30	6,60	80,78	67,23	0,98
B73_S_rep1	54 444 473	44 451 905	148 406 979	44 700 947	30 349 943	10 682 882	82,10	68,28	7,20	82,79	71,01	1,20
B73_S_rep3 (30S)				44 919 371	30 479 139	10 673 699	82,50	68,57	7,19	82,94	71,26	1,19
B73_S_rep3 (34S)				44 012 376	29 482 710	10 621 782	80,84	66,32	7,16	81,50	69,35	1,09
Différence entre les moyennes des S et NS							2,02	3,72	0,48	1,15	2,80	0,12
F2_NS_rep1				21 102 130	15 880 023	5 650 582	72,28	60,25	7,11	75,22	63,95	1,43
F2_NS_rep2				21 252 153	15 760 777	5 717 579	72,79	59,80	7,19	75,41	63,66	1,41
F2_NS_rep3				21 120 943	15 984 812	5 944 071	72,34	60,65	7,47	75,3	64,24	1,53
F2_S_rep1	29 196 265	26 354 921	79 526 312	21 942 281	17 018 447	6 630 886	75,15	64,57	8,34	76,62	67,19	1,77
F2_S_rep2				21 729 506	17 366 222	7 066 281	74,43	65,89	8,89	76,28	67,88	1,94
F2_S_rep3				22 076 692	17 227 232	6 810 068	75,61	65,37	8,56	76,80	67,75	1,79
Différence entre les moyennes des S et NS							2,60	5,04	1,34	1,26	3,66	0,38
F331_NS_rep1				20 907 365	13 485 563	4 682 593	76,32	60,63	6,31	77,1	63,98	1,12
F331_NS_rep2				20 250 985	12 892 744	4 705 009	73,92	57,97	6,34	75,66	62,06	1,04
F331_NS_rep3				20 189 677	13 037 274	4 582 206	73,70	58,62	6,17	75,92	62,5	1,05
F331_S_rep1	27 394 596	22 241 749	74 213 275	20 390 103	13 360 464	4 642 385	74,43	60,07	6,26	75,98	63,58	1,15
F331_S_rep2				20 421 216	13 310 775	4 610 635	74,54	59,85	6,21	75,86	63,4	1,14
F331_S_rep3				20 859 711	13 787 751	5 087 951	76,15	61,99	6,86	76,98	64,95	1,18
Différence entre les moyennes des S et NS							0,39	1,56	0,17	0,05	1,13	0,09

La détection des régions différenciellement méthylées entre plantes du lot « stressé » et plantes du lot « non stressé » montre que le froid induit des variations locales de méthylation pour chaque génotype. Le nombre de DMR détectées dans le génotype B73 est plus élevé (750) que pour F2 (239) et pour F331 (35). La normalisation du nombre de DMR par rapport au nombre de cytosines analysables pour chacun des génotypes (Tableau IV.3), qui prend en compte les différences de quantité des cytosines analysables, montre la même tendance. La comparaison reste cependant limitée, du fait de l'impossibilité d'analyser les régions génomiques spécifiques de ces génotypes. Chez tous les génotypes, on observe à la fois des cas d'hyper- et d'hypométhylation. Mais les proportions relatives de ces deux cas varient dans les différents génotypes. Chez B73 et F331, on observe un nombre équivalent de chaque catégorie (162 contre 168 et 9 contre 11, respectivement pour B73 et F331 pour le contexte CHG). En revanche, chez F2, le contexte CG présente plus de cas d'hypométhylation (79 contre 32) et le contexte CHG présente plus de cas d'hyperméthylation (87 contre 41) (Tableau IV.3).

Pour chacun des trois génotypes, la taille des DMR est plus importante en contexte CHG qu'en contexte CG. Dans le cas de l'hyperméthylation, le pourcentage de différence de méthylation est similaire entre F2 et F331 en CG et CHG qu'entre B73 et F2, à l'inverse à ce que nous avons obtenu avec les niveaux globaux de méthylation. En hypométhylation, la différence est plus similaire entre F2 et F331 pour le CG mais entre F2 et B73 pour le CHG (Tableau IV.3).

Afin d'analyser si les mêmes régions sont modifiées dans les 3 génotypes, une intersection entre ces régions a été réalisée. Aucun chevauchement n'a pu être mis en évidence entre ces régions. Ceci pourrait signifier que les DMR retrouvées sont spécifiques à chaque génotype. Mais ceci pourrait également être dû à une absence de chevauchement des régions analysables avec nos critères pour chaque génotype. Un moyen pour contourner ce problème serait de ne pas restreindre l'analyse aux cytosines analysables entre les 6 répliques de chaque génotype, mais de l'étendre à toutes les cytosines bien couvertes (4x-25x) chez au moins un individu stressé et un individu non stressé. La détection des DMR après alignement des lectures sur les génomes de chaque génotype pourrait également permettre de détecter des DMR dans des régions orthologues, mais ayant divergé en séquence, notamment du fait de l'insertion différentielle des éléments transposables.

Tableau IV.3 : Caractéristiques des DMRs hyperméthylées (**A**) et hypométhylées (**B**) chez les plantes de lot stressé en comparaison au lot non stressé à la sortie du froid chez les 3 génotypes B73, F2 et F331. Les valeurs présentées correspondent à la médiane des valeurs des DMR.

A.

Génotype	CG			CHG		
	B73	F2	F331	B73	F2	F331
Nombre de DMRs	192	32	20	162	87	9
Nombre de DMRs normalisé par nombre de cytosines	3,53	1,10	0,73	3,64	3,30	0,40
Taille de la DMR (bp)	400	400	400	800	800	700
Différence de méthylation (%)	29	32	31	35	27	27
Nombre de cytosines par DMR	21	27	30	34	39	23
Nombre de DMCs par DMR	6	7	7	7	6	6

B.

Génotype	CG			CHG		
	B73	F2	F331	B73	F2	F331
Nombre de DMRs	228	79	25	168	41	11
Nombre de DMRs normalisé par nombre de cytosines	4,19	2,71	0,91	3,78	1,56	0,49
Taille de la DMR (bp)	400	500	400	700	700	500
Différence de méthylation (%)	25	28	29	30	29	26
Nombre de cytosines par DMR	30	28	28	39	38	37
Nombre de DMCs par DMR	7	6	7	7	8	6

2. Conclusions et perspectives

Cette analyse globale préliminaire de la méthylation de l'ADN, à l'échelle du génome met en évidence que le génotype tolérant F331 présente le moins de variations de méthylation entre plantes « stressées » et plantes « non stressées » en comparaison avec B73 et F2. Ceci est prometteur, et pourrait s'expliquer par le fait que le génotype tolérant ne met pas en place une réponse aussi importante que les génotypes sensibles.

Par ailleurs, des variations structurales existent entre lignées de maïs et l'existence d'un seul génome de référence (B73) rend complexe l'alignement des lectures issues d'autres génotypes. Ainsi, beaucoup de séquences générées pour des génotypes de maïs ne sont pas alignées sur le génome de référence et sont jetées et le pourcentage du génome analysé n'est pas le même entre génotypes. De plus, la

comparaison des signaux reste limitée aux régions communes aux trois génomes, ce qui limite la comparaison aux régions non élément transposables (Figure IV.2). La disponibilité des génomes de référence pour les génotypes F2 et F331 (en cours dans le cadre du programme *Amaizing*) va permettre d'affiner l'alignement des lectures pour ces deux génotypes et par la suite une meilleure détection des variations de méthylation à l'échelle de chaque génome. Ceci permettra d'une part de décrire les variations le long des chromosomes de façon plus fiable, mais aussi d'annoter les régions variables en lien avec l'annotation de chaque génome, et non celle de B73, qui est une mauvaise référence, en particulier pour l'annotation de la position des éléments transposables des autres génotypes. Ceci permet de mieux analyser l'annotation des DMR et de dire si la différence de méthylation induite par le froid est liée à la variation structurale observée entre les 3 génotypes, en particulier en lien avec les éléments transposables.

Les analyses sur épigénome et transcriptome complet sont exploratoires et permettent de découvrir de nouveaux locus potentiellement impliqués dans la réponse au froid chez le maïs. Afin de compléter cette approche globale, il serait intéressant d'étudier les variations de méthylation à côté des gènes de réponse au froid décrits dans la littérature, ainsi que celles des gènes localisés dans les régions QTLs révélées *via* des analyses génétiques. Enfin, l'analyse de deux génotypes sensibles et un génotype tolérant donne des pistes sur la question du lien entre méthylation et phénotype, mais celles-ci devront être validées chez d'autres génotypes sensibles et tolérants, afin de tester si les régions détectées sont spécifiques de certains génotypes ou génériques.

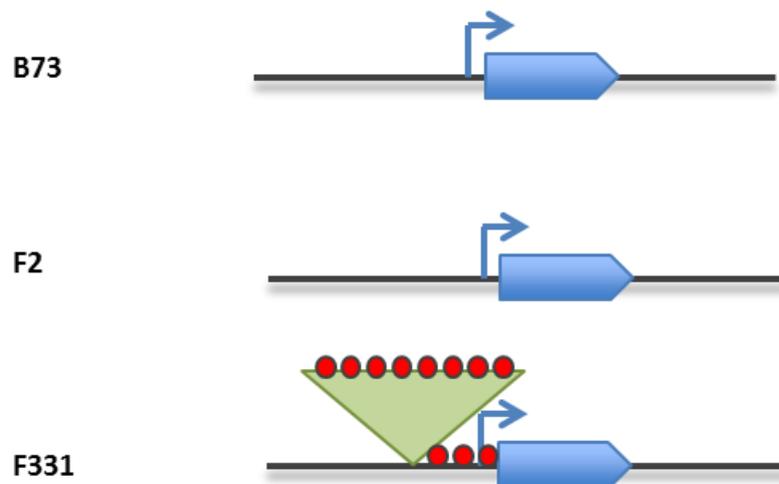


Figure IV.2 : Spreading de la méthylation de l'ADN à côté des ET permettant la détection de DMR liée à des variations structurales entre génotypes

La DMR localisée dans l'ET ne peut pas être révélée entre les trois génotypes, du fait de la présence/absence de l'insertion. Par contre, la DMR induite par le « spreading » de l'élément peut être détectée

Chapitre V.

Discussion générale et perspectives

Mes travaux de thèse constituent les premières études, au sein de l'équipe DyGAP, des variations du méthylome en réponse aux contraintes de l'environnement et de leur implication dans la régulation de l'expression du génome.

Bien que d'autres études aient été réalisées sur ce thème chez le maïs, mon projet de thèse était ambitieux et original du point de vue (i) de la contrainte appliquée : un froid prolongé au début du développement, ayant un intérêt à la fois par rapport à la physiologie associée, et à son implication en agronomie, (ii) de la méthodologie utilisée: une analyse du méthylome genome-wide (BS-seq) avec de fortes profondeurs de séquençage, avec trois répétitions par condition, (iii) du nombre de génotypes analysés : trois génotypes, dont seulement un avait un génome de référence au début de ma thèse, (iv) des stades analysés : c'est la première étude chez le maïs où les changements de méthylome induits par une contrainte abiotique est suivi à travers le développement et sur plusieurs générations. Cette originalité m'a permis d'acquérir des résultats nouveaux et originaux concernant la réponse du méthylome au froid. Mais la contrainte de temps de la thèse ne m'a pas permis de finaliser toutes les questions abordées.

Au commencement de ma thèse, les analyses sur les variations de méthylation de l'ADN après application des contraintes abiotiques étaient essentiellement fondées sur l'analyse des gènes de réponse au stress dans le but d'apporter des explications sur la régulation de la transcription. Les quelques études réalisées sur génome entier ont émergé au cours de ma thèse et avaient pour démarche de rechercher des régions différentiellement méthylées (DMR) entre individus « stressés » et « non stressés », sous la forme de petites régions détectées avec un signal significatif, puis de classier ces DMR en fonction de l'annotation des génomes, et d'analyser le lien avec l'expression des gènes. Alors que la méthylation de l'ADN participe à la régulation des gènes (Roudier *et al.*, 2009), le lien de causalité entre variation de méthylome et de transcriptome induites par une contrainte n'a pas pu être mis en évidence.

Jusqu'en août 2018, j'ai suivi la même démarche. J'ai réalisé plusieurs fois la détection de DMR, en modifiant le filtre des cytosines, le nombre de DMC par DMR, etc. Cela m'a permis de détecter des DMR fiables, que j'ai ensuite analysé en lien avec l'annotation et l'expression des gènes. Avec les données dont je disposais, j'ai également souhaité réaliser une analyse d'exploration à l'échelle globale du génome. Cette analyse m'a permis d'acquérir des résultats nouveaux, qui suggèrent que le méthylome est régulé à différentes échelles génomiques suite à un stress froid chez le maïs.

Augmentation de la méthylation de l'ADN à l'échelle du chromosome suite au froid prolongé

L'étude des variations de la méthylation le long du génome a permis de mettre en évidence une hyperméthylation tout le long des chromosomes. Les cytosines des différents contextes sont affectées, mais pas dans les mêmes régions chromosomiques, et la réponse de chaque contexte suit la

composition du chromosome de manière particulière : le niveau des CG et CHG est augmenté dans les régions péri-centromériques (hétérochromatine fermée), tandis que le niveau des CHH est augmentée dans les séquences d'éléments transposables des régions géniques (hétérochromatine transcriptionnellement active). Ceci suggère que différentes voies régulant le niveau de méthylation (maintenance CG, maintenance CHG ; et RdDM) sont affectées de façon parallèle en réponse au froid et indique que l'hyperméthylation est une réponse importante au froid. Alors que d'autres études ont montré un changement global de méthylation suite à l'exposition à une contrainte abiotique (Steward *et al.*, 2002), notre étude montre pour la première fois la répartition génomique des variations, et ce pour les trois contextes. Cette approche a donné des résultats originaux, et pourrait constituer une nouvelle méthode de détection des variations de méthylation de l'ADN en réponse aux contraintes abiotiques. Le profil observé chez le maïs est la conséquence de l'organisation particulière de son génome, notamment la présence d'ET proches des gènes qui génère l'existence d'une chromatine accessible pour la transcription des gènes, régulée par la voie RdDM. L'observation d'une hyperméthylation chez les trois génotypes de maïs analysés suggère que cette réponse est générale au maïs. L'analyse des profils chromosomiques pour les deux autres lignées permettra de déterminer si les mêmes régions sont affectées par les mêmes mécanismes. Les îlots CHH ont été observés pour la première fois chez le maïs (Gent *et al.*, 2013), et ont depuis été observés chez différentes espèces (Niederhuth *et al.*, 2016), mais son importance semble être variable entre espèces. La comparaison des profils chromosomiques de la différence de méthylation entre individus stressés et non stressés entre plusieurs espèces de composition génomique différente devrait permettre de déterminer si les mêmes mécanismes sont modifiés, et si la composition du génome oriente les profils de méthylation induits par les contraintes environnementales. Le modèle d'étude dont nous disposons pourrait apporter des éléments de réponse à cette question. En effet, les différents génotypes analysés ont des tailles de génome, et peut-être des organisations structurales différentes. L'annotation des variations de méthylation en terme de localisation (gène, ET) et de stabilité entre individus permettra de déterminer si les mêmes séquences et mécanismes sont en jeu dans les différents génotypes, et si les profils chromosomiques sont similaires.

Réflexions sur le rôle biologique de l'hyperméthylation observée

L'augmentation globale du méthylome du maïs suite à l'exposition au froid affecte principalement les séquences d'éléments transposables. La mise en place de plusieurs mécanismes différents indique que cette réponse est importante. Le ciblage des ET suggère que cette hyperméthylation pourrait être mise en place pour éviter les mouvements des ET, en renforçant leur contrôle. Ceci fait penser à une mise en place d'un mécanisme "préventif" au sein du génome pour assurer sa stabilité suite à un stress froid. De façon alternative, l'hyperméthylation pourrait être un effet secondaire de l'activation de certains ET, et l'augmentation globale de la méthylation du génome permettrait de contrôler les nouvelles insertions dès leur intégration dans le génome. L'observation d'une déstabilisation des transcrits des sous-unités PolIII pourrait indiquer que certains éléments transposables ont pu être activés. L'analyse des données de mRNA-seq devrait permettre de détecter si l'expression de certains ET a été augmentée suite au traitement froid, et pourrait permettre de proposer des ET candidats à analyser par des approches ciblées. Une analyse de mobilome (Lanciano *et al.*, 2017) pourrait également permettre de déterminer si des ET ont été activés suite au froid appliqué.

Réflexions sur l'origine des DMR hypométhylées chez les individus stressés

D'après les résultats de l'analyse globale du méthylome, avec l'approche locale (recherche de DMR), nous nous attendions à détecter majoritairement des locus qui augmentent en niveau de méthylation. Cependant, la détection des DMR révèlent l'existence de DMR hyper- et hypométhylées en proportions équivalentes (354 contre 396, respectivement). D'autres études montrent l'existence de ces deux types de DMR (Ganguly *et al.*, 2017, Secco *et al.*, 2015). Dans Secco *et al.* (2015), les hypométhylations sont en nombre moindre par rapport aux hyperméthylations (28 vs. 147). C'était le cas aussi dans l'étude de Ganguly *et al.* (2017) où ils détectent quasiment le même nombre de DMR dans les deux cas (8 hyper vs. 9 hypo) en CG et CHG. Cependant, les auteurs ne proposent pas d'explication pour comprendre le mécanisme de régulation de ce type de locus. Dans Ganguly *et al.* (2017), aucune mention n'est faite sur le sens de méthylation (hyper ou hypo) dans le texte principal. Cette information n'est donnée que dans les tableaux supplémentaires. Ceci suggère que les auteurs n'ont pas d'explication pour l'observation des différents cas. Avec les données dont nous disposons, les DMR que nous avons détectées nous semblent solides, et nous n'avons pas observé de biais particulier qui puisse expliquer la présence de DMR hypométhylées chez les individus stressés. La caractérisation en terme de localisation et de stabilité ne révèle pas de différence majeure entre cas d'hyper- et d'hypo méthylation, ce qui suggère qu'elles proviendraient de mécanismes similaires. Ceci n'est pas attendu puisque les mécanismes impliqués dans la méthylation sont différents de ceux impliqués dans la déméthylation. Les mécanismes pointés par nos observations sont plutôt des mécanismes de méthylation que de déméthylation. Ainsi, une hypothèse serait que ces DMR hypométhylées dans les individus stressés soient en fait des DMR hyperméthylées chez les individus non stressés. En effet, sans générer de « stress », la disposition des deux lots dans des endroits différents (serre et chambre climatisée) pourrait générer des différences suffisamment importantes pour générer des différences de méthylation. Le facteur ayant généré ces DMR reste cependant encore inconnu.

Enfin, une autre hypothèse serait que les mécanismes d'hyperméthylation soient induits par le froid, mais que cette hyperméthylation globale induise une diminution de la méthylation à certains locus dont la méthylation est moins critique, dû au nombre limité de protéines disponibles.

Le suivi des niveaux de méthylation des DMR hypo- et hyperméthylées donnera des informations complémentaires qui permettront de mieux caractériser la biologie de ces régions.

Réflexions sur l'origine de la variation observée entre individus

L'analyse de plusieurs réplicas dans chaque condition nous a permis d'analyser la répétabilité du signal entre individus de chaque condition. Pour cela, nous avons décidé de restreindre l'analyse aux cytosines couvertes entre 4x et 25x dans chacun des 6 réplicas. Ces paramètres sont assez stringents mais ils permettent de comparer le signal issu des mêmes cytosines. J'ai montré que la majorité des DMR entre individus « stressés » et « non stressés » sont variables entre individus d'un même traitement. Dans cette analyse, la stabilité du signal entre individus a été réalisée à partir du niveau de variation de l'ensemble des cytosines de la DMR. Une partie de la variabilité du signal pourrait venir du fait qu'au sein de ces régions, seules certaines cytosines sont stables. Pour caractériser les DMR, j'ai réalisé une étape de filtre qui a nécessité la détection des cytosines différentiellement méthylées (DMC). Nous n'avons pas souhaité analyser séparément les DMC car nous cherchions des différences de méthylation qui pourraient avoir un rôle biologique car à ce jour, le rôle biologique des DMC sur

l'expression des gènes n'est pas décrit et les études physiologiques s'intéressent peu à ce type de variation. Une solution pour mieux étudier la répétabilité des DMR serait d'utiliser ces DMC pour analyser leur variation entre individus, et ainsi déterminer si certaines cytosines ont un niveau stable au sein des DMR.

Le suivi des variations de méthylation au cours du développement et entre générations pourra se faire à l'échelle de la DMR ou de chaque cytosine. Enfin, l'utilisation de DMR à seulement 2 DMC pour le design du SeqCapEpi permettra à la fois de suivre des DMR fortes et plus faibles, et d'avoir accès au suivi de nombreux DMC de différents contextes. Cette analyse devrait donner des résultats nouveaux sur la stabilité des DMC induites. Enfin, si les DMC n'ont pas d'impact direct sur le phénotype, l'apparition dans le génome de nouveaux sites méthylés pourrait augmenter la fréquence des transitions C vers T, et contribuer ainsi à l'évolution du génome. Pour le moment, deux générations ont été réalisées. Il pourra être intéressant de continuer à réaliser des autofécondations sur ce matériel, pour analyser à la fois le devenir des DMR et des DMC.

