



Contrôle épigénétique de la plasticité de l'appareil végétatif du peuplier en réponse à des variations de la disponibilité en eau

Clément Lafon Placette

► To cite this version:

Clément Lafon Placette. Contrôle épigénétique de la plasticité de l'appareil végétatif du peuplier en réponse à des variations de la disponibilité en eau. Biologie végétale. Université d'Orléans, 2012. Français. NNT : 2012ORLE2077 . tel-00840258

HAL Id: tel-00840258

<https://theses.hal.science/tel-00840258>

Submitted on 2 Jul 2013

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.

ÉCOLE DOCTORALE SANTE, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT

UPRES EA 1207 LABORATOIRE de BIOLOGIE des LIGNEUX et des
GRANDES CULTURES - INRA USC1328 ARBRES et REPONSES aux
CONTRAINTES HYDRIQUES et ENVIRONNEMENTALES

THÈSE présentée par :
Clément LAFON PLACETTE

soutenue le : **21 décembre 2012**

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline : Biologies Moléculaire et Cellulaire Végétales

**Contrôle épigénétique de la plasticité de
l'appareil végétatif du peuplier en réponse à
des variations de la disponibilité en eau**

THÈSE dirigée par :

Stéphane MAURY
Franck BRIGNOLAS

Professeur, Université d'Orléans
Professeur, Université d'Orléans

RAPPORTEURS :

Daniel GRIMANELLI
Christophe PLOMION

Directeur de recherche, IRD Montpellier
Directeur de recherche, INRA Bordeaux

JURY :

Etienne DANCHIN

Directeur de recherche, CNRS Toulouse,
Président du jury

Franck BRIGNOLAS
Patricia FAIVRE RAMPANT
Daniel GRIMANELLI
Stéphane MAURY
Christophe PLOMION

Professeur, Université d'Orléans
Chargée de recherche, INRA Evry
Directeur de recherche, IRD Montpellier
Professeur, Université d'Orléans
Directeur de recherche, INRA Bordeaux

Remerciements

Merci à toutes les personnes directement liées à ce travail, indispensables à sa réalisation.

Mes premiers remerciements vont aux membres du jury qui ont accepté d'évaluer mon travail. Merci à **Daniel Grimanelli** et **Christophe Plomion** pour avoir assumé la tâche de rapporteurs et merci à **Patricia Faivre-Rampant** et **Etienne Danchin** pour leur implication en tant qu'examineurs. Leurs regards critiques et leurs interrogations m'ont été très utiles pour améliorer ce manuscrit et prendre du recul sur ce travail. Je tiens également à remercier la Région Centre et l'Agence Nationale de la Recherche pour avoir financé ce projet.

Merci ensuite à **Franck Brignolas** (le directeur) pour m'avoir accueilli au sein du Laboratoire de Biologie des Ligneux et des Grandes Cultures.

Je tiens ensuite à adresser tous mes remerciements à **Stéphane Maury**, encadrant de ma thèse. Merci d'avoir cru en moi, de m'avoir fait confiance et soutenu, et ce malgré mes looks bizarres et mes réveils tardifs. Merci de m'avoir laissé un champ d'action assez libre et d'avoir (au moins) écouté mes idées. Merci également d'avoir toujours gardé un optimisme et un positivisme débordant, toujours indispensable pour remotiver les troupes dans les moments difficiles. Merci enfin pour la découverte des bons restaurants d'Orléans et autres, des bonnes bouteilles et des bases pour bien les apprécier et aussi pour toutes tes blagues sur mon côté sombre... En bref, merci à toi pour tout ce que tu m'as appris et apporté, aussi bien du côté de l'enseignement et de la recherche que des relations humaines.

Un grand merci également à **Franck Brignolas** (l'encadrant de thèse) pour son soutien que j'ai pu mesurer en de multiples occasions : tes encouragements perpétuels, ton regard critique et ton écoute, sans parler du fameux prix du poster Biotechnocentre qui ne serait qu'un doux rêve sans toi... Merci pour ton enthousiasme face à mes résultats de génomique et de biologie moléculaire qui n'ont pourtant pas dû être faciles à ingurgiter. Merci de m'avoir donné goût à l'écophysiologie, en revanche je ne suis toujours pas fan du Ricard !

Mes remerciements vont ensuite à **Claire Hébrard**, ma partenaire de galère. Même encadrant, même bureau, mêmes dates de début et de fin de thèse, stressés aux mêmes périodes. Tout était réuni pour qu'on se tape dessus et pourtant nous sommes tous les deux encore en vie et même plus vivants que jamais. Ces quatre années hautes en couleur n'auraient certainement pas été les mêmes sans toi et je t'en remercie. Je te souhaite bonne route, tu as tout pour ça.

Je remercie **Alain Delaunay** pour son aide et sa fiabilité inégalable au laboratoire. Merci pour son humour caustique voire déroutant sans quoi les journées de manip auraient manqué quelque peu de piquant. Merci pour ta bonne humeur qui a été particulièrement appréciable lors d'une certaine campagne de terrain où nous étions trempés et cernés de peuplier...

Merci à **Mikael Bili** et **Lucie Marchal**, stagiaires M1 (et bien plus !) que j'ai eu la chance d'encadrer. Vous avez fait du bon boulot, et j'espère que vous aurez apprécié autant que moi cette collaboration !

Grand merci à **Régis Fichot** pour ses relectures attentives de ce manuscrit, providentielles à un moment où les nerfs sont proches du bout...

De la même façon, merci à **Cécile Montgiraud** pour sa relecture du manuscrit.

Merci aux membres de mon comité de thèse pour leurs conseils avisés : **Alain Rival**, **Claude Joseph**, **Patricia Faivre-Rampant**.

Merci à **Isabelle Bourgaït** pour sa gentillesse et sa disponibilité. Merci de m'avoir initié à la bioinfo et surtout de m'y avoir donné (bon) goût !

Merci à **Véronique Vidal**, **Valérie Vidal** et **Béline Jesson** pour leur aide précieuse dans les analyses microarray et nos nombreuses demandes bonus en bioinfo.

Enfin, merci à **Patrick Poursat** et **Alain Guichard** pour leur aide dans la culture des peupliers.

Merci à tous ceux qui ont fait de ce long moment (qui a vite passé) une expérience inoubliable.

Un énorme merci à **France Dardeau**, indispensable à bien des égards. Merci pour ta bonne humeur, ta disponibilité et ta tendance à aider tout le monde. Merci aussi d'avoir été le « ciment », l'élément de cohésion entre nous tous, les jeunes du labo, car sans toi l'ambiance n'aurait pas été aussi bonne. Merci pour toutes les bonnes soirées passées en ta compagnie et enfin merci pour ta persévérance dans l'organisation de la fameuse Journée des Doctorants, car malgré vents et marées, on aura pu la faire cette journée, et c'est grâce à toi !

Merci à **Régis Fichot**, **Sylvain Chamaillard**, **Julien Toillon**, **Nicolas Marron** et **Marc Villar** pour leur compagnie toujours agréable et les bons moments passés à Orvieto pour l'IPS V. Grand bravo à Marc pour sa performance « poplar team » à cette occasion !

Merci à **Lucie Bertheau** pour son sens de l'humour et nos discussions, bon courage pour la dernière ligne droite !

Merci à tous les autres djeun's du labo, c'est-à-dire les stagiaires L3, M1 et M2 et puis les CDD. Je pense à **Julien Clémot**, **Maxime**, **Valentin**, **Emilie**, **Géraldine**, **Mariane** (avec un seul « n »), **Fanny**, **Thibault**, **Deki**, **Marsia**, **Flavia**, **Tugçe**, **Selin**, **Guillaume**, **Raphaël**, **Thomas** et j'en oublie. Merci à tous pour avoir apporté un vent de fraîcheur sur le labo, j'ai adoré échanger avec chacun de vous. Le printemps/été 2012 restera inoubliable pour un grand nombre d'entre nous je pense...

Merci ensuite aux membres du LBLGC. Un merci tout particulier à **Richard Bouceffa** pour ses blagues dégueulasses, les bons moments, les « bœufs » et notre performance, un autre

moment inoubliable, et aussi merci pour sa disponibilité et son aide. Un grand merci à **Domenico Morabito** pour son soutien, son aide mais aussi pour nos discussions politico-philo-psycho-théologues démarrées dans son bureau et jamais terminées. Merci à **Valérie Altemayer** pour m'avoir accompagné, encadré et soutenu au cours de mes trois années de monitorat et de mon année d'ATER, et merci à **Carine Martin** pour son aide très appréciable lors de la préparation des TP. Merci à **Sabine Carpin** et sa bonne humeur, **Françoise Chefdor** et ses nombreux conseils, **François Héricourt**, **Sylvain Bourgerie**, **Christiane Depierreux** et son « zen » inébranlable, **Cécile Barbaroux**, **Muriel Feinard-Duranceau**, **Stéphanie Bankhead-Dronnet**, **François Lieutier**, **Aurélien Salle**, **Jean Claude Vala**, **Xavier Pineau** et **Mathilde Baude** pour leur soutien et leurs encouragements lors des dernières phases de la thèse. Merci à **LAURENT JEAN-ALPHONSE** pour son aide administrative.

Merci également aux personnes de l'INRA d'Orléans pour l'intérêt qu'ils ont porté à mon travail : **Catherine** et **Jean-Charles Bastien**, **Gilles Pilate**, **Jean-Charles Leplaie**. Merci à **Marie-Claude Lesage-Descauses** et **Emmanuelle Magnoux** pour leurs conseils et leurs séquençages irréprochables.

Bien évidemment, un grand merci à ma famille, qui m'a toujours soutenu depuis le Sud-Ouest, d'une manière ou d'une autre et du début à la fin. Merci à ma mère, mon père, mes sœurs et mon frère, Corinne, mes nièces infernales, mon parrain. Ma famille c'est aussi Stéphanie, Georges, Ludmila et Anne-Laure qui font que je suis ce que je suis aujourd'hui, et je vous en remercie du fond du cœur.

J'aurai ensuite une pensée infiniment reconnaissante pour Marie. Comme tu l'as dit cette thèse c'est un peu la tienne aussi, et il est évident que sans toi je n'aurais eu ni l'énergie ni l'envie d'arriver là. Merci pour tous ces merveilleux moments, merci pour tout et bonne route.

Merci aux « Psycho », Nyx et Ollie, merci pour tout ce qu'on a vécu ensemble et ce que vous avez fait pour moi, avant et pendant cette thèse, merci pour ces moments « rock'n'roll » ++ ! Et merci à la « team Psycho » : Jeannot, Olivia, Anne, Anaïs-Pome, Laure dont la compagnie a été indispensable.

Merci à **Thibault Saurat** du labo d'à côté pour tous les délires, les danses endiablées dans les clubs d'Orléans, les découvertes culturelles et les longues discussions...

Merci à l'équipe de Radio Campus Orléans pour son ouverture d'esprit et pour m'avoir accueilli à bras ouverts, et merci à Xavier pour m'avoir énormément appris et laissé carte blanche.

Et enfin merci à Selin pour son soutien sans faille, son sourire permanent et cette semaine inoubliable à Dalyan...

A mon père

Sommaire

LISTE DES PRINCIPALES ABREVIATIONS	9
PREAMBULE : EPIGENETIQUE, UN PHENOMENE DE MODE OU UNE REVOLUTION ?	13
INTRODUCTION GENERALE	19
1. Forêts et réponse des arbres face aux changements climatiques, un rôle pour l'épigénétique ? (Bräutigam <i>et al.</i> 2013, <i>Ecology and Evolution</i> : article 1)	20
2. Peuplier et réponse aux variations de disponibilité en eau, un modèle d'étude pour la réponse des arbres aux changements climatiques	24
2.1. Le peuplier, arbre modèle	24
2.2. Problématique, objectifs de la thèse et stratégie	30
ETAT DES CONNAISSANCES : L'EPIGENETIQUE, SES CONCEPTS ET MECANISMES	39
1. Emergence de l'épigénétique	40
1.1. Contexte scientifique ayant amené à la création du concept	40
1.2. L'épigénétique selon Waddington	42
1.3. Idées et perspectives apportées par Waddington	42
1.4. Exceptions aux lois de Mendel et redéfinition de l'épigénétique, exemple de la paramutation	44
2. Mécanismes cellulaires et moléculaires : la notion d'épigénétique aujourd'hui	48
2.1. La chromatine et ses états structuraux	48
2.2. Les modifications d'histone et les variants	52
2.3. La méthylation de l'ADN	54
2.4. ARN non codant	66
2.5. Etats chromatinien et marques épigénétiques	68
3. Implication de la méthylation de l'ADN dans la vie de la plante	72
3.1. Développement	72
3.2. Réponse à l'environnement et plasticité phénotypique	76
3.3. Les variations et leur transmission transgénérationnelle	78
CHAPITRE 1 : METHYLOME DANS UN CONTEXTE DE CHROMATINE NON CONDENSEE : ETUDE SUR LE MERISTEME APICAL CAULINAIRE DU PEUPLIER	83
1. Présentation synthétique des travaux	84
2. Méthylome de la fraction non condensée de la chromatine dans le méristème apical caulinaire chez le peuplier (Lafon-Placette <i>et al.</i> 2013, <i>New Phytologist</i> : article 2)	90
3. Analyses complémentaires	118
4. Limites et perspectives	126

CHAPITRE 2 : IMPACT D'UNE SECHERESSE SUR LA METHYLATION DE L'ADN CHEZ DES HYBRIDES DE PEUPLIER (*P. DELTOIDES* × *P. NIGRA*) 131

1. Présentation synthétique des travaux 132
2. Variations de la méthylation globale de l'ADN de l'apex caulinaire au sein de différents hybrides de peuplier et en réponse à la sécheresse (Gourcilleau *et al.* 2010, *Annals of Forest Science* : article 3) 135
3. Travaux complémentaires 150
 - 3.1. Ontogenèse de la feuille 150
 - 3.2. Stabilité dans le temps de la marque épigénétique chez des hybrides *P. × euramericana* 154
4. Perspectives et limites 156

CHAPITRE 3 : CARACTERISATION DE LOCI AFFECTES POUR LEUR EXPRESSION ET LEUR METHYLATION LORS DE VARIATIONS DE LA DISPONIBILITE EN EAU DU SOL POUR LE GENOTYPE 'CARPACCIO' (*P. DELTOIDES* × *P. NIGRA*) 159

1. Présentation synthétique des travaux 160
2. Réponses du méthylome et du transcriptome à des variations de disponibilité en eau du sol (Lafon-Placette *et al.*, en préparation pour *Plant Cell* : article 4) 164
3. Discussion, limites et perspectives 222

CONCLUSIONS GENERALES, LIMITES ET PERSPECTIVES 235

1. Synthèse des résultats et apports des travaux de la thèse 238
2. Limites de l'étude 242
3. Perspectives 246

REFERENCES 251

ANNEXES ET VALORISATION DES TRAVAUX SCIENTIFIQUES 279

1. Bräutigam *et al.* in *Ecology and Evolution* (révision mineure) : article 1 280
2. Publications et communications 298

Liste des principales abréviations

Lorsque l'abréviation est issue du terme anglais et quand la correspondance en français n'est pas évidente, le terme anglais est indiqué entre parenthèses.

- A** : Adénine
- A** : Assimilation nette de CO₂ par unité de surface
- ABA** : Acide abscissique (« ABscissic Acid »)
- ABRE** : Élément de réponse à l'ABA (« ABA Responsive Element »)
- AcH3** : Histone H3 acétylée
- AcH4** : Histone H4 acétylée
- ADN** : Acide DésoxyriboNucléique
- ADNc** : ADN complémentaire
- ADNg** : ADN génomique
- ADNr** : ADN ribosomique
- ANOVA** : Analyse de variance (« ANalysis Of VAriance »)
- ARN** : Acide RiboNucléique
- ARNm** : ARN messager
- ARNr** : ARN ribosomique
- At** : *Arabidopsis thaliana*
- ATP** : Adénosine TriPhosphate
- BET** : Bromure d'ETHidium
- BS** : Séquençage bisulfite (« Bisulfite Sequencing »)
- C** : Cytosine
- CDS** : Séquence codante (« CoDing Sequence »)
- CG** ou **CpG** : Dinucléotide cytosine-guanine
- CHG** : Trinucléotide cytosine-H-guanine (où H est A, T ou C)
- CHH** : Trinucléotide cytosine-H-H (où H est A, T ou C)
- CMT3** : Chromométhyltransférase 3 (« CHROMOMETHYLASE 3 »)
- CO₂** : Dioxyde de carbone
- cv.** : Cultivar
- Cy3** : Cyanine 3 (Fluorochrome rouge)
- Cy5** : Cyanine 5 (Fluorochrome vert)
- DCL** : Protéine similaire à DICER (« DICER-LIKE protein »)
- ddc** : Triple mutant *drm1 drm2 cmt3*
- DDM1** : Protéine « Méthylation de l'ADN diminuée » (« DECREASED DNA METHYLATION 1 »)
- DME** : ADN glycosylase (« DEMETER »)
- DNase I** : DésoxyriboNucléase I
- DNMT** : ADN méthyltransférase (« DNA METHYLTRANSFERASE »)
- dNTP** : DésoxyNucléoside TriPhosphate
- DRE** : Éléments de réponse à la déshydratation (« Dehydration Responsive Element »)
- DRM** : ADN méthyltransférase aux domaines réarrangés (« DOMAINS-REARRANGED METHYLTRANSFERASE »)
- epiQTL** : QTL épigénétique (« epigenetic QTL »)
- epiRIL** : Lignée recombinante épigénétique (« epigenetic Recombinant Inbred Line »)
- EST** : Etiquette de séquence de gènes exprimés (« Expressed Sequence Tag »)

G : Guanine
gCO₂ : Conductance stomatique au dioxyde de carbone
GO : Ontologie de gène (« Gene Ontology »)
gs : Conductance stomatique de la feuille à la vapeur d'eau
H1 : Histone de liaison
H2A : Histone 'de coeur' H2A
H2B : Histone 'de coeur' H2B
H₂O : Eau
H3 : Histone 'de coeur' H3
H3.2 : variant de l'histone H3 de type 2
H3.3 : variant de l'histone H3 de type 3
H3K27me : histone H3 méthylée sur la lysine 27
H3K36me : histone H3 méthylée sur la lysine 36
H3K4me : histone H3 méthylée sur la lysine 4
H3K9me : histone H3 méthylée sur la lysine 9
H4 : Histone 'de coeur' H4
HAT : Histone Acétylase
HDAC : Histone DésAcétylase
HPLC : Chromatographie liquide à haute performance (« High Pressure Liquid Chromatography »)
HS : Site hypersensible à la DNase I (« Hypersensitive Site »)
MBD : Domaine de liaison à un groupement méthyl (« Methyl Binding Domain »)
mC : MéthylCytosine
McrBC : Enzyme de restriction de méthylcytosine BC (« Methylcytosine restriction enzyme BC »)
MeDIP : Immunoprécipitation de l'ADN (« MethylDNA ImmunoPrecipitation »)
MeDIP-chip : Immunoprécipitation de l'ADN suivie d'une hybridation microarray (« MethylDNA ImmunoPrecipitation chip »)
MeDIP-SEQ : Immunoprécipitation de l'ADN suivie d'un séquençage haut débit (« MethylDNA ImmunoPrecipitation SEQuencing »)
MET1 : ADN méthyltransférase 1 de plantes (« DNA METHYLTRANSFERASE 1 »)
MS-PCR : PCR sensible à la méthylation (« MethylSensitive-PCR »)
P : Probabilité (*p*-value)
PCA : Analyse en composante principale (« Principal Component Analysis »)
PCR : Réaction en chaîne de polymérisation (« Polymerase Chain Reaction »)
Pt : *Populus trichocarpa*
QTL : Loci contrôlant un caractère quantitatif (« Quantitative Trait Loci »)
RdDM : Méthylation de l'ADN dirigée par l'ARN (« RNA-directed DNA Methylation »)
RDR1 : ARN polymérase ARN-dépendante 1 (« RNA-Dependent RNA polymerase 1 »)
RDR2 : ARN polymérase ARN-dépendante 2 (« RNA-Dependent RNA polymerase 2 »)
RGR : Taux de croissance relatif (« Relative Growth Rate »)
ROS : Forme active de l'oxygène (« Reactive Oxygen Species »)
ROS1 : ADN glycosylase (« REPRESSOR OF SILENCING 1 »)
r_P : Coefficient de Pearson
RT-PCR : PCR après transcription inverse (« Reverse Transcription – PCR »)
SAM : S-Adénosyl-L-Méthionine
SAM : méristème apical caulinaire (« Shoot Apical Meristem »)

SE : Erreur Standard (« Standard Error »)
siRNA : petit ARN interférence (« short interfering RNA »)
SLA : Surface massique de la feuille (« Specific Leaf Area »)
sp : Espèce (« species »)
sRNA : petit ARN (« small RNA »)
SWI : Protéine de remodelage de la chromatine (« SWItching mating type protein »)
T : Thymine
TAE : Tris Acétate EDTA
TBE : Tris Borate EDTA
TCR : Taillis à Courte Rotation
TE : Élément transposable (« Transposable Element »)
T_m : Température de demi-dénaturation (« melting Temperature »)
TSS : site d'initiation de la transcription (« Transcription Start Site »)
TTCR : Taillis à Très Courte Rotation
TTS : site de terminaison de la transcription (« Transcription Termination Site »)
U : Uracile
UTR : région non traduite (« UnTranslated Region »)
UV : Ultra-Violet
WUE : Efficience d'utilisation de l'eau (« Water-Use Efficiency »)
Ψ : Potentiel hydrique
Ψ_{pd} : Potentiel hydrique foliaire de base

**Préambule : Epigénétique, un phénomène de mode ou une
révolution ?**

Aujourd'hui, le mot « épigénétique » apparaît dans de multiples contextes (Figure I.1). Pourquoi donc un tel engouement médiatique pour cette discipline de la biologie ?

Probablement parce que certains faits marquants laissent penser que l'épigénétique pourrait compléter de façon notable notre manière de penser l'évolution ou encore la psychologie humaine. Par exemple, le cas de la *Hongerwinter*, famine ayant eu lieu en 1944 aux Pays-Bas suite à un rationnement extrême par le régime Nazi, est souvent cité dans les articles de vulgarisation scientifique. Lors de cette famine et parmi les adultes sous-alimentés, des femmes ont accouché de nouveau-nés en mauvaise santé et plus petits que la normale. Les embryons exposés à cette famine lors de leur conception présentaient une fois adultes des risques accrus de maladies cardiovasculaires ou de schizophrénie, contrairement à ceux exposés lors du dernier trimestre de gestation (Heijmans *et al.*, 2008). Ainsi, l'exposition très tôt dans l'embryogenèse à la famine de 1944 a généré des changements de nature non génétique persistant tout au long de la vie adulte. Des travaux récents ont montré que ces individus, soixante ans après leur naissance, possédaient des changements épigénétiques (profils de méthylation de l'ADN) dans certains gènes (Heijmans *et al.*, 2008) et transmis aux générations suivantes.

De nombreux autres cas ont été décrits, aussi bien chez les animaux que chez les plantes, suggérant qu'en plus de la génétique, une autre information influence le phénotype d'un individu, est inductible par l'environnement et transmissible au cours des générations. De Conrad Waddington en 1942 aux biologistes actuels, le concept d'« épigénétique » a beaucoup évolué, laissant une place de plus en plus grande aux mécanismes moléculaires. Les recherches des dernières décennies ont en effet révélé que l'information épigénétique correspondait à des modifications de la structure de la chromatine (association ADN-protéines dans le noyau) sous le contrôle de différentes marques sur la chromatine (méthylation de l'ADN, modifications post-traductionnelles des histones), induisant des modifications d'expression de gènes et transmissibles par division cellulaire (Figure I.2). Cette découverte remet en cause la séquence d'ADN comme *seul* support de l'information héréditaire. Deux citations résument bien cette idée (Allis *et al.*, 2007) : « We are more than the sum of our genes » (Klar) et « You can inherit something beyond the DNA sequence. That's where the real excitement in genetics is now » (Watson).

L'influence de l'environnement sur les mécanismes épigénétiques et donc la part de variance transmise non-génétique questionne également notre vision et définition de l'évolution. En effet, les variations épigénétiques se placeraient ainsi comme facteur transmis



Figure I.1. Exemples témoignant de l'intérêt croissant pour l'épigénétique aussi bien dans le grand public, l'industrie ou la recherche en biologie.

des variations de traits phénotypiques aux côtés de l'effet parental, de facteurs écologiques ou sociaux (Danchin *et al.*, 2011). Cependant, l'épigénétique est une discipline récente, en plein essor depuis les années 2000, et des études sont encore indispensables pour estimer l'ampleur du phénomène et son rôle dans l'évolution des espèces. Néanmoins, sans être une révolution, les découvertes passées, présentes et futures en épigénétique pourraient bien être une avancée majeure dans la biologie.