Applications potentielles des variations du méthylome induites par le froid

Un des buts de mes travaux était de faire un lien entre les DMR et les variations phénotypiques observées entre les trois génotypes. Les résultats que j'ai présentés sont préliminaires, mais montrent une réponse différente des trois génotypes à l'échelle du génome. De façon intéressante, le génotype tolérant F331 montre une variation plus faible que les deux autres. Cette observation est prometteuse, et l'analyse des profils de méthylation le long des chromosomes en utilisant un alignement des lectures de BS-seq sur les génomes de chaque lignée devrait permettre de révéler les similarités et différences de réponse entre les trois génotypes. Ceci devrait également permettre de mieux estimer les niveaux de méthylation en condition standard, et ainsi de mieux comprendre si c'est la variation de méthylation entre stressés et non stressés, ou le niveau de base plus la variation qui est important. L'analyse de la position des DMR par rapport aux gènes permettra d'analyser les ontologies des gènes proches des DMR, ainsi que leur niveau d'expression dans les trois génotypes, et devrait permettre d'appréhender la question de l'intégration de l'information apportée par l'épigénotype dans le phénotype, comme proposé par (Springer, 2013) (Figure 1).

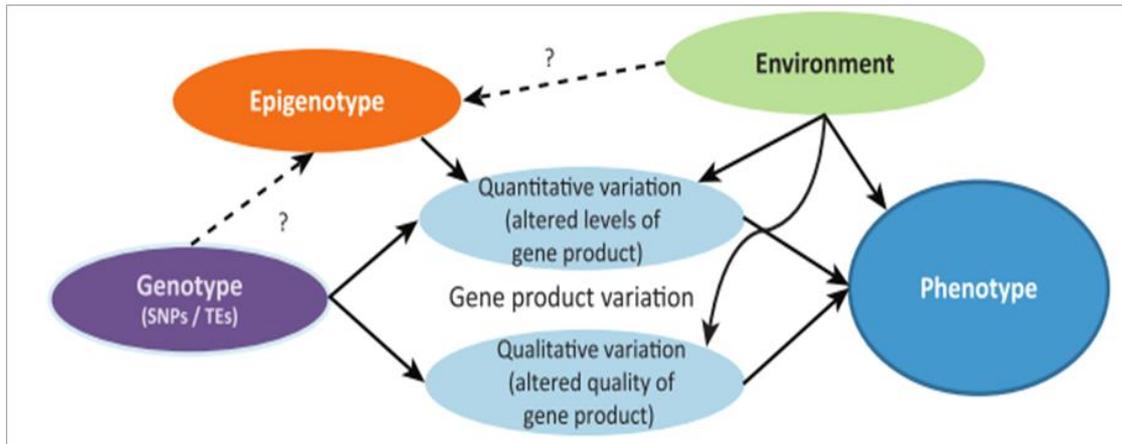


Figure 1 : Le phénotype d'un organisme est contrôlé par les variations du génotype, de l'épigénotype et de l'environnement (Springer, 2013)

Les variations du génotype incluent les variations de type SNPs, les sites InDels et les ETs.

L'environnement peut directement affecter le phénotype, il affecte les niveaux d'expression des gènes ou bien résulte dans des modifications post-traductionnelles du produit du gène lui-même.

L'épigénotype pourrait résulter d'interactions avec le génotype (comme par exemple la proximité aux ET) ou des signaux environnementaux. Les lignes en pointillés indiquent les interactions qui pourraient influencer l'épigénotype.

La détection des DMR dans chaque génotype indépendamment, après alignement des séquences sur chaque génome de référence, permettra de mieux estimer la quantité de DMR par génotype, et de déterminer si ce nombre est associé aux phénotypes observés. Ceci pourrait permettre de caractériser des locus de la réponse au froid, qui pourraient être utilisés pour cribler des panels pour choisir des lignées d'intérêt. En particulier, il sera intéressant d'analyser le lien entre DMR et ET, afin de caractériser des épiallèles facilités ou obligatoires, qui ont une probabilité de stabilité intergénérationnelle. Les deux générations de maïs que nous avons générées vont permettre d'étudier la transmission de ces épiallèles. La caractérisation d'épiallèles avec une transmission élevée pourrait avoir une potentielle application dans les programmes d'amélioration variétale comme source de nouvelle variation phénotypique ou comme marqueur.

Annexes

Annexe 1.

Analyse bibliographique des régions génomiques impliquées dans la réponse aux contraintes abiotiques chez le maïs

Les études relatives au froid sont en blanc. Les études relatives à d'autres contraintes abiotiques que le froid sont indiquées en gris clair.

	Régions génomiques	Contrainte abiotique	Conditions	Méthode de détection	Référence
Protéines	ZmCDPK1 Adh1	Froid (sévère)	Rapide passage à 5°C au stade 4-5 feuilles	Isolation ADNc - Hybridation Northern - Séquençage Sanger	(Kusano <i>et al.</i> , 1995)
	ZmMAPKKK ZmCLC-D ZmRLK	Froid (sévère)	6°C pendant 2, 6 et 12 h au stade 3 feuilles	Isolation ADNc – RT PCR – qRT PCR	(Yang <i>et al.</i> , 2010)
	bZIP	Froid (sévère) ABA NaCl	Rapide passage à 5°C au stade 4-5 feuilles	Isolation ADNc - Hybridation Northern - Séquençage Sanger	(Kusano <i>et al.</i> , 1995)
	ZmDREB1A (facteur de transcription)	Froid NaCl Sécheresse ABA	4 °C pour des plantules âgées de 7 jours depuis le semis	Isolation ADNc – phylogénie – hybridation Southern et Northern	(Qin <i>et al.</i> , 2004)

	ZmMPK17	Froid ABA /AS/ AJ Ehylene H ₂ O ₂	4 °C sur des plantules âgées de 2 semaines	Isolation du gene MAPK – hybridation Northern	(Pan <i>et al.</i> , 2012)
	ZmCBF3 (expression spécifique dans la racine)	Froid Chaleur (37°C) Sécheresse ABA	4°C pendant 6h	Expression de genes GUS	(Liu <i>et al.</i> , 2015)
	ZmMPK3	NaCl Sécheresse Métaux lourds UV Blessures		Phylogénie- RT PCR – expression YFP	(Wang <i>et al.</i> , 2010)
Gènes	GRMZM2G30107 GRMZM2G078839 GRMZM2G146475 GRMZM2G138907	Froid	8/6 °C pendant 14h au stade 3 feuilles (ETH-DH7 (génotype tolérant) et ETH-DL3 (génotype sensible))	Isolation ADNc - Gene Ontology - RTPCR	(Sobkowiak <i>et al.</i> , 2014)
	10 ZmRCI2 (tous ces gènes contiennent les éléments cis, connus pour leur réponse aux contraintes abiotiques : ABRE,	Sécheresse		Blastp (<i>Arabidopsis</i> et <i>Oryza</i>) pour trouver les gènes RCI2 chez le maïs - Structure du gène et localisation sur le	(Zhao <i>et al.</i> , 2014d)

	LTRE ou DRE dans leur région promotrice (2 Kb)			génomique (V2) – analyses phylogénétiques - Validation qRt PCR	
Transcrits	Transcrits liés à différentes fonctions physiologiques (rythme circadien, métabolisme secondaire, signaux de transduction)	Froid (modéré)	Froid court : 14°C/12°C pendant 4h et Froid prolongé : 14°C/12°C pendant 28h au stade 3 feuilles (<i>Z. mays indentata</i> : dent CM 109 inbred (modèle pour les études physiologiques chez le maïs)	Puce d'ADNc – Gene Ontology – RT PCR	(Trzcinska-Danielewicz et al., 2009)
	Transcrits liés à différentes fonctions physiologiques (protéines de transport, métabolisme de la photosynthèse, signaux de transduction)	Froid Sécheresse NaCl Blessure Chaleur (37°C) H ₂ O ₂	6 °C pendant 48 h au stade 3 feuilles	Isolation ADNc – RT PCR	(Nguyen <i>et al.</i> , 2009)
	Transcrits impliqués dans la régulation du métabolisme des gibbérellines (des gènes communs et des gènes propres à la réponse de chaque contrainte sont identifiés. Les mêmes gènes peuvent être régulés de manière différentes pour les 3 contraintes)	Froid Sécheresse NaCl	4 °C pendant 24 h au stade 3 feuilles	Séquençage Illumina – analyses RNAseq	(Shan <i>et al.</i> , 2013)

QTLs	2 QTLs sur le chromosome 3 et 8 pour le caractère mesuré de fluorescence du PSII (Fv/Fm)	Froid	10 °C/1h pour des plantules âgés de 14 jours depuis le semis (famille issue du croisement entre lignées de maïs doux W6786 su1Se1 et IL731A su1se1).	Interval mapping - critère Akaike (PLABQTL)	(Pimentel <i>et al.</i> , 2005)
	40 QTLs répartis sur les 10 chromosomes contrôlant la vigueur des racines et des plantules	Froid	16°C pendant 12 h au moment de la germination (famille issue du croisement entre deux parents dentés Lo964 (tolérant au froid au moment de la germination) et Lo1016 (sensible au froid au moment de la germination))	Composite Interval Mapping	(Hund <i>et al.</i> , 2004)
	3 QTLs sur les chromosomes 2, 4 et 8 contrôlant les caractères de fluorescence du PSII, poids de la matière sèche et le taux d'anthocyanine, sous conditions de froid	Froid	14°C/14h et 8°C /10h au moment du semis (famille issue du croisement entre un parent corné EP42 (intermédiaire) et un parent denté A661 (sensible au froid))	Composite Interval Mapping (metaQTLs)	(Rodríguez <i>et al.</i> , 2013)
ARN non	ARNInc	Sécheresse Salinité		Séquençage ARN - Gene Ontology – identification de IncARN – validation RT PCR	(Forestan <i>et al.</i> , 2016)

		Sécheresse et salinité			
	ARNs (23 et 24 nucléotides issus des gènes de type RMR-6) en amont des gènes hautement transcrits	Sécheresse Salinité Sécheresse et salinité		Séquençage ARN et ARNs - Gene Ontology	(Lunardon <i>et al.</i> , 2016)
Eléments transposables	Les éléments transposables des familles joemon, Naiba, gyma, etug et Zm00346 sont enrichis dans le gènes activés par le froid	Froid UV (50°C) Chaleur NaCL	5°C pendant 16h au stade plantule	Séquençage ARN et Q PCR	(Makarevitch <i>et al.</i> , 2015)
	Eléments transposables de la famille hAT	Culture des tissus et traitement à l'agent 5-aza-2-deoxycytidine		Isolation ADNc – phylogénie	(Smith <i>et al.</i> , 2012)
	Eléments transposables des familles: Gypsy, Copia, LTR, DTA/hAT	Sécheresse Salinité Sécheresse et salinité		Séquençage ARN – validation RT PCR	(Forestan <i>et al.</i> , 2016)

Annexe 2.

Détail des dispositifs expérimentaux

1. Détails des différentes étapes de la culture

Après une germination (4 jours) dans des boîtes de pétri en conditions standard (serre), les plantes ont été mises à pousser dans des pots de 4,5L. Au stade levée, deux plantules ont été placées par pot afin d'être sûrs d'avoir une plante par pot après le séjour au froid. Une étape d'éclaircissage est réalisée à la sortie du froid pour ne garder qu'une plante par pot pour la suite de la culture.

Un filet à grandes mailles a été placé dans chaque pot avant de le remplir de terreau, afin d'éviter d'abimer les racines lors de la transplantation des plantes dans des pots plus grands (18L) pour finir leur développement en serre (Figure 1). Ce dispositif a aussi été mis en place pour le lot « non stressé » afin d'obtenir des conditions comparables entre les deux lots.



Figure 1 : Test de transplantation des plantes en motte à partir du pot de 4,4L au pot de 18L au moment de la mise en commun des deux lot "stressé" et "non stressé" en serre

Le filet bleu est installé dans le pot de 4,5L pour éviter d'abimer les racines.

Lot « stressé »

Le lot « stressé » a été placé dans des conditions de culture « froid », en chambre climatisée (jour : 16h à 14°C ; nuit : 8h à 10°C) pour une durée de 21 jours. Les plantes ont été réparties en 3 blocs de 24 pots chacun (8 pots x 3 géotypes, Figure 2). Le nom des pots est indiqué par une lettre alphabétique (de A à F) respectivement pour les différents sous blocs. La durée et le commencement du jour dans la chambre de culture ont été calés sur celles de la serre. Afin de limiter l'effet de l'hétérogénéité spatiale de l'éclairage et de la température dans la chambre de culture, la position des pots a été modifiée manuellement tous les 4 jours selon un plan précis (Figure 3).



Figure 2 : Détail de la disposition du « lot froid » à l'entrée en chambre climatisée
 Les lettres montrent les noms des pots pour chaque génotype B73 (vert), F2 (jaune) et F331 (violet).
 Les lettres en gras montrent les plantes avec des thermocouples. La photo montre l'état des plantes au stade levée à l'entrée en chambre climatisée (30 mars 2015).

Lot « non stressé »

Le lot « non stressé » a été installé en serre (jour : 16h à 24°C ; nuit : 8h à 18°C). L'ensemble du lot a été répété trois fois, avec trois dates de semis décalés. Les dates de semis ont été estimées à partir de calculs de temps thermiques selon la méthode de Parent *et al.* (2010): il faut compter environ 3 jours équivalents à 20°C ($D_{20^{\circ}\text{C}}$) pour faire une feuille en serre chaude. Suite à ces calculs, il aurait fallu théoriquement semer le lot « non stressé » 16 jours après le lot « non stressé ». En pratique, nous avons réalisé trois semis décalés, à +14, +16 et +21 jours. Pour chaque date, 3 blocs de 18 pots ont été installés (6 pots x 3 génotypes) (Figure 4).

Appariement des lots « stressé » et « non stressé »

Après 21 jours de traitement froid, les conditions de la chambre climatisée ont été modifiées afin de permettre aux plantes du lot « stressé » de s'acclimater aux conditions de la serre. Cette phase a duré 3 jours, avec une augmentation progressive de la température : jour 1 : 15,5 °C J./14 °C N., jour 2 : 18 °C J./16 °C N., jour 3 : 20,5 °C J./18 °C N., matin du jour 4 (jour du repiquage) : 22 °C J./20 °C N.

Avant le commencement de cette phase d'acclimatation, le nombre de feuilles dégainées, le nombre de feuilles visibles et la longueur de la dernière feuille dégainée ont été notés pour les plantes des trois lots « non stressé » et du lot « stressé », afin de déterminer le lot « non stressé » ayant l'état de développement le plus proche du lot « stressé ». La date 3 était très décalée du « lot stressé », et a été écartée de suite (Figure 5). Le tableau 1 détaille les notations des lots « stressé » 1 et 2. Le développement des plantes du lot « non stressé » 1 était quasi synchrone avec celui du lot de la chambre climatisée. Cependant, nous savons que la phase d'acclimatation, ainsi que l'état de stress des plantes « stressées » induiraient un retard supplémentaire. De plus, les plantes de la date1 étaient aussi très étioilées (Figure 5). Nous avons ainsi sélectionné le lot « non stressé » 2 pour la mise en commun avec le lot « stressé » et les prélèvements à la sortie du froid.

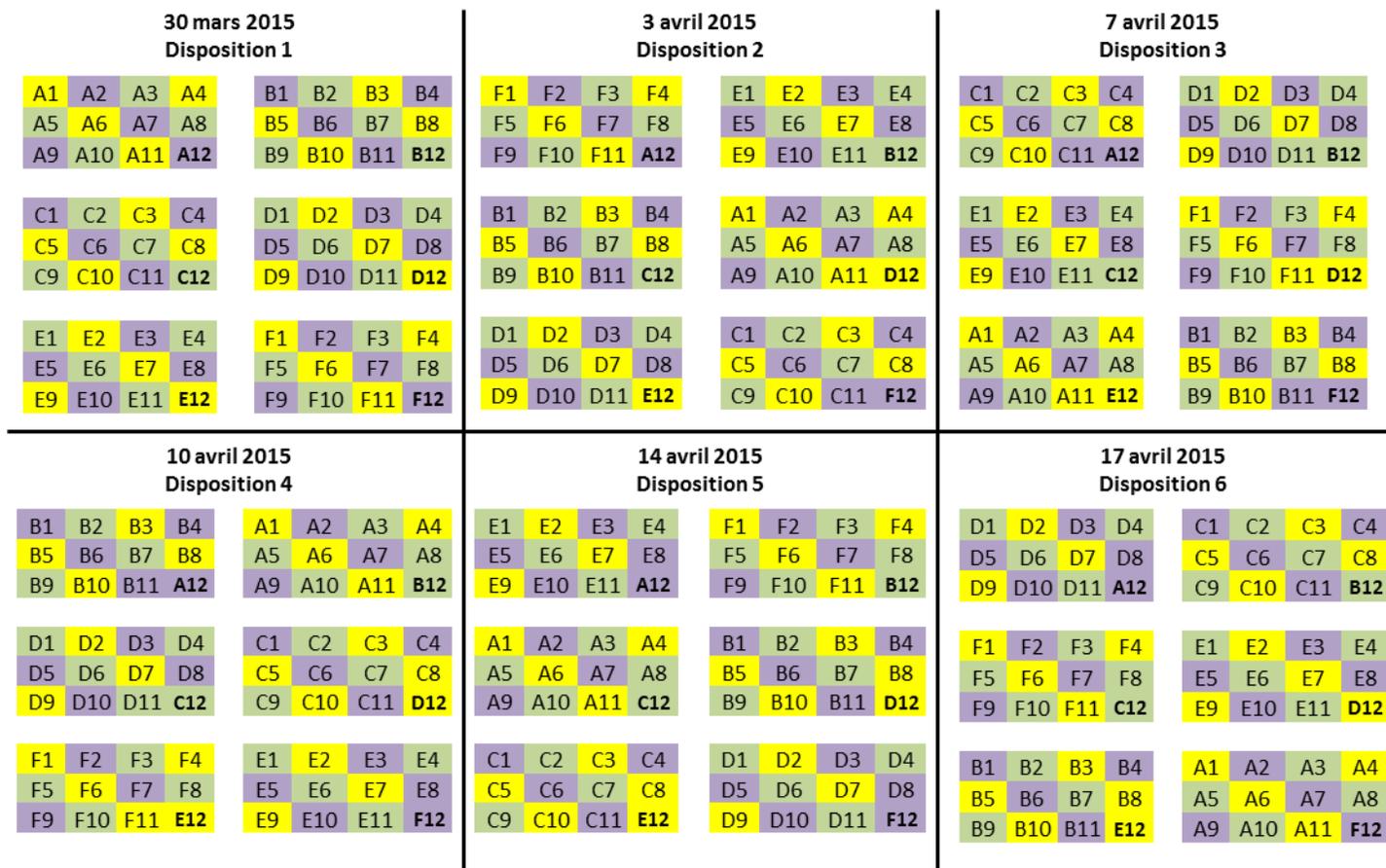


Figure 3 : Détail de la position des génotypes après chaque rotation dans la chambre de culture
Les plantes avec les thermocouples (en gras) ne changent pas de position

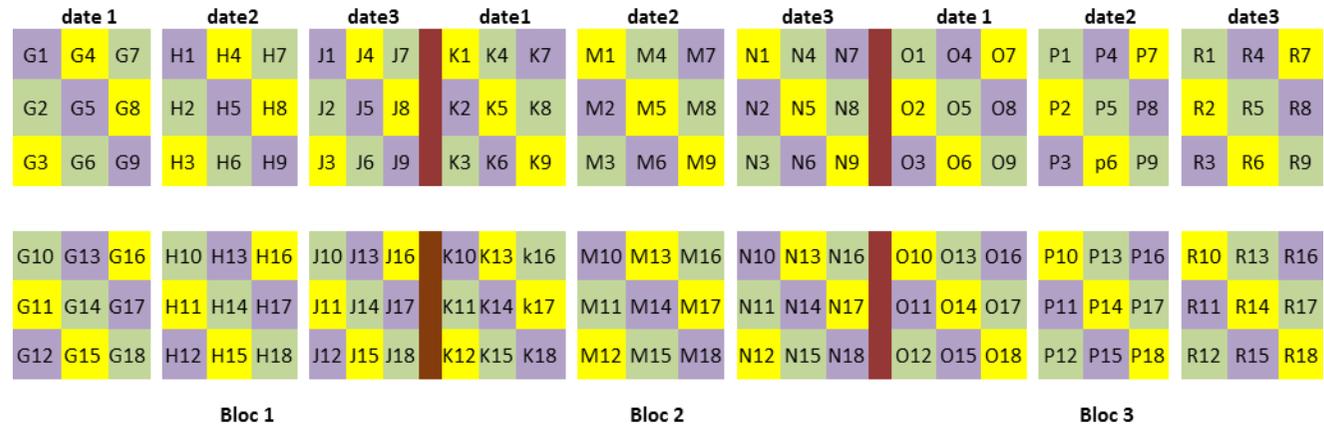


Figure 4: Détail de la disposition du lot « non stressé » suivant 3 dates différentes de semis

Les numéros montrent les noms des plantes pour chaque génotype B73 (vert), F2 (jaune) et F331 (violet).

Date 1: semis du 06/04/2015, date 2: semis du 08/04/2015, date 3 : semis du 13/04/2015.

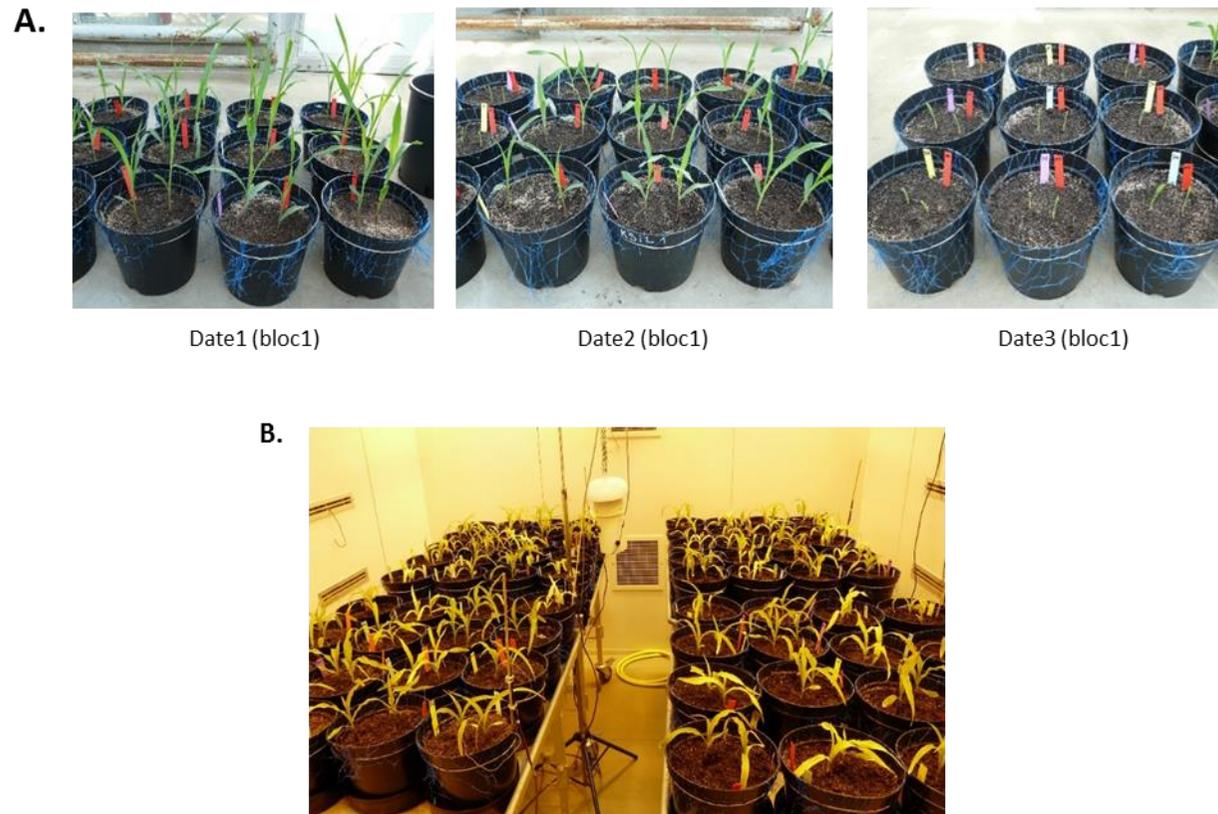


Figure 5: Etat du développement des plantes en serre pour les 3 dates de semis au moment de l'arrêt du traitement froid en chambre climatisée (A) et en chambre climatisée (B) après 21 jours de traitement froid

Tableau 1 : Notation des feuilles dégainées (fd), des feuilles visibles (fv) et de la longueur de la dernière feuille dégainée (lodf) pour les génotypes B73, F2 et F331. Les valeurs indiquent les médianes pour chaque bloc.

	Bloc	Serre date1			Serre date2			Chambre climatisée		
		fd	fv	lodf	fd	fv	lodf	fd	fv	lodf
B73	1	2	4	132,5	1	3	121	2	4	76,5
	2	2	4	70	1	3	77	2	4	56
	3	2	3	115	1	3	57,5	2	4	72
F2	1	2	4	162,5	1	3	108	2	4	102
	2	2	4	80	1	3	72	2	4	86,5
	3	2	4	52,5	1	3	47	2	4	86
F331	1	2	4	120	2	4	30	2	5	47,5
	2	2	4	142,5	1	4	17	2	5	34
	3	2	4	95	1	3	61,5	2	5	50

Ces notations nous ont également permis de choisir les plantes à éliminer lors de l'éclaircissage. Pour chaque génotype, l'écart de la longueur de la dernière feuille dégainée de chaque plante à la médiane du bloc a été calculé, et pour chaque pot, la plante avec la valeur la plus faible a été conservée. Enfin, les blocs des lots « stressés » et « non stressés » ont été appariés en fonction de leur valeur médiane. Ainsi, le bloc 1 a été apparié avec le bloc 2, le bloc 2 avec le bloc 3 et le bloc 3 avec le bloc 1, respectivement de la chambre de culture et la serre. Des micro-parcelles contenant des plantes « stressées » et « non stressées » d'un même génotype ont ainsi été générées. Chaque micro-parcelle contient 12 plantes (6 plantes du lot « stressé » et 6 plantes du lot « non stressé ») et est séparée en 3 blocs (108 plantes au total pour les trois génotypes mis en commun en serre) (Figure 6). Les prélèvements réalisés au cours du développement sont décrits dans le chapitre 2.

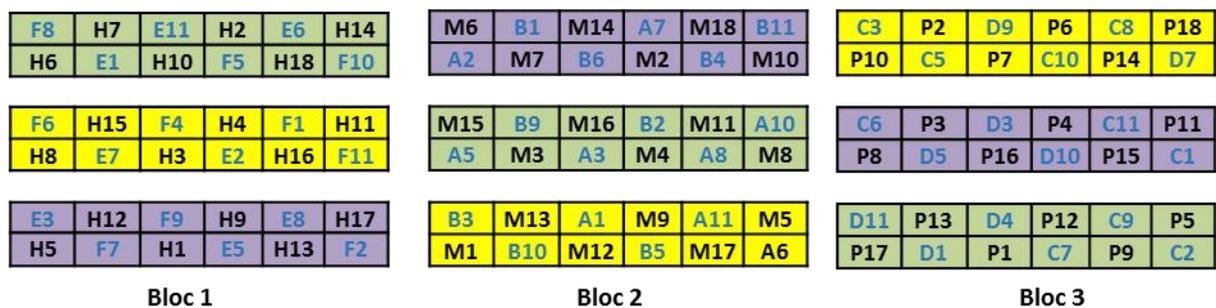


Figure 6 : Dispositif apparié entre les plantes du lot « stressé » et les plantes du lot « non stressé » en serre après la fin du traitement froid.

Les lettres en bleu indiquent les plantes du lot « stressé », en noir celles du lot « non stressé »

La photo présente les plantes du bloc 3. Elle montre l'état de développement des plantes « stressées » et « non stressées » lors de la mise en commun des deux lot après l'arrêt de l'exposition au froid (à gauche :F2; au milieu: F331 et à droite: B73).

2. Semis de la génération G1

Des autofécondations et des croisements réciproques entre plantes du lot « froid » et du lot « contrôle » ont été réalisés. Cependant, le développement paniculaire et l'émission des soies n'étaient pas souvent synchronisés entre plantes « non stressées » et plantes « stressées », ce qui a limité la réalisation des croisements. De plus, les conditions en serre n'ont pas favorisé le développement d'épis même en autofécondation. Sur F331, aucun des deux types de fécondation n'a réussi à produire des épis, probablement à cause de la tardivité liée à la sensibilité à la photopériode pour cette lignée. Le tableau 2 détaille le nombre de fécondations ayant conduit à l'obtention d'un épi. La figure 7 montre les épis récoltés pour B7 et F2. Après récolte, les épis ont été conservés en chambre froide et le semis de la génération G1 a été réalisé en mars 2017 au sein du site INRA d'Estrées-Mons (Picardie) entièrement en serre en respectant les mêmes conditions de la génération d'avant (jour : 16h à 24°C ; nuit : 8h à 18°C).

Tableau 2 : Autofécondations et croisements entre plantes stressées et plantes non stressées

Génotype	croisements	croisements réciproques	Autofécondation chez les plantes contrôle	Autofécondation chez les plantes froid"
B73	5 (2 réussis)	0	4 (3 réussis)	1 réussi
F2	20 (11 réussis)	4 (1 réussi dans les 2 sens)	2 réussis	4 (1 réussi)
F331	6 (0 réussis)	0	0	0

Les semis sont réalisés en des pots de 18L suivant un dispositif aléatoire. Au total, pour le génotype F2, 50 pots ont été installés (12 autofécondations, 20 croisements de type femelle « non stressé » X mâle « stressé » (Cm), 10 croisements de type croisement de type femelle « stressé » X mâle « non stressé » (Cf), 4 croisements réciproques (Cm_R et Cf_R). Ces pots sont disposés en 4 blocs pour le génotype F2. Les répétitions dans les blocs correspondent à des graines provenant du même épi de la génération G0 (Figure 8).

Pour le génotype B73, 18 pots ont été installés dans un seul bloc (10 autofécondations, 4 croisements de type femelle « non stressé » X mâle « stressé » (Cm), 4 croisement de type femelle « stressé » X mâle « non stressé » (Cf)). Le nombre de blocs différents entre B73 et F2 est dû au nombre de fécondations réussies dont nous disposons de la génération d'avant (Figure 8).

Comme auparavant, nous avons semé deux graines par pot après la germination des graines en boîte de pétri (à part pour un croisement de type Cf pour F2 et une autofécondation pour B73 où nous ne disposons que de 2 graines). Au stade 3 feuilles visibles, nous avons réalisé l'éclaircissage des plantules en laissant une plante par pot jusqu'à la floraison. Des autofécondations ont été réalisées sur ces plantes et les épis ont été récoltés. Les grains ont été stockés en chambre froide et sont disponibles pour le semis de la génération G2.

Les prélèvements sur la génération G1 sont décrits dans le chapitre 2.

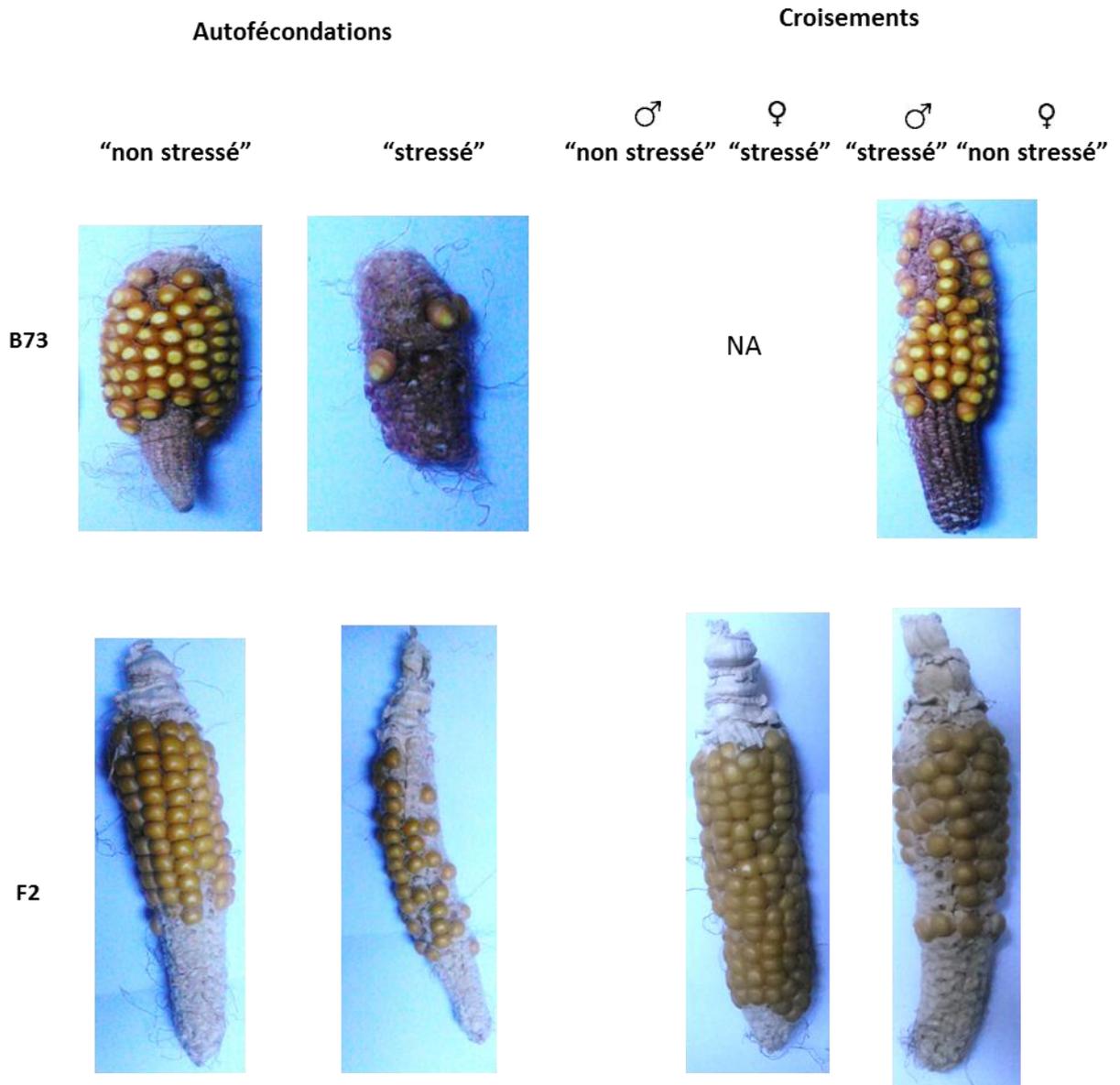


Figure 7 : Photos d'épis de maïs issus d'autofécondation de croisement pour les génotypes B73 et F2

Cf2_1	Cm1_1	Cf3_1	Autof3_1	Cm2_2	Cm4_2	Autof2_2	Autof3_3	Cm5_3	Cm1_3	Cm1_4	Autof2_4	Cf1_4	Cf4_1	Autof5_2	Cm6_3	Cf4_3	Autof4_4
Cm5_1	Autof1_1	Cm3_1	Cm2_1	Autof1_2	Cf2_2	Cm5_2	Cm_R_3	Cf1_3	Autof2_3	Autof1_4	Cm3_4	Cm4_4	Autof5_1	Autof6_1	Autof4_3	Cm6_4	Cf4_4
	Cm_R_1	Cf1_1	Autof2_1	Cm3_2	Autof3_2	Cm1_2	Cf3_3	Cm4_3	Cm3_3	Cm_R_4	Cf2_4	Cf_R_4	Autof4_1	Cm6_2	Autof5_3	Autof6_2	
	Cf_R_1	Cm4_1	Cf2_2	Cf1_2	Cm_R_2	Cf_R_2	Cm2_3	Cf_R_3	Autof1_3	Autof3_4	Cm2_4	Cm5_4	Cm6_1	Autof4_2	Cf4_2	Autof5_4	
Bloc 1				Bloc 2			Bloc 3			Bloc 4			Bloc 5				

Figure 8 : Dispositif expérimental de la génération (G1) pour le génotype F2 (en jaune) et B73 (en vert)

Le code C/Autof indique croisement ou autofécondation, le code m/f indique que le parent « stressé » a été utilisé en mâle ou en femelle. Le code R indique que pour ce croisement, un deuxième croisement a eu lieu dans l'autre sens, à partir des mêmes parents.

Le premier chiffre après le nom du croisement indique le numéro du croisement (épi) et le deuxième chiffre la répétition dans les blocs (graines provenant du même épi). Pour les autofécondation, le code rouge indique que le parent était de type « stressé » et noir « non stressé ». Noter que les cas Cf3_1 et Cf3_2 (pour F2) et Autof6 (pour B73) (en gras) sont présents uniquement en deux répétitions faute de présences de graines.

Annexe 3.

Mise au point de la détection de régions différentiellement méthylées entre échantillons de maïs après un stress abiotique

L'analyse bioinformatique des données BS-seq est assez récente (le premier méthylome, obtenu pour *Arabidopsis thaliana*, date de 2008 (Cokus *et al.*, 2008, Lister *et al.*, 2008). Plusieurs méthodes d'analyse ont été développées pour la détection de régions différentiellement méthylées, mais elles sont essentiellement dédiées à l'analyse d'échantillons humains, qui sont dépourvus de méthylation en contextes CHG et CHH, et pour des cas de variations fortes (notamment pour l'analyse de tissus cancéreux). Enfin, la majorité des études réalisées à ce jour compare deux méthylomes sans répétition biologique (en comparant deux échantillons correspondant à un individu chacun ou à un « pool » d'individus chacun), et le développement de tels pipelines est en train d'émerger, mais n'est pas encore stabilisé.

La détection de régions différentiellement méthylées (DMR) entre individus « stressés » et « non stressés » à partir de nos jeu de données comprend trois contraintes : (i) bien que les séquençages réalisés soient de forte profondeur (20x global), la complexité du génome du maïs rend les alignements assez fragmentés le long du génome, en particulier lorsque l'on applique des seuils de qualité, ce qui ne permet pas toujours une contiguité des cytosines analysées le long du génome, (ii) le signal attendu n'est pas forcément très fort, et donc vraisemblablement difficile à détecter, (iii) pour chaque condition, nous avons des données de méthylome pour trois individus indépendants.

Afin de détecter et de caractériser les DMR entre nos deux lots d'échantillons, j'ai réalisé une étape de mise au point permettant de répondre, au mieux, à ces trois contraintes.

1. Test entre méthode de segmentation et méthode fondée sur l'analyse de fenêtres

Il existe deux grandes classes de logiciels dédiés à la détection de DMR : ceux fondés sur la segmentation du génome et ceux fondés sur la comparaison des signaux au sein de fenêtres de taille définie. Au commencement de ma thèse, les logiciels methpipe (Song *et al.*, 2013) et methylkit (Akalın *et al.*, 2012) étaient les plus utilisés de chacune de ces méthodes.

Le principe de methpipe est de parcourir les données de chaque échantillon le long du génome afin de détecter des régions hyperméthylées (« HyperMethylated Regions », ou HMR) en se basant sur l'approche HMM. En parallèle, des DMC sont détectés entre les échantillons à comparer. La significativité des sites méthylés à l'intérieur des HMR est évaluée par une distribution beta-binomiale en tenant compte de la couverture.

Methylkit se base sur une analyse statistique au sein de cytosines individuelles (pour les DMC) ou de fenêtres (pour les DMR), qui peuvent être glissantes ou non. La taille, le pas et la différence de méthylation sont des paramètres à fixer *a priori*. Le test statistique employé est le test de Fischer.

A priori, nous pensions que les approches de type segmentation permettraient mieux de détecter les limites des DMR, puisque celles-ci sont fondées sur le signal lui-même, et non sur des fenêtres dont la taille doit être définie indépendamment du signal. Par contre, le logiciel methpipe ayant été construit pour l'analyse des différences de méthylation liées au cancer chez l'homme, donc des différences très marquées et au niveau d'ilôts CpG, nous ne savions pas s'il serait adapté à mes données.

En attendant l'arrivée de mes données pour l'analyse de la réponse du méthylome au froid, j'ai testé ces deux logiciels sur un jeu de données de méthylomes de différentes lignées de maïs, avec un seul réplica par lignée. Dans ce cas, les deux logiciels ont montré une détection correcte, avec quelques variations mineures. Comme attendu, j'ai observé une meilleure estimation des bornes avec methpipe. A noter que dans ces échantillons, les différences de méthylation, quand elles existent, sont importantes, ce qui facilite la détection.

J'ai donc testé ce logiciel sur des données BS-seq issues de plantes ayant poussé ou non en condition de déficit hydrique. J'ai observé un faible nombre de DMR. Leur analyse dans IGV (Integrative Genomics Viewer) (<http://software.broadinstitute.org/software/igv/>) a révélé un signal très faible, qui nous a paru plutôt de l'ordre du bruit que du signal biologique. Notamment, j'ai identifié un grand nombre de DMC qui auraient vraisemblablement dues être considérées comme des faux-positifs (Figure 1). Dans le cadre du Réseau thématique pluridisciplinaire Eco-Evo-Epigen (RTP3E), nous avons organisé des séances de réflexion autour des logiciels de détection des DMR. Les statisticiens nous ont déconseillé l'utilisation de méthodes de segmentation dans le cas de signaux faibles, car ils sont difficilement détectables. Suite à ces discussions, nous avons décidé de nous focaliser sur l'utilisation du logiciel methylkit, qui est largement utilisé dans le domaine végétal.