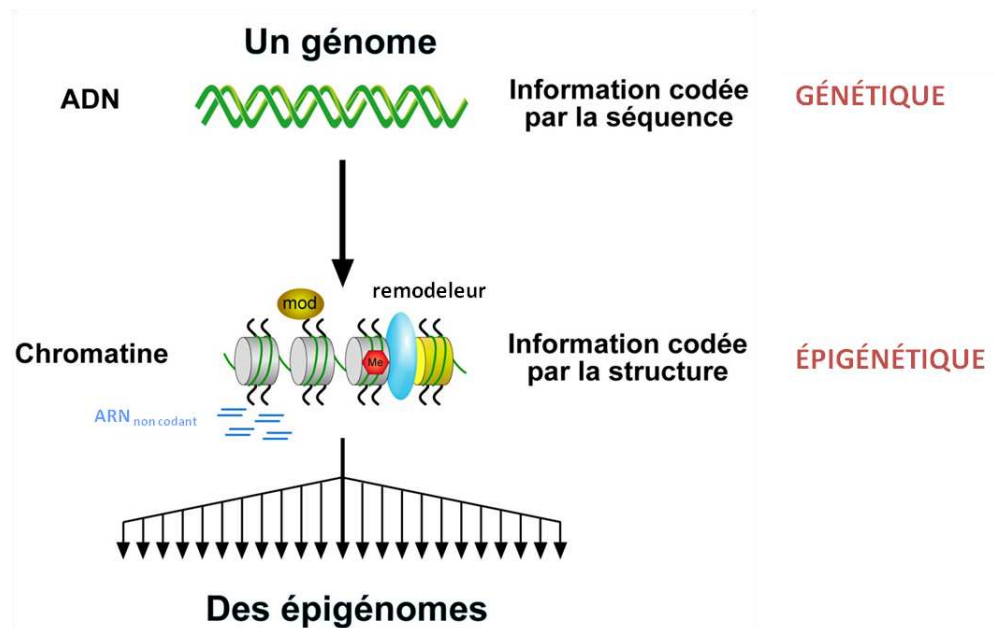


Figure I.2. Schéma représentant la définition moderne de l'épigénétique. A partir d'un génome, plusieurs états de la chromatine, transmissibles par mitose et/ou méiose, peuvent exister (Allis *et al.*, 2007).

Introduction générale

1. Forêts et réponse des arbres face aux changements climatiques, un rôle pour l'épigénétique ? (Bräutigam *et al.* 2013, *Ecology and Evolution* : article 1)

De par leur nature pérenne et sessile, les arbres et leur réponse aux changements climatiques globaux prédits représentent un modèle biologique en matière de stratégies de plasticité phénotypique (au niveau de l'individu) et adaptatives (au niveau des populations). Les changements climatiques globaux attendus, avec des périodes de sécheresse de durée et d'intensité sans précédents, pourraient atteindre voire dépasser les limites des capacités de réponses des arbres (Intergovernmental Panel on Climate Change, 2007; Allen *et al.*, 2010; Figure II.1A). Les conséquences en seraient aussi bien écologiques (santé des forêts) qu'économiques (production de bois), la limitation en eau étant un des contributeurs majeurs du déclin des forêts (Bigler *et al.*; van Mantgem *et al.*, 2009; Allen *et al.*, 2010; Hamanishi & Campbell, 2011).

Une centaine de cas de mortalité de régions forestières liée à des stress climatiques (chaleur et/ou sécheresse) a été recensée depuis les années 1970 (Allen *et al.*, 2010), touchant tous les continents. Des épisodes de sécheresse sévère au début des années 2000 ont notamment augmenté drastiquement le taux de mortalité d'essences de pin en Turquie (Semerci *et al.*, 2008) et en Grèce, affectant même *Pinus halepensis* sub. *brutia*, la plus tolérante à la sécheresse des espèces de pin méditerranéen (Körner *et al.*, 2005). **Trois causes majeures** ont été avancées pour expliquer la mortalité liée aux événements de sécheresse et/ou fortes températures (McDowell *et al.*, 2008; Figure II.1B). Tout d'abord, une sécheresse intense combinée à une forte demande évaporative peut entraîner une **embolie généralisée** des vaisseaux conducteurs, aboutissant à l'arrêt du flux d'eau dans la plante (Tyree & Zimmermann, 2002) ; la formation de nouvelles pousses, la resaturation du xylème ou la tolérance des cellules à la dessiccation peuvent toutefois limiter ce problème (McDowell *et al.*, 2008). Cependant, la fermeture des stomates en réponse à la sécheresse limite également la photosynthèse et l'assimilation de CO₂, pouvant aboutir à un **épuisement des ressources en carbone** de la plante. De plus, des températures plus élevées en hiver augmentent la respiration et les pertes en carbone associées (Damesin, 2003). Enfin, des températures plus élevées modifient et accélèrent le cycle de vie de certains **agents biotiques** (insectes, pathogènes). L'épuisement en carbone, de par la diminution du métabolisme secondaire, est susceptible d'accroître la sensibilité des arbres aux stress biotiques (McDowell *et al.*, 2008; Sala & Hoch, 2009).

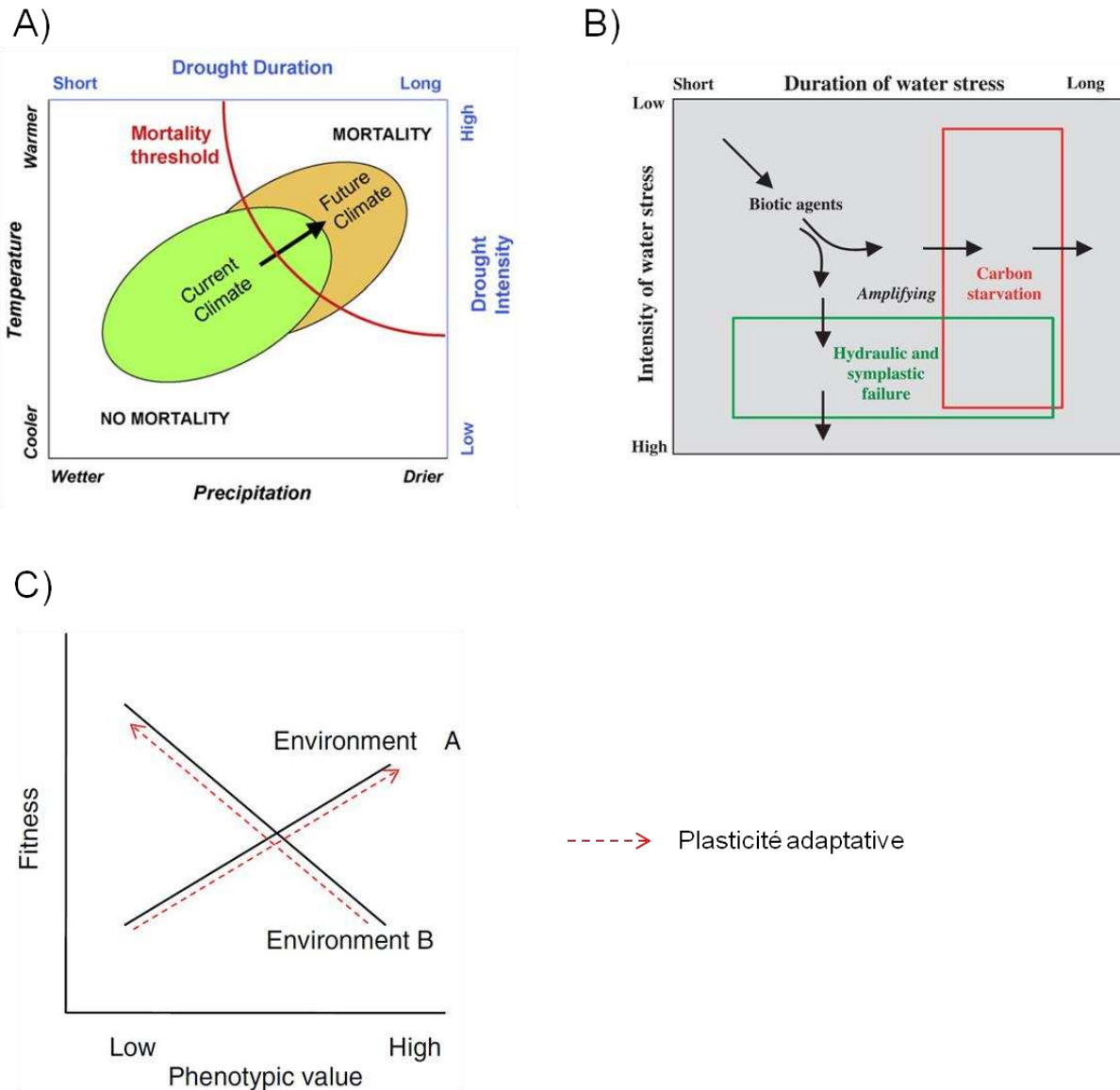


Figure II.1. Facteurs influençant la survie des arbres forestiers face aux sécheresses extrêmes prédites dans les prochaines années. A) Relation théorique entre durée de sécheresse, son intensité et trois mécanismes hypothétiques de la mortalité chez les arbres. L'épuisement de carbone peut avoir lieu lorsque la sécheresse inhibe la photosynthèse plus longtemps que la capacité de stockage du carbone ne le permet pour le maintien du métabolisme. La cavitation peut se produire si l'intensité de la sécheresse du sol et la demande évaporative sont suffisantes pour créer une embolie généralisée dans les vaisseaux conducteurs, aboutissant à l'arrêt du flux d'eau dans la plante. L'effet des agents biotiques, tels que les insectes ou pathogènes, peut amplifier ou être amplifié par l'épuisement de carbone ou la cavitation (Allen *et al.*, 2010). B) Diagramme montrant la gamme de variabilité des précipitations et température (ou alternativement de la durée et de l'intensité des épisodes de sécheresse) du climat actuel, avec une faible portion de cette gamme dépassant le seuil de mortalité espèce-spécifique des arbres. A l'inverse, le climat futur prédit, avec des événements de sécheresse et de température extrêmes, présente un risque accru de mortalité pour les arbres (McDowell *et al.*, 2008). C) Exemple de plasticité adaptative pour un trait phénotypique donné. Ce trait, selon l'environnement (A ou B), a un rôle positif ou négatif sur la valeur adaptative d'un génotype (Nicotra *et al.*, 2010).

Face à ces menaces, les arbres peuvent ajuster leur développement *via* la plasticité phénotypique, s'adapter *via* la sélection naturelle et/ou migrer (leur descendance) dans des régions où le climat est celui auquel ils sont adaptés (Lande, 2009; Nicotra *et al.*, 2010). Le rôle de la plasticité phénotypique, définie comme la gamme de phénotypes que peut exprimer un génotype en réponse aux variations de l'environnement, a longtemps été sous-estimé dans le processus d'évolution des espèces. Cependant, elle a été documentée pour de nombreux caractères incluant entre autres la phénologie, la durée de vie de la feuille ou la morphologie du xylème en réponse à des facteurs liés aux changements climatiques (Metcalf *et al.*, 2003; Fichot *et al.*, 2009; Nicotra *et al.*, 2010). De plus, il est généralement accepté que la variabilité génétique au sein d'une population participe à sa capacité d'adaptation à de nouveaux environnements (Jump *et al.*, 2009) et la capacité à percevoir les variations environnementales et produire une réponse plastique possède une telle variabilité chez les arbres (Monclus *et al.*, 2006; Chamaillard *et al.*, 2011). Ainsi, la plasticité phénotypique pourrait à la fois tamponner les risques liés aux changements climatiques à l'échelle de l'individu mais également être un caractère soumis à la sélection naturelle à l'échelle de l'espèce (Nicotra *et al.*, 2010). En outre, certains auteurs définissent la « plasticité adaptative » comme la plasticité phénotypique augmentant la valeur sélective d'un génotype (Nicotra *et al.*, 2010). Cette plasticité adaptative (Figure II.1C) est ainsi envisagée comme un déterminant de l'adaptation des arbres aux changements climatiques globaux.

Compte tenu du rôle de l'épigénétique dans la plasticité développementale en réponse à des stress environnementaux d'une part (Blödnér *et al.*, 2007), et des évidences croissantes de transmission *via* la méiose de certaines marques épigénétiques d'autre part (Feng *et al.*, 2010b; Saze, 2012), il est important de comprendre la part que ces mécanismes auraient à jouer dans l'adaptation des arbres aux changements climatiques annoncés. Cette question a été abordée par un groupe de travail euro-nord américain réuni à l'initiative d'un programme de recherche européen appelé FoResTTTrac (*Forest ecosystem genomics Research: supporTing Transatlantic Cooperation* ; www.foresttrac.eu) visant à favoriser les collaborations scientifiques entre Europe et Amérique du Nord concernant l'adaptation des arbres aux changements climatiques. Les réflexions de ce groupe ont été valorisées sous forme de revue, à laquelle j'ai participé, des connaissances actuelles en épigénétique des plantes, de son rôle potentiel dans l'adaptation des arbres aux changements climatiques et des efforts scientifiques à fournir pour mieux le comprendre.

Pour ne pas déstabiliser la structure de l'introduction de cette thèse et sa lecture, nous avons décidé de placer l'article en annexe.

2. Peuplier et réponse aux variations de disponibilité en eau, un modèle d'étude pour la réponse des arbres aux changements climatiques

2.1. Le peuplier, arbre modèle

Pourquoi le peuplier ?

Une aire de répartition très large dans des écosystèmes variés

Les peupliers sont des plantes pérennes Angiospermes Eucotylédones de la famille des Salicacées. Cette famille comprend le genre *Salix*, regroupant entre autres les espèces de saules et d'osiers, et le genre *Populus*, regroupant les espèces de peuplier. Les espèces de ces deux genres sont considérées comme espèces pionnières, possédant à ce titre une croissance extrêmement rapide et une forte capacité de multiplication végétative (Dickmann & Kuzovkina, 2008).

Taxonomiquement, le genre *Populus* est composé d'environ trente espèces réparties en 6 sections botaniques (*Abaso*, *Aigeiros*, *Leucoides*, *Populus*, *Tacamahaca* et *Turanga*) sur la base de critères à la fois morphologiques et écologiques (Dickmann & Kuzovkina, 2008).

Actuellement, l'aire naturelle de répartition du genre *Populus* couvre la quasi-totalité de l'hémisphère Nord (Amérique du Nord, Europe, Afrique du Nord, Asie) ; aucun peuplier ne pousse à l'état sauvage dans l'hémisphère Sud bien qu'il soit largement planté par l'Homme dans certaines régions tempérées (Dickmann & Kuzovkina, 2008). De façon générale, les peupliers sont présents aux abords des terrains humides en bordures de rivières ou de fleuves, voire dans les vallées et plaines alluviales temporairement inondées. On trouve toutefois quelques espèces adaptées aux régions désertiques de par leur comportement phréatophytique (*Populus euphratica* Oliv. et *Populus pruinosa* Schrenk dans la section *Turanga*).

Ressources génomiques et outils moléculaires développés

Avant même le séquençage de son génome, fait qui a définitivement entériné le choix du peuplier comme un des arbre modèles, de nombreux groupes de recherche avaient contribué à l'effort de séquençage de cet organisme, ayant permis d'accumuler 376 565 séquences d'EST en 2006, sur des espèces, organes ou conditions variés (Jansson & Douglas, 2007). Un nombre important de bases de données moléculaires était déjà disponible (Tuskan *et al.*, 2004; Sterky *et al.*, 2004), nombre continuant de croître jusqu'à

aujourd'hui, où des outils comme *PopGenIE* intègrent les données de l'ensemble de la communauté « peuplier » (Sjödin *et al.*, 2009). Par ailleurs, quelques équipes américaines et européennes avaient déjà acquis une expertise de la transformation génétique et régénération *in vitro* du peuplier (Busov *et al.*, 2005).

En 2006, le séquençage du génome du peuplier a été publié (Tuskan *et al.*, 2006), obtenu à partir du clone « Nisqually-1 » de l'espèce *Populus trichocarpa*, alors utilisé dans les programmes de sélection américains. Pour une plante, le peuplier est relativement simple d'un point de vue génétique : il est diploïde ($2n = 38$) et son génome ne mesure que 423 Mb (v3, dernier assemblage en date de 2012), taille relativement faible pour un arbre (pour comparaison, le génome de *Pinus tadea* fait 21,7 Gb ; Morse *et al.*, 2009). Le peuplier présente également un intérêt du point de vue de l'évolution des génomes puisque phylogénétiquement proche d'*Arabidopsis*, il partage avec la plante modèle un évènement de duplication de génome (datant d'environ 100 Ma) mais a également subi un autre évènement du même type (~ 60 Ma) coïncidant avec l'émergence des genres *Salix* et *Populus* (Tuskan *et al.*, 2006; Abrouk *et al.*, 2010). Ainsi, il contient 41 335 gènes prédits codant pour des protéines, ce qui signifie que chaque gène chez *Arabidopsis* possède en moyenne 1,4 homologues chez *P. trichocarpa*. Plus intéressant encore, l'évolution génomique se fait beaucoup plus lentement chez le peuplier que chez *Arabidopsis*, probablement dû à son plus long cycle de vie mais aussi à sa forte capacité de multiplication végétative, permettant à d'« anciens » génotypes de persister sur des millénaires et de contribuer génétiquement de manière importante à la descendance (Mitton & Grant, 1996). La lenteur de l'horloge moléculaire du génome du peuplier laisse penser qu'il ressemble fortement au génome de l'Euroside ancestral (Tuskan *et al.*, 2006).

Un modèle pour l'étude du développement

Sa croissance rapide et sa forte propension à la multiplication végétative font du peuplier un modèle d'étude intéressant pour le développement, autant sur le plan fondamental qu'appliqué (Dickmann & Kuzovkina, 2008). De plus, dans un contexte de changements climatiques, sa forte demande en eau et sa faible tolérance à la sécheresse en font un cas d'étude pertinent.

Par ailleurs, le peuplier présente une maturité sexuelle précoce (4-6 ans) comparé à d'autres arbres, permettant des études génétiques, de croisements et de localisation de QTL plus rapides (Yin *et al.*, 2004). Les peupliers sont des espèces dioïques dont la détermination du sexe est de type ZW, les mâles étant WW et les femelles ZW (Yin *et al.*,

2008). De manière intéressante, la région déterminant le sexe est située sur le chromosome 19 et inclut un cluster de gènes *NBS-LRR*, impliqués dans la résistance aux pathogènes. Cependant, même si le sexe est déterminé de manière génétique, un grand nombre de « mutants » hermaphrodites a été observé au sein de plusieurs espèces. Ils présentaient un taux de réversion important laissant penser à une implication de mécanismes épigénétiques (Santamour, 1956; Chardenon, 1968; Stettler, 1971; Rowland *et al.*, 2002). Enfin, La grande particularité du genre *Populus* réside dans la quasi-inexistence de barrières entre espèces (Stanton, 2009). Cette caractéristique quasi-unique chez les arbres explique notamment les difficultés rencontrées par les botanistes et les systématiciens (Eckenwalder, 1996). Certaines hybridations sont toutefois impossibles ou unidirectionnelles, l'exemple typique étant celui des hybrides *Populus deltoides* × *P. nigra* (syn. *P. × euramericana*) qui n'est possible que dans le sens *P. deltoides* (♀) × *P. nigra* (♂). Ces caractéristiques apportent des perspectives intéressantes pour les études d'évolution des espèces et de spéciation, la barrière interspécifique étant un phénomène encore mal compris. Ce processus possède une composante épigénétique (Kinoshita, 2007) et le cas des croisements unidirectionnels supposent probablement une empreinte parentale, laissant penser que barrière interspécifique et empreinte parentale partageraient des mécanismes épigénétiques communs.

Une plante d'intérêt agronomique

L'hybridation interspécifique constitue la base des programmes de création et de sélection variétale en populiculture (Stanton, 2009). Les descendants F_1 issus des croisements présentent généralement une forte vigueur hybride (phénomène d'hétérosis) en terme de productivité (Marron & Ceulemans, 2006; Marron *et al.*, 2006; Dillen *et al.*, 2009).

En Europe, la plupart des cultivars commercialisés sont des hybrides interspécifiques issus de croisements entre trois espèces, deux originaires des Etats-Unis (*Populus deltoides* Bartr. ex Marsh. et *Populus trichocarpa* Torr. & Gray) et une originaire d'Europe (*Populus nigra* L.). Les hybrides de type *Populus deltoides* × *P. trichocarpa* ou *Populus trichocarpa* × *P. deltoides* (syn. *P. × interamericana*), qui présentaient une résistance totale au champignon pathogène *Melampsora larici-populina* (agent de la rouille foliaire), ont dominé les ventes de cultivars jusqu'en 1997, date à laquelle de nombreux dépérissements se sont produits en raison du contournement des résistances qualitatives par de nouvelles souches virulentes. Depuis, les ventes de cultivars sont dominées par les hybrides de type *Populus deltoides* × *P. nigra* (Thivolle-Cazat, 2003). Aujourd'hui, l'objectif des programmes d'amélioration est de fournir aux propriétaires forestiers le matériel génétique le mieux adap-

-té aux exigences stationnelles et au mode de culture pour la production de bois. Jusqu'à présent, les critères de sélection ont été centrés sur la productivité et la résistance aux pathogènes (rouille foliaire, brunissure des feuilles et chancre bactérien) notamment (Berthelot *et al.*, 2005).

En France, les surfaces plantées en peuplier représentent plus de 235 000 ha, soit moins de 2% des surfaces boisées. Toutefois, la forte productivité, la faible durée de rotation et l'utilisation de techniques productivistes permettent de compenser nettement les faibles surfaces populicoles, le peuplier représentant la deuxième essence feuillue récoltée en volume de bois d'œuvre (1,5 millions de m³ par an), derrière le chêne. La création du Groupement d'Intérêt Scientifique (GIS) Peuplier en 2001, regroupant l'AFOCEL (aujourd'hui FCBA), le Cemagref et l'INRA, a permis de structurer le programme français d'amélioration génétique du peuplier. Le bois produit est essentiellement utilisé pour la confection d'emballages légers, et dans une moindre mesure pour la production de pâte à papier.

Un arbre au cœur des problématiques de changements climatiques

En parallèle des conduites culturales classiques sur des rotations de 15 à 20 ans, les conduites culturales en haute densité de type taillis à courte ou très courte rotation (TCR et TTCR, respectivement) reçoivent actuellement un regain d'intérêt. Ces plantations pourraient notamment jouer le rôle de puits de carbone et donc mitiger les effets du changement global tout en assurant la production d'énergie renouvelable (Liberloo *et al.*, 2009; Dillen *et al.*, 2010).

2.2. Problématique, objectifs de la thèse et stratégie

Jusqu'à présent, la populiculture était concentrée majoritairement dans les plaines alluviales, là où la ressource en eau n'est pas un facteur limitant à la production. Toutefois, l'augmentation des événements climatiques extrêmes de type canicule et sécheresse prévue à moyen et long termes (Meehl & Tebaldi, 2004; Intergovernmental Panel on Climate Change, 2007) pourrait affecter la productivité des plantations futures, particulièrement dans les zones hors vallées (Loustau *et al.*, 2005), et également menacer le compartiment sauvage (peuplier noir).

Dans ce contexte et dans une volonté de gestion raisonnée des ressources en eau, identifier de nouveaux critères de sélection semble indispensable afin d'ajuster les cultivars aux exigences stationnelles et de faire face aux modifications climatiques prédites. En outre,

en plus du caractère productivité, d'autres critères sont à considérer tels que la capacité de produire une même quantité de biomasse à moindre coût en eau, autrement dit l'efficacité d'utilisation de l'eau, ou la capacité de limiter l'impact de contraintes édaphiques ponctuelles et modérées sur le niveau de productivité, autrement dit la tolérance à la sécheresse. En cas d'épisodes de sécheresse plus intenses, la résistance à la cavitation (voir chapitre précédent) pourrait également être un caractère important à prendre en compte (Brignolas *et al.*, 2012).

La thématique de recherche de l'équipe « Arbres et Réponses aux Contraintes Hydriques et Environnementales » (INRA USC1328) du Laboratoire de Biologie des Ligneux et des Grandes Cultures de l'Université d'Orléans (LBLGC, UPRES-EA 1207) vise à développer une approche intégrative afin d'identifier chez le peuplier les bases écophysiologiques et moléculaires impliquées dans les variations génétique et phénotypique de caractères complexes d'intérêt comme le potentiel de croissance, l'efficacité d'utilisation de l'eau (WUE), la tolérance à la sécheresse ou encore la résistance à la cavitation. Cette thématique est développée en combinant des approches complémentaires (écophysiologie, génétique quantitative, épigénétique, protéomique et transcriptomique) sur différents fonds génétiques [compartiment cultivé (hybrides interspécifiques *P. × euramericana* et *P. × interamericana*) *versus* compartiment sauvage (*Populus nigra*)]. Sur le plan appliqué, les travaux ont pour objectif final d'une part de juger de la possibilité de sélectionner du matériel génétique à la fois performant et adapté aux exigences stationnelles (compartiment cultivé), et d'autre part de fournir des informations utiles à la conservation des ressources génétiques naturelles (compartiment sauvage). Sur un plan plus fondamental, il s'agit de comprendre comment ces caractères peuvent influencer sur la plasticité phénotypique des espèces ligneuses en réponse aux changements climatiques annoncés, voire sur leur adaptation. Ces travaux sont réalisés en étroite collaboration avec les centres INRA d'Orléans (Unité 'Amélioration, Génétique et Physiologie Forestières'), de Clermont-Ferrand (équipe HΨDRO de l'UMR INRA-UBP 'Physique et Physiologie Intégratives de l'Arbre Fruitier et Forestier'), de Nancy (équipe Bioclimatologie et Ecophysiologie de l'UMR INRA-UHP 'Ecologie et Ecophysiologie Forestière') et de Bordeaux (équipe Génétique de l'UMR INRA-Bordeaux I 'Biodiversité, Gènes et Communautés').