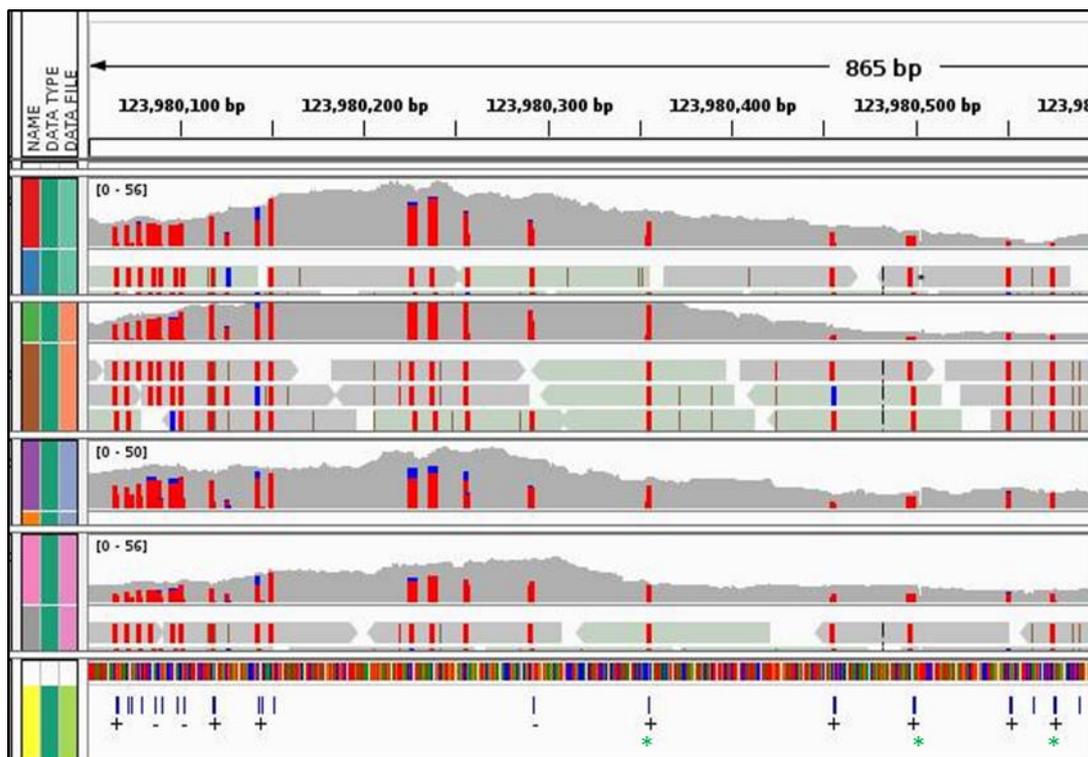


Figure 1 : Détection des DMCs par Methpipe (En haut deux individus stressés et en bas deux individus témoin)

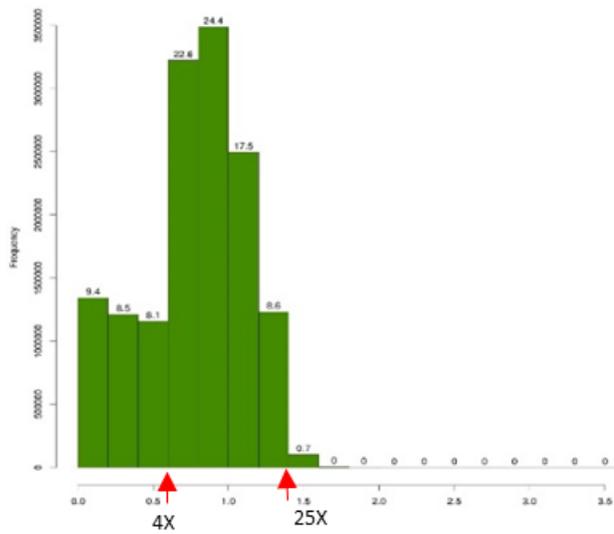
*Présence de faux positifs.

2. Choix du seuil de couverture par cytosine

Les données d'alignement de séquences DNA-seq montrent souvent une assez grande variation en terme de couverture le long du génome. Dans le cas particulier du BS-seq, cela pose le problème d'une estimation erronée du niveau quantitatif de la méthylation des cytosines. Ces problèmes sont d'autant plus marqués quand on analyse un génome complexe avec une forte proportion d'éléments transposables comme c'est le cas du maïs. *A priori*, puisque nous cherchions à détecter des variations de 10%, nous pensions utiliser les cytosines couvertes à au moins 10X. Cependant, ce seuil nous paraissait très stringent, d'autant que nous souhaitions conserver que les régions couvertes dans les 6 réplicas. Afin de choisir de façon moins arbitraire des seuils de profondeur permettant de filtrer les cytosines trop peu et trop couvertes, je me suis basée sur l'analyse des histogrammes de couverture des cytosines par brin, comme proposé par C. Grunau (Figure 2). Cette distribution montre la séparation de deux populations de cytosines. Une première avec une continuité dans la distribution de la couverture et une deuxième avec un mode présentant des ruptures. En se basant sur les limites inférieures et supérieures de cette deuxième distribution, j'ai fixé des seuils de couverture minimum et maximum à 4X et 25X par cytosine et par brin.

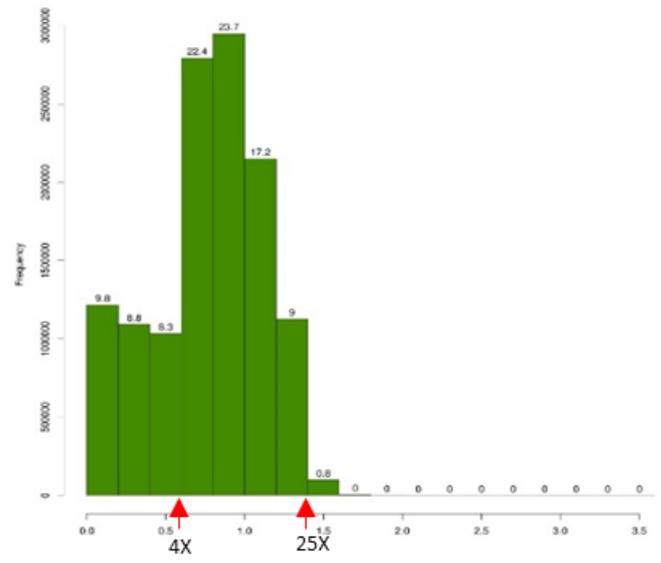
Dans le cadre de mon projet de thèse, une question importante était d'analyser la stabilité de la méthylation entre individus d'un même traitement, afin de tester ensuite si les locus avec différents niveaux de stabilité étaient maintenus de façon différente au cours du développement. Je souhaitais donc pouvoir analyser des DMR ayant un signal sur chacun des individus. Afin de répondre au mieux, à cette question, j'ai choisi d'analyser uniquement les cytosines couvertes entre 4X et 25X dans les 6 réplicas. Ce critère est assez stringent, et diminue beaucoup le nombre de cytosines analysables. Mais il permet de ne pas biaiser les pourcentages de méthylation par une couverture plus forte chez un individu. En effet, bien que methylkit prévoie l'utilisation de réplicas, en réalité il combine l'information obtenue par l'ensemble des réplicas. En fait, il utilise la somme des lectures de séquence de l'ensemble des individus, et ne fait même pas une moyenne pondérée par la profondeur obtenue pour chaque individu. L'utilisation de données de couverture similaires entre individus permet d'atténuer ce biais. Ceci est d'autant plus important que l'on peut s'attendre à une variation assez forte entre individus. En outre, ceci permet d'analyser ensuite le niveau de méthylation de chaque individu, et ainsi de déterminer le niveau de stabilité de la région.

CG



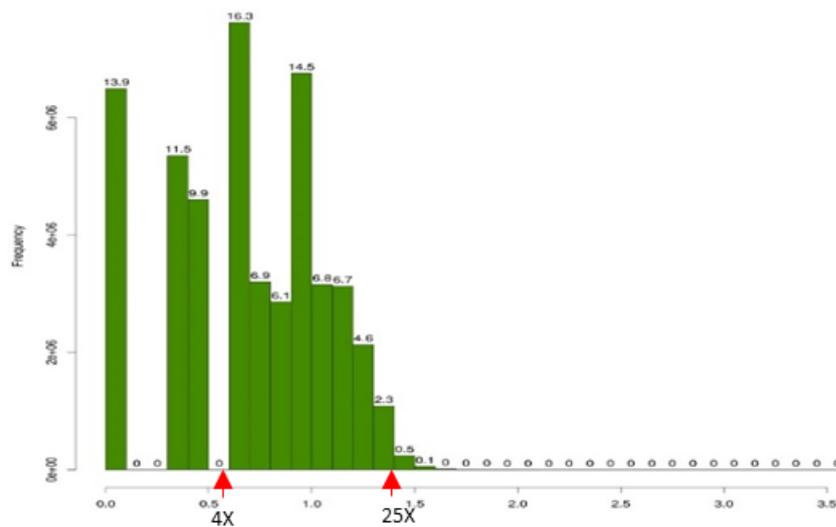
Couverture par base (log10)

CHG



Couverture par base (log10)

CHH



Couverture par base (log10)

Figure 2 : Distribution de la couverture des cytosines sur le chromosome 5 de B73.

La distribution montre des cytosines couvertes de manière stochastique (à gauche) et d'autres de manière centrée (à droite). Un seul individu est représenté, pour les trois contextes. Une distribution similaire est observée pour tous les jeux de données analysés (6 réplicas pour chacun de géotypes B73, F2 et F331). Les flèches indiquent les profondeurs de séquençage choisies pour le filtre.

3. Détection et filtres sur DMR analysables

La méthode de détection des DMR s'est reposée sur l'utilisation du package R MethyKit (Akalin *et al.*, 2012). Je l'ai appliqué sur les cytosines de chaque contexte séparément. Comme exposé ci-dessus, la taille des fenêtres est choisie *a priori*. Ce choix est critique, car cette taille définit à la fois la capacité de détection (si un signal est petit dans une grande fenêtre, elle ne passe pas le test statistique), et la taille des régions détectées (si les fenêtres sont petites, le signal risque d'être morcelé). Un autre paramètre important est le niveau de méthylation retenu pour la fenêtre, et le nombre de cytosines différentiellement méthylées (« Differentially Methylated Cytosines », ou DMC) par région.

Afin de pouvoir détecter un grand nombre de régions, j'ai décidé de faire tourner ce package sur trois tailles de fenêtres différentes : 200pb, 500pb et 1000pb. Dans chaque cas, j'ai utilisé des fenêtres chevauchantes avec un pas de 100pb, et une différence de méthylation de 10% (q-value =0,01). Une fois cette première détection réalisée, j'ai assemblé les DMR contiguës (chevauchant à 1 nucléotide) pour chaque taille séparément. La figure 3.A présente la distribution de la taille des DMR CG et CHG sans aucun filtre pour chaque taille de fenêtre. Elle montre l'impact du choix de la taille de fenêtrage sur la taille finale de la DMR : par exemple pour la taille de détection de 200pb, la taille finale des DMR varie de 200 à 400pb. Cependant, cette figure montre que la majorité des DMR sont trouvées dans la taille de la fenêtre utilisée respectivement pour les tailles de 200, 500 et 1000pb.

Pour le contexte CG et CHG, des cytosines individuelles (DMC) sont aussi détectées en utilisant un seuil de différence de méthylation de 10% et une q-value de 0,01. Les DMC n'ont pas été détectées pour le contexte CHH car les niveaux de variation sont très faibles et les résultats étaient peu fiables. Le contexte CHH a été utilisé uniquement pour typer les niveaux de méthylation dans les DMR détectées pour les contextes CG et CHG.

Pour l'analyse des différences de méthylation entre plantes « stressées » et « non stressées » à la sortie du traitement froid (chapitre 1), nous cherchions à détecter des régions avec un potentiel rôle biologique, donc nous avons opté pour un niveau de variation assez fort (10% de variation sur chaque fenêtres, et 5 DMC sur la DMR finale), comme proposé dans Secco *et al.* (2015). La figure 3.B montre, pour chacune des 3 tailles utilisées (200, 500 et 1000pb), la distribution de taille des DMR détectées et possédant au moins 5 DMC. Elle montre un décalage de la taille des DMR de chaque fenêtre de détection vers la droite indiquant que les petites régions sont éliminées dans chaque cas. Elle montre également que la majorité des cas récupérés correspondent à des DMR qui sont chevauchantes pour une même taille de détection. Ceci montre que les DMC ne sont pas groupés pour une même fenêtre de détection et que le signal est trouvé dans une continuité de régions chevauchantes ayant une différence de méthylation.

Afin de remédier aux problèmes de bords des DMR détectées et afin d'éliminer la redondance entre les DMR retrouvées par les 3 tailles de fenêtres, nous avons décidé de merger les DMR se chevauchant entre les 3 tailles. En effet, pour les DMR qui se chevauchent entre 2 ou 3 tailles de détection et présentant le même nombre de DMC, celle ayant la taille la plus petite est récupérée. A un nombre différent de DMC, la DMR ayant le plus de DMC est gardée (Figure 4). La distribution des tailles des DMR à au moins 5 DMC après le merge entre les 3 fenêtres de détection est illustrée dans la figure 3.C. elle montre que, majoritairement, les cas de DMR récupérés sont ceux détectés en utilisant la fenêtre 200. Ce qui indique l'existence de signal fort sur des petites régions locales dans le génome.

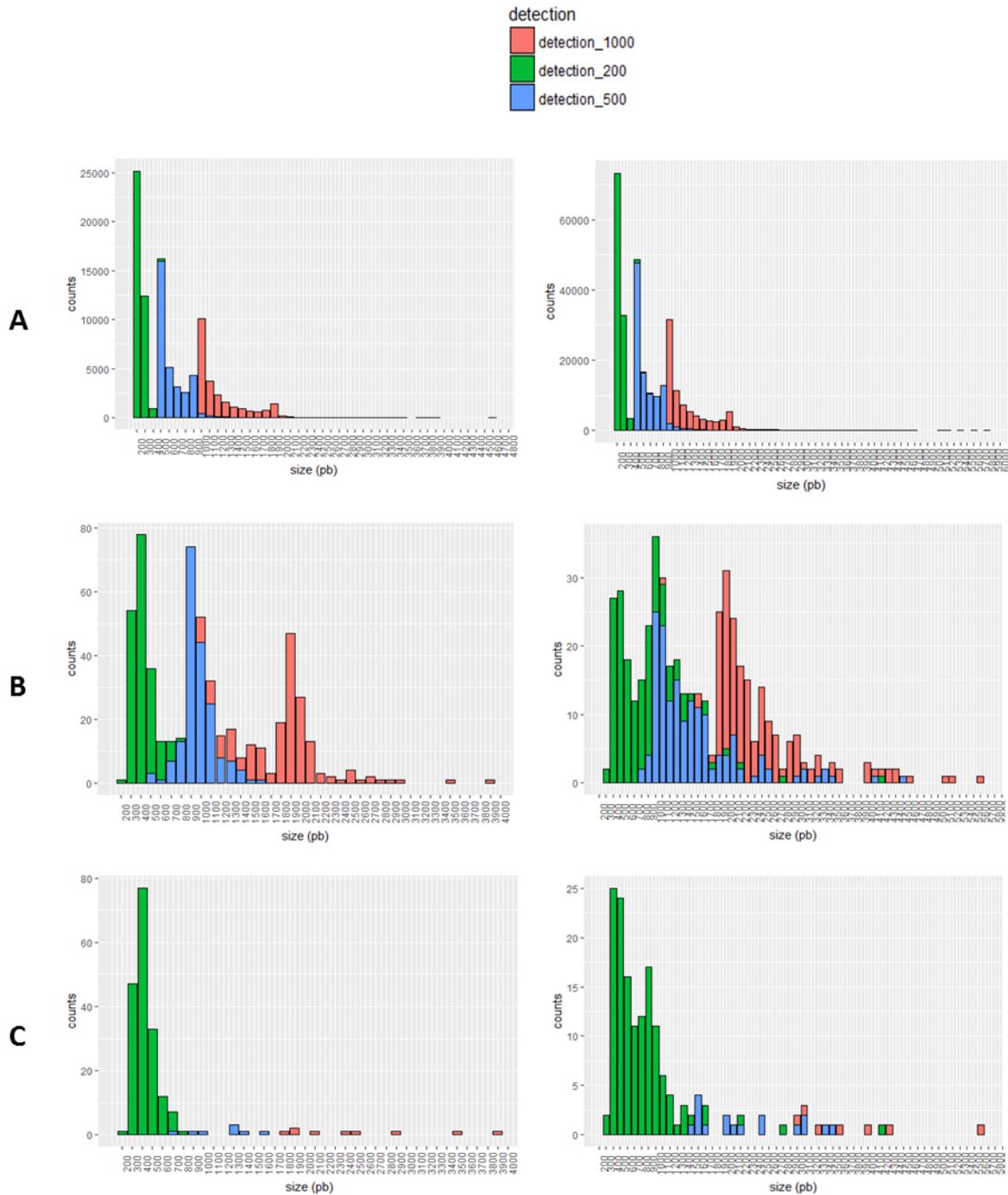


Figure 3 : Tailles des DMR CG (à gauche) et CHG (à droite)

A. Taille des DMR pour chaque fenêtre de détection après merge à 1 nucléotide d'overlap. **B.** Taille des DMR avec au moins 5 DMC pour chaque fenêtre de détection. **C.** Taille des DMR avec au moins 5 DMC après merge entre les 3 tailles de détection. Les couleurs montrent la taille de fenêtre utilisée pour définir la DMR (pour A et B), et la taille de fenêtre de la DMR ayant été retenue après le merge (pour C).

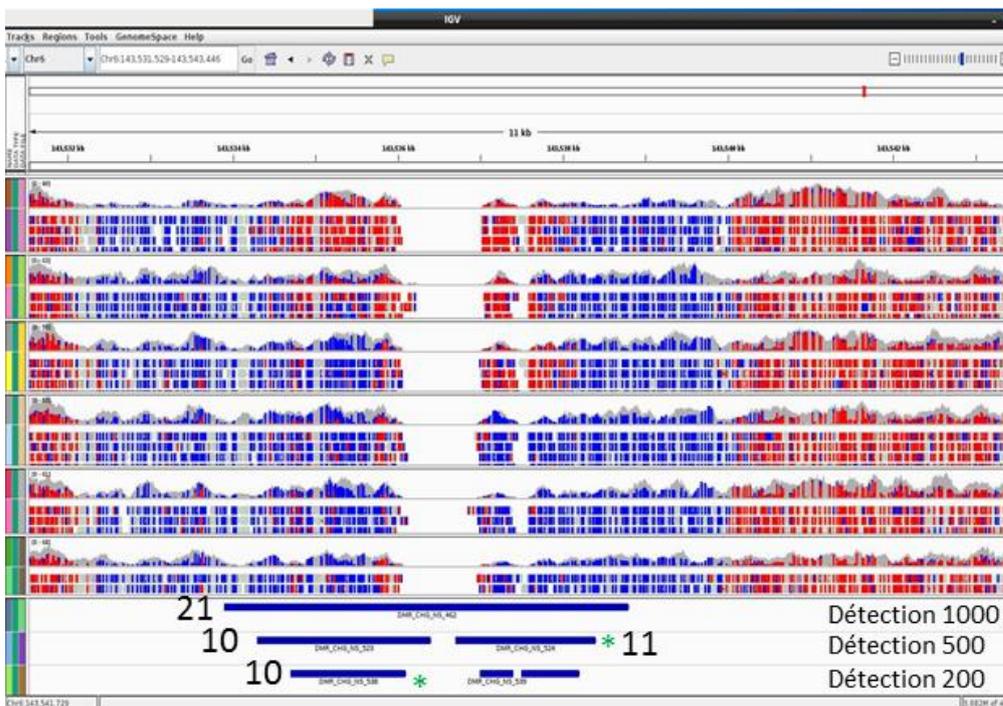
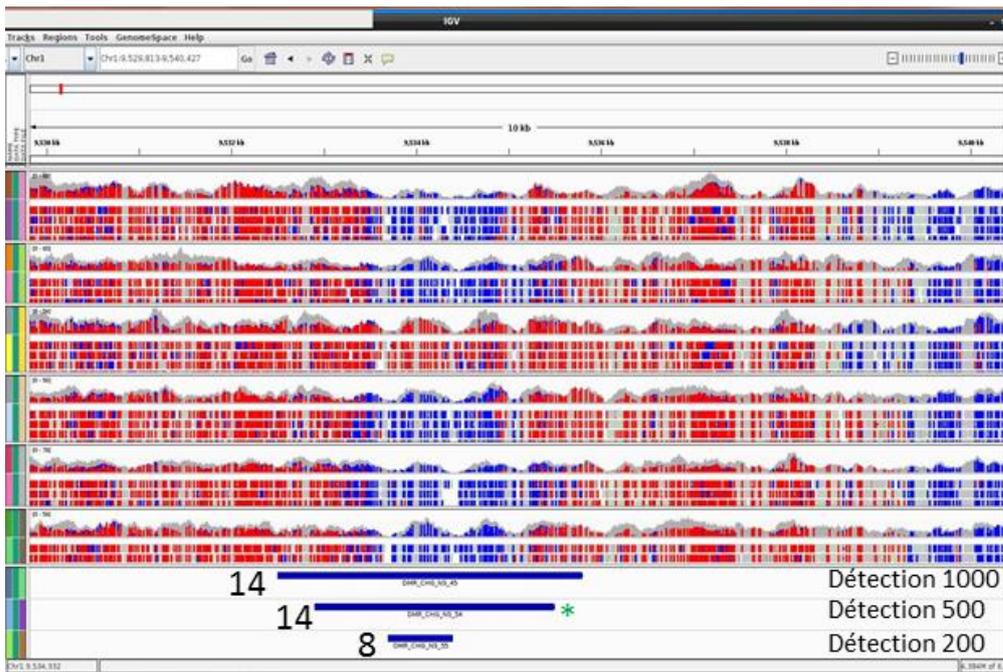


Figure 4: Exemples de choix de DMR à au moins 5 DMCs en cas de chevauchement entre les 3 tailles de détection

En haut les 3 individus non stressés et en bas les individus stressés. Le nombre de DMC est représenté à côté de chaque DMR.