L'axe « épigénétique » de l'équipe (S. Maury) base ses travaux sur les résultats du laboratoire et notamment ceux obtenus en écophysiologie (F. Brignolas). Ces derniers, portant sur les principaux cultivars commercialisés en Europe (*Populus deltoides* × *P. nigra*) mais également plus récemment sur des populations d'espèces naturelles (*P. nigra*), ont permis de mettre en évidence une large gamme de variations génotypiques pour des

caractères tels que le potentiel de croissance, WUE, la tolérance à la sécheresse et la résistance à la cavitation (Monclus *et al.*, 2006; Marron *et al.*, 2006; Fichot *et al.*, 2010; Chamaillard *et al.*, 2011). Ces résultats ont également montré que la plasticité de ces traits en réponse à des variations de la disponibilité en eau possède une gamme de variations génotypiques tout aussi importante. Les travaux en protéomique ont révélé que les variations de ces traits pouvaient être reliées à des variations à l'échelle du protéome, notamment au niveau de l'accumulation de protéines impliquées dans la photosynthèse (Bonhomme *et al.*, 2009a,b). Au regard des connaissances existantes chez les plantes sur le rôle de l'épigénétique dans la plasticité phénotypique et dans les variations génotypiques de traits quantitatifs (Boyko & Kovalchuk, 2008; Johannes *et al.*, 2009), l'axe épigénétique se propose d'étudier dans quelle mesure les variations génotypiques et phénotypiques observées pourraient avoir une composante épigénétique.

La feuille est considérée comme un niveau d'étude privilégié pour la compréhension des relations hydriques et carbonées de l'arbre puisqu'elle est à la fois le site des processus de transpiration et de photosynthèse. Cependant, pour rechercher les déterminants de caractères liés à la morphogenèse, il semble plus pertinent de s'intéresser aux régions de la plante responsable de sa trajectoire développementale et donc de la formation des organes, c'est-à-dire le méristème apical (dans le cas de la tige feuillée). Qui plus est, la notion d'épigénétique implique tacitement la transmission au moins par mitose de modifications de la chromatine.

L'hypothèse de travail de cet axe de recherche est donc que des variations épigénétiques dans le méristème apical *(i)* au sein de différents génotypes, *(ii)* en réponse à des variations de l'environnement et *(iii)* de cette réponse au sein de différents génotypes, induisent des variations d'expression de gènes et sont transmises aux futurs organes, aboutissant à des variations de la morphogenèse de la tige feuillée (Figure II.2). Cette hypothèse pourrait par exemple participer à expliquer le fait que la morphogenèse des feuilles produites pendant un déficit hydrique est modifiée (surface foliaire réduite, feuilles plus épaisses), phénomène réversible lorsque le régime hydrique redevient non limitant (Marron *et al.*, 2002, 2003).

Bien que l'axe épigénétique existe au laboratoire depuis 2000 (S. Maury), ses activités sur le modèle peuplier n'ont démarré qu'en 2007 avec les travaux de Master 2 de Delphine Gourcilleau. Aussi, les résultats présentés dans ce manuscrit correspondent à la première thèse au laboratoire sur cette thématique. Au démarrage de celle-ci, rien n'était connu en épigénétique chez le peuplier et les données concernant son rôle dans la réponse à l'environnement chez les arbres étaient très parcellaires. Ce champ était donc prometteur

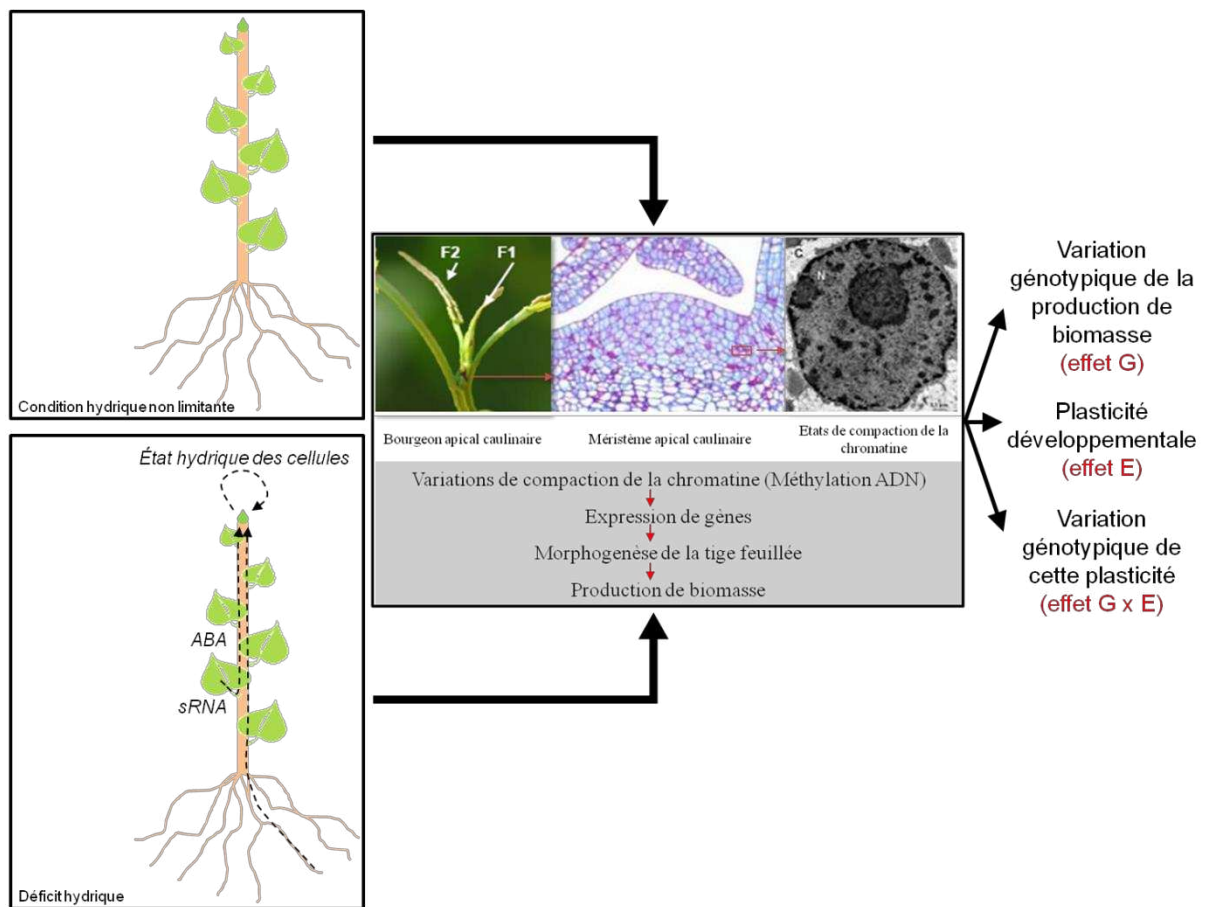


Figure II.2. Hypothèse de travail de l'axe épigénétique. Cette hypothèse postule que des variations de compaction de la chromatine dans le méristème apical caulinaire, affectant l'expression de certains gènes, sont transmises aux futurs organes, modifiant leur morphogenèse.

même si le nombre croissant des modifications de la chromatine connues rendait nécessaire de focaliser les recherches sur une, voire quelques marques épigénétiques. Compte tenu des résultats obtenus par Delphine Gourcilleau et de la stabilité de la méthylation de l'ADN, transmissible par mitose voire méiose, cette marque a été choisie comme candidat en tant que potentiel déterminant des variations phénotypiques et génotypiques observées chez le peuplier.

Dans ce contexte, l'objectif général de cette thèse était tout d'abord de caractériser chez le peuplier la méthylation de l'ADN dans l'apex caulinaire puis d'évaluer son rôle dans la plasticité développementale de la tige feuillée en réponse à la sécheresse. Cet objectif général a été décliné en trois objectifs plus spécifiques, donnant ainsi naissance à trois parties structurant la thèse :

1. La caractérisation du méthylome de la chromatine non condensée dans le méristème apical caulinaire en relation avec l'expression des gènes.
2. L'étude de l'impact d'une sécheresse sur la méthylation globale de l'ADN dans l'apex caulinaire dans différents génotypes et sur leur production de biomasse.
3. La caractérisation de loci affectés pour leur expression et leur méthylation lors de variations de disponibilité en eau du sol pour un génotype hybride (*P. deltoides* × *P. nigra*).

Afin de mieux appréhender cette problématique et les objectifs précis de ce travail, un état de l'art sur l'épigénétique et son implication dans la vie de la plante est proposé. Le manuscrit s'organise ensuite en trois chapitres, chacun contenant un article accepté ou en préparation. Certains résultats non publiés sont présentés à titre complémentaire. Enfin, les principaux résultats acquis dans le cadre de ce travail sont synthétisés, avec leurs limites et accompagnés de perspectives de recherche.

Dans le cadre du 1^{er} objectif, le clone '101-74' de l'espèce *Populus trichocarpa* a été étudié en condition hydrique non limitante. Après isolement de la chromatine non condensée des cellules du méristème apical à l'aide d'une digestion ménagée par la DNase I, les fragments d'ADN méthylés ont été isolés par immunoprécipitation puis séquencés (technique Illumina/Solexa). Les lectures de 36-mer générées ont ensuite été repositionnées sur le génome du peuplier (v2.0). Les loci ainsi mappés ont été analysés et caractérisés en particulier pour leur expression.

Dans le cadre du 2nd objectif, six clones hybrides (*Populus deltoides* × *P. nigra*) ont été étudiés en condition hydrique non limitante et en condition de déficit hydrique modéré court. Des caractères phénotypiques liés à leur potentiel de croissance ont été mesurés,

différentes marques épigénétiques ont été étudiées à l'échelle globale ainsi que l'activité et l'accumulation d'enzymes responsables de l'établissement de ces marques. Des corrélations statistiques ont ensuite été testées entre les caractères liés à la production de biomasse et les marques épigénétiques.

Enfin, dans le cadre du dernier objectif, un des clones précédemment étudiés ('Carpaccio', *Populus deltoides* × *P. nigra*) a été suivi en condition hydrique non limitante, en condition de déficit hydrique modéré court avec en supplément une modalité de ré-irrigation après le déficit hydrique. Des caractères phénotypiques indicateurs de la réponse au déficit hydrique ont été mesurés pour les trois conditions. La réponse du méristème apical a été suivie au niveau transcriptomique et épigénomique (méthylation de l'ADN) à l'aide d'une approche microarray (expression et MeDIP-chip).

Etat des connaissances : L'épigénétique, ses concepts et mécanismes

1. Emergence de l'épigénétique

On attribue à juste titre l'invention du terme « épigénétique » à Conrad Hal Waddington (1905-1975). Ses travaux en tant qu'embryologiste et biologiste du développement ont en effet influencé la pensée scientifique actuelle même si ce qu'il entendait par épigénétique était assez différent du terme utilisé aujourd'hui.

1.1. Contexte scientifique ayant amené à la création du concept

Waddington a commencé à travailler sur l'embryologie dans les années 1930. Convaincu que la génétique était à la base de la biologie, il allait rapidement introduire ses concepts dans l'embryologie (Waddington, 1939). La génétique était alors une discipline naissante qui définissait un gène comme « unité de l'hérédité » non matérialisée.

Ainsi, le concept de gène n'était pas utilisé en embryologie pour expliquer le développement d'un embryon. Celui-ci était expliqué à l'aide de concepts biochimiques peu précis comme le 'champ embryonnaire' (région spatiale initiant le développement d'une structure complète). Waddington allait enrichir et clarifier ce modèle explicatif en y ajoutant le concept de gène, sans pour autant exclure d'autres facteurs, notamment liés au cytoplasme.

Pour arriver à un modèle cohérent incluant à la fois génétique et connaissances existantes en embryologie, Waddington s'est inspiré d'un débat scientifique ayant eu lieu aux 16^e-17^e siècles. Ce débat opposait deux théories, la préformation, selon laquelle tous les caractères de l'adulte étaient présents dès l'ovule fécondé et n'avaient besoin que de grandir et l'épigenèse, ancien terme désignant la croissance et la différenciation embryonnaires, se focalisant sur les interactions entre les constituants du zygote (Van Speybroeck *et al.*, 2002). Pour Waddington, ces deux théories étaient complémentaires : d'une part, la préformation pouvait signifier que les caractères adultes étaient représentés par « quelque chose » dans l'ovule fécondé, même sous forme potentielle et d'autre part, l'épigenèse soutenait que le zygote contenait des éléments qui interagissaient durant le développement pour produire un grand nombre de caractéristiques adultes non présentes précédemment (Waddington, 1966). Ainsi, selon Waddington, la préformation pouvait s'assimiler à la génétique, considérant les gènes comme éléments préformés alors que l'épigenèse traduisait les idées de la biologie du développement, mettant l'accent sur les interactions entre les éléments du zygote, y compris les gènes. Sa théorie devait porter un nom, il emploierait un néologisme :

$$\text{Epigenesis} + \text{Genetics} = \text{Epigenetics}$$

Le terme « épigénétique » était né.

1.2. L'épigénétique selon Waddington

Cependant, cette équation ne convenait pas à Waddington puisque notamment, elle ne pouvait expliquer les différences de phénotype observées au sein d'un organisme (entre organes par exemple). D'après lui, le phénotype était le produit du génotype, de l'ensemble des constituants d'une cellule ou d'un tissu et de leurs interactions avec le génotype (= la fraction épigénétique) tout au long du développement et en réaction à l'environnement. Ainsi, Waddington allait définir l'épigénétique comme « la branche de la biologie qui étudie les interactions causales entre les gènes et leurs produits établissant le phénotype » (Waddington, 1942). Par gène, il entendait ici « séquence d'ADN ».

Cette définition s'articulait autour de la cellule, du *gene action system* (un gène conduit à une protéine qui conduit à un phénotype via des éléments cytoplasmiques) et d'une action de gène à gène (Waddington, 1962). Ce système était cyclique puisque les gènes régulaient la fabrication des éléments cytoplasmiques, qui eux-mêmes régulaient l'activité des gènes à différents niveaux (Figure III.1).

En d'autres termes, alors que la génétique classique s'intéressait uniquement au comportement des gènes lors de l'hérédité, l'épigénétique de Waddington décrivait leur comportement lors du développement de l'ovule fécondé jusqu'à l'adulte, décrivant un nouveau composant du phénotype, l'épigénotype :

$$\text{Genotype} + \text{Epigenotype} + \text{Environment} = \text{Particular Phenotype}$$

└─▶ = la composante épigénétique d'une cellule/d'un tissu

1.3. Idées et perspectives apportées par Waddington

Waddington a participé à la modernisation de l'embryologie, discipline héritière de la théorie de l'épigenèse du 17^e siècle, en y ajoutant des concepts de génétique, domaine présentant peu d'intérêt pour les embryologistes des années 1940. Il l'a subtilement fait en réactualisant le vieux débat scientifique opposant préformation et épigenèse et rassemblant ces deux théories sous un nouveau concept, l'épigénétique. Cependant, ce concept ne sera jamais vraiment utilisé en embryologie et il faudra attendre les années 1990 pour que le terme soit repris, bien qu'ayant changé de signification.

En outre, l'épigénétique de Waddington a été inventée pour expliquer le développement embryonnaire à une époque où le gène n'avait pas de définition stricte, excepté « unité d'hérédité », définition vague pouvant décrire tout ce qui était qualifié d'héréditaire. Néanmoins, conscient de ce flou, Waddington définissait, dans son système épigénétique, le gène comme « séquence d'ADN », précision permettant de définir les

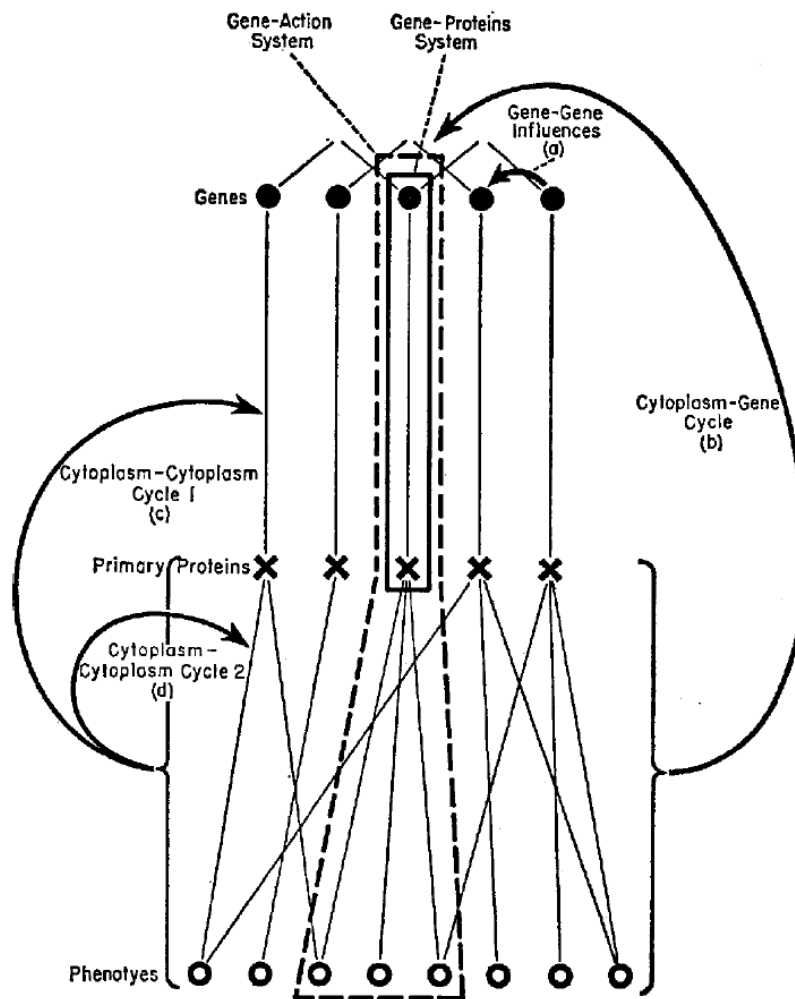


Figure III.1. Fonctionnement du système épigénétique au sein d'une cellule (Waddington, 1962). Les différentes influences envisagées sont représentées dans le schéma, (a) de gène à gène, (b) du cytoplasme vers les gènes, (c) du cytoplasme vers le système [gène→protéine] et (d) du cytoplasme vers le système [protéine primaire→phénotype]. L'ensemble constitue le contexte génique.

composants présents « autour du gène », nommés « épi-génétiques ».

Ainsi, selon lui, lors du développement, sous l'action des gènes et des constituants épigénétiques, une population de cellules se différenciait dans différents types pour former un organisme (Waddington, 1966). Le célèbre *epigenetic landscape*, avec sa bille et ses sillons bifurquant à certains endroits, traduisait cette idée mais pas uniquement : cette métaphore illustre le concept de « canalisation », selon lequel un embryon, malgré les perturbations environnementales, pouvait se développer selon un plan établi (la bille revenant dans le sillon grâce la pente des versants de la vallée). Dans ce « paysage », les gènes et leurs interactions causales avec leurs produits, autrement dit l'épigénétique, modelaient le relief (Figure III.2).

Malgré tout, le concept d'épigénétique ne revêtait pas le sens entendu aujourd'hui. Certes, Waddington était conscient qu'au sein d'un individu, donc à partir d'un même génotype, différents organes se développaient, quel que soit l'environnement, et donc que quelque chose d'autre, son « épigénotype » avait un rôle à jouer dans la différenciation des cellules. Certes, Waddington était convaincu que la structure du chromosome jouait un rôle dans l'expression des gènes (Waddington, 1962), remettant au goût du jour certaines découvertes comme la variégation liée à l'effet position chez la drosophile (Sturtevant, 1925) ou la paramutation (interaction entre les allèles d'un gène menant à un changement héritable par mitose et/ou méiose dans l'expression d'un des deux allèles) chez le maïs (McClintock, 1950).

Cependant, l'épigénétique de Waddington avait pour but de décrire de manière dynamique le comportement des gènes au cours du développement. Autrement dit, il montrait l'importance de la génétique dans l'embryologie et centrait sa discipline autour des gènes, « séquences d'ADN », tandis que l'épigénétique moderne allait nuancer cette importance.

1.4. Exceptions aux lois de Mendel et redéfinition de l'épigénétique, exemple de la paramutation

Dans les années 1990, le nombre de découvertes de cas d'hérédité ne se conformant pas aux lois de Mendel ne cessant de croître, un champ de la biologie ne se focalisant plus sur la séquence d'ADN allait naître. Il allait prendre un nom inventé quelque 50 ans plus tôt : l'épigénétique. Le cas de la paramutation est décrit ci-après.

La paramutation a été définie dans les années 1950 par Brink comme une interaction entre les allèles d'un gène qui menait à un changement héritable par mitose et ou méiose dans l'expression d'un des deux allèles (Chandler, 2007). Brink a décrit pour la première fois

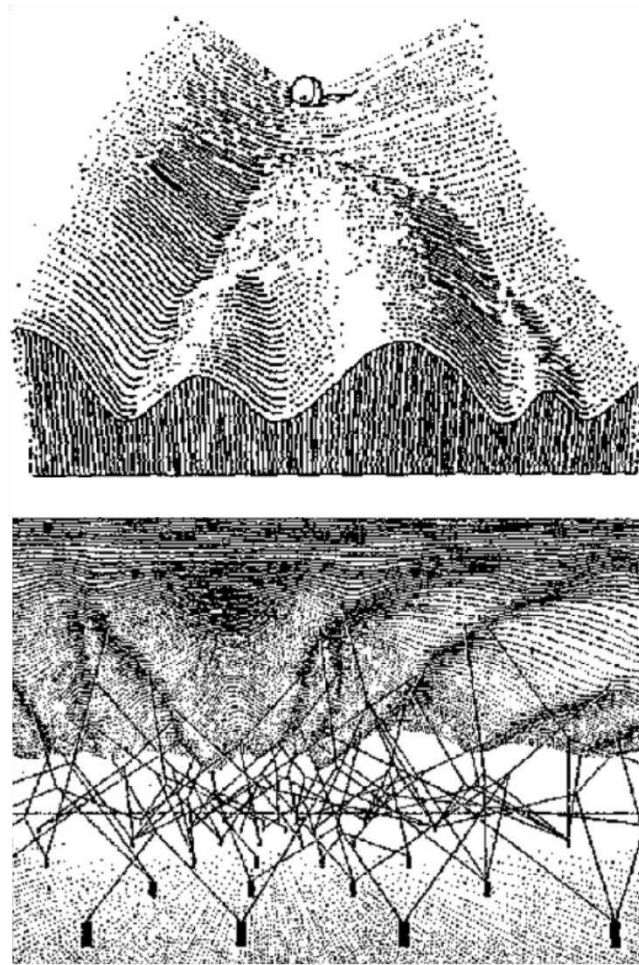


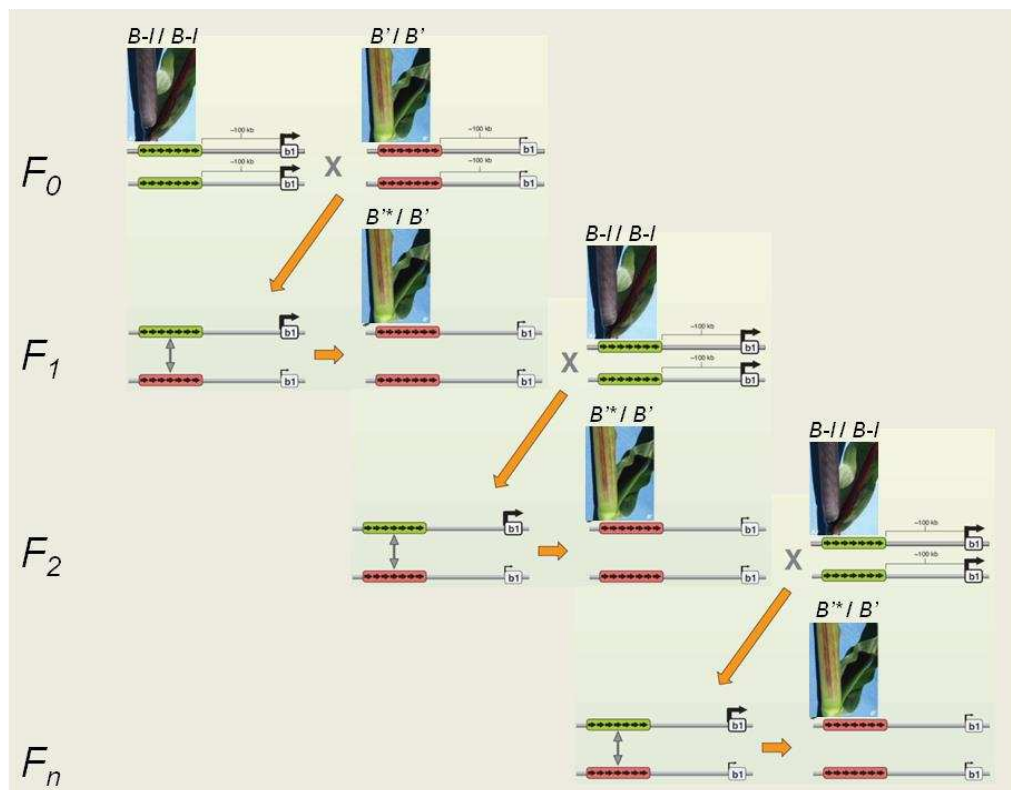
Figure III.2. Paysage épigénétique (*epigenetic landscape*) schématisé par Waddington pour expliquer les phénomènes de « canalisation » et de différenciation cellulaire (Waddington, 1957). Selon lui, les gènes (rectangles noirs en bas) gouvernent le relief de ce paysage, et donc du devenir cellulaire, sous l'influence du système épigénétique décrit dans la Figure III.1.

la paramutation chez le maïs en 1956. En croisant l'allèle *R-r*, conférant une couleur violet foncé aux graines, avec l'allèle *R-stippled*, donnant des graines tachetées de violet, l'allèle *R-r* était changé de manière héritable puisqu'il conférait une couleur faiblement pigmentée aux graines, alors appelé *R-r'* (Brink, 1956). Ce phénomène ne respectait donc pas les lois de l'hérédité de Mendel. Cette observation a permis d'identifier un allèle susceptible à la paramutation, dit « paramutable », et un allèle facilitant ce changement, appelé « paramutagène » (Chandler & Stam, 2004). L'allèle paramuté devient par la suite et de manière héritable allèle paramutagène et ainsi de suite. Les cas similaires des gènes *plant color1/booster1 (b1)* et *purple plant1 (pl1)*, tous deux impliqués dans la voie de biosynthèse des anthocyanes (Dooner *et al.*, 1991), traduit bien ce changement héritable (Figure III.3). Cependant, la stabilité de ce changement au cours des générations est variable (Martienssen, 1996).