* DMR retenue pour l'analyse

Le détail de cette méthodologie d'analyse est illustré dans la figure 5. C'est cette méthodologie qui a été utilisée pour obtenir les DMR du **chapitre 1**.

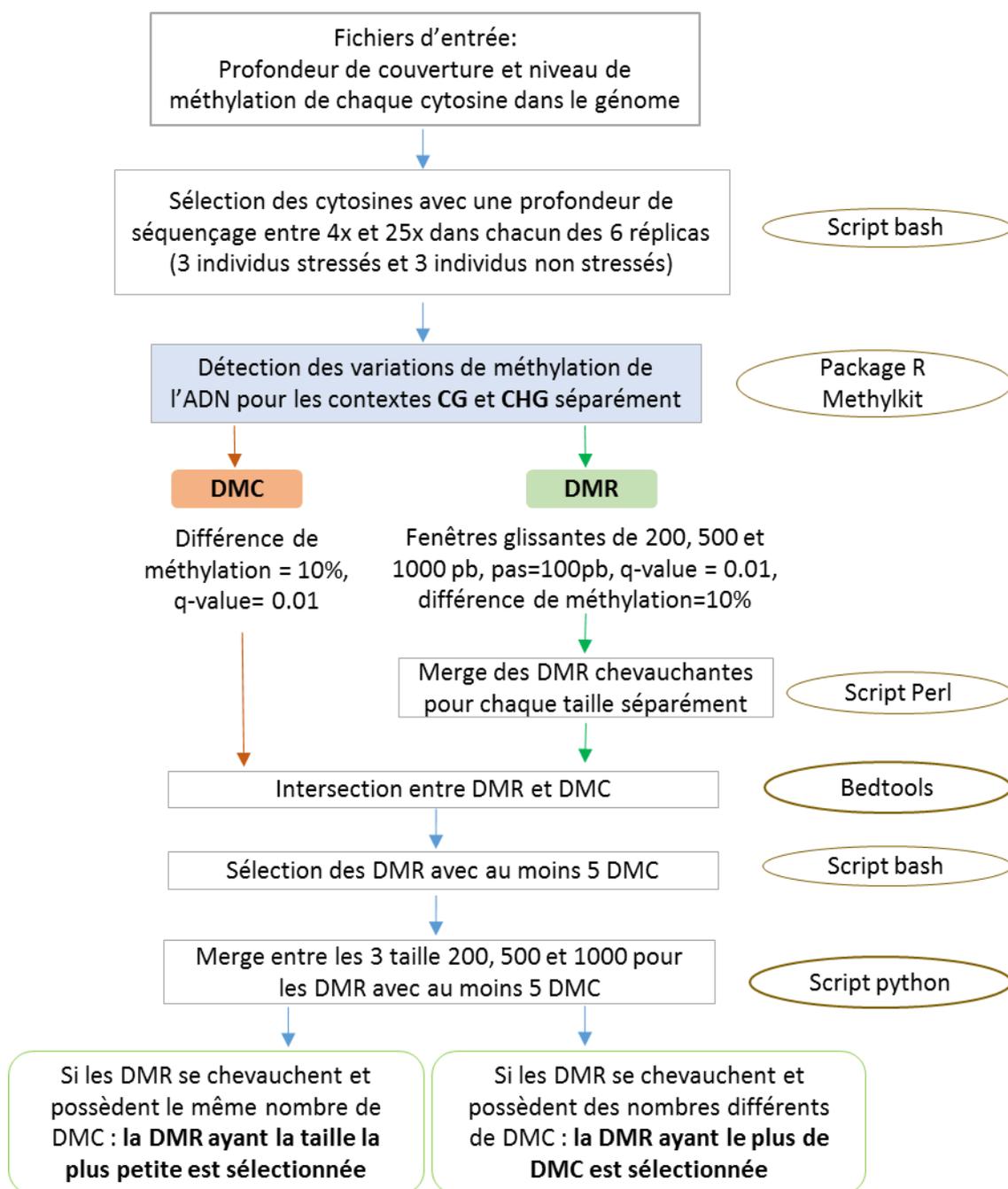


Figure 5 : Workflow de détection des DMR entre plantes non stressées et plantes stressées

Les cercles indiquent les pipelines utilisés. DMR: Differentially Methylated Regions, DMC : Differentially Methylated Cytosines

Les régions génomiques utilisées pour le suivi des niveaux de méthylation au cours du développement et de la génération G1 pour le design par la méthode SeqCapEpi et qui ont fait l'objet du chapitre 2 sont différentes de celles qui ont été analysées à la sortie du froid (chapitre 1).

En effet, pour des raisons de calendrier intrinsèque à la collaboration avec Biogemma, cette analyse a dû être réalisée avant la stabilisation des critères qui nous ont permis d'obtenir les DMR finales. Alors que les DMR du chapitre 1 ont été filtrées pour contenir au moins 5 DMC, le seuil utilisé pour les DMR du SeqCapEpi est moins stringent : 2 DMC par DMR CG et CHG. De plus, à l'inverse du crible réalisé

pour le chapitre 1, ici lorsqu'une région est détectée par les 3 tailles, la limite inférieure et supérieure des régions chevauchantes est retenue. Si le chevauchement concerne les tailles de détection 200, 500 et 1000 ou bien 200 et 500, la taille 500 est retenue. En effet pour ces deux catégories de merge, nous avons remarqué que les bords de DMR étaient mal estimés en prenant le merge, la majorité des cas montre des DMR imbriquées où le signal est le mieux représentée par la détection 500pb, en regardant sur IGV. La distribution de la taille du merge entre ces tailles est illustrée dans la figure 6. Pour les DMR CHH où la détection de DMC n'était pas possible, le filtre s'est basé uniquement sur la taille où nous avons éliminé toute les DMR à 200 et 300pb (Figure 6).

Pour cette analyse, nous avons aussi détecté des régions en utilisant les cytosines dans les 3 contextes ensemble et en appliquant des fenêtres glissantes de 500pb avec un pas de 100pb et une différence de méthylation de 10%. Ceci permet d'analyser des DMR montrant des variations statistiques dans les 3 contextes à la fois. Afin d'éliminer la redondance entre ces DMR et les DMR détectées séparément dans chaque contexte, une intersection est faite entre les DMR CG, CHG et CHH. Ensuite, une intersection inverse de ces DMR et des DMR détectées dans les 3 contexte ensemble est effectuée.

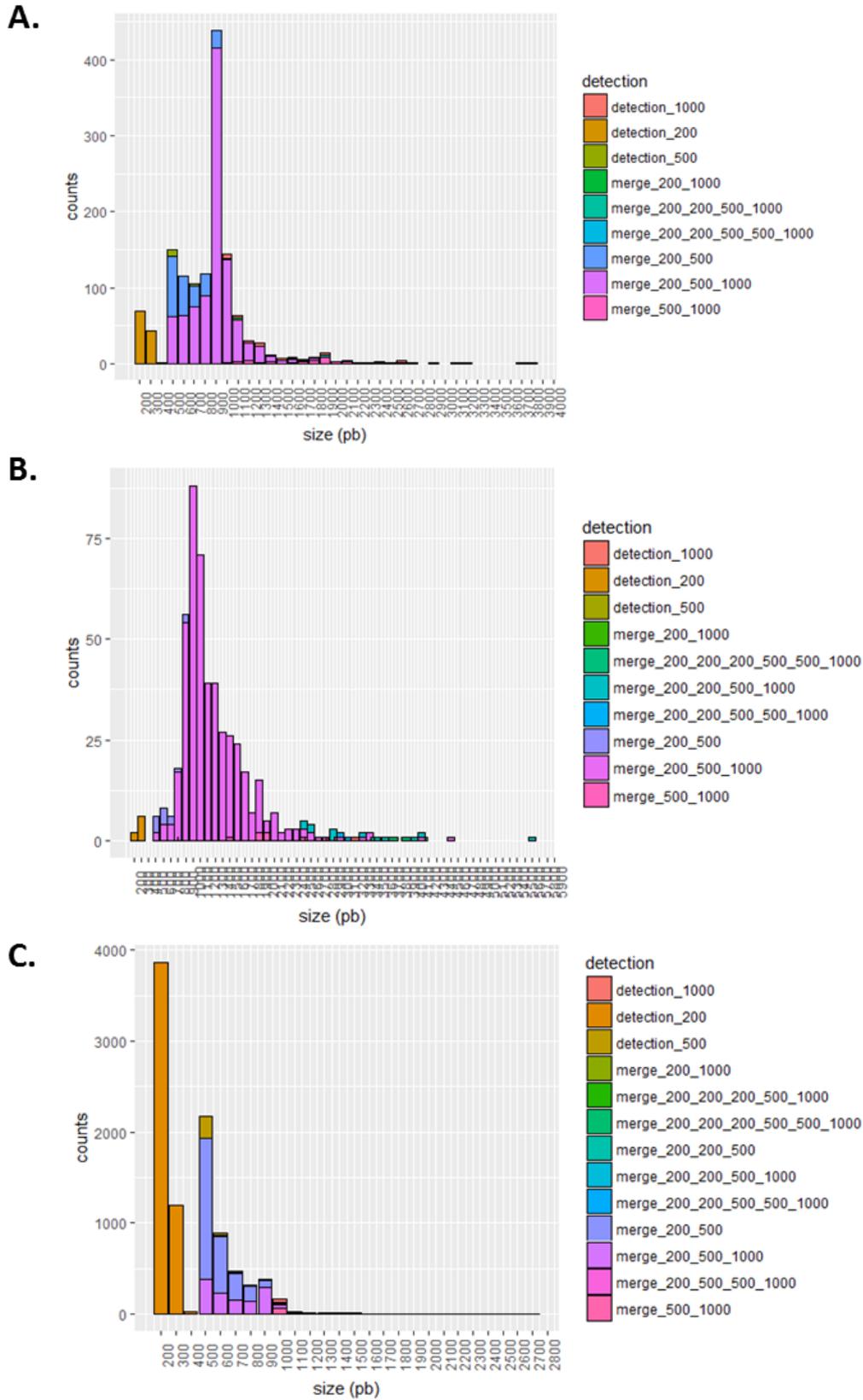


Figure 6 : Distribution de la taille des DMR après merge entre les DMR détectées par methylkit avec trois tailles de fenêtres (200, 500 et 1000)

Pour CG (A) et CHG (B) le merge est réalisé sur les DMR détectées avec au moins 2 DMC Pour CHH (C), le merge est fait sur les DMR non filtrées suivant le nombre de DMC

Annexe 4.

Formations suivies au cours de ma thèse et valorisation de mes travaux

Formations suivies

Connaissance de l'Entreprise

- Forum BioTechno, 09 juin 2016. Centre de recherche des Cordeliers-Paris.

Communication

- Fête de la science, 14-16 octobre 2016. Préparation et présentation de l'atelier : Les tribulations du maïs : du Mexique à notre quotidien. Gif-sur Yvette.

- MT180 - Ma thèse en 180 secondes, 21 mars 2016. Faculté des sciences. Orsay.

- S'exprimer devant un groupe, 14 décembre 2015. UFR des Sciences, Versailles.

Formation scientifique

- Détection de variants structuraux par reséquantage, 24 septembre 2015. GQE Le Moulon-Gif_sur_Yvette

- Epigénétique et signalisation cellulaire, 5 et 7 avril 2016. Muséum d'Histoire Naturelle. Paris

- Ecole thématique 3E (Epigénétique en Ecologie et Evolution). 18-22 mai 2015. Saint Pierre d'Oléron

- Modélisation statistiques et analyse de données biologiques, 20-24 mars 2017. AgroParisTech, Paris.

Outils bio-informatiques

- Initiation à Python. Unité MaiAGE, 30-31 mars 2016 centre INRA Jouy-en-Josas, Domaine de Vilvert.

- Python avancé. Unité MaiAGE, 1^{er} avril 2016 centre INRA Jouy-en-Josas, Domaine de Vilvert.

Rédaction

- Writing a scientific paper. 3 octobre 2016 (demi-journée organisée par la directrice de l'ED, Jacqui SHYKOFF).

Participation à des enseignements

- Cours sur l'épigénétique des plantes. Niveau : 2^{ème} année ingénieur d'AgroParisTech. UE : "Ce que nous apprennent les génomes", 3 novembre 2015 et 8 novembre 2016.

- TP en bio-informatique. Niveau : Master 2 - Sciences du végétal. UE : "Génomique et amélioration des plantes". Université Paris-Sud, 20 septembre 2016.

Communications dans des colloques

Zeineb ACHOUR, Catherine GIAUFFRET, Hélène SELLIER, Jean-Philippe PICHON, Magalie LEVEUGLE, Hélène CORTI, Agnès ROUSSELET, Clémentine VITTE. Meeting annuel du projet Amaizing. Impact of cold on maize methylome. 22 et 23 octobre 2015. Pau, France (Présentation orale).

Zeineb ACHOUR, Catherine GIAUFFRET, Hélène SELLIER, Jean-Philippe PICHON, Hélène CORTI, Agnès ROUSSELET, Clémentine VITTE. Journées des doctorants de l'ED Sciences du végétal. Impact of cold on the maize methylome. 5 et 6 novembre 2015. Faculté des sciences, Orsay, France (Poster).

Zeineb ACHOUR, Catherine GIAUFFRET, H  l  ne SELLIER, Jean-Philippe PICHON, Magalie LEVEUGLE, H  l  ne CORTI, Agn  s ROUSSELET, Cl  mentine VITTE. Impact of cold on maize leaves methylome. VIII European Plant Science Retreat (EPSR). 20-23 juin 2016. Facult   de biologie, Barcelone, Espagne (Poster).

Zeineb ACHOUR, Sara CASTELLETTI, Jean-Philippe PICHON, Johann JOETS, Magalie LEVEUGLE, Agn  s ROUSSELET, Cl  mentine VITTE. Interconnection between genome and methylome variation in maize. DynaGEV. 7 et 8 juillet 2016. AgroParisTech, Paris, France (Poster).

Zeineb ACHOUR, Catherine GIAUFFRET, H  l  ne SELLIER, Jean-Philippe PICHON, H  l  ne CORTI, Agn  s ROUSSELET, Cl  mentine VITTE. Impact of cold on the maize methylome. 17 et 18 novembre 2016. Journ  es des doctorants de l'ED Sciences du v  g  tal. Facult   des sciences, Orsay, France (Pr  sentation orale).

Zeineb ACHOUR, Catherine GIAUFFRET, H  l  ne SELLIER, Jean-Philippe PICHON, Herv   DUBORJAL, Magalie LEVEUGLE, H  l  ne CORTI, Agn  s ROUSSELET and Cl  mentine VITTE. Cold induces large methylome changes in maize. 22 et 23 juin 2017. Epiplant meeting on plant epigenetics. IPS2, Orsay, France (Poster).

Zeineb ACHOUR, Catherine GIAUFFRET, H  l  ne SELLIER, Jean-Philippe PICHON, Herv   DUBORJAL, Magalie LEVEUGLE, H  l  ne CORTI, Agn  s ROUSSELET and Cl  mentine VITTE. Cold induces specific methylome changes in maize. 12 et 15 septembre 2017. 40th New Phytologist Symposium "Plant epigenetics : From mechanisms to ecological relevance", Universit   de Vienne, Vienne, Autriche (Poster).

Zeineb ACHOUR, Catherine GIAUFFRET, H  l  ne SELLIER, Jean-Philippe PICHON, Herv   DUBORJAL, Magalie LEVEUGLE, H  l  ne CORTI, Agn  s ROUSSELET and Cl  mentine VITTE. Cold induces transcriptional and methylation changes in the sensitive line B73. 60th Annual Maize Genetics Conference. 22-25 mars 2018, Palais du Grand Large, Saint-Malo, France (Poster).

Références bibliographiques

- Ahmad, F., X. Huang, H. X. Lan, T. Huma, Y. M. Bao, J. Huang and H. S. Zhang, 2014: Comprehensive gene expression analysis of the DNA (cytosine-5) methyltransferase family in rice (*Oryza sativa* L.). *Genet. Mol. Res*, **13**, 5159-5172.
- Ahmed, I., A. Sarazin, C. Bowler, V. Colot and H. Quesneville, 2011: Genome-wide evidence for local DNA methylation spreading from small RNA-targeted sequences in *Arabidopsis*. *Nucleic Acids Research*, **39**, 6919-6931.
- Akalin, A., M. Kormaksson, S. Li, F. E. Garrett-Bakelman, M. E. Figueroa, A. Melnick and C. E. Mason, 2012: methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*, **13**, R87.
- Allis, C. D., T. Jenuwein and D. Reinberg, 2007: *Epigenetics*. CSHL Press.
- Alonso, C., R. Perez, P. Bazaga and C. M. Herrera, 2015: Global DNA cytosine methylation as an evolving trait: phylogenetic signal and correlated evolution with genome size in angiosperms. *Front Genet*, **6**, 4.
- Andrews, J. R., M. J. Fryer and N. R. Baker, 1995: Characterization of chilling effects on photosynthetic performance of maize crops during early season growth using chlorophyll fluorescence. *Journal of Experimental Botany*, **46**, 1195-1203.
- Atkinson, N. J. and P. E. Urwin, 2012: The interaction of plant biotic and abiotic stresses: from genes to the field. *Journal of experimental botany*, **63**, 3523-3543.
- Bailey, A. O., T. Panchenko, K. M. Sathyan, J. J. Petkowski, P.-J. Pai, D. L. Bai, D. H. Russell, I. G. Macara, J. Shabanowitz and D. F. Hunt, 2013: Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proceedings of the National Academy of Sciences*, **110**, 11827-11832.
- Barkan, A. and R. A. Martienssen, 1991: Inactivation of maize transposon Mu suppresses a mutant phenotype by activating an outward-reading promoter near the end of Mu1. *Proceedings of the National Academy of Sciences*, **88**, 3502-3506.
- Baubec, T., A. Finke, O. Mittelsten Scheid and A. Pecinka, 2014: Meristem-specific expression of epigenetic regulators safeguards transposon silencing in *Arabidopsis*. *EMBO Rep*, **15**, 446-452.
- Baucom, R. S., J. C. Estill, C. Chaparro, N. Upshaw, A. Jogi, J. M. Deragon, R. P. Westerman, P. J. Sanmiguel and J. L. Bennetzen, 2009: Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. *PLoS Genet*, **5**, e1000732.
- Baulcombe, D. C. and C. Dean, 2014: Epigenetic regulation in plant responses to the environment. *Cold Spring Harb Perspect Biol*, **6**, a019471.
- Bechoux, N., G. Bernier and P. Lejeune, 2000: Environmental effects on the early stages of tassel morphogenesis in maize (*Zea mays* L.). *Plant, Cell & Environment*, **23**, 91-98.
- Bewick, A. J., L. Ji, C. E. Niederhuth, E. M. Willing, B. T. Hofmeister, X. Shi, L. Wang, Z. Lu, N. A. Rohr, B. Hartwig, C. Kiefer, R. B. Deal, J. Schmutz, J. Grimwood, H. Stroud, S. E. Jacobsen, K. Schneeberger, X. Zhang and R. J. Schmitz, 2016: On the origin and evolutionary consequences of gene body DNA methylation. *Proc Natl Acad Sci U S A*, **113**, 9111-9116.
- Bewick, A. J. and R. J. Schmitz, 2017: Gene body DNA methylation in plants. *Curr Opin Plant Biol*, **36**, 103-110.
- Bird, A., 2007: Perceptions of epigenetics. *Nature*, **447**, 396.
- Blevins, T., R. Podicheti, V. Mishra, M. Marasco, J. Wang, D. Rusch, H. Tang and C. S. Pikaard, 2015: Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in *Arabidopsis*. *Elife*, **4**.
- Blum, A., 2015: Stress, strain, signaling, and adaptation—not just a matter of definition. *Journal of experimental botany*, **67**, 562-565.