Par ailleurs, des expériences plus récentes ont montré que le changement de l'allèle paramutable pouvait être transitoire et que la paramutation pouvait être réversible (Gross & Hollick, 2007; Figure III.3B). Enfin, les paramutations affectent aussi les animaux comme la souris et la drosophile (Chandler, 2007; de Vanssay *et al.*, 2012). Les mécanismes de la paramutation ne sont pas clairement élucidés à ce jour.

La paramutation présente donc des caractéristiques particulières : elle provoque des changements d'expression de gènes héritable par mitose et méiose sans modification de séquence de l'ADN, son hérédité échappe aux lois de Mendel et elle est réversible.

A)



B)

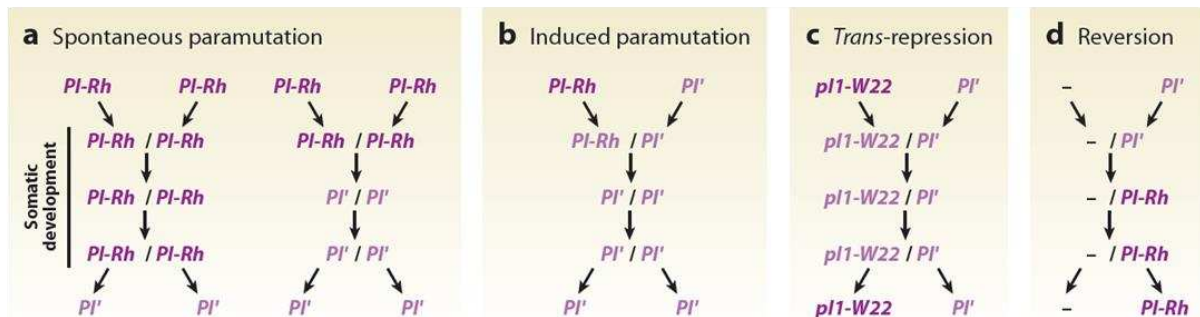


Figure III.3. Paramutation chez le maïs (Chandler, 2010; Hollick, 2010). A) Cas du gène *b1*, impliqué dans la voie de biosynthèse des anthocyanes. Le croisement en F_0 entre un homozygote $B-I/B-I$ (plantes colorées) et un homozygote B'/B' (plantes faiblement colorées) donne une descendance 100% B'/B' . B^* correspond à un allèle $B-I$ modifié suite à l'influence de l'allèle paramutagène B' . B) Exemple du gène *p1*, également impliqué dans la voie de biosynthèse des anthocyanes. $PI-Rh$ est l'allèle exprimé (phénotype coloré) et PI' est faiblement transcrit (phénotype peu coloré). La paramutation spontanée (a) se produit lors du développement somatique, transmise par mitose, ou se produit lors de la méiose, transmise aux gamètes. La paramutation facilitée (b) est le résultat de l'action paramutagénique de l'allèle PI' sur l'allèle $PI-Rh$, celui-ci devenant à son tour paramutagène. La répression *trans* (c) est similaire excepté que l'allèle *pI1-W22* paramuté sous l'influence de PI' est ré-exprimé dans la descendance si PI' est absent (Gross & Hollick, 2007). La réversion de la paramutation (d) est possible dans le cas d'hémizygotie, l'allèle PI' redevenant $PI-Rh$.

Suite à la découverte croissante de nouvelles exceptions aux lois de l'hérédité de Mendel, Holliday propose en 1990 une nouvelle définition de l'épigénétique comme étant « l'étude des mécanismes spatio-temporels contrôlant l'activité des gènes au cours du développement d'un organisme » (Holliday, 1990). Il y ajoute en 1994 la notion d'héritabilité, définissant alors l'épigénétique comme « l'étude du changement de l'expression des gènes dans des organismes dont les cellules sont différenciées, et de l'héritabilité mitotique des profils d'expression des gènes » (Holliday, 1994).

En 1996, Russo trouve une définition plus précise de l'épigénétique, à savoir « l'étude des changements héréditaires mitotiques et/ou méiotiques affectant la fonction d'un gène et qui ne peuvent pas être justifiés par des changements de séquences de l'ADN » (Russo *et al.*, 1996).

2. Mécanismes cellulaires et moléculaires : la notion d'épigénétique aujourd'hui

Comme le pensait McClintock dans les années 1950, les données moléculaires acquises ces dernières décennies montrent que la chromatine, sa structure et son organisation spatiale sont impliquées dans de nombreux processus essentiels au développement (McClintock, 1951; Roudier *et al.*, 2009).

2.1. La chromatine et ses états structuraux

2.1.1. Composition

La chromatine, support du matériel génétique, est une structure nucléoprotéique complexe constituée d'ADN et d'histones (Exner & Hennig, 2008). Longtemps restreinte à un simple « emballage » de l'ADN, la chromatine est désormais considérée comme une structure dynamique présentant différents états de condensation et qui affecte aussi bien la transcription, la réplication, la réparation de l'ADN, la recombinaison et la transposition que la ségrégation des chromosomes (Cuylen & Haering, 2011; Wright *et al.*, 2012; Soria *et al.*, 2012).

L'unité de base de la chromatine est le nucléosome dont le cœur, composé d'un octamère d'histones, est entouré de 147 pb d'ADN (Luger, 2006). Cet octamère comporte une paire de 4 types histone : H2A, H2B, H3 et H4 (Figure III.4). Chaque nucléosome est relié à un autre par une région de liaison, composée de 50 pb d'ADN sur laquelle se fixe l'histone H1.

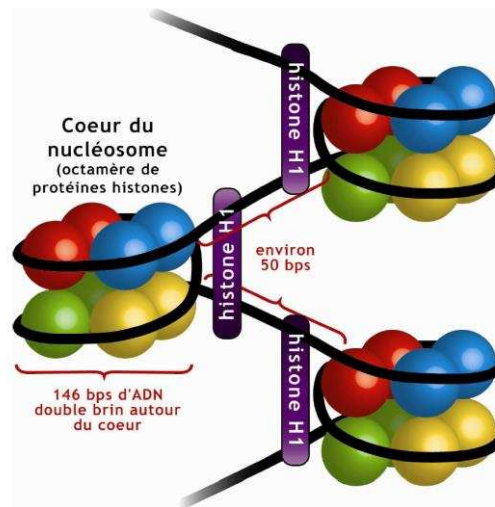


Figure III.4. Nucléosomes, composés du cœur octamérique histone et reliés entre eux par une région de liaison (50 pb d'ADN et histone H1). D'après Buchanan *et al.*, 2000.

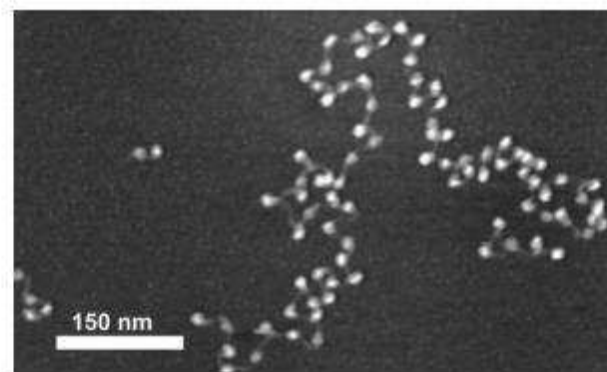


Figure III.5. Photographie d'un nucléofilament (érythrocyte de poulet) réalisée par cryo-microscopie à force atomique (Shao, 1999). Chaque nucléosome, séparé de son voisin par l'ADN de liaison, mesure ~12-15 nm (tâches blanches).

2.1.2. Condensation de la chromatine

Le chromosome métaphasique est le résultat de différents niveaux de condensation de la chromatine à l'aide d'une armature protéique (condensines). Le premier niveau, le nucléofilament ou « collier de perle », de 11 nm de diamètre, est une succession de nucléosomes (Figure III.5). Le deuxième, appelé fibre chromatinienne de 30 nm, correspond à l'enroulement hélicoïdal du nucléofilament en solénoïde (Figure III.6). Enfin, le dernier correspond au surenroulement de la fibre chromatinienne sur elle-même (Felsenfeld & Groudine, 2003; Kamakaka & Biggins, 2005).

Dans le noyau interphasique, on distingue par microscopie électronique deux états de condensation de la fibre chromatinienne : l'euchromatine, structure relâchée et lieu d'une forte transcription, et l'hétérochromatine, structure compacte où l'activité transcriptionnelle est faible (Weiler & Wakimoto, 1995; Grewal & Elgin, 2007). Dans cette dernière, on distingue l'hétérochromatine constitutive de l'hétérochromatine facultative (Trojer & Reinberg, 2007). La première, dont l'état de condensation est relativement stable, regroupe les régions centromériques, péri-centromériques et télomériques, est riche en séquences répétées et pauvre en gènes. La seconde, en revanche, correspond aux domaines chromatinien capables d'adopter de façon réversible des caractères structuraux et fonctionnels proches de l'hétérochromatine constitutive (Trojer & Reinberg, 2007). Ce changement de conformation s'effectue sous l'action combinée de complexes de remodelage de la chromatine ATP-dépendants, de protéines du groupe Polycomb (PcG) ou trithorax (trxG). Les complexes de remodelage de la chromatine ATP-dépendants sont des complexes multi-protéiques qui utilisent l'énergie produite par l'hydrolyse de l'ATP pour déstabiliser les contacts histone-ADN, déplacer les octamères d'histones ou catalyser l'incorporation de variants d'histones spécifiques, impactant l'accessibilité de l'ADN à d'autres protéines (Clapier & Cairns, 2009). Ces complexes sont classés en plusieurs familles, dont la plus étudiée est celle des SWI/SNF. Les complexes de remodelage de la chromatine ATP-dépendants, en association avec des modifications d'histones, participent à l'établissement rapide, réversible et/ou transmissible de différents profils d'expression de gènes impliqués dans la régulation de processus développementaux essentiels chez les animaux et les plantes, tels que le développement embryonnaire ou la floraison (Ho & Crabtree, 2010). Les protéines PcG et trxG sont présentes chez les plantes et les animaux (Buzas *et al.*, 2012). Les premières, associées en complexes, induisent un état chromatinien répressif via leur activité enzymatique d'histone méthyltransferase (Schwartz & Pirrotta, 2008; Buzas *et al.*, 2012) alors que les secondes ont un effet antagoniste sur la transcription (Papp & Müller, 2006). Les protéines PcG et trxG jouent un rôle fondamental dans le développement, aussi bien chez les animaux que chez les plantes (Buzas *et al.*, 2012).

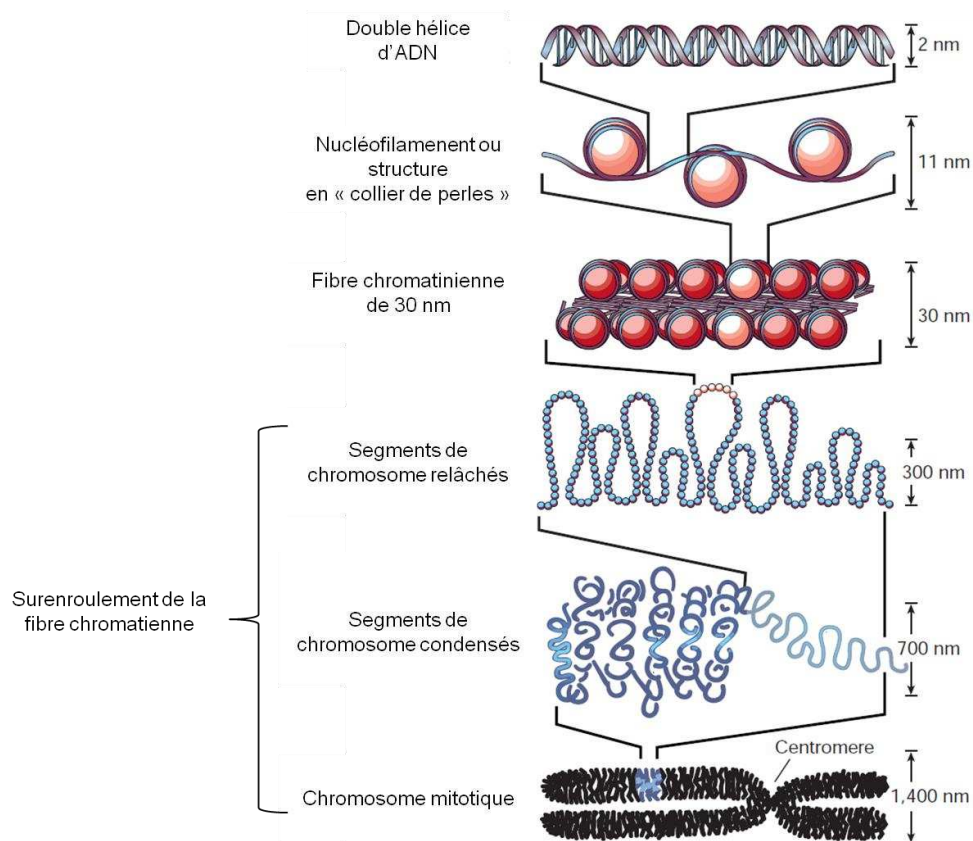


Figure III.6. Les différents niveaux de condensation de la chromatine (Felsenfeld & Groudine, 2003).

2.2. Les modifications d'histone et les variants

Les modifications ajoutées au cœur du nucléosome ou l'incorporation de variants d'histone comptent parmi les processus définissant l'état de la chromatine et modulant l'accessibilité à l'ADN, que ce soit directement ou indirectement par le recrutement d'effecteurs additionnels (Roudier *et al.*, 2009).

2.2.1. Marques histones

Les histones peuvent subir une gamme importante de modifications post-traductionnelles, soit avant soit après incorporation dans le nucléosome (Roudier *et al.*, 2009; Tan *et al.*, 2011). Ces modifications post-traductionnelles affectent les 4 histones du cœur du nucléosome et plus particulièrement la queue N-terminal des histones H3 et H4. Elles incluent l'acétylation, la méthylation, l'ubiquitination ou la sumoylation des lysines, la méthylation des arginines ou encore la phosphorylation des sérines et thréonines (Zhang *et al.*, 2007).

Ces modifications peuvent altérer, de façon différentielle, la charge des histones, les interactions protéine-protéine au sein du nucléosome et donc sa stabilité ou encore les interactions histone-ADN, modifiant la mobilité des nucléosomes le long de l'ADN (Cosgrove *et al.*, 2004). Au sein d'une histone, l'acide aminé modifié, le type et le degré de modification (par exemple la mono-, di- ou triméthylation) affectent de manière différente ces propriétés (Pfluger & Wagner, 2007). Enfin, les marques histones interagissent mutuellement, certaines étant antagonistes et n'étant jamais retrouvées ensemble alors que d'autres agissent de manière synergétique (Roudier *et al.*, 2009; Lee *et al.*, 2010). Il a été suggéré que les différentes modifications des histones et leurs interactions constitueraient le « code histone » qui détermine l'état transcriptionnel d'un locus. Cependant, des études récentes semblent indiquer que les modifications des histones ont un effet dépendant du contexte, rendant délicat l'association d'un « code » bien précis à un niveau transcriptionnel donné (Lee *et al.*, 2010).

Ainsi, la combinaison de modifications d'histone peut modifier l'accessibilité de l'ADN ou la stabilité des nucléosomes ou encore être le point de liaison d'autres facteurs chromatinien, aboutissant à un rôle répresseur ou activateur sur l'expression des gènes (Pfluger & Wagner, 2007). Par exemple, alors que l'acétylation des histones H3/H4 et les di-triméthylation des lysines 4 des histones H3 (H3K4me2/me3) sont présentes dans gènes activement transcrits, H3K27me3 et H3K9me3 se retrouvent plutôt dans des gènes réprimés (Figure III.7). Enfin, bien que certaines modifications soient labiles et transitoires, d'autres peuvent être transmises par mitose (Roudier *et al.*, 2009).

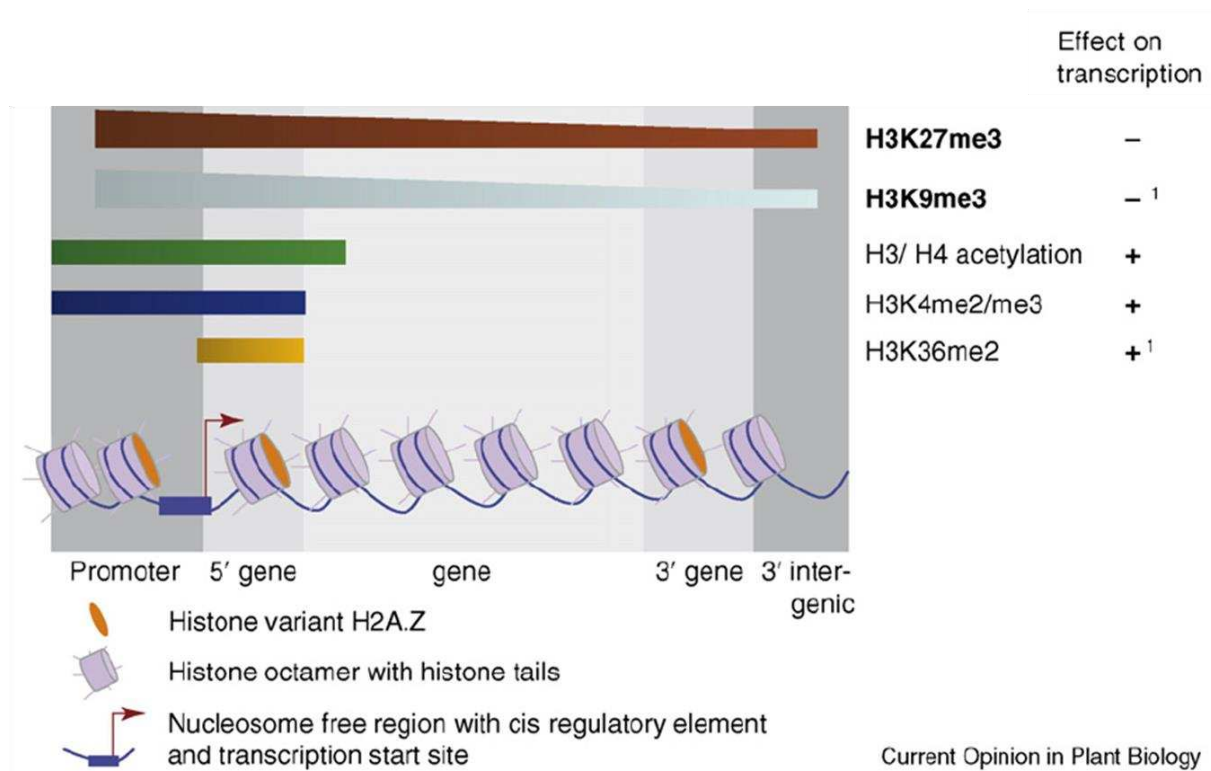


Figure III.7. Répartition de quelques modifications d'histone le long d'un gène d'*Arabidopsis* et leur effet sur la transcription. ¹ H3K9me3 et H3K36me2 peuvent avoir différents rôles sur la transcription selon l'endroit du gène où elles se trouvent (Pfluger & Wagner, 2007).

2.2.2. Variants d'histone

Chez les plantes et les mammifères, chacune des quatre histones majeures H1, H2A, H2B et H3 est codée par plusieurs gènes (Kiefer, 2007), permettant l'existence pour une même histone de plusieurs variants aux séquences protéiques et propriétés biophysiques propres. Le nombre de variants dépend de l'histone et des organismes. Par exemple, l'histone H2A est celle possédant le plus de variants, tels que H2A.X et H2A.Z présents chez les Eucaryotes et H2A.Bbd et MacroH2A uniquement présents chez les vertébrés.

Le rôle des variants dépend de leurs propriétés physico-chimiques. Notamment, l'interaction des variants H2A.Z et H3.3 avec les autres histones est faible, altérant la stabilité du nucléosome (Jin *et al.*, 2009). L'accessibilité à l'ADN en est facilitée et c'est pourquoi ces variants se retrouvent dans les régions fortement transcrites et notamment au niveau du *Transcription Start Site* (TSS) des gènes (Figure III.7).

2.3. La méthylation de l'ADN

2.3.1. Définition et découverte

La méthylation de l'ADN consiste en l'addition d'un groupement méthyle ($-CH_3$) provenant de la S-adenosyl-L-méthionine sur le carbone 5 d'un résidu cytosine (Figure III.8). La 5-méthylcytosine a été identifiée pour la première fois comme un acide nucléique en 1925, à partir d'ADN de *Mycobacterium tuberculosis* (Johnson & Coghill, 1925). En 1950, la présence et l'abondance de la 5-méthylcytosine avaient déjà été décrites dans l'ADN de plusieurs animaux et dans l'ADN du germe de blé (Wyatt, 1950, 1951). Durant la décennie suivante, la 5-méthylcytosine fut dosée chez de nombreuses plantes telles que la betterave, le riz, le maïs, le tournesol et le coton (Thomas & Sherratt, 1956 ; Ergle & Katterman, 1961). La méthylation de l'ADN est donc présente à la fois chez les Procaryotes et chez la grande majorité des Eucaryotes, incluant les plantes, les champignons et les animaux. Elle est impliquée dans une myriade de processus biologiques incluant l'embryogénèse et le développement, l'empreinte parentale, l'inhibition des éléments transposables et la régulation transcriptionnelle des gènes (Li *et al.*, 1992 ; Bestor, 1998 ; Lippman *et al.*, 2004 ; Zhang *et al.*, 2006 ; Zilberman *et al.*, 2007 ; Weber & Schübeler, 2007 ; Lister *et al.*, 2008).

La méthylation de l'ADN est régulée par deux voies enzymatiques distinctes appelées méthylation « *de novo* » et « de maintenance », régies par les ADN méthyltransférases. La méthylation *de novo* correspond à la méthylation de séquences antérieurement non méthylées *via* l'ajout de groupements méthyle sur de nouveaux résidus cytosine en dehors de la phase de réplication, induisant l'apparition de nouveaux patrons de méthylation. La

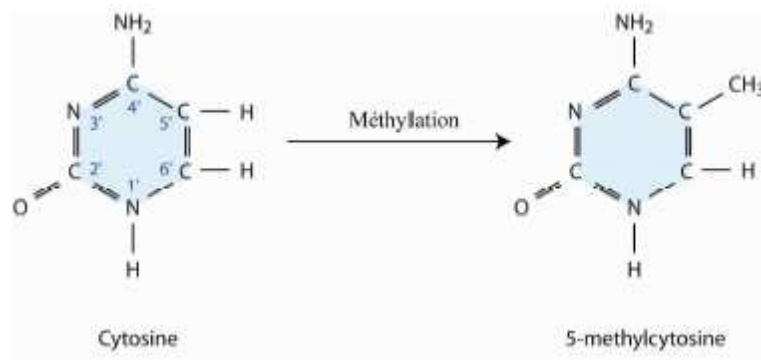


Figure III.8. Cytosine et méthylcytosine.

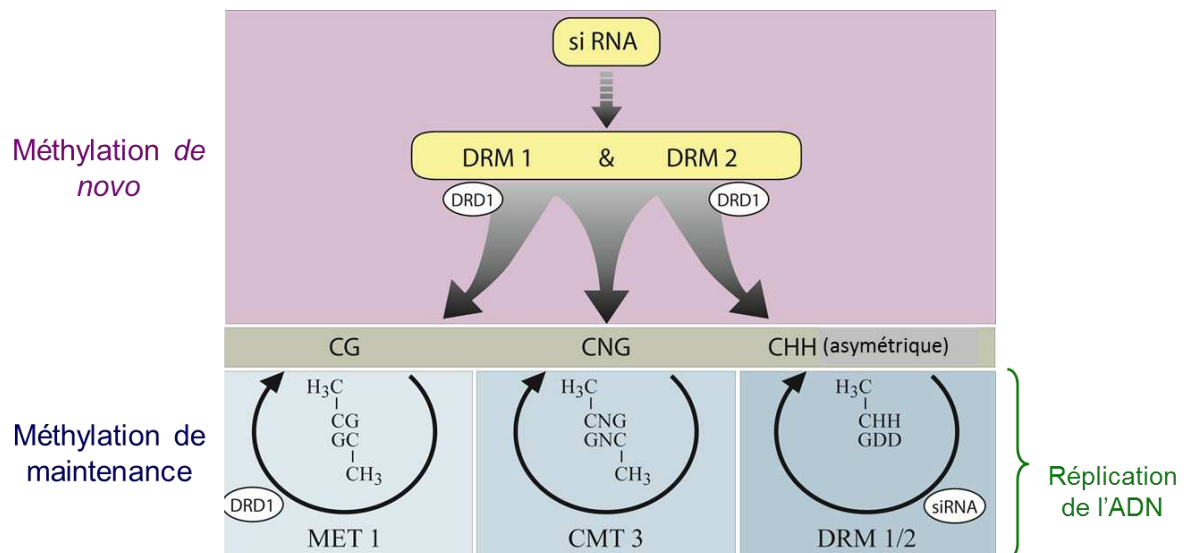


Figure III.9. Etablissement et maintenance de la méthylation chez Arabidopsis (Trap-Gentil, 2009).

méthylation de maintenance est le processus par lequel les patrons de méthylation préexistants sont maintenus après réplication de l'ADN, guidé par la méthylation présente sur le brin d'ADN parental (Zhang *et al.*, 2010). Ainsi, le profil de méthylation de l'ADN est stable et transmissible par mitose voire par méiose (Holliday, 1990).

2.3.2. Les processus de méthylation et déméthylation de l'ADN chez les plantes

Chez les plantes, la méthylation de l'ADN a lieu dans les contextes CG, CHG et CHH (où H = A, C ou T). La méthylation des sites CG est la plus abondante, ces sites pouvant atteindre un niveau de méthylation de 80 à 100%, tandis que la méthylation des sites CHH excède rarement 10%. Le niveau de méthylation des sites CHG varie de 20 à 100% (Cokus *et al.*, 2008). La méthylation des sites CG est maintenue par l'ADN méthyltransferase MET1, et par un cofacteur reconnaissant l'ADN hémi-méthylé lors de la réplication appelé VIM (VARIATION IN METHYLATION) chez *Arabidopsis* (Law & Jacobsen, 2010). La méthylation des sites CHG est maintenue par l'enzyme CMT3 spécifique aux plantes (Figure III.9), et une histone méthyltransferase appelée KRYPTONITE (KYP). La méthylation des sites CHH est maintenue par les ADN méthyltransferases DRM 1 et 2 via une voie appelée *RNA-directed DNA methylation* (RdDM ; voir paragraphe « ARN non codant » ; Law & Jacobsen, 2010). La voie RdDM est aussi responsable de la méthylation *de novo*, et ce quel que soit le contexte nucléotidique. MET1 serait par ailleurs aussi impliquée dans la méthylation *de novo* des sites CG via la voie RdDM (Matzke *et al.*, 2007).

Bien que la méthylation de l'ADN soit une marque épigénétique stable, des taux réduits de méthylation sont observés au cours du développement des plantes (Feng *et al.*, 2010b). Cette perte de méthylation peut être passive, les cytosines méthylées étant alors remplacées par des cytosines non méthylées durant la réplication de l'ADN en l'absence d'une méthylation de maintenance, ou active. La déméthylation active résulte de l'activité des ADN glycosylases en combinaison avec la voie de réparation de l'ADN nommée *base excision repair* (Ikeda & Kinoshita, 2008 ; Zhu, 2009). DEMETER (DME) et REPRESSOR OF SILENCING 1 (ROS1) sont les membres de la famille des ADN glycosylases chez *Arabidopsis*, qui inclut également DEMETER-LIKE 2 et 3 (DML2 et DML3). Chez *Arabidopsis*, les glycosylases reconnaissent les cytosines méthylées et les excisent de l'ADN double-brin quel que soit le contexte nucléotidique, une mutation dans l'un des gènes codant pour ces enzymes entraînant une hausse de la méthylation de l'ADN à des loci spécifiques (Law & Jacobsen, 2010). DME et ROS1 ont une forte préférence pour les sites CG mais possèdent aussi une activité au niveau des sites CHG et CHH (Vanyushin & Ashapkin, 2011).

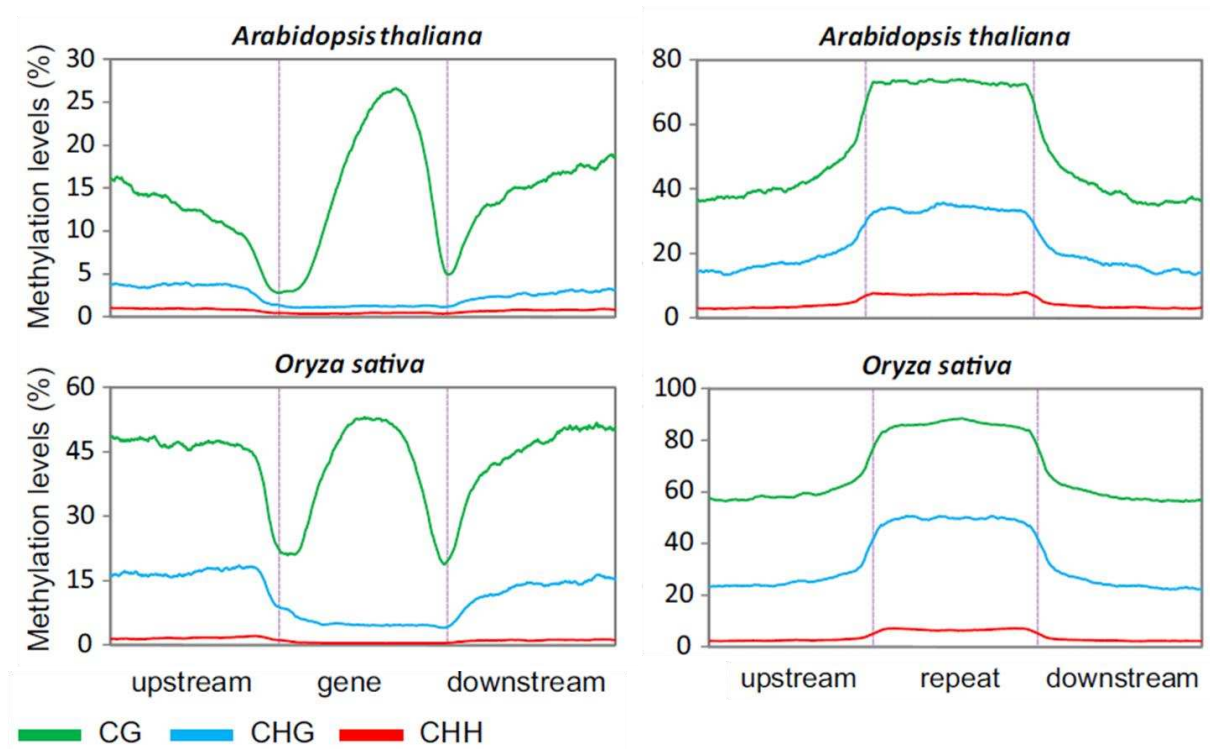


Figure III.10. Profil de méthylation des gènes (gauche) et des séquences répétées (droite) chez *Arabidopsis* et le riz (Feng *et al.*, 2010a).

2.3.3. Distribution génomique et rôle de la méthylation de l'ADN dans l'expression des gènes chez les plantes

Dès la fin des années 1980, bien que les mécanismes épigénétiques contrôlant l'expression des gènes soient encore mal compris, un nombre croissant d'éléments met en évidence l'importance de la méthylation de l'ADN dans ce processus (Holliday, 1990). Par exemple, l'inactivité transcriptionnelle de certains gènes a été associée à la présence de cytosines méthylées dans la région du promoteur de ces gènes (Doerfler, 1983 ; Dennis & Brettell, 1990). La déméthylation de l'ADN par application de 5-azacytidine a également été corrélée avec l'induction de la transcription de gènes chez certains animaux et plantes (Klaas *et al.*, 1989; Jones *et al.*, 1990).

Les études de la distribution de la méthylation de l'ADN à l'échelle du génome (méthylome), rendues possibles ces dernières années grâce à des techniques haut-débit (microarray, séquençages de nouvelle génération), ont permis une avancée importante des connaissances sur le rôle de cette marque épigénétique (Zhang *et al.*, 2006 ; Lister *et al.*, 2009). Un fait déjà connu, confirmé par ces travaux, est que les zones centromériques et péracentromériques sont méthylées dans les sites CG, CHG et CHH chez les végétaux, tout comme les séquences répétées et les transposons (Zhang *et al.*, 2006; Weber & Schübeler, 2007; Feng *et al.*, 2010a). La fonction primaire de la méthylation de l'ADN serait d'inhiber la transcription et le mouvement des éléments transposables pour protéger le génome de leurs effets délétères (Gehring & Henikoff, 2007). En effet, chez les mutants *ddm1* (*decreased cytosine methylation 1*) d'*Arabidopsis*, qui possèdent un niveau de méthylation faible, le taux de transposition est accru par rapport à une lignée sauvage et l'insertion de transposons dans les gènes engendre des phénotypes altérés (Singer *et al.*, 2001 ; Miura *et al.*, 2001).

De manière plus surprenante, la méthylation de l'ADN dans l'euchromatine et particulièrement dans le corps des gènes est loin d'être négligeable, touchant entre 30 et 45% des sites CG selon l'organisme, correspondant à *Arabidopsis* et au riz respectivement (Figure III.10 ; Cokus *et al.*, 2008; Li *et al.*, 2008; Lister *et al.*, 2008; Feng *et al.*, 2010a; Zemach *et al.*, 2010). A l'inverse, la plupart des régions promotrices sont dépourvues de méthylation, tout comme l'extrémité 3' des gènes (Zhang *et al.*, 2006; Feng *et al.*, 2010a; Zemach *et al.*, 2010). Dans le corps des gènes, la méthylation de l'ADN ne semble pas jouer un rôle de répression de la transcription. Au contraire, la méthylation dans le corps des gènes montre une relation parabolique avec leur expression, les gènes exprimés à un niveau intermédiaire étant les plus méthylés (Zhang *et al.*, 2006 ; Zilberman *et al.*, 2007 ; Zemach *et al.*, 2010 ; Figure III.11). De plus, chez *Arabidopsis*, les gènes ciblés par cette méthylation ont une tendance à être exprimés de manière ubiquiste alors que les gènes méthylés dans le

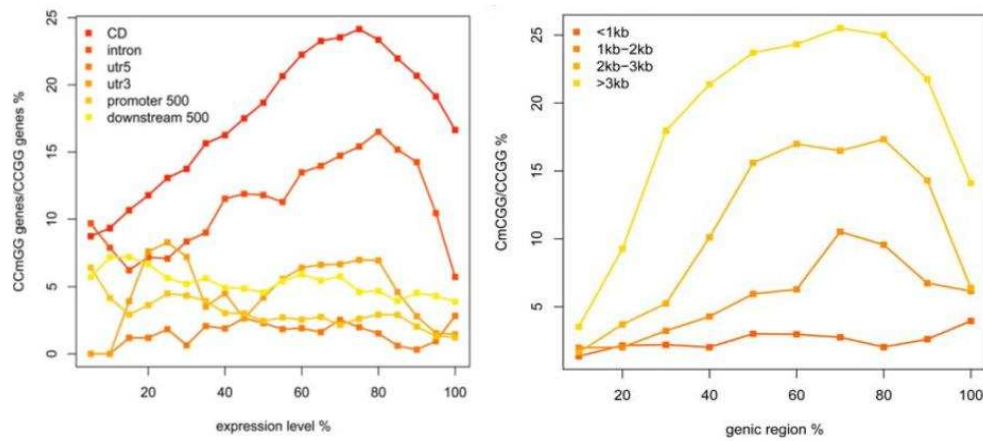


Figure III.11. Méthylation de l'ADN chez *Arabidopsis* en fonction du niveau d'expression des gènes (à gauche) et de leur longueur (à droite) (Zhang *et al.*, 2008).

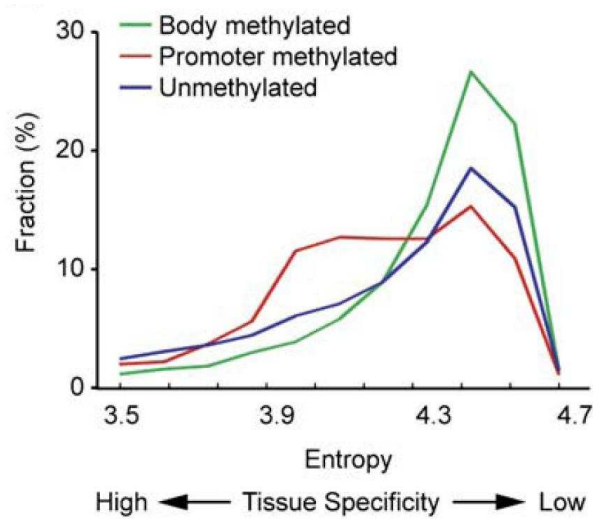


Figure III.12. Tissu-spécificité des gènes méthylés, non-méthylés et méthylés dans le promoteur chez *Arabidopsis* (Zhang *et al.*, 2006).

promoteur ont une tendance à être exprimés de façon tissu-spécifique (Zhang *et al.*, 2006 ; Figure III.12).

Aussi, chez les plantes, la relation entre méthylation de l'ADN et expression des gènes s'avère complexe et suggère un rôle de cette marque dans des processus différents, bien que liés à la transcription (Zilberman *et al.*, 2007). Dans le promoteur, les résultats existant actuellement semblent indiquer un lien négatif entre méthylation de l'ADN et fixation de facteurs de transcription, expliquant notamment l'expression de gènes tissu-spécifiques (Zhang *et al.*, 2006, 2012a). Pour le corps des gènes, l'augmentation de la méthylation avec la taille des gènes chez *Arabidopsis* a laissé penser à un rôle de protection contre les sites cryptiques d'initiation, dont la probabilité d'apparition augmente logiquement avec la longueur du gène (Zilberman *et al.*, 2007 ; Zhang *et al.*, 2008 ; Figure III.11). Bien que cette observation ait été faite à plusieurs reprises (Zilberman *et al.*, 2007; Zhang *et al.*, 2008; Roudier *et al.*, 2011), ce rôle putatif n'a encore jamais été vérifié. D'autres études faites chez *Arabidopsis* montrent que les exons sont enrichis à la fois en méthylation de l'ADN et en nucléosomes et l'ARN Pol II s'accumule plus particulièrement dans les exons par rapport aux introns. Cette forte densité de nucléosomes méthylés dans les exons induirait ainsi l'arrêt de l'ARN Pol II au niveau des exons pour augmenter la précision de l'épissage des introns (Chodavarapu *et al.*, 2010). Là encore cependant, cette idée reste une hypothèse non confirmée par des approches expérimentales. Néanmoins, une étude statistique récente de l'ensemble des données de méthylation des gènes disponibles chez *Arabidopsis* confirme que les gènes les plus méthylés en leur corps sont plus longs, plus importants fonctionnellement (d'après le phénotype des mutants connus) et évoluent moins vite (taux de mutations plus faible). Cette étude suggère donc que la méthylation de l'ADN jouerait un rôle central dans le fonctionnement de la transcription des gènes et leur évolution (Takuno & Gaut, 2012), sans qu'il soit pour l'instant bien déterminé.

2.3.4. Méthylation de l'ADN au sein des Eucaryotes : approche phylogénétique

La méthylation des cytosines est présente chez une grande partie des Eucaryotes (Feng *et al.*, 2010a; Zemach *et al.*, 2010). Dans ce contexte, il semble intéressant de connaître dans quelle mesure ses caractéristiques et son rôle sont conservés au sein de ce domaine. Grâce au développement des études de type méthylome, effectuées à ce jour sur plus de 20 espèces eucaryotes, une approche phylogénétique de la méthylation de l'ADN a pu voir le jour (Feng *et al.*, 2010a; Zemach *et al.*, 2010; Jeltsch, 2010 ; Figure III.13).

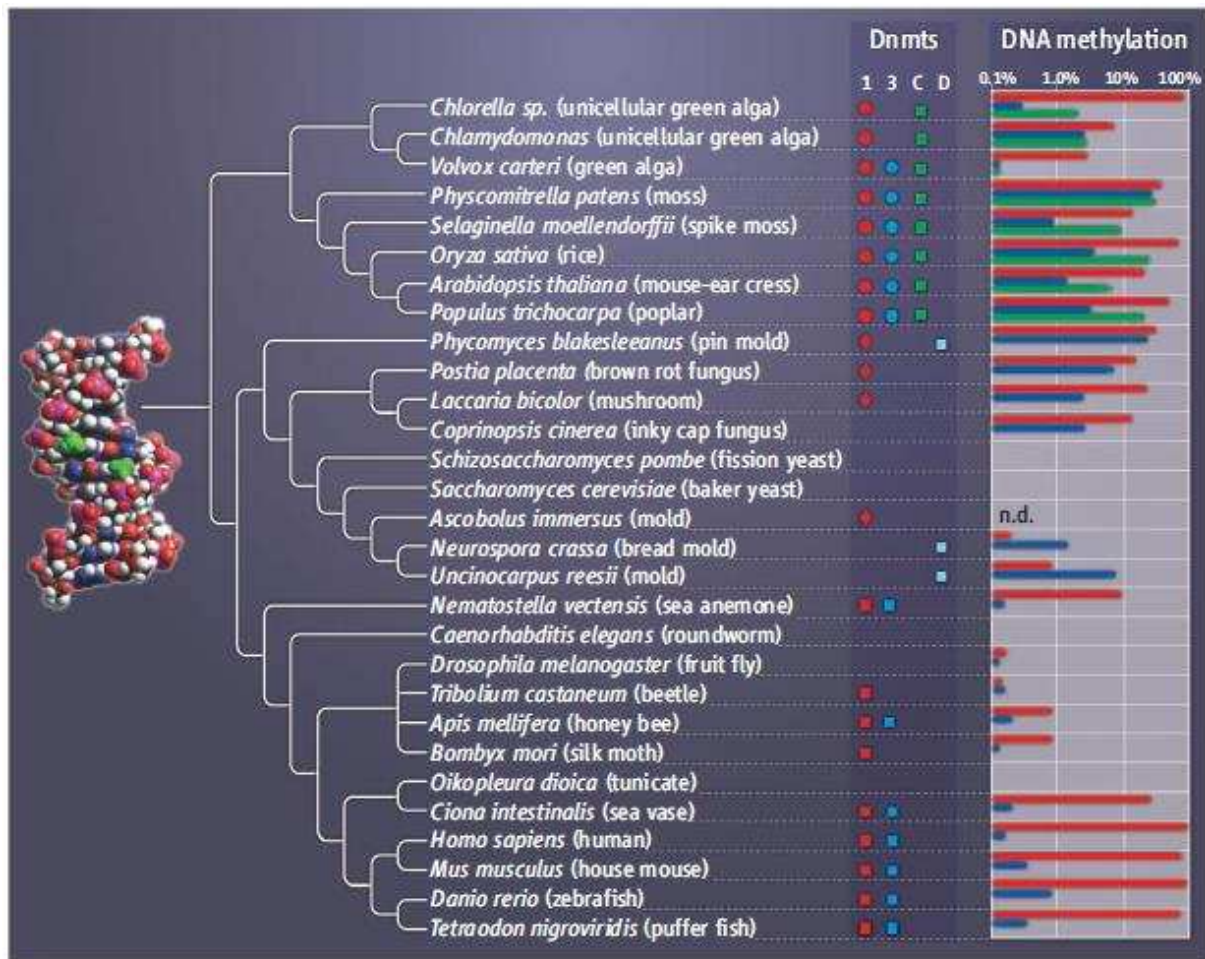


Figure III.13. Méthylomes et ADN méthyltransférases (DNMT) dans 20 espèces modèles. Les différentes familles de DNMT sont représentées : en rouge DNMT1, en bleu DNMT3, en vert les CHROMOMETHYLASES et en bleu clair DIM-2 (DEFECTIVE IN DNA METHYLATION 2). À droite sont représentés les différents contextes de méthylation associés aux DNMT : en rouge CG, en vert CHG et en bleu CHH (Jeltsch, 2010).

En ce qui concerne les ADN méthyltransférases, la MET1 des plantes possède un homologue dans la plupart des Eucaryotes étudiés (Jeltsch, 2010 ; Figure III.13), la DNMT1. Ainsi, de manière logique, la méthylation en CG est retrouvée chez l'homme, la souris, la ciconie, le poisson zèbre, l'abeille ou encore les champignons (Suzuki & Bird, 2008 ; Lister *et al.*, 2009 ; Feng *et al.*, 2010a ; Zemach *et al.*, 2010). Cependant, une grande variabilité dans le niveau et la distribution de cette méthylation est observée au sein de ces organismes. En effet, chez les mammifères ou les poissons, la méthylation touche 70-80% des dinucléotides CG dans tout le génome (Smallwood & Kelsey, 2012). Les îlots CpG (loci à forte densité en dinucléotides CG) sont très peu touchés par la méthylation de l'ADN et sont retrouvés dans 40% des promoteurs et des exons (Larsen *et al.*, 1992). A l'inverse, les insectes tels que l'abeille ou encore les champignons présentent une méthylation de l'ADN assez faible (Zemach *et al.*, 2010).

Les autres types d'ADN méthyltransférase sont moins répandus au sein des Eucaryotes, expliquant une plus grande disparité de la méthylation de l'ADN non-CG dans ce domaine (Jeltsch, 2010). L'enzyme DRM, responsable de la méthylation *de novo* des sites CG et non-CG chez les plantes, possède un homologue chez une partie des animaux uniquement (notamment les vertébrés), la DNMT3. Cependant, contrairement aux plantes, la méthylation non-CG chez les vertébrés est très faible et restreinte aux cellules souches embryonnaires (Ramsahoye *et al.*, 2000 ; Lister *et al.*, 2009), bien que la raison en soit inconnue à ce jour. Enfin, CMT3 n'est présente que chez les végétaux, expliquant chez ces organismes une forte méthylation des sites CHG. Néanmoins, DIM-2, une ADN méthyltransférase propre aux champignons, affecte indifféremment sites CG et CH et présente des caractéristiques communes à CMT3 suggérant une origine monophylétique de ces deux enzymes (Zemach *et al.*, 2010).

De manière très générale, alors que la méthylation de l'ADN chez les vertébrés touche l'ensemble du génome, les invertébrés, les plantes et les champignons présentent quant à eux une mosaïque de méthylation caractérisée par des domaines méthylés et non méthylés dispersés (Suzuki & Bird, 2008 ; Figure III.14). Plus précisément, les zones centromériques et péri-centromériques sont, comme chez les plantes, hyperméthylées chez les vertébrés et les champignons, tout comme les transposons, dont l'activité est réprimée par la méthylation de l'ADN (Zhang *et al.*, 2006; Weber & Schübeler, 2007; Feng *et al.*, 2010a; Zemach *et al.*, 2010 ; Figure III.15). Cependant, les insectes comme l'abeille et le ver à soie constituent une exception puisque les séquences répétées y sont hypométhylées, indiquant que les gènes sont la cible préférentielle de la méthylation (Feng *et al.*, 2010a ; Xiang *et al.*, 2010). De manière intéressante, la méthylation des sites CG du corps des gènes a été observée aussi bien chez les animaux que les végétaux (Cokus *et al.*, 2008 ;

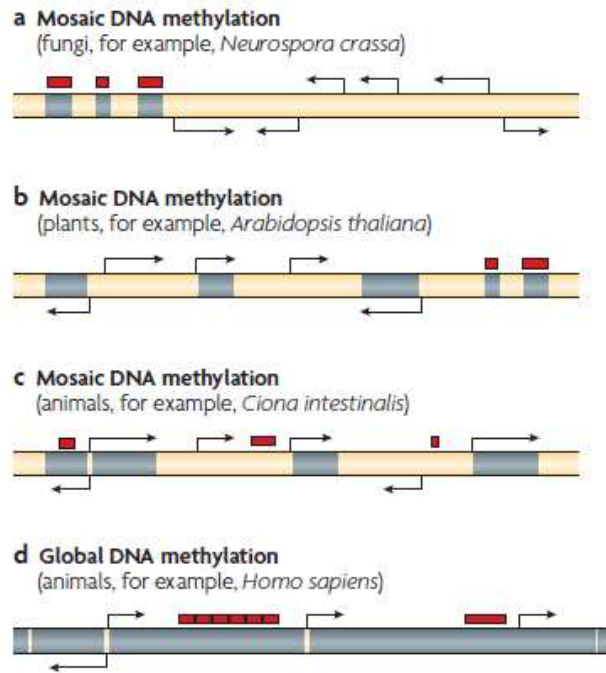


Figure III.14. Répartition de la méthylation de l'ADN dans le génome de divers organismes. (a) Chez un champignon, *Neurospora crassa* ; (b) chez une plante, *Arabidopsis thaliana* ; (c) chez un animal invertébré, *Ciona intestinalis* ; (d) chez l'humain. Les domaines méthylés sont représentés en gris, les domaines non-méthylés en jaune et les éléments transposables en rouge (Suzuki & Bird, 2008).