- Bock, C., 2012: Analysing and interpreting DNA methylation data. *Nature reviews. Genetics*, **13**, 705.
- Bolger, A. M., M. Lohse and B. Usadel, 2014: Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.
- Boyko, A. and I. Kovalchuk, 2011: Genome instability and epigenetic modification--heritable responses to environmental stress? *Curr Opin Plant Biol*, **14**, 260-266.
- Brandenburg, J.-T., T. Mary-Huard, G. Rigaille, S. J. Hearne, H. Corti, J. Joets, C. Vitte, A. Charcosset, S. D. Nicolas and M. I. Tenaillon, 2017: Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts. *PLoS genetics*, **13**, e1006666.
- Calarco, J. P., F. Borges, M. T. Donoghue, F. Van Ex, P. E. Jullien, T. Lopes, R. Gardner, F. Berger, J. A. Feijo, J. D. Becker and R. A. Martienssen, 2012: Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell*, **151**, 194-205.
- Calatayud, P.-A., J.-P. Garrec and M. Nicole, 2013: Adaptation des plantes aux stress environnementaux. *N. Sauvion, P.-A. Calatayud, D. Thiéry, & F. Marion-Poll (Vol. Eds.), Interactions insectesplantes*, 229e245.
- Camus-Kulandaivelu, L., J.-B. Veyrieras, D. Madur, V. Combes, M. Fourmann, S. Barraud, P. Dubreuil, B. Gouesnard, D. Manicacci and A. Charcosset, 2006: Maize adaptation to temperate climate: relationship between population structure and polymorphism in the Dwarf8 gene. *Genetics*, **172**, 2449-2463.
- Cavrak, V. V., N. Lettner, S. Jamge, A. Kosarewicz, L. M. Bayer and O. Mittelsten Scheid, 2014: How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genet*, **10**, e1004115.
- Chan, F. L. and L. H. Wong, 2012: Transcription in the maintenance of centromere chromatin identity. *Nucleic acids research*, **40**, 11178-11188.
- Chan, S. W., I. R. Henderson and S. E. Jacobsen, 2005: Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat Rev Genet*, **6**, 351-360.
- Choi, C. S. and H. Sano, 2007: Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. *Mol Genet Genomics*, **277**, 589-600.
- Choo, K. A., 2001: Domain organization at the centromere and neocentromere. *Developmental cell*, **1**, 165-177.
- Cokus, S. J., S. Feng, X. Zhang, Z. Chen, B. Merriman, C. D. Haudenschild, S. Pradhan, S. F. Nelson, M. Pellegrini and S. E. Jacobsen, 2008: Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature*, **452**, 215-219.
- Coleman-Derr, D. and D. Zilberman, 2012: Deposition of histone variant H2A. Z within gene bodies regulates responsive genes. *PLoS genetics*, **8**, e1002988.
- Cortijo, S., R. Wardenaar, M. Colomé-Tatché, A. Gilly, M. Etcheverry, K. Labadie, E. Caillieux, J.-M. Aury, P. Wincker and F. Roudier, 2014: Mapping the epigenetic basis of complex traits. *Science*, **343**, 1145-1148.
- Cramer, G. R., K. Urano, S. Delrot, M. Pezzotti and K. Shinozaki, 2011: Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biol*, **11**, 163.
- Cubas, P., C. Vincent and E. Coen, 1999: An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*, **401**, 157.
- Cuerda-Gil, D. and R. K. Slotkin, 2016: Non-canonical RNA-directed DNA methylation. *Nature plants*, **2**, 16163.
- Daccord, N., J.-M. Celton, G. Linsmith, C. Becker, N. Choisine, E. Schijlen, H. van de Geest, L. Bianco, D. Micheletti and R. Velasco, 2017: High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nature genetics*, **49**, 1099.
- Darracq, A., C. Vitte, S. Nicolas, J. Duarte, J.-P. Pichon, T. Mary-Huard, C. Chevalier, A. Bérard, M.-C. Le Paslier and P. Rogowsky, 2018: Sequence analysis of European maize inbred line F2 provides new insights into molecular and chromosomal characteristics of presence/absence variants. *BMC Genomics*, **19**, 119.

- Ding, H., J. Gao, C. Qin, H. Ma, H. Huang, P. Song, X. Luo, H. Lin, Y. Shen, G. Pan and Z. Zhang, 2014: The dynamics of DNA methylation in maize roots under Pb stress. *Int J Mol Sci*, **15**, 23537-23554.
- Dodsworth, S., A. R. Leitch and I. J. Leitch, 2015: Genome size diversity in angiosperms and its influence on gene space. *Curr Opin Genet Dev*, **35**, 73-78.
- Doebley, J., J. D. Wendel, J. Smith, C. W. Stuber and M. M. Goodman, 1988: The origin of cornbelt maize: the isozyme evidence. *Economic botany*, **42**, 120-131.
- Du, Y., C. N. Topp and R. K. Dawe, 2010: DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS genetics*, **6**, e1000835.
- Dubin, M. J., P. Zhang, D. Meng, M.-S. Remigereau, E. J. Osborne, F. P. Casale, P. Drewe, A. Kahles, G. Jean and B. Vilhjálmsson, 2015: DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *Elife*, **4**, e05255.
- Duhl, D. M., H. Vrieling, K. A. Miller, G. L. Wolff and G. S. Barsh, 1994: Neomorphic agouti mutations in obese yellow mice. *Nature genetics*, **8**, 59.
- Eagles, H. and J. Lothrop, 1994: Highland maize from Central Mexico—its origin, characteristics, and use in breeding programs. *Crop Science*, **34**, 11-19.
- Eichten, S. R., R. Briskine, J. Song, Q. Li, R. Swanson-Wagner, P. J. Hermanson, A. J. Waters, E. Starr, P. T. West, P. Tiffin, C. L. Myers, M. W. Vaughn and N. M. Springer, 2013: Epigenetic and genetic influences on DNA methylation variation in maize populations. *Plant Cell*, **25**, 2783-2797.
- Eichten, S. R., N. A. Ellis, I. Makarevitch, C. T. Yeh, J. I. Gent, L. Guo, K. M. McGinnis, X. Zhang, P. S. Schnable, M. W. Vaughn, R. K. Dawe and N. M. Springer, 2012: Spreading of heterochromatin is limited to specific families of maize retrotransposons. *PLoS Genet*, **8**, e1003127.
- Eichten, S. R., R. J. Schmitz and N. M. Springer, 2014: Epigenetics: Beyond Chromatin Modifications and Complex Genetic Regulation. *Plant Physiol*, **165**, 933-947.
- Eichten, S. R. and N. M. Springer, 2015: Minimal evidence for consistent changes in maize DNA methylation patterns following environmental stress. *Front Plant Sci*, **6**, 308.
- Eichten, S. R., R. A. Swanson-Wagner, J. C. Schnable, A. J. Waters, P. J. Hermanson, S. Liu, C. T. Yeh, Y. Jia, K. Gendler, M. Freeling, P. S. Schnable, M. W. Vaughn and N. M. Springer, 2011: Heritable epigenetic variation among maize inbreds. *PLoS Genet*, **7**, e1002372.
- Ernst, J. and M. Kellis, 2010: Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol*, **28**, 817-825.
- Evan, W., 2004: Why should we constrain stress and limitation? Why conceptual terms deserve broad definitions. *Journal of vegetation science*, **15**, 569-571.
- Fei, Y., Y. Xue, P. Du, S. Yang and X. Deng, 2017: Expression analysis and promoter methylation under osmotic and salinity stress of TaGAPC1 in wheat (*Triticum aestivum* L). *Protoplasma*, **254**, 987-996.
- Feng, S. and S. E. Jacobsen, 2011: Epigenetic modifications in plants: an evolutionary perspective. *Curr Opin Plant Biol*, **14**, 179-186.
- Forestan, C., R. Aiese Cigliano, S. Farinati, A. Lunardon, W. Sanseverino and S. Varotto, 2016: Stress-induced and epigenetic-mediated maize transcriptome regulation study by means of transcriptome reannotation and differential expression analysis. *Sci Rep*, **6**, 30446.
- Foyer, C. H. and G. Noctor, 2005: Oxidant and antioxidant signalling in plants: a re-evaluation of the concept of oxidative stress in a physiological context. *Plant, Cell & Environment*, **28**, 1056-1071.
- Foyer, C. H., H. Vanacker, L. D. Gomez and J. Harbinson, 2002: Regulation of photosynthesis and antioxidant metabolism in maize leaves at optimal and chilling temperatures. *Plant Physiology and Biochemistry*, **40**, 659-668.
- Fracheboud, Y., P. Haldimann, J. Leipner and P. Stamp, 1999: Chlorophyll fluorescence as a selection tool for cold tolerance of photosynthesis in maize (*Zea mays* L.). *Journal of experimental botany*, **50**, 1533-1540.
- Franz, P. and H. de Jong, 2011: From nucleosome to chromosome: a dynamic organization of genetic information. *Plant J*, **66**, 4-17.
- Frommer, M., L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy and C. L. Paul, 1992: A genomic sequencing protocol that yields a positive display of 5-methylcytosine

- residues in individual DNA strands. *Proceedings of the National Academy of Sciences*, **89**, 1827-1831.
- Furner, I. J. and M. Matzke, 2011: Methylation and demethylation of the Arabidopsis genome. *Curr Opin Plant Biol*, **14**, 137-141.
- Ganguly, D. R., P. A. Crisp, S. R. Eichten and B. J. Pogson, 2017: The Arabidopsis DNA Methylome Is Stable under Transgenerational Drought Stress. *Plant Physiol*, **175**, 1893-1912.
- Gehring, M. and S. Henikoff, 2007: DNA methylation dynamics in plant genomes. *Biochim Biophys Acta*, **1769**, 276-286.
- Gent, J. I., Y. Dong, J. Jiang and R. K. Dawe, 2011: Strong epigenetic similarity between maize centromeric and pericentromeric regions at the level of small RNAs, DNA methylation and H3 chromatin modifications. *Nucleic acids research*, **40**, 1550-1560.
- Gent, J. I., N. A. Ellis, L. Guo, A. E. Harkess, Y. Yao, X. Zhang and R. K. Dawe, 2013: CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. *Genome Res*, **23**, 628-637.
- Gent, J. I., T. F. Madzima, R. Bader, M. R. Kent, X. Zhang, M. Stam, K. M. McGinnis and R. K. Dawe, 2014: Accessible DNA and relative depletion of H3K9me2 at maize loci undergoing RNA-directed DNA methylation. *Plant Cell*, **26**, 4903-4917.
- Gerstein, M. B., Z. J. Lu, E. L. Van Nostrand, C. Cheng, B. I. Arshinoff, T. Liu, K. Y. Yip, R. Robilotto, A. Rechtsteiner, K. Ikegami, P. Alves, A. Chateigner, M. Perry, M. Morris, R. K. Auerbach, X. Feng, J. Leng, A. Vielle, W. Niu, K. Rhrissorrakrai, A. Agarwal, R. P. Alexander, G. Barber, C. M. Brdlik, J. Brennan, J. J. Brouillet, A. Carr, M. S. Cheung, H. Clawson, S. Contrino, L. O. Dannenberg, A. F. Dernburg, A. Desai, L. Dick, A. C. Dose, J. Du, T. Egelhofer, S. Ercan, G. Euskirchen, B. Ewing, E. A. Feingold, R. Gassmann, P. J. Good, P. Green, F. Gullier, M. Gutwein, M. S. Guyer, L. Habegger, T. Han, J. G. Henikoff, S. R. Henz, A. Hinrichs, H. Holster, T. Hyman, A. L. Iniguez, J. Janette, M. Jensen, M. Kato, W. J. Kent, E. Kephart, V. Khivansara, E. Khurana, J. K. Kim, P. Kolasinska-Zwierz, E. C. Lai, I. Latorre, A. Leahey, S. Lewis, P. Lloyd, L. Lochovsky, R. F. Lowdon, Y. Lubling, R. Lyne, M. MacCoss, S. D. Mackowiak, M. Mangone, S. McKay, D. Mecnas, G. Merrihew, D. M. Miller, 3rd, A. Muroyama, J. I. Murray, S. L. Ooi, H. Pham, T. Phippen, E. A. Preston, N. Rajewsky, G. Ratsch, H. Rosenbaum, J. Rozowsky, K. Rutherford, P. Ruzanov, M. Sarov, R. Sasidharan, A. Sboner, P. Scheid, E. Segal, H. Shin, C. Shou, F. J. Slack, et al., 2010: Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science*, **330**, 1775-1787.
- Giauffret, C., R. Bonhomme and M. Derieux, 1995: Genotypic differences for temperature response of leaf appearance rate and leaf elongation rate in field-grown maize. *Agronomie*, **15**, 123-137.
- Gibney, E. R. and C. M. Nolan, 2010: Epigenetics and gene expression. *Heredity (Edinb)*, **105**, 4-13.
- González, R. M., M. M. Ricardi and N. D. Iusem, 2011: Atypical epigenetic mark in an atypical location: cytosine methylation at asymmetric (CNN) sites within the body of a non-repetitive tomato gene. *BMC plant biology*, **11**, 94.
- Grobman, A., D. Bonavia, T. D. Dillehay, D. R. Piperno, J. Iriarte and I. Holst, 2012: Pre-ceramic maize from Paredones and Huaca Prieta, Peru. *Proceedings of the National Academy of Sciences*, **109**, 1755-1759.
- Gu, T., S. Ren, Y. Wang, Y. Han and Y. Li, 2016: Characterization of DNA methyltransferase and demethylase genes in *Fragaria vesca*. *Mol Genet Genomics*, **291**, 1333-1345.
- Hagmann, J., C. Becker, J. Müller, O. Stegle, R. C. Meyer, G. Wang, K. Schneeberger, J. Fitz, T. Altmann and J. Bergelson, 2015: Century-scale methylome stability in a recently diverged Arabidopsis thaliana lineage. *PLoS genetics*, **11**, e1004920.
- Hashida, S. N., T. Uchiyama, C. Martin, Y. Kishima, Y. Sano and T. Mikami, 2006: The temperature-dependent change in methylation of the Antirrhinum transposon Tam3 is controlled by the activity of its transposase. *Plant Cell*, **18**, 104-118.
- Herr, A. J., M. B. Jensen, T. Dalmay and D. C. Baulcombe, 2005: RNA polymerase IV directs silencing of endogenous DNA. *Science*, **308**, 118-120.
- Hirsch, C. D. and N. M. Springer, 2017: Transposable element influences on gene expression in plants. *Biochim Biophys Acta*, **1860**, 157-165.

- Hirsch, C. N., J. M. Foerster, J. M. Johnson, R. S. Sekhon, G. Muttoni, B. Vaillancourt, F. Peñagaricano, E. Lindquist, M. A. Pedraza and K. Barry, 2014: Insights into the maize pan-genome and pan-transcriptome. *The Plant Cell*, **26**, 121-135.
- Hofmeister, B. T., K. Lee, N. A. Rohr, D. W. Hall and R. J. Schmitz, 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.
- Holliday, R., 1987: The inheritance of epigenetic defects. *Science*, **238**, 163-170.
- Holliday, R., 1994: Epigenetics: an overview. *genesis*, **15**, 453-457.
- Hosaka, A. and T. Kakutani, 2018: Transposable elements, genome evolution and transgenerational epigenetic variation. *Curr Opin Genet Dev*, **49**, 43-48.
- Hu, L., N. Li, C. Xu, S. Zhong, X. Lin, J. Yang, T. Zhou, A. Yuliang, Y. Wu and Y.-R. Chen, 2014: Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proceedings of the National Academy of Sciences*, **111**, 10642-10647.
- Hu, Y., L. Zhang, S. He, M. Huang, J. Tan, L. Zhao, S. Yan, H. Li, K. Zhou, Y. Liang and L. Li, 2012: Cold stress selectively unsilences tandem repeats in heterochromatin associated with accumulation of H3K9ac. *Plant Cell Environ*, **35**, 2130-2142.
- Hu, Y., L. Zhang, L. Zhao, J. Li, S. He, K. Zhou, F. Yang, M. Huang, L. Jiang and L. Li, 2011: Trichostatin A selectively suppresses the cold-induced transcription of the ZmDREB1 gene in maize. *PLoS One*, **6**, e22132.
- Hund, A., Y. Fracheboud, A. Soldati, E. Frascaroli, S. Salvi and P. Stamp, 2004: QTL controlling root and shoot traits of maize seedlings under cold stress. *Theor Appl Genet*, **109**, 618-629.
- Ibarra, C. A., X. Feng, V. K. Schoft, T.-F. Hsieh, R. Uzawa, J. A. Rodrigues, A. Zemach, N. Chumak, A. Machlicova and T. Nishimura, 2012: Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science*, **337**, 1360-1364.
- Irizarry, R. A., C. Ladd-Acosta, B. Carvalho, H. Wu, S. A. Brandenburg, J. A. Jeddelloh, B. Wen and A. P. Feinberg, 2008: Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome research*, **18**, 780-790.
- Ito, H., H. Gaubert, E. Bucher, M. Mirouze, I. Vaillant and J. Paszkowski, 2011: An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature*, **472**, 115.
- Jackson, R. D., 1986: Remote sensing of biotic and abiotic plant stress. *Annual review of phytopathology*, **24**, 265-287.
- Janda, T., G. Szalai, I. Tari and E. Paldi, 1999: Hydroponic treatment with salicylic acid decreases the effects of chilling injury in maize (*Zea mays* L.) plants. *Planta*, **208**, 175-180.
- Janowiak, F., B. Maas and K. Dörffling, 2002: Importance of abscisic acid for chilling tolerance of maize seedlings. *Journal of Plant Physiology*, **159**, 635-643.
- Ji, L., D. A. Neumann and R. J. Schmitz, 2015: Crop Epigenomics: Identifying, Unlocking, and Harnessing Cryptic Variation in Crop Genomes. *Mol Plant*, **8**, 860-870.
- Jiao, Y., P. Peluso, J. Shi, T. Liang, M. C. Stitzer, B. Wang, M. S. Campbell, J. C. Stein, X. Wei, C. S. Chin, K. Guill, M. Regulski, S. Kumari, A. Olson, J. Gent, K. L. Schneider, T. K. Wolfgruber, M. R. May, N. M. Springer, E. Antoniou, W. R. McCombie, G. G. Presting, M. McMullen, J. Ross-Ibarra, R. K. Dawe, A. Hastie, D. R. Rank and D. Ware, 2017: Improved maize reference genome with single-molecule technologies. *Nature*, **546**, 524-527.
- Johannes, F., E. Porcher, F. K. Teixeira, V. Saliba-Colombani, M. Simon, N. Agier, A. Bulski, J. Albuissou, F. Heredia, P. Audigier, D. Bouchez, C. Dillmann, P. Guerche, F. Hospital and V. Colot, 2009: Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet*, **5**, e1000530.
- Jones, P. A., 2012: Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*, **13**, 484-492.
- Jullien, P. E., D. Susaki, R. Yelagandula, T. Higashiyama and F. Berger, 2012: DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Current Biology*, **22**, 1825-1830.

- Karan, R., T. DeLeon, H. Biradar and P. K. Subudhi, 2012: Salt stress induced variation in DNA methylation pattern and its influence on gene expression in contrasting rice genotypes. *PLoS One*, **7**, e40203.
- Kawakatsu, T., S.-s. C. Huang, F. Jupe, E. Sasaki, R. J. Schmitz, M. A. Urich, R. Castanon, J. R. Nery, C. Barragan and Y. He, 2016: Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell*, **166**, 492-505.
- Kawashima, T. and F. Berger, 2014: Epigenetic reprogramming in plant sexual reproduction. *Nature Reviews. Genetics*, **15**, 613.
- Kharchenko, P. V., A. A. Alekseyenko, Y. B. Schwartz, A. Minoda, N. C. Riddle, J. Ernst, P. J. Sabo, E. Larschan, A. A. Gorchakov, T. Gu, D. Linder-Basso, A. Plachetka, G. Shanower, M. Y. Tolstorukov, L. J. Luquette, R. Xi, Y. L. Jung, R. W. Park, E. P. Bishop, T. K. Canfield, R. Sandstrom, R. E. Thurman, D. M. MacAlpine, J. A. Stamatoyannopoulos, M. Kellis, S. C. Elgin, M. I. Kuroda, V. Pirrotta, G. H. Karpen and P. J. Park, 2011: Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature*, **471**, 480-485.
- Kim, M. Y. and D. Zilberman, 2014: DNA methylation as a system of plant genomic immunity. *Trends Plant Sci*, **19**, 320-326.
- Koo, D.-H., F. Han, J. A. Birchler and J. Jiang, 2011: Distinct DNA methylation patterns associated with active and inactive centromeres of the maize B chromosome. *Genome research*, **21**, 908-914.
- Kopylova, E., L. Noé and H. Touzet, 2012: SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, **28**, 3211-3217.
- Körner, C., 2007: The use of 'altitude' in ecological research. *Trends in ecology & evolution*, **22**, 569-574.
- Kou, H. P., Y. Li, X. X. Song, X. F. Ou, S. C. Xing, J. Ma, D. Von Wettstein and B. Liu, 2011: Heritable alteration in DNA methylation induced by nitrogen-deficiency stress accompanies enhanced tolerance by progenies to the stress in rice (*Oryza sativa* L.). *J Plant Physiol*, **168**, 1685-1693.
- Krueger, F. and S. R. Andrews, 2011: Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *bioinformatics*, **27**, 1571-1572.
- Kusano, T., T. Berberich, M. Harada, N. Suzuki and K. Sugawara, 1995: A maize DNA-binding factor with a bZIP motif is induced by low temperature. *Molecular and General Genetics MGG*, **248**, 507-517.
- Labra, M., A. Ghiani, S. Citterio, S. Sgorbati, F. Sala, C. Vannini, M. Ruffini-Castiglione and M. Bracale, 2002: Analysis of cytosine methylation pattern in response to water deficit in pea root tips. *Plant Biology*, **4**, 694-699.
- Lacoste, N. and J. Côté, 2003: Le code épigénétique des histones. *médecine/sciences*, **19**, 955-959.
- Lanciano, S., M. C. Carpentier, C. Llauro, E. Jobet, D. Robakowska-Hyzorek, E. Lasserre, A. Ghesquiere, O. Panaud and M. Mirouze, 2017: Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. *PLoS Genet*, **13**, e1006630.
- Lauria, M. and V. Rossi, 2011: Epigenetic control of gene regulation in plants. *Biochim Biophys Acta*, **1809**, 369-378.
- Law, J. A., J. Du, C. J. Hale, S. Feng, K. Krajewski, A. M. Palanca, B. D. Strahl, D. J. Patel and S. E. Jacobsen, 2013: Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature*, **498**, 385-389.
- Law, J. A. and S. E. Jacobsen, 2010: Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews. Genetics*, **11**, 204.
- Law, J. A., A. A. Vashisht, J. A. Wohlschlegel and S. E. Jacobsen, 2011: SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet*, **7**, e1002195.
- Le, T. N., Y. Miyazaki, S. Takuno and H. Saze, 2015: Epigenetic regulation of intragenic transposable elements impacts gene transcription in *Arabidopsis thaliana*. *Nucleic Acids Res*, **43**, 3911-3921.
- Lee, E., M. Staebler and M. Tollenaar, 2002a: Genetic variation in physiological discriminators for cold tolerance—early autotrophic phase of maize development. *Crop Science*, **42**, 1919-1929.