Zhang *et al.*, 2008 ; Ball *et al.*, 2009 ; Lister *et al.*, 2009 ; Feng *et al.*, 2010a ; Zemach *et al.*, 2010 ; Figure III.15). De plus, les exons sont globalement enrichis en méthylation dans les sites CG par rapport aux introns chez les animaux comme les végétaux, particulièrement chez l'abeille et *Chlamydomonas* et de façon plus discrète chez les vertébrés (Lister *et al.*, 2009 ; Feng *et al.*, 2010a). La colocalisation des nucléosomes, de la méthylation de l'ADN et de l'ARN Pol II observée chez *Arabidopsis* au niveau des exons semble aussi vraie chez les animaux (Andersson *et al.*, 2009; Chodavarapu *et al.*, 2010), tout comme la relation parabolique entre méthylation du corps des gènes et leur expression (Zemach *et al.*, 2010). Aussi, le rôle de la méthylation du corps des gènes dans le fonctionnement de la transcription, possiblement *via* son influence sur l'épissage, semble indiquer que ce phénomène est une propriété ancienne commune aux génomes eucaryotes (Tran *et al.*, 2005; Zilberman *et al.*, 2007; Feng *et al.*, 2010a; Zemach *et al.*, 2010; Jeltsch, 2010 ; Figure III.15). Par ailleurs, la plupart des régions promotrices sont dépourvues de méthylation chez les plantes comme les animaux (Zhang *et al.*, 2006 ; Weber & Schübeler, 2007 ; Feng *et al.*, 2010a ; Zemach *et al.*, 2010), suggérant également un rôle de la méthylation des promoteurs dans l'expression des gènes similaire au sein des Eucaryotes.

Enfin, l'absence de la méthylation de l'ADN chez des organismes modèles tels que *Saccharomyces cerevisiae* ou *Caenorhabditis elegans* ou son niveau très faible chez la drosophile suggère que cette marque épigénétique n'est pas essentielle pour le fonctionnement d'un organisme (Zemach & Zilberman, 2010). Elle peut même s'avérer délétère, les cytosines méthylées étant plus sujettes à la mutation par déamination (C→T) que les cytosines non modifiées (Shen *et al.*, 1994). Néanmoins, la répression de l'activité des éléments transposables paraît particulièrement nécessaire aux organismes pluricellulaires et à reproduction sexuée, présentant une méthylation de l'ADN importante et ciblée sur les séquences répétées (Feng *et al.*, 2010a; Zemach *et al.*, 2010). En effet, le mode de reproduction sexuée semble avoir favorisé la sélection d'éléments transposables particulièrement actifs, nécessitant d'autant plus leur répression pour l'intégrité du génome (Zemach & Zilberman, 2010).

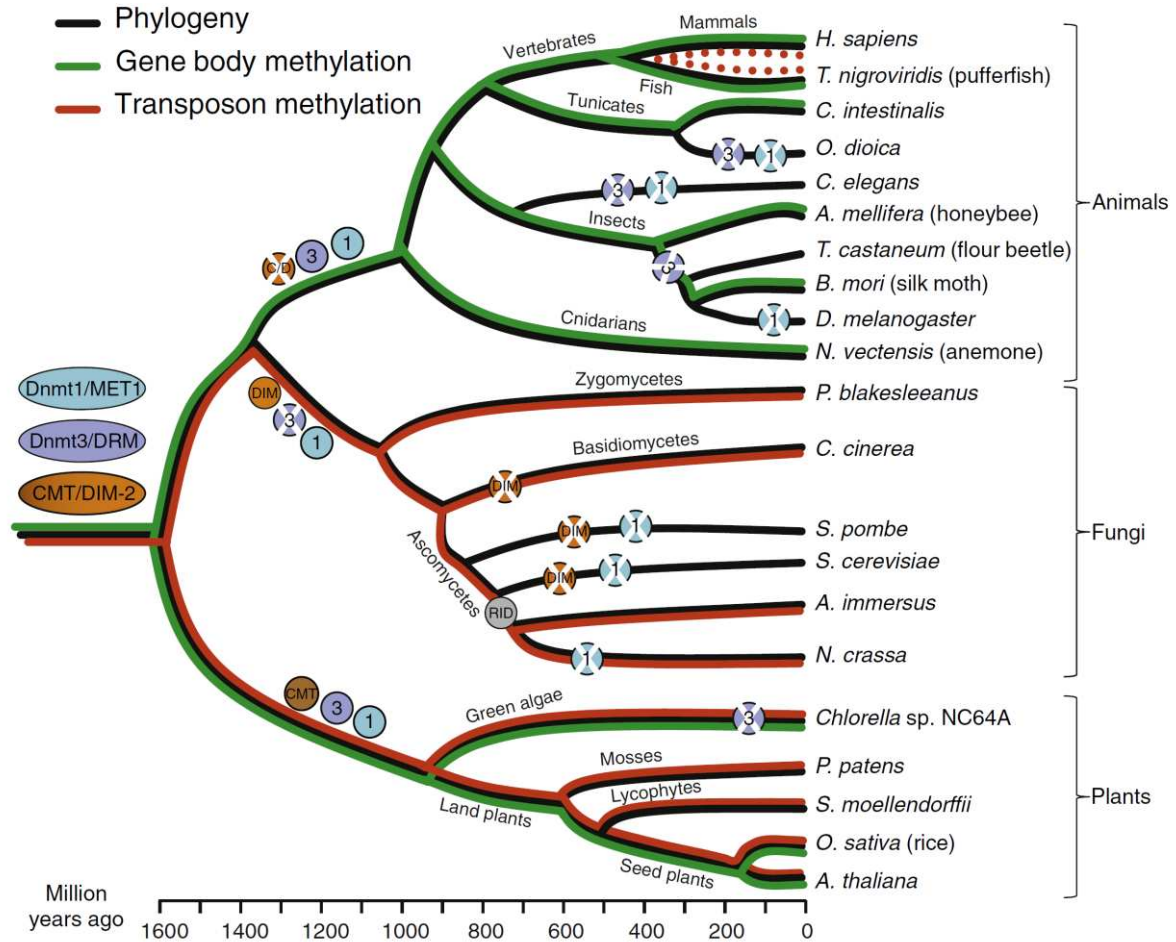


Figure III.15. Evolution de la méthylation de l'ADN chez les Eucaryotes. Les cercles contenant « 1 », « 3 », « CMT », « DIM » et « C/D » représentent les ADN méthyltransférases DNMT1, DNMT3, CMT, DIM-2 et CMT/DIM-2 respectivement. Les cercles avec une croix représentent la perte de la famille du gène indiqué. Les lignes verte et marron montrent la trajectoire évolutive de la méthylation du corps des gènes et des éléments transposables (TE), respectivement. Les lignes marrons en pointillés représentent la méthylation des TE chez les vertébrés (Zemach & Zilberman, 2010).

2.4. ARN non codant

Un autre niveau de contrôle épigénétique implique les ARN non codant (Esteller, 2011). Des phénomènes tels que l'inactivation du chromosome X chez les mammifères, l'empreinte parentale et certains cas de paramutation sont contrôlés par des ARN non codant qui entraînent des modifications covalentes de la chromatine aboutissant à un état répressif stable et héritable (Lejeune & Allshire, 2011). Cependant, tous les ARN non codant n'induisent pas ces modifications épigénétiques et n'ont un rôle que dans la régulation post-transcriptionnelle, tels les micro-ARN (Esteller, 2011). Parmi ceux induisant des modifications de la chromatine stables, les *PIWI-interacting RNA* (piRNA) chez les animaux et les *small interfering RNA* (siRNA) chez les plantes ont été abondamment décrits (Teixeira & Colot, 2010; Esteller, 2011). Ces deux voies de répression épigénétique ont une partie de leur fonctionnement commun. En effet chez les deux règnes, ces ARN non codants sont transcrits à partir de séquences cibles (transposons en particulier), reconnus et liés par des protéines de la famille ARGONAUTE. Ces protéines liées aux ARN non codants vont ensuite se fixer sur les séquences cibles et induire leur répression chromatinienne *via* le recrutement d'enzymes de méthylation de l'ADN ou de modification des histones (Teixeira & Colot, 2010; Esteller, 2011). Chez les plantes, ces ARN non codants sont synthétisés (simple brin) par l'ARN Pol IV, propre aux plantes, à partir desquels des ARN double brin sont générés par la RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) puis clivés par DICER-LIKE 3 en ARN de 24 nucléotides (nt), les siRNA (Lahmy *et al.*, 2010; Teixeira & Colot, 2010). Ces siRNA sont ensuite fixés à ARGONAUTE 4 (AGO4) et ce complexe s'associe à des enzymes de modification des histones, de méthylation de l'ADN et de remodelage de la chromatine aboutissant à la répression des séquences cibles (Figure III.16).

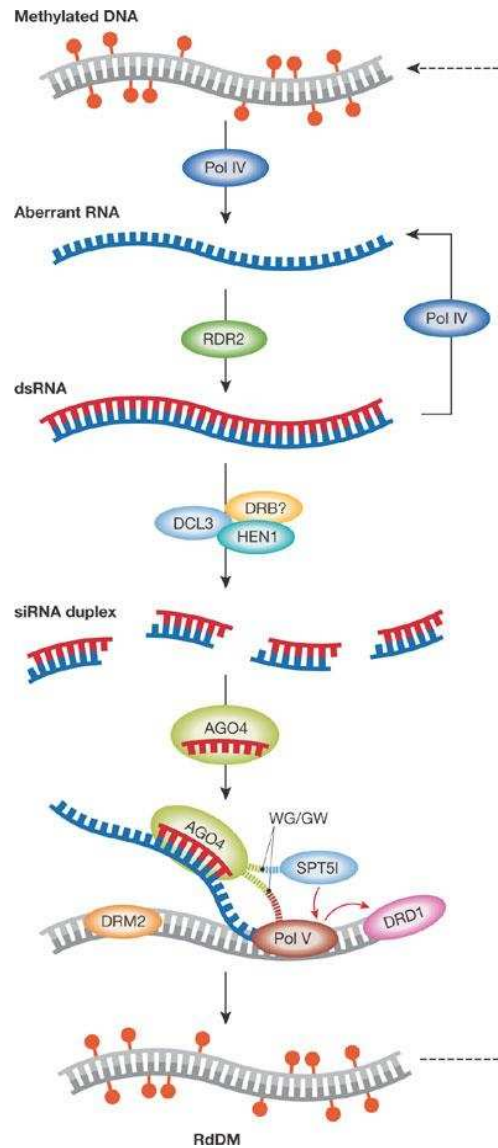


Figure III.16. Voie *RNA-directed DNA Methylation* (RdDM) chez les plantes. AGO4, ARGONAUTE 4; DCL3, DICER-LIKE 3; DRB, DOUBLE-STRANDED RNA-BINDING PROTEIN; DRD1, DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1; DRM2, DOMAIN REARRANGED METHYLASE 2; dsRNA, *double-stranded RNA*; HEN1, HUA ENHANCER 1; nt, *nucleotide*; Pol IV, *RNA polymerase IV*; RdDM, *RNA-directed DNA methylation*; RDR2, *RNA-dependent RNA polymerase II*; siRNA, *small-interfering RNA*; SPT5I, *suppressor of Ty insertion 5-like*; WG/GW, *neighbouring tryptophan–glycine/glycine–tryptophan residues* (Wang & Dennis, 2009).

2.5. Etats chromatinien et marques épigénétiques

Il n'est pas toujours pertinent d'étudier les différentes marques épigénétiques séparément. En effet, deux études récentes de la distribution d'un grand nombre de marques dans le génome, l'une chez la drosophile et l'autre chez *Arabidopsis*, ont démontré qu'un état donné de la chromatine est le résultat d'une combinaison de marques, plutôt que le fait d'une seule (Lee *et al.*, 2010; Filion *et al.*, 2010; Roudier *et al.*, 2011). Malgré le grand nombre de facteurs étudiés (53 protéines associées à la chromatine, 11 marques histones, méthylation de l'ADN), seulement 5 combinaisons, donc 5 états chromatinien chez la drosophile et 4 chez *Arabidopsis* ont été trouvés. Les conclusions sont similaires, montrant qu'il n'existe pas l'hétéro- et l'euchromatine, mais des hétérochromatines et des euchromatines, dont les marques épigénétiques et leurs combinaisons sont spécifiques à chaque état, et dont la répartition dans le génome n'est pas fixe, influencée de manière tissu-dépendante notamment. Enfin, ces états chromatinien présentent énormément de caractéristiques communes entre la drosophile et *Arabidopsis* (Roudier *et al.*, 2011).

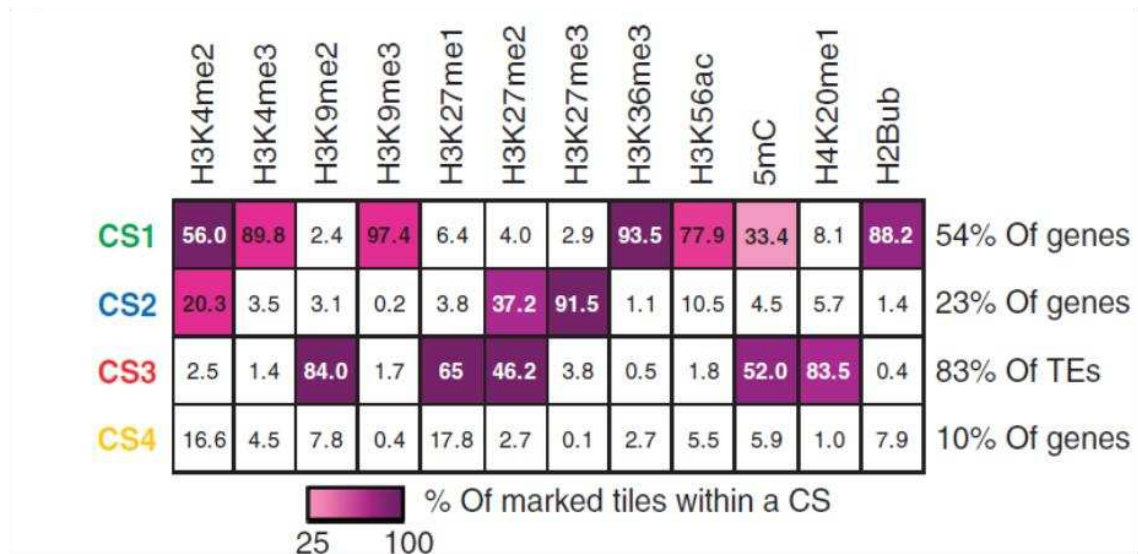
Plus particulièrement, chez *Arabidopsis*, chacun des 4 états chromatinien a pu être associé à des caractéristiques propres (Figure III.17A) :

- le 1^{er}, riche en de nombreuses marques histone dont H3K4me2-3, H3K9me3 ou H3K36me3 et avec un niveau moyen en méthylation de l'ADN, contient la majorité des gènes et est associé à un niveau d'expression important.
- le 2^e, riche en H3K4me2 et H3K27me2-3 et sans méthylation de l'ADN, est essentiellement composé de gènes, réprimés en majorité.
- le 3^e, riche en H3K9me2 et H4K20me1 notamment et avec un fort niveau de méthylation de l'ADN, contient la grande majorité des éléments transposables.
- le dernier, associé à aucune modification particulière, est plutôt situé dans l'intergénique.

Ainsi, en plus de l'hétérochromatine constitutive, riche en éléments transposables et en méthylation de l'ADN entre autres, un autre état répressif de la chromatine a été décrit, spécifique des gènes, sans méthylation de l'ADN et caractéristique d'une répression par les protéines du groupe Polycomb (Köhler & Hennig, 2010). Les 4 états chromatinien répartis en domaines successifs chez *Arabidopsis* forment un épigénome mosaïque (Figure III.17B).

Ces études, bien que permettant une vision d'ensemble de la chromatine et de ses états, ne permettent pas une étude précise de l'état chromatinien au sein des différentes parties géniques, notamment des promoteurs.

A)



B)

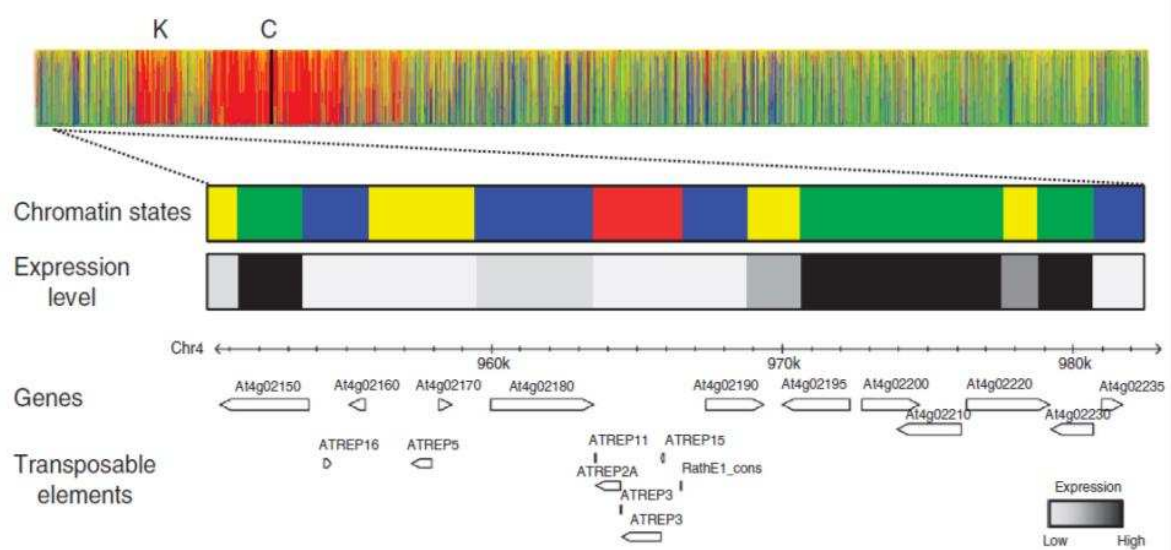


Figure III.17. Les 4 états principaux de la chromatine chez Arabidopsis (Roudier *et al.*, 2011). A) Le tableau indique la composition des 4 états chromatinien (CS) principaux. La distribution des 12 marques de la chromatine au sein de chaque CS est indiquée par une *heat map* représentant le pourcentage de sondes contenant les marques, de 25% (rose clair) à 100% (violet foncé). Les chiffres à l'intérieur des cases représentent, pour chaque marque épigénétique, le pourcentage de sondes associées à un des 4 états chromatinien. Le pourcentage de gènes contenus dans les états CS1, CS2 et CS4 et le pourcentage de TE contenus dans l'état CS3 sont aussi montrés. B) Distribution des 4 CS le long du chromosome 4. Chaque domaine est coloré selon son état chromatinien. K : *heterochromatic knob*. La partie non séquencée du centromère (C) est représentée par la ligne noire verticale.

Au vu de l'avancée des connaissances des mécanismes moléculaires, une nouvelle définition de l'épigénétique a été proposée comme « les altérations de la chromatine qui établissent et propagent différents profils d'expression de gènes à partir d'un même génome. » (Figure III.18 ; Allis *et al.*, 2007). Roudier *et al.* (2009) souligneront l'importance de la propagation par mitose/méiose de ces altérations, n'étant autrement que des mécanismes de régulation chromatinienne de la transcription des gènes (Figure III.19).

Malgré la progression des connaissances en épigénétique, un grand nombre de phénomènes restent encore mal compris, de la paramutation au rôle de la méthylation dans l'expression des gènes en passant par les mécanismes coordonnant l'ensemble des marques épigénétiques qui se regroupent dans des combinaisons bien précises. La découverte de nouvelles marques comme l'hydroxyméthylcytosine laisse penser que les phénomènes régissant la stabilité et l'organisation du génome sont encore bien mal appréhendés...

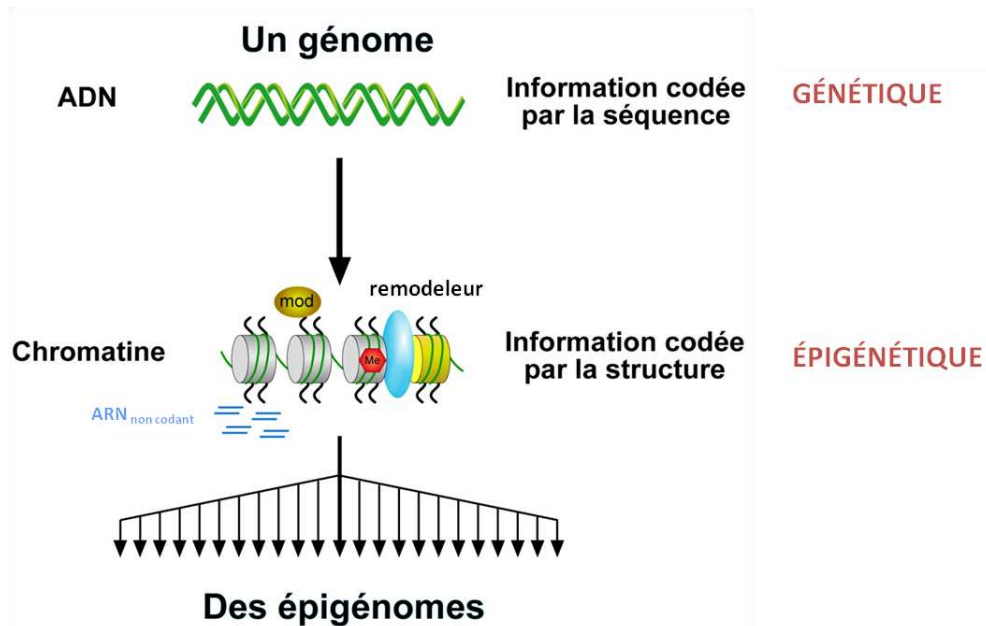


Figure III.18. Schéma représentant la définition moderne de l'épigénétique. A partir d'un génome, plusieurs états de la chromatine, transmissibles par mitose/méiose, peuvent exister (Allis *et al.*, 2007). Cette figure est déjà présentée dans le préambule mais pour faciliter la lecture, elle est rajoutée ici.

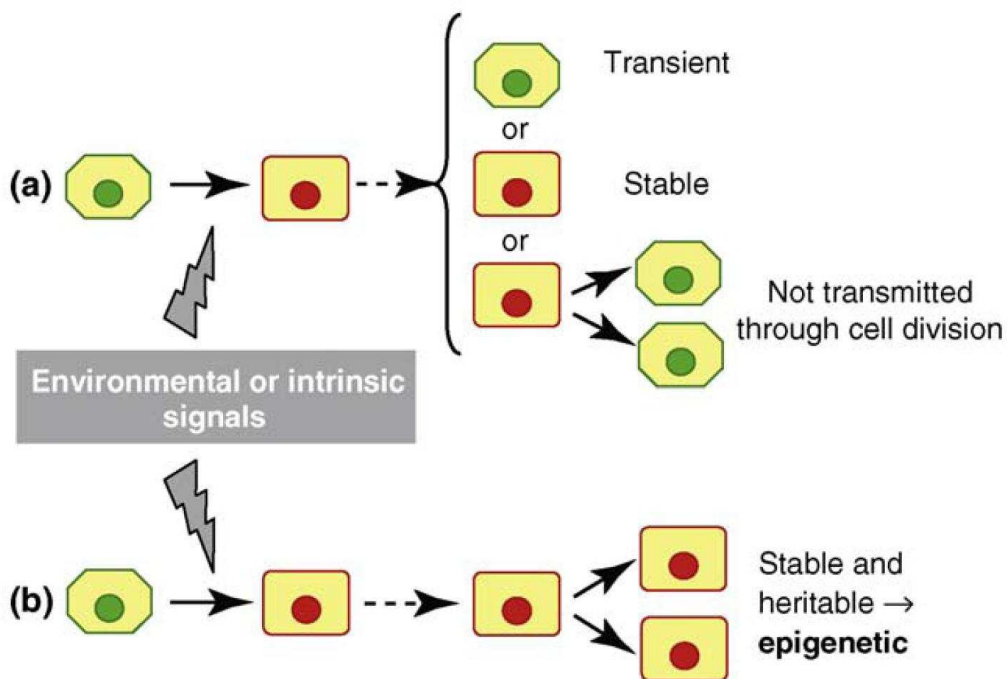


Figure III.19. Régulation chromatinienne ou réponse épigénétique après un signal environnemental ou intrinsèque (Roudier *et al.*, 2009). (a) Les états chromatinien peuvent être soit inversés en l'absence du signal d'induction ou stables mais non transmis aux cellules filles (= régulation chromatinienne). (b) Les états chromatinien peuvent être transmis aux cellules filles en l'absence du signal d'induction (= épigénétique).

3. Implication de la méthylation de l'ADN dans la vie de la plante

3.1. Développement

A l'heure actuelle, l'implication de la méthylation des cytosines dans le contrôle développemental des plantes n'est que partiellement connue. Néanmoins, l'utilisation de mutants de la méthylation (mutation des gènes codant des enzymes liées à la mise en place de la méthylation) a permis une progression dans ce domaine (Zhang *et al.*, 2010). La première preuve de l'implication de la méthylation dans le développement des végétaux provient de l'étude du mutant anti-sens *met1* d'*Arabidopsis* (Figure III.20), qui révèle des phénotypes remarquables tels que la réduction de la dominance apicale, le décalage du délai de floraison, des anomalies florales et des feuilles enroulées (Finnegan *et al.*, 1996). Ces effets sont transmissibles au cours des divisions cellulaires et se maintiennent pendant plusieurs générations. Comme les plantes-mères, les plantes-filles présentent de faibles taux de méthylation, même en l'absence du transgène (Finnegan *et al.*, 1996). Chez le tabac, les mutants anti-sens *met1* présentent des phénotypes semblables à ceux observés chez *Arabidopsis* (Nakano *et al.*, 2000), suggérant que la méthylation des cytosines a des effets pléiotropiques et affecte de nombreux gènes impliqués dans le développement des végétaux.

Chez les végétaux, plusieurs études comparant les niveaux de méthylation globale de différents tissus ont révélé des différences. Ainsi, chez la tomate des tissus immatures tels que des tiges, des feuilles et des racines sont moins méthylés que des feuilles matures, des fruits ou des graines (Messeguer *et al.*, 1991). Ces résultats sont en accord avec une étude plus récente (Teyssier *et al.*, 2008) montrant que les taux de méthylation varient dans les différents tissus du fruit de tomate de façon distincte au cours du développement et de la maturation du fruit. De plus, des régions différenciellement méthylées ont été identifiées entre plusieurs organes (Ruiz-García *et al.*, 2005) chez *Arabidopsis* et le peuplier (Vining *et al.*, 2012), entre des plantules et des plantes adultes chez le riz (Xiong *et al.*, 1999; Sha *et al.*, 2005) et entre des types de feuilles particulier chez le maïs (Lu *et al.*, 2008). Ainsi, la méthylation de l'ADN à l'échelle du génome pourrait jouer un rôle dans l'orientation et la maintenance de l'expression différentielle des gènes entre tissus (Vining *et al.*, 2012). Des variations de méthylation globale et des séquences différenciellement méthylées ont aussi été observées chez la betterave sucrière lors de la différenciation et organogenèse *in vitro* à partir de cals (Causevic *et al.*, 2005, 2006; Maury *et al.*, 2012).



Figure III.20. Effets d'une mutation dans le gène *MET1* chez *Arabidopsis*. A gauche le phénotype sauvage et à droite les différents phénotypes obtenus à partir de transformations indépendantes du gène *MET1* (Vaucheret & Fagard, 2001).

Chez les végétaux, les variations de méthylation de l'ADN des gènes soumis à l'empreinte parentale font partie de leur cycle de vie en conditions normales. L'empreinte parentale est un phénomène bien connu chez les plantes comme chez les mammifères correspondant à l'expression différentielle des allèles maternels et paternels selon leur origine parentale. Chez les Angiospermes, l'empreinte parentale prend place essentiellement dans l'albumen, qui sert de réserve nutritive à l'embryon au cours du développement de la graine. Pour se faire, une phase de déméthylation dans la cellule centrale, le précurseur de l'albumen, est requise (Choi *et al.*, 2002 ; Gehring *et al.*, 2006 ; Jullien *et al.*, 2008). Cette déméthylation serait à la fois passive *via* l'inactivation de MET1 et l'absence de méthylation de maintenance qui en résulte (Jullien *et al.*, 2008), et active *via* l'activité de déméthylation de DME. Cette déméthylation induit l'expression de l'allèle maternel dans l'endosperme, tandis que l'allèle paternel reste méthylé et inactif. Dernièrement, des études à l'échelle génomique ont permis d'isoler entre 100 et 300 gènes régulés par l'empreinte parentale chez *Arabidopsis*, le riz et le maïs (Gehring *et al.*, 2011; McKeown *et al.*, 2011; Hsieh *et al.*, 2011; Luo *et al.*, 2011; Wolff *et al.*, 2011; Waters *et al.*, 2011; Zhang *et al.*, 2011)

La méthylation de l'ADN est impliquée dans le passage du stade végétatif au stade reproducteur chez les végétaux. En effet, la vernalisation, processus correspondant à l'acquisition ou l'accélération de l'aptitude à fleurir après une exposition au froid de plusieurs semaines (Chouard, 1960), induit des variations de méthylation globale de l'ADN (Burn *et al.*, 1993; Trap-Gentil *et al.*, 2011). De plus, le traitement de plantes avec de la 5-azacytidine, un inhibiteur des DNMT, induit la floraison chez l'écotype C24 d'*Arabidopsis* (Burn *et al.*, 1993 ; Dennis *et al.*, 1998 ; Finnegan *et al.*, 1998) et le blé (Brock & Davidson, 1994) mais pas chez la betterave sucrière (Trap-Gentil *et al.*, 2011). Des études de génétique inverse montrent que des mutations dans les gènes *DECREASED DNA METHYLATION1* (*DDM1*, un facteur de remodelage de la chromatine), *MET1* (Vongs *et al.*, 1993 ; Kakutani *et al.*, 1996), *KYP* (Jackson *et al.*, 2004) et *ROS1* (Agius *et al.*, 2006) affectent le niveau global de méthylation, menant à des variations du délai de floraison chez *Arabidopsis*. Chez la betterave sucrière des cibles des remaniements de la méthylation de l'ADN au cours de la vernalisation ont été identifiées (Hébrard *et al.*, 2013). La méthylation de l'ADN est donc impliquée dans la régulation de nombreux processus développementaux chez les plantes mais également dans la réponse à des variations de l'environnement, comme en témoigne l'exemple de la vernalisation.

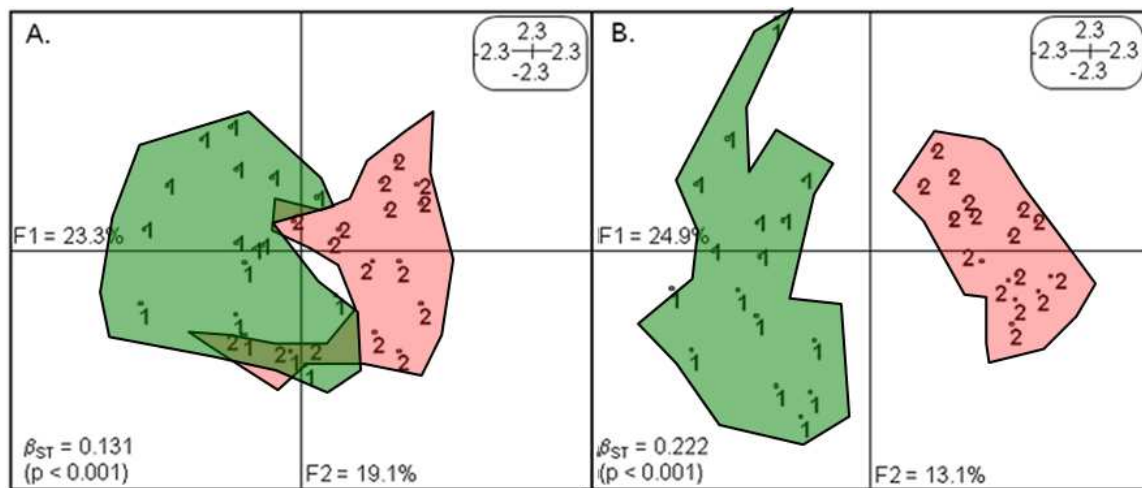


Figure III.21. Analyses en composantes principales des profils (A) génétiques et (B) épigénétiques dans les populations naturelles de *Laguncularia racemosa*. Les individus retrouvés sur les bords de rivières sont entourés de vert et les individus retrouvés dans les marais salants sont entourés de rose (Lira-Medeiros *et al.*, 2010).

3.2. Réponse à l'environnement et plasticité phénotypique

Les végétaux sont constamment exposés aux variations de l'environnement. Parce que ce sont des organismes sessiles, ils ne peuvent pas fuir ces conditions stressantes. Pour cette raison, les plantes ont développé des mécanismes leur permettant de répondre et de s'adapter aux stress biotiques et abiotiques récurrents (Blödner *et al.*, 2007). Les végétaux peuvent modérer les effets du stress en régulant l'expression de gènes impliqués dans diverses voies, notamment *via* la méthylation de l'ADN.

Chez *Arabidopsis*, une baisse de méthylation réduit la tolérance au stress salin. En effet, les mutants *met1*, dont la méthylation est quasiment abolie, sont hypersensibles au stress salin. De plus, le transporteur de sodium *AtHKT1*, qui est indispensable à la tolérance au stress salin, est surexprimé dans les racines de ce mutant. *AtHKT1* présente une région hyperméthylée en amont de son codon *start*, qui serait la cible de petits ARN impliqués dans la méthylation des sites non-CG. MET1 compléterait la méthylation de cette région dans les sites CG. La méthylation de cette zone mène donc à l'inhibition d'*AtHKT1* et à une tolérance accrue au stress salin (Baek *et al.*, 2011). A l'inverse chez *Laguncularia racemosa*, une plante spécifique de la mangrove, une hypométhylation à l'échelle du génome est associée à une tolérance au stress salin accrue (Lira-Medeiros *et al.*, 2010). Cette plante présente une grande variabilité morphologique selon l'endroit où elle vit. Lorsqu'elle est implantée en bord de rivière (BR), *Laguncularia racemosa* est plus haute et ses feuilles sont plus développées que dans les marais salants (MS), où la salinité est plus élevée. Les plantes MS présentent une proportion de loci méthylés moindre par comparaison aux plantes BR après analyse MSAP (*Methylation Sensitive Amplification Polymorphism*). De plus, les individus analysés présentent des profils génétiques similaires mais des profils épigénétiques distincts caractéristiques des deux environnements étudiés (Figure III.21).

La méthylation de l'ADN intervient dans la réponse à des stress abiotiques mais aussi dans la réponse aux agents pathogènes. Chez *Arabidopsis*, des mutants déficients pour la méthylation de maintenance des sites CG (*met1*) et non-CG (*ddc* pour *drm1/drm2/cmt3*) exposés à la bactérie pathogène *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*) se montrent résistants à la colonisation bactérienne et ne développent pas le phénotype chlorotique caractéristique d'une infection pathogène. Des mutants partiellement affectés pour la méthylation des sites CG (*ddm1*) ou non-CG (*rdr1*, *rdr2*, *rdr6*, *drd1*, *nrdp1a* et *dcl2/3/4*) présentent une résistance à *Pst* modestement accrue (Figure III.22). La croissance bactérienne de souches non-virulentes ou non-pathogènes de *P. syringae* est également réduite chez les mutants *met1-3* et *ddc*, indiquant que la baisse de méthylation associée à toutes ces mutations augmente de façon non-spécifique la résistance aux bactéries

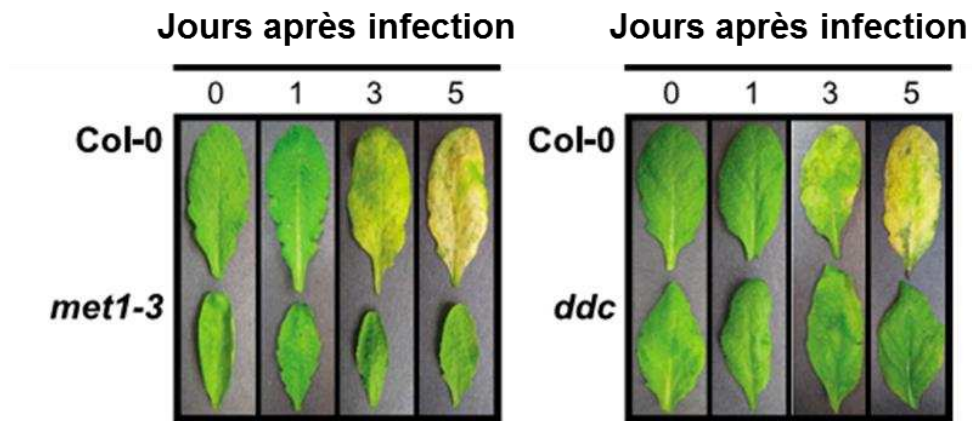


Figure III.22. Feuilles issues de lignées sauvages Col-0 ou des mutants *met1-3* ou *ddc* après infection par *Pseudomonas syringae* chez *Arabidopsis* (Dowen *et al.*, 2012).

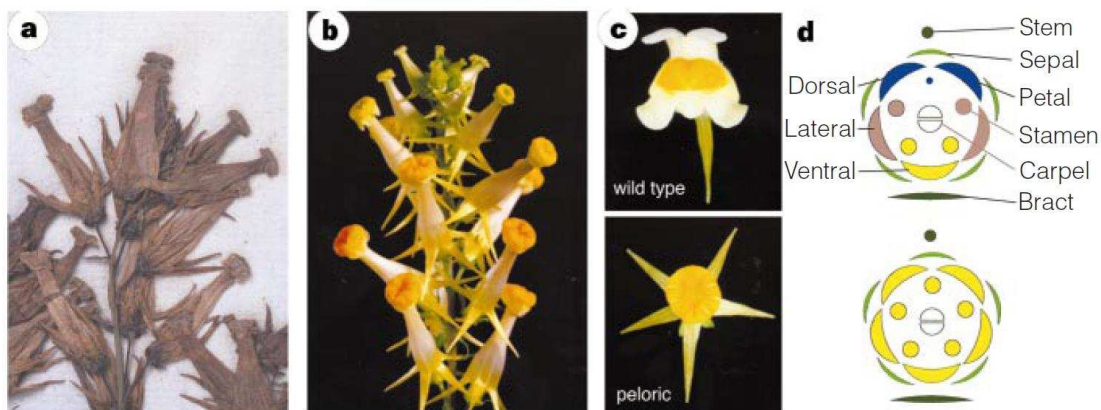


Figure III.23. Fleurs du type sauvage et du mutant pélorique chez *Linnea vulgaris*. (a) spécimen de fleurs du mutant pélorique issu de l'herbier originel élaboré par Linné. (b) Inflorescence d'un spécimen vivant du mutant pélorique. (c) Vue de face de fleur de type sauvage et du mutant pélorique. (d) Diagramme floral du type sauvage et du mutant pélorique (Cubas *et al.*, 1999).

(Downen *et al.*, 2012). De plus, l'exposition à *Pst* mène à des variations d'expression de gènes de réponse à des pathogènes chez les mutants par rapport à la lignée sauvage, indiquant que les réseaux transcriptionnels induits par *Pst* interfèrent avec la méthylation de l'ADN pour réguler l'expression des gènes (Downen *et al.*, 2012).

Ainsi, des altérations de la méthylation de l'ADN sont donc une source possible de variations phénotypiques chez les plantes en l'absence de changement dans la séquence d'ADN. Des études ont montré que cette marque épigénétique pouvait être transmise *via* la méiose, laissant supposer que les variations phénotypiques associées pourraient aussi être transmises.

3.3. Les variations et leur transmission transgénérationnelle

Pour expliquer l'évolution des espèces, Darwin postulait que « les variations, quelque faibles qu'elles soient et de quelque cause qu'elles proviennent, (...) se transmettent ordinairement à leur descendance, pourvu qu'elles soient utiles à ces individus » (Darwin, 1859). Il se demandait par ailleurs « qui pourrait soutenir que (...) l'épaississement de la fourrure d'un animal arctique n'est pas héréditaire pendant quelques générations tout au moins ? », rejoignant les idées d'Aristote et Lamarck sur la transmission des caractères acquis. Remettant ces idées au goût du jour, la transmission transgénérationnelle de marques épigénétiques comme la méthylation de l'ADN, sensibles à l'environnement, est devenue le sujet d'un débat passionné aussi bien pour la communauté scientifique que pour le grand public (Saze, 2008; Nestler, 2012). Ce débat implique diverses questions : les variations phénotypiques, sont-elles « de quelque cause que ce soit », y compris épigénétique ? Si oui, sont-elles transmises « ordinairement à leur descendance » ? Enfin les variations épigénétiques induites par l'environnement sont-elles héréditaires pour « quelques générations tout au moins » ? La méthylation de l'ADN, modification covalente du support de l'hérédité, est l'un des candidats les plus logiques de la transmission épigénétique transgénérationnelle. Appuyant cette idée, le cas de la « mutation » pélorique chez *Linnea vulgaris* décrite il y a plus de 250 ans, transmissible au cours des générations et induisant un phénotype floral altéré (Figure III.23), a été montrée pour être en réalité une épimutation liée à la méthylation de l'ADN, constituant ainsi le premier exemple historique du rôle de cette marque épigénétique dans l'évolution d'une espèce (Cubas *et al.*, 1999).

Plusieurs études chez *Arabidopsis* ont été menées ces dernières années afin d'évaluer la variabilité de la méthylation de l'ADN et son hérédité au sein de populations naturelles ou artificielles (Vaughn *et al.*, 2007; Zhang *et al.*, 2008; Schmitz *et al.*, 2011; Becker *et al.*, 2011). Globalement, ces études apportent des conclusions similaires : le polymorphisme de méthylation se situe essentiellement au niveau des sites CG et au sein

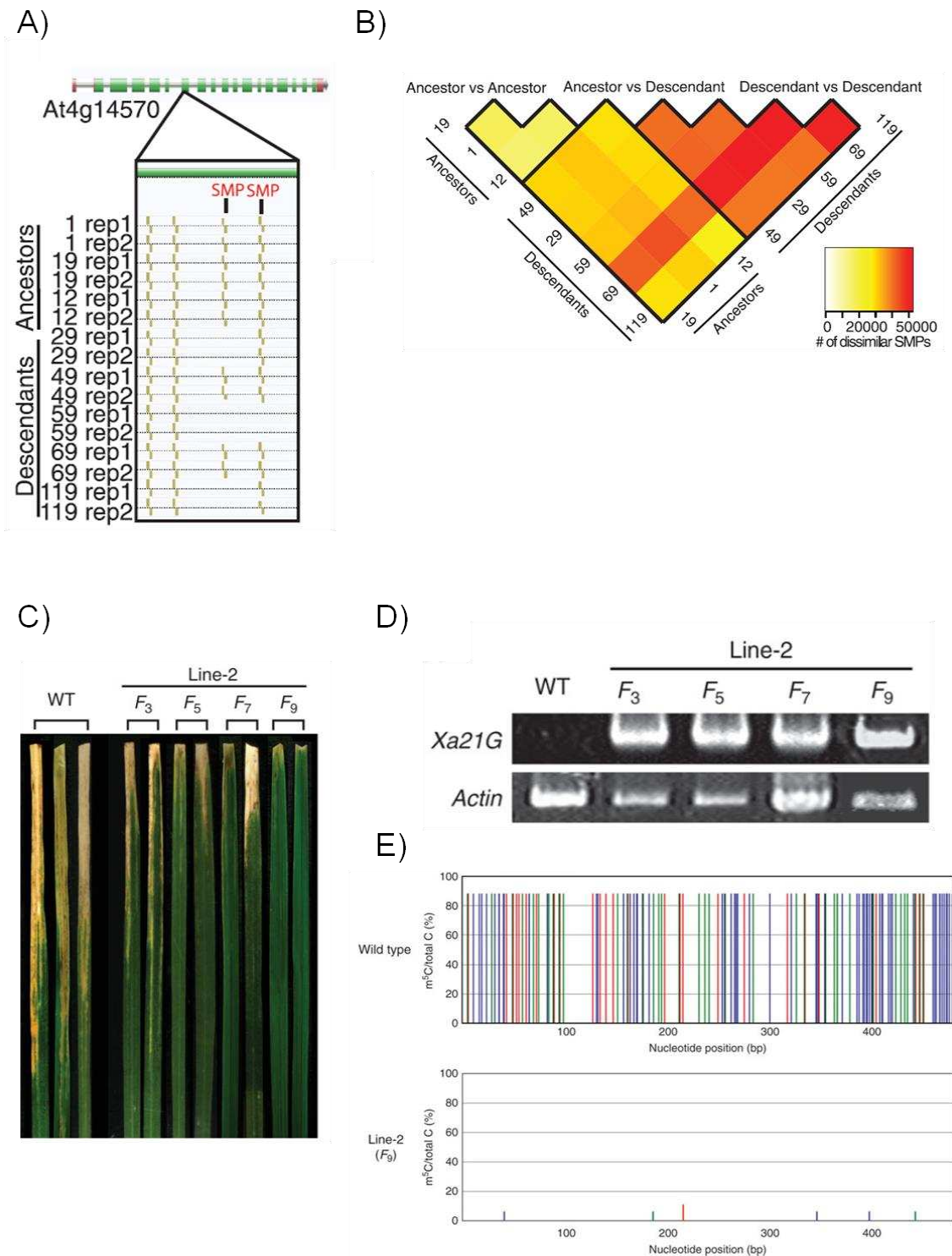


Figure III.24. Etudes des variations de méthylation de l'ADN et leur hérédité. A) Profil de méthylation d'un gène chez des parents et leurs descendants après 30 générations chez *Arabidopsis*. Les barres jaunes indiquent la méthylation en CG. SMP = *Single Methylation Polymorphism*. B) Représentations par *heat map* des différences de méthylation entre les lignées, exprimées en nombre de SMP (Schmitz *et al.*, 2011). C) Feuilles de riz soumises à *Xanthomonas* chez les plants contrôles et la lignée (Line-2) descendante des individus traités par 5-azadeoxycytidine. D) Expression de *Xa21G*, gène de résistance aux pathogènes, chez les mêmes individus. E) Profil de méthylation de *Xa21G* chez les contrôles et les descendants de la lignée Line-2 après 9 générations (Akimoto *et al.*, 2007).

des gènes. A l'inverse, la méthylation dans les éléments transposables et dans les sites CHG et CHH est peu variable parmi les populations étudiées. Ces études montrent également qu'un nombre important d'épimutations spontanées apparaissent au cours des générations. Notamment, après 30 générations, les descendants présentent une variabilité de méthylation plus importante entre eux que par rapport à leurs ancêtres (Figure III.24A-B). Le taux d'épimutation a ainsi été estimé à $4,46 \times 10^{-4}$ *Single Methylation Polymorphisms* (SMP) par site CG et par génération chez *Arabidopsis*, bien plus important que le taux de mutation génétique, estimé à 7×10^{-9} substitutions de base par site et par génération (Ossowski *et al.*, 2009). Enfin, certains gènes présentant des variations de méthylation étaient aussi différenciellement exprimés, même si ce cas n'était pas généralisé (Schmitz *et al.*, 2011; Becker *et al.*, 2011). Il semble donc que certaines épimutations soient « silencieuses ».

Les évidences les plus marquantes de la transmission transgénérationnelle de la méthylation de l'ADN proviennent d'études génétiques chez *Arabidopsis* (Reinders *et al.*, 2009; Johannes *et al.*, 2009). A partir d'un croisement entre un individu sauvage et un mutant altéré pour la méthylation de l'ADN (*met1* ou *ddm1*), ces deux études ont produit des lignées isogéniques différant uniquement par la répartition de la méthylation de l'ADN dans leur génome, appelées lignées recombinantes épigénétiques (epiRIL). Non seulement ces lignées étaient épigénétiquement stables sur au moins 8 générations, mais également une variation entre ces lignées a été observée pour des traits quantitatifs comme le temps de floraison ou la résistance à des pathogènes (Reinders *et al.*, 2009; Johannes *et al.*, 2009). Ainsi la variation dans ces traits phénotypiques pourrait être expliquée par la méthylation de l'ADN, même si une activité importante des transposons a été détectée dans ces lignées, induisant des variations génétiques (Reinders *et al.*, 2009).