- Lee, M., N. Sharopova, W. D. Beavis, D. Grant, M. Katt, D. Blair and A. Hallauer, 2002b: Expanding the genetic map of maize with the intermated B73× Mo17 (IBM) population. *Plant molecular biology*, **48**, 453-461.
- Leipner, J. and E. Mayer, 2008: QTL mapping in maize seedlings reveals little relevance of C4 cycle enzymes and antioxidants for genotypic differences in chilling tolerance of photosynthesis. *Maydica*, **53**, 269-277.
- Leipner, J. and P. Stamp, 2009: Chilling stress in maize seedlings. *Handbook of Maize: Its Biology*. Springer.
- Leung, J., C. Valon, B. Moreau, M. Boeglin, C. Lefoulon, A. Joshi-Saha and I. Chérel, 2012: Mécanisme moléculaire d'action de l'acide abscissique en réponse à la sécheresse chez les végétaux. *Biologie Aujourd'hui*, **206**, 301-312.
- Lev Maor, G., A. Yearim and G. Ast, 2015: The alternative role of DNA methylation in splicing regulation. *Trends Genet*, **31**, 274-280.
- Levitt, J., 1972: Response of plants to environmental stresses. *Water, radiation, Salt and other stresses*, **2**.
- Lewandowska-Gnatowska, E., L. Polkowska-Kowalczyk, J. Szczegielniak, M. Barciszewska, J. Barciszewski and G. Muszynska, 2014: Is DNA methylation modulated by wounding-induced oxidative burst in maize? *Plant Physiol Biochem*, **82**, 202-208.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing, 2009: The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.
- Li, Q., S. R. Eichten, P. J. Hermanson and N. M. Springer, 2014a: Inheritance patterns and stability of DNA methylation variation in maize near-isogenic lines. *Genetics*, **196**, 667-676.
- Li, Q., S. R. Eichten, P. J. Hermanson, V. M. Zaunbrecher, J. Song, J. Wendt, H. Rosenbaum, T. F. Madzima, A. E. Sloan, J. Huang, D. L. Burgess, T. A. Richmond, K. M. McGinnis, R. B. Meeley, O. N. Danilevskaya, M. W. Vaughn, S. M. Kaeppler, J. A. Jeddelloh and N. M. Springer, 2014b: Genetic perturbation of the maize methylome. *Plant Cell*, **26**, 4602-4616.
- Li, Q., J. Song, P. T. West, G. Zynda, S. R. Eichten, M. W. Vaughn and N. M. Springer, 2015: Examining the Causes and Consequences of Context-Specific Differential DNA Methylation in Maize. *Plant Physiol*, **168**, 1262-1274.
- Li, X., J. Zhu, F. Hu, S. Ge, M. Ye, H. Xiang, G. Zhang, X. Zheng, H. Zhang, S. Zhang, Q. Li, R. Luo, C. Yu, J. Yu, J. Sun, X. Zou, X. Cao, X. Xie, J. Wang and W. Wang, 2012: Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics*, **13**, 300.
- Lisch, D., 2009: Epigenetic regulation of transposable elements in plants. *Annu Rev Plant Biol*, **60**, 43-66.
- Lisch, D., 2013: How important are transposons for plant evolution? *Nat Rev Genet*, **14**, 49-61.
- Lister, R., R. C. O'Malley, J. Tonti-Filippini, B. D. Gregory, C. C. Berry, A. H. Millar and J. R. Ecker, 2008: Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell*, **133**, 523-536.
- Liu, J., F. Wang, G. Yu, X. Zhang, C. Jia, J. Qin and H. Pan, 2015: Functional Analysis of the Maize C-Repeat/DRE Motif-Binding Transcription Factor CBF3 Promoter in Response to Abiotic Stress. *Int J Mol Sci*, **16**, 12131-12146.
- Liu, T., A. Rechtsteiner, T. A. Egelhofer, A. Vielle, I. Latorre, M. S. Cheung, S. Ercan, K. Ikegami, M. Jensen, P. Kolasinska-Zwierz, H. Rosenbaum, H. Shin, S. Taing, T. Takasaki, A. L. Iniguez, A. Desai, A. F. Dernburg, H. Kimura, J. D. Lieb, J. Ahringer, S. Strome and X. S. Liu, 2011: Broad chromosomal domains of histone modification patterns in *C. elegans*. *Genome Res*, **21**, 227-236.
- Long, S. P. and A. K. Spence, 2013: Toward cool C4 crops. *Annual Review of Plant Biology*, **64**, 701-722.
- Louarn, G., B. Andrieu and C. Giauffret, 2010: A size-mediated effect can compensate for transient chilling stress affecting maize (*Zea mays*) leaf extension. *New Phytol*, **187**, 106-118.

- Lu, X., W. Wang, W. Ren, Z. Chai, W. Guo, R. Chen, L. Wang, J. Zhao, Z. Lang, Y. Fan, J. Zhao and C. Zhang, 2015: Genome-Wide Epigenetic Regulation of Gene Transcription in Maize Seeds. *PLoS One*, **10**, e0139582.
- Lu, Y., T. Rong and M. Cao, 2008: Analysis of DNA methylation in different maize tissues. *Journal of Genetics and Genomics*, **35**, 41-48.
- Lunardon, A., C. Forestan, S. Farinati, M. J. Axtell and S. Varotto, 2016: Genome-Wide Characterization of Maize Small RNA Loci and Their Regulation in the required to maintain repression6-1 (rnr6-1) Mutant and Long-Term Abiotic Stresses. *Plant Physiol*, **170**, 1535-1548.
- Maiti, R. K. and P. Satya, 2014: Research advances in major cereal crops for adaptation to abiotic stresses. *GM crops & food*, **5**, 259-279.
- Makarevitch, I., A. J. Waters, P. T. West, M. Stitzer, C. N. Hirsch, J. Ross-Ibarra and N. M. Springer, 2015: Transposable elements contribute to activation of maize genes in response to abiotic stress. *PLoS Genet*, **11**, e1004915.
- Malvar, R. A., M. Cartea, P. Revilla and A. Ordás, 1997: Identification of field corn inbreds adapted to Europe to improve agronomic performance of sweet corn hybrids. *Crop science*, **37**, 1134-1141.
- Manning, K., M. Tor, M. Poole, Y. Hong, A. J. Thompson, G. J. King, J. J. Giovannoni and G. B. Seymour, 2006: A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat Genet*, **38**, 948-952.
- Marks, P. A., R. A. Rifkind, V. M. Richon, R. Breslow, T. Miller and W. K. Kelly, 2001: Histone deacetylases and cancer: causes and therapies. *Nature Reviews Cancer*, **1**, 194.
- Marocco, A., C. Lorenzoni and Y. Francheboud, 2005: Chilling stress in maize. *Maydica*, **50**, 571.
- Maron, L. G., M. A. Piñeros, L. V. Kochian and S. R. McCouch, 2016: Redefining 'stress resistance genes', and why it matters. *Journal of experimental botany*, **67**, 5588.
- Martienssen, R. and A. Baron, 1994: Coordinate suppression of mutations caused by Robertson's mutator transposons in maize. *Genetics*, **136**, 1157-1170.
- Martin, A., C. Troadec, A. Boualem, M. Rajab, R. Fernandez, H. Morin, M. Pitrat, C. Dogimont and A. Bendahmane, 2009: A transposon-induced epigenetic change leads to sex determination in melon. *Nature*, **461**, 1135.
- Matsuoka, Y., Y. Vigouroux, M. M. Goodman, J. Sanchez, E. Buckler and J. Doebley, 2002: A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences*, **99**, 6080-6084.
- Matzke, M. A. and R. A. Mosher, 2014: RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics*, **15**, 394-408.
- Melkonian, J., L.-X. Yu and T. L. Setter, 2004: Chilling responses of maize (*Zea mays* L.) seedlings: root hydraulic conductance, abscisic acid, and stomatal conductance. *Journal of Experimental Botany*, **55**, 1751-1760.
- Meyer, P., 2015: Epigenetic variation and environmental change. *J Exp Bot*, **66**, 3541-3548.
- Mirouze, M. and C. Vitte, 2014: Transposable elements, a treasure trove to decipher epigenetic variation: insights from Arabidopsis and crop epigenomes. *J Exp Bot*, **65**, 2801-2812.
- Morgan, H. D., H. G. Sutherland, D. I. Martin and E. Whitelaw, 1999: Epigenetic inheritance at the agouti locus in the mouse. *Nature genetics*, **23**, 314-318.
- Naydenov, M., V. Baev, E. Apostolova, N. Gospodinova, G. Sablok, M. Gozmanova and G. Yahubyan, 2015a: High-temperature effect on genes engaged in DNA methylation and affected by DNA methylation in Arabidopsis. *Plant Physiol Biochem*, **87**, 102-108.
- Naydenov, M., V. Baev, E. Apostolova, N. Gospodinova, G. Sablok, M. Gozmanova and G. Yahubyan, 2015b: High-temperature effect on genes engaged in DNA methylation and affected by DNA methylation in Arabidopsis. *Plant Physiology and Biochemistry*, **87**, 102-108.
- Nguyen, H. T., J. Leipner, P. Stamp and O. Guerra-Peraza, 2009: Low temperature stress in maize (*Zea mays* L.) induces genes involved in photosynthesis and signal transduction as studied by suppression subtractive hybridization. *Plant Physiol Biochem*, **47**, 116-122.

- Niederhuth, C. E., A. J. Bewick, L. Ji, M. S. Alabady, K. Do Kim, Q. Li, N. A. Rohr, A. Rambani, J. M. Burke and J. A. Udall, 2016: Widespread natural variation of DNA methylation within angiosperms. *Genome biology*, **17**, 194.
- Niederhuth, C. E. and R. J. Schmitz, 2017: Putting DNA methylation in context: from genomes to gene expression in plants. *Biochim Biophys Acta*, **1860**, 149-156.
- O'Malley, R. C., S. C. Huang, L. Song, M. G. Lewsey, A. Bartlett, J. R. Nery, M. Galli, A. Gallavotti and J. R. Ecker, 2016: Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell*, **165**, 1280-1292.
- Ong-Abdullah, M., J. M. Ordway, N. Jiang, S. E. Ooi, S. Y. Kok, N. Sarpan, N. Azimi, A. T. Hashim, Z. Ishak, S. K. Rosli, F. A. Malike, N. A. Bakar, M. Marjuni, N. Abdullah, Z. Yaakub, M. D. Amiruddin, R. Nookiah, R. Singh, E. T. Low, K. L. Chan, N. Azizi, S. W. Smith, B. Bacher, M. A. Budiman, A. Van Brunt, C. Wischmeyer, M. Beil, M. Hogan, N. Lakey, C. C. Lim, X. Arulandoo, C. K. Wong, C. N. Choo, W. C. Wong, Y. Y. Kwan, S. S. Alwee, R. Sambanthamurthi and R. A. Martienssen, 2015: Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*, **525**, 533-537.
- Orlando, V., 2003: Polycomb, epigenomes, and control of cell identity. *Cell*, **112**, 599-606.
- Osabe, K., J. D. Clement, F. Bedon, F. A. Pettolino, L. Ziolkowski, D. J. Llewellyn, E. J. Finnegan and I. W. Wilson, 2014: Genetic and DNA methylation changes in cotton (*Gossypium*) genotypes and tissues. *PLoS one*, **9**, e86049.
- Pan, J., M. Zhang, X. Kong, X. Xing, Y. Liu, Y. Zhou, Y. Liu, L. Sun and D. Li, 2012: ZmMPK17, a novel maize group D MAP kinase gene, is involved in multiple stress responses. *Planta*, **235**, 661-676.
- Parent, B., O. Turc, Y. Gibon, M. Stitt and F. Tardieu, 2010: Modelling temperature-compensated physiological rates, based on the co-ordination of responses to temperature of developmental processes. *Journal of Experimental Botany*, **61**, 2057-2069.
- Pecinka, A., H. Q. Dinh, T. Baubec, M. Rosa, N. Lettner and O. Mittelsten Scheid, 2010: Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. *Plant Cell*, **22**, 3118-3129.
- Pecinka, A. and O. Mittelsten Scheid, 2012: Stress-induced chromatin changes: a critical view on their heritability. *Plant and Cell Physiology*, **53**, 801-808.
- Penterman, J., D. Zilberman, J. H. Huh, T. Ballinger, S. Henikoff and R. L. Fischer, 2007: DNA demethylation in the *Arabidopsis* genome. *Proc Natl Acad Sci U S A*, **104**, 6752-6757.
- Perry, L., D. H. Sandweiss, D. R. Piperno, K. Rademaker, M. A. Malpass, A. Umire and P. De la Vera, 2006: Early maize agriculture and interzonal interaction in southern Peru. *Nature*, **440**, 76.
- Pikaard, C. S. and O. Mittelsten Scheid, 2014: Epigenetic regulation in plants. *Cold Spring Harb Perspect Biol*, **6**, a019315.
- Pimentel, C., P. A. Davey, J. A. Juvik and S. P. Long, 2005: Gene loci in maize influencing susceptibility to chilling dependent photoinhibition of photosynthesis. *Photosynthesis Research*, **85**, 319-326.
- Piperno, D. R., A. J. Ranere, I. Holst, J. Iriarte and R. Dickau, 2009: Starch grain and phytolith evidence for early ninth millennium BP maize from the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences*, **106**, 5019-5024.
- Probst, A. V. and O. Mittelsten Scheid, 2015: Stress-induced structural changes in plant chromatin. *Curr Opin Plant Biol*, **27**, 8-16.
- Prokhortchouk, E. and P. A. Defossez, 2008: The cell biology of DNA methylation in mammals. *Biochim Biophys Acta*, **1783**, 2167-2173.
- Qian, Y., Y. Xi, B. Cheng and S. Zhu, 2014: Genome-wide identification and expression profiling of DNA methyltransferase gene family in maize. *Plant Cell Rep*, **33**, 1661-1672.
- Qin, F., Y. Sakuma, J. Li, Q. Liu, Y.-Q. Li, K. Shinozaki and K. Yamaguchi-Shinozaki, 2004: Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant and Cell Physiology*, **45**, 1042-1052.
- Qiu, J., 2006: Epigenetics: unfinished symphony. *Nature*, **441**, 143-145.

- Quadrana, L., A. Bortolini Silveira, G. F. Mayhew, C. LeBlanc, R. A. Martienssen, J. A. Jeddelloh and V. Colot, 2016a: The Arabidopsis thaliana mobilome and its impact at the species level. *Elife*, **5**.
- Quadrana, L. and V. Colot, 2016: Plant Transgenerational Epigenetics. *Annu Rev Genet*, **50**, 467-491.
- Quadrana, L., A. B. Silveira, G. F. Mayhew, C. LeBlanc, R. A. Martienssen, J. A. Jeddelloh and V. Colot, 2016b: The Arabidopsis thaliana mobilome and its impact at the species level. *Elife*, **5**, e15716.
- Quinlan, A. R. and I. M. Hall, 2010: BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841-842.
- Rebourg, C., M. Chastanet, B. Gouesnard, C. Welcker, P. Dubreuil and A. Charcosset, 2003: Maize introduction into Europe: the history reviewed in the light of molecular data. *Theoretical and applied genetics*, **106**, 895-903.
- Regulski, M., Z. Lu, J. Kendall, M. T. Donoghue, J. Reinders, V. Llaca, S. Deschamps, A. Smith, D. Levy, W. R. McCombie, S. Tingey, A. Rafalski, J. Hicks, D. Ware and R. A. Martienssen, 2013: The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res*, **23**, 1651-1662.
- Reinders, J., B. B. Wulff, M. Mirouze, A. Marí-Ordóñez, M. Dapp, W. Rozhon, E. Bucher, G. Theiler and J. Paszkowski, 2009: Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes & Development*, **23**, 939-950.
- Rey, O., E. Danchin, M. Mirouze, C. Loot and S. Blanchet, 2016: Adaptation to Global Change: A Transposable Element-Epigenetics Perspective. *Trends Ecol Evol*, **31**, 514-526.
- Reyna-Lopez, G., J. Simpson and J. Ruiz-Herrera, 1997: Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. *Molecular and General Genetics MGG*, **253**, 703-710.
- Richards, E. J., 2006: Inherited epigenetic variation--revisiting soft inheritance. *Nature reviews. Genetics*, **7**, 395.
- Rigaill, G., S. Balzergue, V. Brunaud, E. Blondet, A. Rau, O. Rogier, J. Caius, C. Maugis-Rabusseau, L. Soubigou-Taconnat and S. Aubourg, 2016: Synthetic data sets for the identification of key ingredients for RNA-seq differential analysis. *Briefings in bioinformatics*, **19**, 65-76.
- Riva-Roveda, L., B. Escale, C. Giauffret and C. Périlleux, 2016: Maize plants can enter a standby mode to cope with chilling stress. *BMC plant biology*, **16**, 212.
- Riva-Roveda, L. and C. Périlleux, 2015: Effets du froid sur les stades précoces de développement du maïs (*Zea mays* L.)(synthèse bibliographique). *Biotechnologie, Agronomie, Société et Environnement*, **19**, 42.
- Rodríguez, V. M., A. Butrón, M. O. A. Rady, P. Soengas and P. Revilla, 2013: Identification of quantitative trait loci involved in the response to cold stress in maize (*Zea mays* L.). *Molecular Breeding*, **33**, 363-371.
- Roessler, K., S. Takuno and B. S. Gaut, 2016: CG Methylation Covaries with Differential Gene Expression between Leaf and Floral Bud Tissues of *Brachypodium distachyon*. *PLoS One*, **11**, e0150002.
- Roudier, F., I. Ahmed, C. Berard, A. Sarazin, T. Mary-Huard, S. Cortijo, D. Bouyer, E. Caillieux, E. Duvernois-Berthet, L. Al-Shikhley, L. Giraut, B. Despres, S. Drevensek, F. Barneche, S. Derozier, V. Brunaud, S. Aubourg, A. Schnittger, C. Bowler, M. L. Martin-Magniette, S. Robin, M. Caboche and V. Colot, 2011: Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J*, **30**, 1928-1938.
- Roudier, F., F. K. Teixeira and V. Colot, 2009: Chromatin indexing in Arabidopsis: an epigenomic tale of tails and more. *Trends Genet*, **25**, 511-517.
- Roy, S., J. Ernst, P. V. Kharchenko, P. Kheradpour, N. Negre, M. L. Eaton, J. M. Landolin, C. A. Bristow, L. Ma and M. F. Lin, 2010: Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science*, **330**, 1787-1797.
- Ruiz-Garcia, L., M. Cervera and J. Martinez-Zapater, 2005: DNA methylation increases throughout Arabidopsis development. *Planta*, **222**, 301-306.
- Santos, A. P., L. J. Ferreira and M. M. Oliveira, 2017: Concerted Flexibility of Chromatin Structure, Methylome, and Histone Modifications along with Plant Stress Responses. *Biology (Basel)*, **6**.

- Saze, H. and T. Kakutani, 2007: Heritable epigenetic mutation of a transposon-flanked Arabidopsis gene due to lack of the chromatin-remodeling factor DDM1. *The EMBO Journal*, **26**, 3641-3652.
- Schmitz, R. J., Y. He, O. Valdes-Lopez, S. M. Khan, T. Joshi, M. A. Urich, J. R. Nery, B. Diers, D. Xu, G. Stacey and J. R. Ecker, 2013a: Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome Res*, **23**, 1663-1674.
- Schmitz, R. J., M. D. Schultz, M. A. Urich, J. R. Nery, M. Pelizzola, O. Libiger, A. Alix, R. B. McCosh, H. Chen, N. J. Schork and J. R. Ecker, 2013b: Patterns of population epigenomic diversity. *Nature*, **495**, 193-198.
- Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton and T. A. Graves, 2009: The B73 maize genome: complexity, diversity, and dynamics. *science*, **326**, 1112-1115.
- Schueler, M. G. and B. A. Sullivan, 2006: Structural and functional dynamics of human centromeric chromatin. *Annu. Rev. Genomics Hum. Genet.*, **7**, 301-313.
- Scott, K. C., 2013: Transcription and ncRNAs: at the cent (rome) re of kinetochore assembly and maintenance. *Chromosome research*, **21**, 643-651.
- Secco, D., C. Wang, H. Shou, M. D. Schultz, S. Chiarenza, L. Nussaume, J. R. Ecker, J. Whelan and R. Lister, 2015: Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *Elife*, **4**.
- Seisenberger, S., J. R. Peat, T. A. Hore, F. Santos, W. Dean and W. Reik, 2013: Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **368**, 20110330.
- Seymour, D. K. and C. Becker, 2017: The causes and consequences of DNA methylome variation in plants. *Curr Opin Plant Biol*, **36**, 56-63.
- Shan, X., Y. Li, Y. Jiang, Z. Jiang, W. Hao and Y. Yuan, 2013: Transcriptome Profile Analysis of Maize Seedlings in Response to High-salinity, Drought and Cold Stresses by Deep Sequencing. *Plant Molecular Biology Reporter*, **31**, 1485-1491.
- Sharma, R., R. K. Mohan Singh, G. Malik, P. Deveshwar, A. K. Tyagi, S. Kapoor and M. Kapoor, 2009: Rice cytosine DNA methyltransferases - gene expression profiling during reproductive development and abiotic stress. *FEBS J*, **276**, 6301-6311.
- Shi, J. and R. K. Dawe, 2006: Partitioning of the maize epigenome by the number of methyl groups on histone H3 lysines 9 and 27. *Genetics*, **173**, 1571-1583.
- Slotkin, R. K. and R. Martienssen, 2007: Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*, **8**, 272-285.
- Smith, A. M., C. N. Hansey and S. M. Kaeppler, 2012: TCUP: A Novel hAT Transposon Active in Maize Tissue Culture. *Front Plant Sci*, **3**, 6.
- Sobkowiak, A., M. Jonczyk, E. Jarochowska, P. Biecek, J. Trzcinska-Danielewicz, J. Leipner, J. Fronk and P. Sowinski, 2014: Genome-wide transcriptomic analysis of response to low temperature reveals candidate genes determining divergent cold-sensitivity of maize inbred lines. *Plant Mol Biol*, **85**, 317-331.
- Song, Q., B. Decato, E. E. Hong, M. Zhou, F. Fang, J. Qu, T. Garvin, M. Kessler, J. Zhou and A. D. Smith, 2013: A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. *PLoS one*, **8**, e81148.
- Song, Y., D. Ji, S. Li, P. Wang, Q. Li and F. Xiang, 2012: The dynamic changes of DNA methylation and histone modifications of salt responsive transcription factor genes in soybean. *PLoS One*, **7**, e41274.
- Springer, N. M., 2013: Epigenetics and crop improvement. *Trends Genet*, **29**, 241-247.
- Springer, N. M., D. Lisch and Q. Li, 2016: Creating Order from Chaos: Epigenome Dynamics in Plants with Complex Genomes. *Plant Cell*, **28**, 314-325.
- Springer, N. M. and R. J. Schmitz, 2017: Exploiting induced and natural epigenetic variation for crop improvement. *Nat Rev Genet*, **18**, 563-575.

- Stein, J. C., Y. Yu, D. Copetti, D. J. Zwickl, L. Zhang, C. Zhang, K. Chougule, D. Gao, A. Iwata and J. L. Goicoechea, 2018: Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nature Genetics*, **1**.
- Steward, N., M. Ito, Y. Yamaguchi, N. Koizumi and H. Sano, 2002: Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *J Biol Chem*, **277**, 37741-37746.
- Steward, N., T. Kusano and H. Sano, 2000: Expression of ZmMET1, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in actively proliferating cells, but also with altered DNA methylation status in cold-stressed quiescent cells. *Nucleic Acids Research*, **28**, 3250-3259.
- Stroud, H., T. Do, J. Du, X. Zhong, S. Feng, L. Johnson, D. J. Patel and S. E. Jacobsen, 2014: Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nature structural & molecular biology*, **21**, 64-72.
- Stroud, H., M. V. Greenberg, S. Feng, Y. V. Bernatavichute and S. E. Jacobsen, 2013: Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell*, **152**, 352-364.
- Stuart, T., S. R. Eichten, J. Cahn, Y. Karpievitch, J. O. Borevitz and R. Lister, 2016: Population scale mapping of novel transposable element diversity reveals links to gene regulation and epigenomic variation. *bioRxiv*, 039511.
- Swigonova, Z., J. Lai, J. Ma, W. Ramakrishna, V. Llaca, J. L. Bennetzen and J. Messing, 2004: On the tetraploid origin of the maize genome. *Comp Funct Genomics*, **5**, 281-284.
- Takuno, S. and B. S. Gaut, 2012: Body-methylated genes in Arabidopsis thaliana are functionally important and evolve slowly. *Mol Biol Evol*, **29**, 219-227.
- Takuno, S. and B. S. Gaut, 2013: Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proceedings of the National Academy of Sciences*, **110**, 1797-1802.
- Takuno, S., P. Ralph, K. Swarts, R. J. Elshire, J. C. Glaubitz, E. S. Buckler, M. B. Hufford and J. Ross-Ibarra, 2015: Independent molecular basis of convergent highland adaptation in maize. *Genetics*, **200**, 1297-1312.
- Takuno, S., J. H. Ran and B. S. Gaut, 2016: Evolutionary patterns of genic DNA methylation vary across land plants. *Nat Plants*, **2**, 15222.
- Talbert, P. B., R. Masuelli, A. P. Tyagi, L. Comai and S. Henikoff, 2002: Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. *The Plant Cell*, **14**, 1053-1066.
- Teixeira, F. K. and V. Colot, 2009: Gene body DNA methylation in plants: a means to an end or an end to a means? *The EMBO journal*, **28**, 997-998.
- Tenaillon, M. I. and A. Charcosset, 2011: A European perspective on maize history. *Comptes rendus biologies*, **334**, 221-228.
- Teyssier, E., G. Bernacchia, S. Maury, A. H. Kit, L. Stammitti-Bert, D. Rolin and P. Gallusci, 2008: Tissue dependent variations of DNA methylation and endoreduplication levels during tomato fruit development and ripening. *Planta*, **228**, 391.
- Thieme, M., S. Lanciano, S. Balzergue, N. Daccord, M. Mirouze and E. Bucher, 2017: Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. *Genome Biol*, **18**, 134.
- Till, B. J., S. H. Reynolds, C. Weil, N. Springer, C. Burtner, K. Young, E. Bowers, C. A. Codomo, L. C. Enns and A. R. Odden, 2004: Discovery of induced point mutations in maize genes by TILLING. *BMC plant biology*, **4**, 12.
- Topp, C. N., C. X. Zhong and R. K. Dawe, 2004: Centromere-encoded RNAs are integral components of the maize kinetochore. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15986-15991.
- Tran, R. K., J. G. Henikoff, D. Zilberman, R. F. Ditt, S. E. Jacobsen and S. Henikoff, 2005: DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. *Curr Biol*, **15**, 154-159.

- Trzcinska-Danielewicz, J., A. Biliska, J. Fronk, P. Zielenkiewicz, E. Jarochovska, M. Roszczyk, M. Jończyk, E. Axentowicz, M. Skoneczny and P. Sowiński, 2009: Global analysis of gene expression in maize leaves treated with low temperature. *Plant Science*, **177**, 648-658.
- Tsuchiya, T. and T. Eulgem, 2013: An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. *Proceedings of the National Academy of Sciences*, **110**, E3535-E3543.
- Turco, G. M., K. Kajala, G. Kunde-Ramamoorthy, C. Y. Ngan, A. Olson, S. Deshpande, D. Tolkunov, B. Waring, S. Stelpflug, P. Klein, J. Schmutz, S. Kaeppler, D. Ware, C. L. Wei, J. P. Etchells and S. M. Brady, 2017: DNA methylation and gene expression regulation associated with vascularization in *Sorghum bicolor*. *New Phytol*, **214**, 1213-1229.
- Underwood, C. J., I. R. Henderson and R. A. Martienssen, 2017: Genetic and epigenetic variation of transposable elements in Arabidopsis. *Curr Opin Plant Biol*, **36**, 135-141.
- Van Hooser, A. A., I. I. Ouspenski, H. C. Gregson, D. A. Starr, T. J. Yen, M. L. Goldberg, K. Yokomori, W. C. Earnshaw, K. F. Sullivan and B. Brinkley, 2001: Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *Journal of cell science*, **114**, 3529-3542.
- Vaughn, M. W., M. Tanurdžić, Z. Lippman, H. Jiang, R. Carrasquillo, P. D. Rabinowicz, N. Dedhia, W. R. McCombie, N. Agier and A. Bulski, 2007: Epigenetic natural variation in Arabidopsis thaliana. *PLoS biology*, **5**, e174.
- Verhoeven, K. J., J. J. Jansen, P. J. van Dijk and A. Biere, 2010: Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol*, **185**, 1108-1118.
- Vidalis, A., D. Zivkovic, R. Wardenaar, D. Roquis, A. Tellier and F. Johannes, 2016: Methylome evolution in plants. *Genome Biol*, **17**, 264.
- Vitte, C., M. A. Fustier, K. Alix and M. I. Tenaillon, 2014: The bright side of transposons in crop evolution. *Brief Funct Genomics*, **13**, 276-295.
- Vriet, C., L. Hennig and C. Laloi, 2015: Stress-induced chromatin changes in plants: of memories, metabolites and crop improvement. *Cellular and Molecular Life Sciences*, **72**, 1261-1273.
- Waddington, C. H., 1942a: Canalization of development and the inheritance of acquired characters. *Nature*, **150**, 563-565.
- Waddington, C. H., 1942b: The epigenotype. *Endeavour*, **1**, 18-20.
- Wang, J., H. Ding, A. Zhang, F. Ma, J. Cao and M. Jiang, 2010: A novel mitogen-activated protein kinase gene in maize (*Zea mays*), ZmMPK3, is involved in response to diverse environmental cues. *J Integr Plant Biol*, **52**, 442-452.
- Wang, P., C. Gao, X. Bian, S. Zhao, C. Zhao, H. Xia, H. Song, L. Hou, S. Wan and X. Wang, 2016a: Genome-Wide Identification and Comparative Analysis of Cytosine-5 DNA Methyltransferase and Demethylase Families in Wild and Cultivated Peanut. *Front Plant Sci*, **7**, 7.
- Wang, Q. and H. K. Dooner, 2006: Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. *Proc Natl Acad Sci U S A*, **103**, 17644-17649.
- Wang, X., A. A. Elling, X. Li, N. Li, Z. Peng, G. He, H. Sun, Y. Qi, X. S. Liu and X. W. Deng, 2009: Genome-Wide and Organ-Specific Landscapes of Epigenetic Modifications and Their Relationships to mRNA and Small RNA Transcriptomes in Maize. *The Plant Cell Online*, **21**, 1053-1069.
- Wang, X., L. Hu, X. Wang, N. Li, C. Xu, L. Gong and B. Liu, 2016b: DNA Methylation Affects Gene Alternative Splicing in Plants: An Example from Rice. *Mol Plant*, **9**, 305-307.
- Wei, F., E. Coe, W. Nelson, A. K. Bharti, F. Engler, E. Butler, H. Kim, J. L. Goicoechea, M. Chen, S. Lee, G. Fuks, H. Sanchez-Villeda, S. Schroeder, Z. Fang, M. McMullen, G. Davis, J. E. Bowers, A. H. Paterson, M. Schaeffer, J. Gardiner, K. Cone, J. Messing, C. Soderlund and R. A. Wing, 2007: Physical and genetic structure of the maize genome reflects its complex evolutionary history. *PLoS Genet*, **3**, e123.
- Wendte, J. M. and R. J. Schmitz, 2017: Specifications of targeting heterochromatin modifications in plants. *Molecular plant*.
- Weng, Y.-I., T. H.-M. Huang and P. S. Yan, 2009: Methylated DNA immunoprecipitation and microarray-based analysis: detection of DNA methylation in breast cancer cell lines. *Molecular Endocrinology: Methods and Protocols*, 165-176.

- West, P. T., Q. Li, L. Ji, S. R. Eichten, J. Song, M. W. Vaughn, R. J. Schmitz and N. M. Springer, 2014: Genomic distribution of H3K9me2 and DNA methylation in a maize genome. *PLoS One*, **9**, e105267.
- Wibowo, A., C. Becker, G. Marconi, J. Durr, J. Price, J. Hagmann, R. Papareddy, H. Putra, J. Kageyama and J. Becker, 2016: Hyperosmotic stress memory in Arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity. *Elife*, **5**.
- Woo, H. R., O. Pontes, C. S. Pikaard and E. J. Richards, 2007: VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev*, **21**, 267-277.
- Wu, L., H. Zhou, Q. Zhang, J. Zhang, F. Ni, C. Liu and Y. Qi, 2010: DNA methylation mediated by a microRNA pathway. *Mol Cell*, **38**, 465-475.
- Yan, H., S. Kikuchi, P. Neumann, W. Zhang, Y. Wu, F. Chen and J. Jiang, 2010: Genome-wide mapping of cytosine methylation revealed dynamic DNA methylation patterns associated with genes and centromeres in rice. *Plant J*, **63**, 353-365.
- Yang, G., H. Zou, Y. Wu, H. Liu and Y. Yuan, 2010: Identification and characterisation of candidate genes involved in chilling responses in maize (*Zea mays* L.). *Plant Cell, Tissue and Organ Culture (PCTOC)*, **106**, 127-141.
- Ye, R., Z. Chen, B. Lian, M. J. Rowley, N. Xia, J. Chai, Y. Li, X.-J. He, A. T. Wierzbicki and Y. Qi, 2016: A Dicer-independent route for biogenesis of siRNAs that direct DNA methylation in Arabidopsis. *Molecular cell*, **61**, 222-235.
- Yong, W.-S., F.-M. Hsu and P.-Y. Chen, 2016a: Profiling genome-wide DNA methylation. *Epigenetics & chromatin*, **9**, 26.
- Yong, W. S., F. M. Hsu and P. Y. Chen, 2016b: Profiling genome-wide DNA methylation. *Epigenetics Chromatin*, **9**, 26.
- Yu, Y., X. Yang, H. Wang, F. Shi, Y. Liu, J. Liu, L. Li, D. Wang and B. Liu, 2013: Cytosine methylation alteration in natural populations of *Leymus chinensis* induced by multiple abiotic stresses. *PLoS one*, **8**, e55772.
- yyGreaves, J. A., 1996: Improving suboptimal temperature tolerance in maize- the search for variation. *Journal of Experimental Botany*, **47**, 307-323.
- Zemach, A., M. Y. Kim, P.-H. Hsieh, D. Coleman-Derr, L. Eshed-Williams, K. Thao, S. L. Harmer and D. Zilberman, 2013: The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*, **153**, 193-205.
- Zemach, A., M. Y. Kim, P. Silva, J. A. Rodrigues, B. Dotson, M. D. Brooks and D. Zilberman, 2010: Local DNA hypomethylation activates genes in rice endosperm. *Proceedings of the National Academy of Sciences*, **107**, 18729-18734.
- Zhai, J., S. Bischof, H. Wang, S. Feng, T.-f. Lee, C. Teng, X. Chen, S. Y. Park, L. Liu and J. Gallego-Bartolome, 2015: A one precursor one siRNA model for Pol IV-dependent siRNA biogenesis. *Cell*, **163**, 445-455.
- Zhang, W., H. R. Lee, D. H. Koo and J. Jiang, 2008: Epigenetic modification of centromeric chromatin: hypomethylation of DNA sequences in the CENH3-associated chromatin in Arabidopsis thaliana and maize. *Plant Cell*, **20**, 25-34.
- Zhao, L., P. Wang, H. Hou, H. Zhang, Y. Wang, S. Yan, Y. Huang, H. Li, J. Tan, A. Hu, F. Gao, Q. Zhang, Y. Li, H. Zhou, W. Zhang and L. Li, 2014a: Transcriptional regulation of cell cycle genes in response to abiotic stresses correlates with dynamic changes in histone modifications in maize. *PLoS One*, **9**, e106070.
- Zhao, L., P. Wang, S. Yan, F. Gao, H. Li, H. Hou, Q. Zhang, J. Tan and L. Li, 2014b: Promoter-associated histone acetylation is involved in the osmotic stress-induced transcriptional regulation of the maize ZmDREB2A gene. *Physiol Plant*, **151**, 459-467.
- Zhao, M.-T., J. J. Whyte, G. M. Hopkins, M. D. Kirk and R. S. Prather, 2014c: Methylated DNA immunoprecipitation and high-throughput sequencing (MeDIP-seq) using low amounts of genomic DNA. *Cellular Reprogramming (Formerly Cloning and Stem Cells)*, **16**, 175-184.

- Zhao, Y., H. Tong, R. Cai, X. Peng, X. Li, D. Gan and S. Zhu, 2014d: Identification and characterization of the RCI2 gene family in maize (*Zea mays*). *Journal of genetics*, **93**, 655-666.
- Zheng, B., Z. Wang, S. Li, B. Yu, J.-Y. Liu and X. Chen, 2009: Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes & development*, **23**, 2850-2860.
- Zhong, C. X., J. B. Marshall, C. Topp, R. Mroczek, A. Kato, K. Nagaki, J. A. Birchler, J. Jiang and R. K. Dawe, 2002: Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. *The Plant Cell*, **14**, 2825-2836.
- Zhong, S., Z. Fei, Y. R. Chen, Y. Zheng, M. Huang, J. Vrebalov, R. McQuinn, N. Gapper, B. Liu, J. Xiang, Y. Shao and J. J. Giovannoni, 2013: Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat Biotechnol*, **31**, 154-159.
- Zilberman, D., M. Gehring, R. K. Tran, T. Ballinger and S. Henikoff, 2007: Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet*, **39**, 61-69.

Titre : Réponse du méthylome suite à l'exposition au froid chez une espèce à génome complexe : le maïs (*Zea mays* ssp. *mays*)

Mots clés : Méthylation de l'ADN, froid, éléments transposables, BS-seq, DMR, maïs

Résumé : La caractérisation moléculaire de la réponse des plantes aux contraintes environnementales permet de mieux comprendre les bases de l'adaptation des plantes à leur milieu, et pourrait aider à l'amélioration des plantes cultivées.

L'épigénome est constitué de l'ensemble des marques présentes sur la chromatine et participe à la régulation de l'expression du génome, notamment au cours du développement. Il contribue aussi à la stabilité des génomes, notamment en empêchant la transposition des éléments transposables (ET). L'épigénome varie en fonction des contraintes environnementales et une meilleure compréhension de ces variations pourrait permettre d'apporter une vision nouvelle de l'interaction entre la plante et son environnement. Cependant, l'étendue des modifications de l'épigénome, le type de séquences affectées et les mécanismes impliqués restent à déterminer.

Dans ce cadre, j'ai analysé l'impact du froid sur le méthylome du maïs, une plante à génome complexe riche en ET.

Dans un premier axe, j'ai analysé le méthylome d'un génotype sensible au froid, B73, par séquençage haut-débit d'ADN traité au bisulfite de sodium (BS-seq).

Cette analyse comparative entre plantes « stressées » et « non stressées » a été menée (i) à l'échelle chromosomique, sans *a priori* sur le niveau de variation de méthylation de l'ADN et (ii) à l'échelle locale (régions différenciellement méthylées, ou « DMR ») en fixant des niveaux de variations forts (>10%). Ces deux types d'analyses ont permis de montrer que le froid déclenche une hyperméthylation à l'échelle du génome, à laquelle se superposent des hyper- et hypométhylations à l'échelle locale. Ces variations sont observées pour les trois contextes de cytosine et dans différentes régions génomiques associées aux gènes et aux ET. Ceci suggère l'activation parallèle de plusieurs mécanismes de régulation de la méthylation de l'ADN en réponse au froid.

Dans un second axe, j'ai suivi ces DMR au cours du développement et dans la descendance afin d'étudier leur transmission, en lien avec leur localisation génomique et les contextes de cytosine affectés.

Dans un troisième axe, j'ai étudié le lien entre variations de méthylation et sensibilité au froid en comparant la réponse du méthylome chez trois génotypes de maïs (B73, F2 et F331) présentant une réponse phénotypique contrastée pour ce caractère.

Title: Methylome response following cold exposure in a complex genome species: maize (*Zea mays* ssp. *mays*)

Keywords: DNA methylation, cold, transposable elements, BS-seq, DMR, maize

Abstract : Molecular characterization of plant response to environmental constraints allows to both better understand plant adaptation and help crop improvement. The epigenome is composed of chromatin marks and participates to the regulation of genome expression, notably through development. It is also involved in genome stability, essentially by preventing the transposition of transposable elements (TE). The epigenome can be modified by environmental cues and better understanding this variation could give new insights on the interaction between the plant and its environment. However, the extent of this modification, targeted sequences and underlying mechanisms remain to be elucidated.

In this context, I analyzed the impact of cold on the methylome of maize, a plant with a complex genome with high proportion of TEs.

In a first part, I analyzed the methylome of a cold-sensitive genotype, B73, using whole genome bisulfite sequencing (BS-seq). This comparative analysis between "stressed" and "unstressed" plants was carried out (i) at the chromosome scale, without a priori definition of a

(ii) at a localized scale (Differentially Methylated regions, « DMRs ») using high minimum methylation difference rate (10%). These two types of analysis revealed that cold triggers hypermethylation at the genome scale, as well as hyper- and hypo-methylation at the local scale. These variations were observed in the 3 contexts of cytosine and occur in different genomic regions associated with genes and TEs. This suggests the parallel activation of different regulatory pathways in response to cold.

In a second part, I focused on following-up methylation changes through development and in the progeny in conjunction with the genomic sequences and the cytosine context involved

In a third part, I studied the relationship between methylome variations and cold sensitivity by comparing the methylomes of three maize genotypes (B73, F2 and F331) with a contrasted phenotypic response to cold.