Une autre question fondamentale est la transmission des variations de la méthylation de l'ADN induites par un stress environnemental et la transmission du caractère acquis de tolérance à ce stress. Les travaux abordant cette question sont nombreux (Boyko *et al.*, 2010; Lang-Mladek *et al.*, 2010; Bilichak *et al.*, 2012; Ou *et al.*, 2012). Leurs conclusions sont similaires, à savoir que les descendants de la génération ayant subi un stress sont tolérants au stress appliqué à leurs parents et présentent des modifications de profil de méthylation de l'ADN de certains gènes, parfois impliqués dans la réponse aux stress. Cependant, au bout de 2 ou 3 générations, soit ces études s'arrêtent, soit elles montrent que la tolérance acquise et les modifications épigénétiques s'estompent. Une étude très intéressante sur le riz montre néanmoins une transmission à plus long terme (Akimoto *et al.*, 2007). Dans ce travail, la descendance de parents soumis à un traitement hypométhylant (5-azadeoxycytidine) a été suivie pendant 9 générations. Au bout de 9 générations, certains gènes de résistance aux

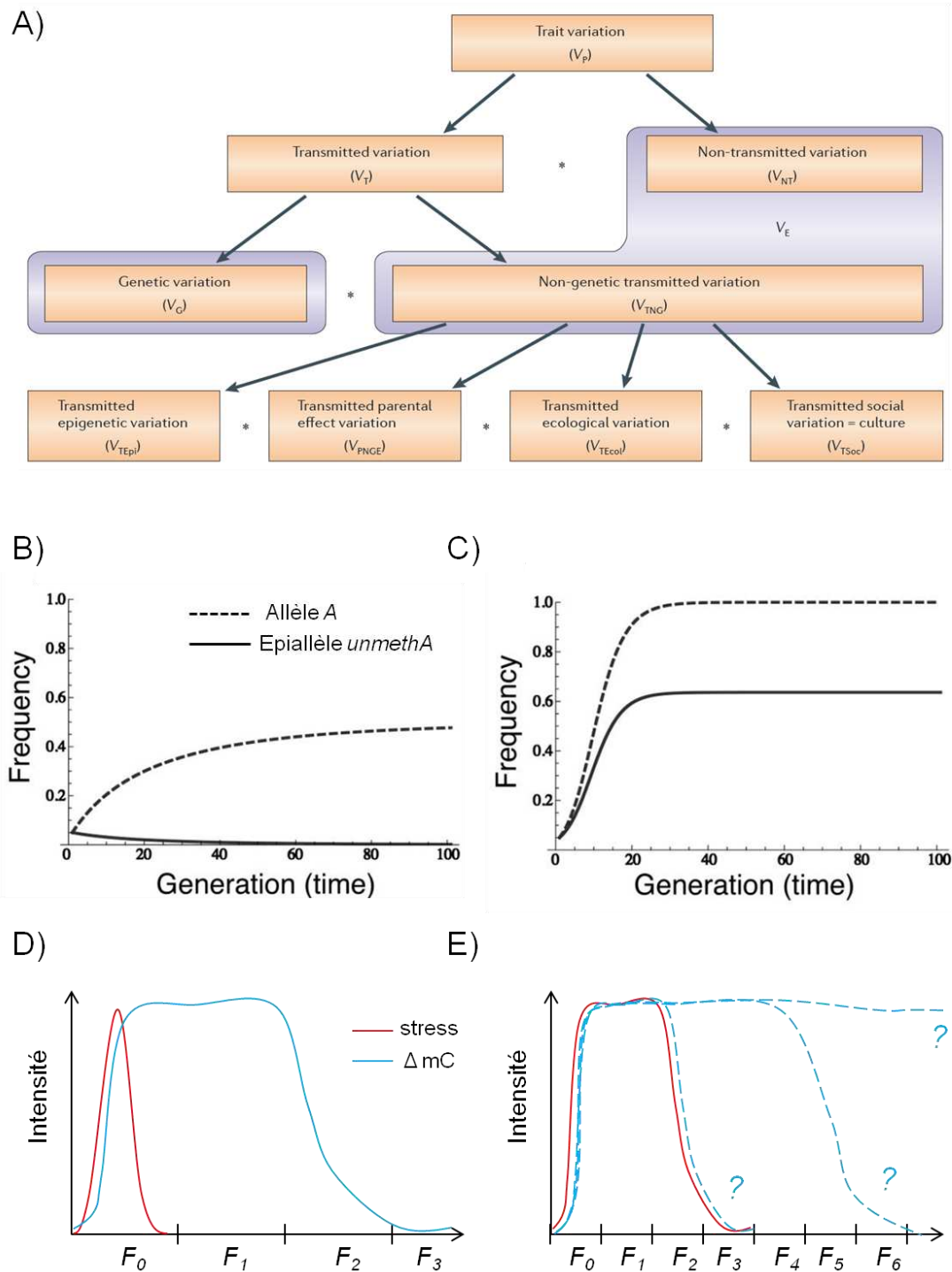


Figure III.25. Modèles d'évolution incluant l'hérédité épigénétique. A) Composantes génétique et non-génétique de la variation des traits phénotypiques et sa transmission (Danchin *et al.*, 2011). B) et C) Evolution de la fréquence d'un allèle A conférant un avantage évolutif, et de son épiallèle *unmethA* non méthylé. Dans ce modèle, méthylé = réprimé vs non méthylé = exprimé. En B) la probabilité que la méthylation de A soit transmise à chaque génération est forte, en C) elle est faible (Day & Bonduriansky, 2011). D) Stress subi par une génération F_0 induisant des variations de méthylation transmises aux générations suivantes, rapidement estompées. Inspiré de Boyko *et al.*, 2010; Lang-Mladek *et al.*, 2010; Ou *et al.*, 2012. E) Stress subi par plusieurs générations successives et possibles répercussions sur la stabilité générationnelle des variations.

pathogènes étaient toujours hypométhylés et activés, conférant une résistance accrue à une souche de *Xanthomonas* en comparaison avec les plants contrôles (Figure III.24C-E).

Contrairement aux mammifères où l'ensemble du génome est déméthylé successivement lors de la gamétogenèse et de l'embryogenèse, un tel phénomène global chez les plantes semble absent aussi bien dans les gamètes que dans l'embryon (Feng *et al.*, 2010b; Ibarra *et al.*, 2012; Calarco *et al.*, 2012). On pourrait donc s'attendre à ce que les profils de méthylation de l'ADN d'une génération G_0 soient transmis aux générations suivantes grâce aux mécanismes enzymatiques de propagation. L'observation de faibles stabilités générationnelles de changements de méthylation de l'ADN induits par l'environnement (Boyko *et al.*, 2010; Ou *et al.*, 2012) interrogent néanmoins sur les mécanismes définissant leur transmissibilité ou au contraire leur effacement (Saze, 2012). Certains considèrent même que les changements épigénétiques sont dans la grande majorité transitoires et ne sont que très rarement transmis à la descendance (Slatkin, 2009; Pecinka & Mittelsten Scheid, 2012), jouant par exemple un rôle mineur dans « l'hérédité manquante » (= facteurs hérités de risque de maladies génétiques complexes non identifiés par des études de génétique d'association).

Malgré cela, certains biologistes de l'évolution incluent désormais l'épigénétique comme facteur transmis de la variation de traits phénotypiques aux côtés de l'effet parental, de facteurs écologiques ou sociaux (Danchin *et al.*, 2011), rappelant que « nous sommes plus que la somme de nos gènes » (Figure III.25A). D'autres encore ont modélisé l'influence des interactions entre allèles génétiques, leurs épiallèles et leur risque de réversibilité sur l'évolution de la variabilité phénotypique, génétique et épigénétique d'une population (Day & Bonduriansky, 2011; Figure III.25B-C).

Ainsi, des preuves montrent que des variations phénotypiques impliquent la méthylation de l'ADN et si celles-ci sont héréditaires, ceci ouvre le champ à des perspectives évolutives ou à des applications pour la sélection agronomique. Cependant, même si certaines variations de la méthylation de l'ADN semblent transmissibles aux générations suivantes, il reste à comprendre dans quelle mesure elles le sont, en termes de stabilité et d'ampleur, afin de pouvoir statuer sur le rôle de l'épigénétique dans l'adaptation à l'environnement et l'évolution. Pour cela, la caractérisation des mécanismes définissant le « choix » de l'épiallèle qui sera transmis ou effacé lors du passage à la génération suivante est une étape clé. Enfin, si suite à stress subi par une génération, la transmission de variations épigénétiques et de caractères associés s'estompent rapidement, on peut s'interroger sur l'effet d'un même stress persistant au cours de plusieurs générations : la stabilité des marques épigénétiques serait-elle similaire ou renforcée (Figure III.25D-E) ?

Chapitre 1 : Méthylome dans un contexte de chromatine non condensée : étude sur le méristème apical caulinaire du peuplier

1. Présentation synthétique des travaux

Lorsque ce travail a été initié en 2008, la première étude de méthylation de l'ADN à l'échelle d'un génome entier (méthylome), effectuée sur la plante modèle *Arabidopsis*, était encore récente (Zhang *et al.*, 2006). Cette publication, avancée majeure pour l'épigénétique, montrait qu'une proportion non négligeable des gènes (33%) étaient méthylés dans leur corps mais surtout que cette méthylation de l'ADN affectait l'expression des gènes suivant une relation plus complexe que la simple inhibition d'expression. Cette relation allait devenir l'objet de nombreuses études mais peu de données étaient alors disponibles sur d'autres organismes pour évaluer la conservation ou non de ces observations chez d'autres plantes ou Eucaryotes. C'est le développement des techniques haut-débit (séquençage et microarrays) et leur accessibilité en termes de coût qui allaient procurer un outil puissant pour les études épigénétiques à l'échelle génomique ou épigénomique (Beck & Rakyar, 2008).

Chez le peuplier, notre modèle d'étude, aucune publication en épigénomique n'était disponible. Par ailleurs, le séquençage du génome de *P. trichocarpa* était encore récent (Tuskan *et al.*, 2006) et son assemblage n'en était qu'à sa première version. De ce fait, de nombreuses zones d'ombre étaient encore présentes, que ce soit sous forme de séquences isolées et non assemblées dans l'un des 19 groupes de liaison ou tout simplement sous forme de fragments dont la séquence était inconnue. Ces zones mal caractérisées correspondaient pour une grande partie à des séquences répétées (Tuskan *et al.*, 2006), loci qui posent toujours problème aux approches de séquençage de génomes, rendant compliquée une approche méthylome globale chez le peuplier.

De manière intéressante, des études menées chez l'Homme avaient montré l'utilité de la désoxyribonucléase I (DNase I) dans certaines approches génomiques. En effet, cette enzyme digère préférentiellement les zones appauvries en nucléosome et coupe difficilement les zones condensées de la chromatine. Il est alors possible d'isoler certaines séquences liées aux gènes, pauvres en nucléosomes, comme leurs éléments cis-régulateurs (fixation de facteurs de transcription), leur site d'initiation de la transcription (TSS) ou corps du gène, et ce au détriment des séquences répétées (Crawford *et al.*, 2004, 2006; Boyle *et al.*, 2008).

Dans ce contexte, une stratégie méthylome focalisée sur la partie non condensée de la chromatine, connue pour être riche en gènes et appauvrie en séquences répétées, avait été choisie afin de (i) disposer des premières données de méthylome chez le peuplier et initier des travaux en épigénétique sur cet arbre, (ii) mieux caractériser la méthylation de

l'ADN affectant les gènes et son implication dans le contrôle de leur expression dans un contexte de chromatine non condensée chez le peuplier et (iii) mener une étude de la méthylation de l'ADN simplifiée permettant d'éliminer les séquences répétées. Cette étude s'est focalisée sur le méristème apical caulinaire pour les raisons précisées dans l'introduction.

Ainsi, après isolement de la partie non condensée de la chromatine (préférentiellement digérée par la DNase I) du méristème apical chez *P. trichocarpa*, une immunoprécipitation de l'ADN méthylé (Methyl DNA ImmunoPrecipitation ou MeDIP) couplée à un séquençage haut-débit (Illumina/Solexa) a été effectuée (Figure IV.1). Le séquençage a généré de courtes séquences (appelées 36-mères) qui ont ensuite été repositionnées par bioinformatique sur le génome de *P. trichocarpa* v2.0. Cela a permis une cartographie de la méthylation de l'ADN de la partie non condensée de la chromatine du méristème apical caulinaire. En parallèle, à l'aide du séquençage bisulfite, la mesure de niveaux de méthylation contexte-spécifiques (CG, CHG et CHH) a été menée sur des loci contenus dans la fraction séquencée (appelée fraction DNase I-MeDIP-SEQ).

Les résultats (disponibles sur <http://www.popgenie.org/>) ont montré que la fraction DNase I-MeDIP-SEQ couvrait uniformément les 19 scaffolds de type chromosome chez *P. trichocarpa*. En revanche, la densité des 36-mères indiquait 4 scaffolds plus particulièrement ciblés (n°8, 9, 14 et 17). Bien que couvrant 1,9 % du génome du peuplier, la fraction DNase I-MeDIP-SEQ contenait des séquences appartenant à 74 % des modèles de gènes v2.0. Ainsi un fort enrichissement en gènes et particulièrement en exons a été démontré dans notre fraction. La méthylation du corps des gènes dans cette fraction variait en fonction des caractéristiques des gènes (taille, expression tissu-spécifique et redondance) et un appauvrissement en intron des gènes exprimés de manière tissu-spécifique a été constaté, suggérant un possible rôle régulateur de la méthylation de l'ADN dans ces zones. De plus et contrairement aux autres types de séquences répétées, les loci d'ADN ribosomique (ADNr) étaient particulièrement enrichis dans la fraction DNase I-MeDIP-SEQ expliquant la forte densité sur 4 scaffolds, ce qui suggérait une structure chromatinienne particulière de ces loci. Enfin, l'étude des deux gènes *MET1* (*ADN METHYLTRANSFERASE 1*) présentant 83% de similarité a montré une variation génotypique de la méthylation de l'ADN et de l'expression entre les trois génotypes étudiés. Ce résultat ouvre la voie à de prochaines études comparatives de variabilité génétique, épigénétique et transcriptomique.

En conclusion, cette étude a permis d'avancer dans notre compréhension de la méthylation des gènes dans le cadre de la chromatine non condensée dont la configuration

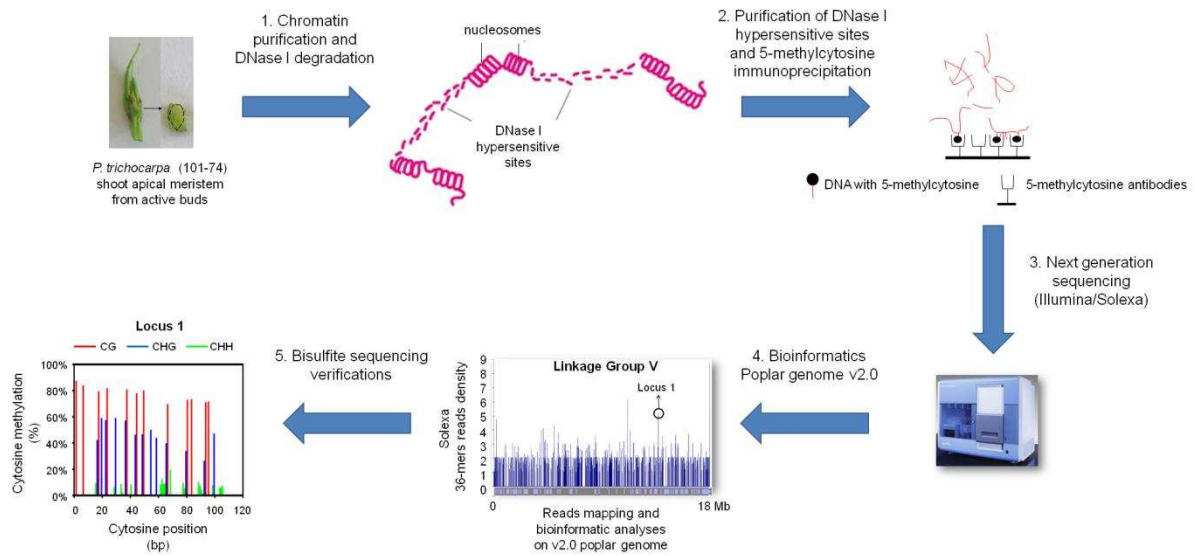


Figure IV.1. Présentation de la stratégie DNase I – MeDIP – SEQ. Adapté d'un poster présenté lors du colloque *Bioenergy Trees* (26^e symposium *New Phytologist*, Nancy 2011).

est dépendante du tissu étudié mais aussi de stimuli environnementaux (Valouev *et al.*, 2011; Zhang *et al.*, 2012a). Cette étude indique également un moyen simple et peu coûteux de cibler le méthylome des gènes, ce qui pourrait être particulièrement intéressant dans le cas d'organismes dont le génome est complexe ou non définitivement assemblé comme le pin ou le maïs (Morse *et al.*, 2009; Schnable *et al.*, 2009). Enfin, ce travail a fourni une base de données originale sur le méthylome du peuplier (<http://www.popgenie.org/>) utilisable par la communauté scientifique et comparable à des données récentes (Vining *et al.*, 2012 ; <http://poplar.cgrb.oregonstate.edu/cgi-bin/gbrowse/Populus/>).

Cet article a été publié dans *New Phytologist*.

2. Méthylome de la fraction non condensée de la chromatine dans le méristème apical caulinaire chez le peuplier (Lafon-Placette *et al.* 2013, *New Phytologist* : article 2)



Methylome of DNase I sensitive chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach revealing characteristics of gene-body DNA methylation in open chromatin state

Clément Lafon-Placette^{1,2}, Patricia Faivre-Rampant³, Alain Delaunay^{1,2}, Nathaniel Street⁴, Franck Brignolas^{1,2} and Stéphane Maury^{1,2}

¹Université d'Orléans, Faculté des Sciences, Laboratoire de Biologie des Ligneux et des Grandes Cultures (LBLGC), UPRES EA 1207, Orléans, 45067, France; ²INRA, USC1328 Arbres et Réponses aux Contraintes Hydriques et Environnementales (ARCHE), Orléans, 45067, France; ³INRA, UMR1165, UMR INRA/Université d'Evry, Unité de Recherche en Génomique Végétale, Centre de Recherche de Versailles-Grignon, Evry, 91057, France; ⁴Department of Plant Physiology, Umeå Plant Science Centre, Umeå University, Umeå, 901 87, Sweden

Summary

Author for correspondence:
Stéphane Maury
Tel: +33 (0)2 38 41 70 22
Email: stephane.maury@univ-orleans.fr

Received: 7 August 2012
Accepted: 2 October 2012

New Phytologist (2013) **197**: 416–430
doi: 10.1111/nph.12026

Key words: DNA methylation, genes, methylome, open chromatin, *Populus trichocarpa* (black cottonwood), shoot apical meristem.

- DNA methylation is involved in the control of plant development and adaptation to the environment through modifications of chromatin compaction and gene expression. In poplar (*Populus trichocarpa*), a perennial plant, variations in DNA methylation have been reported between genotypes and tissues or in response to drought. Nevertheless, the relationships between gene-body DNA methylation, gene expression and chromatin compaction still need clarification.
- Here, DNA methylation was mapped in the noncondensed chromatin fraction from *P. trichocarpa* shoot apical meristematic cells, the center of plant morphogenesis, where DNA methylation variations could influence the developmental trajectory. DNase I was used to isolate the noncondensed chromatin fraction. Methylated sequences were immunoprecipitated, sequenced using Illumina/Solexa technology and mapped on the v2.0 poplar genome. Bisulfite sequencing of candidate sequences was used to confirm mapping data and to assess cytosine contexts and methylation levels.
- While the methylated DNase I hypersensitive site fraction covered 1.9% of the poplar genome, it contained sequences corresponding to 74% of poplar gene models, mostly exons. The level and cytosine context of gene-body DNA methylation varied with the structural characteristics of the genes.
- Taken together, our data show that DNA methylation is widespread and variable among genes in open chromatin of meristematic cells, in agreement with a role in their developmental trajectory.

Introduction

In plants, epigenetic reprogramming occurs at key developmental stages and in response to environmental stimuli (Feng *et al.*, 2010b). These epigenomic events play important roles in genome protection, control of gene expression and inheritance via chromatin structural reworking (Teixeira & Colot, 2010). Among epigenetic marks, DNA methylation shows high stability through mitosis and meiosis and has been thoroughly investigated. In plants, methylated cytosines are found mainly in CG dinucleotides and to a lesser extent in CHG and asymmetric CHH contexts (where H is A, T, or C). Non-CG methylation is specific to plants and fungi, with the exception of mammalian embryonic stem cells (Suzuki & Bird, 2008; Lister *et al.*, 2009). The enzymatic machinery of methylation (Goll & Bestor, 2005;

Teixeira & Colot, 2010) and demethylation (Penterman *et al.*, 2007), driven by both developmental and environmental stimuli, shapes a methylome that is the set of nucleic acid methylation modifications in an organism's genome or in a particular cell.

The first methylome characterization was performed in *Arabidopsis* (Zhang *et al.*, 2006) and revealed that DNA methylation is concentrated in heterochromatin and repeats but is also present on 30% of genes. Several other approaches have confirmed this distribution across the genomes of other species, including animals, and raised the problem of repeated sequence data processing, as a consequence of both their repetitive nature and their high DNA methylation (Beck & Rakan, 2008; Cokus *et al.*, 2008). Indeed, repeats are affected by high CG and non-CG methylation, while the gene body is only methylated in CG

(Cokus *et al.*, 2008). Despite its high conservation in many organisms, the functional role of gene-body methylation needs to be clarified.

DNase I is a pancreatic enzyme that preferentially digests with no sequence-based bias (Crawford *et al.*, 2006) nucleosome-depleted DNA, while tightly packaged chromatin is more resistant to cleavage (Crawford *et al.*, 2004). In plants, genome-scale assays revealed tissue-specific DNase hypersensitive sites in promoters and potential regulatory elements, but also within genes, allowing the preferential isolation of genes and the elimination of repeated regions (Zhang *et al.*, 2012a,b). In addition, while DNA methylation and nucleosomes co-localize on exons, some nucleosome-depleted regions are also methylated (Chodavarapu *et al.*, 2010; Pecinka *et al.*, 2010).

Draft sequencing of the poplar (*Populus trichocarpa*; western poplar) genome has led to the construction of 19 scaffolds which contain 370 Mb of sequence (out of 403 Mb) and the identification of >40 000 protein-coding transcripts (<http://www.phytozome.net/poplar>; Tuskan *et al.*, 2006). As a consequence of its genome sequence becoming available, along with various molecular tools, and its substantial genetic and phenotypic variation, *Populus* spp. has become widely used as a model tree (Tuskan *et al.*, 2006; Jansson & Douglas, 2007). As perennial plants with long life-spans and generation times, trees have to face and acclimate to changing environments and are therefore models of interest for epigenetic studies (Hamanishi & Campbell, 2011). Interestingly, genetic variability of global DNA methylation, ranging from 4% to 12%, has been observed in *Populus* × *euramericana* hybrid shoot apices (Gourcilleau *et al.*, 2010) and has been positively correlated with biomass production. Furthermore, differences in global DNA methylation paralleled transcriptome level trends in leaves of poplar genotypes during drought, suggesting an epigenomic basis for the clone history-dependent transcriptome divergence (Raj *et al.*, 2011). Recently, drafts of genome-wide *P. trichocarpa* methylome revealed several particularities of this model, such as a high CHG methylation (Feng *et al.*, 2010a), as well as a negative correlation between promoter and, particularly, gene-body DNA methylation and gene expression level (Vining *et al.*, 2012). However, the relationships between gene-body DNA methylation and tissue-specific gene expression remain to be clarified.

In this context, our aim was to characterize the methylome of noncondensed chromatin to investigate gene-body DNA methylation characteristics in an open chromatin state. Moreover, the elimination of heterochromatic repeated loci would reduce the complexity of the whole-genome analysis. For this purpose, the chromatin fraction hypersensitive to DNase I was isolated from *P. trichocarpa* shoot apical meristem (SAM) cells and was used in MethylDNA Immunoprecipitation followed by Illumina/Solexa (Fasteris, Plan-les-Ouates, Switzerland) sequencing (MeDIP-SEQ). The mapping on the *P. trichocarpa* genome was confirmed by bisulfite sequencing of candidate mapped loci to assess the cytosine contexts and the corresponding methylation levels. Taken together, our results showed that methylated DNase I hypersensitive sites covered 2% of the genome but

were regularly distributed along poplar scaffolds and strongly enriched in genes, particularly in exons. Furthermore, gene-body DNA methylation in the noncondensed chromatin fraction was dependent on structural gene characteristics, redundancy in the genome, tissue-specific pattern of expression and among poplar genotypes.

Materials and Methods

Plant material and growth conditions

Unrooted dormant stem cuttings of *Populus trichocarpa* (Torr. & Gray) clone 101-74 (used for the DNase-MeDIP-SEQ procedure and subsequent manipulations) and *Populus* hybrid (*Populus deltoides* × *Populus nigra*) clones Carpaccio and Soligo (only used for RT-PCR and bisulfite sequencing analysis of the *PtMET1a* (POPTR_0019s00240) and *PtMET1b* (POPTR_0004s14140) genes) were obtained from National Institute for Agronomic Research (INRA) nurseries (Orléans, France in 2008 and Nancy, France in 2006, respectively). Ten-year-old mother plants are cut off each spring. Cuttings were 22 cm long and 0.5–2 cm wide and were harvested in February from plants grown in INRA nurseries. Unrooted cuttings (20 for 101-74 and 12 for each hybrid clone) were then stored for 1 month in a cold (2°C) dark room. Before potting, cuttings were soaked in water in the vertical position to encourage rooting. For Carpaccio and Soligo cuttings, growth conditions have been previously described (Gourcilleau *et al.*, 2010). Rooted cuttings were planted in 20-l pots containing a mixture of peat and sand (50/50, v/v) amended with magnesian chalk (60 g per 100 l) and fertilized with a slow-release fertilizer (Nutricote T100; 13/13/13/2, N/P/K/Mg + trace elements; FERTIL SA, Boulogne Billancourt, France) according to (Gourcilleau *et al.*, 2010). All cuttings were placed in a glasshouse where the temperature was maintained in the range 15–27°C and where they were exposed to natural daylight (ranging from 350 to 900 $\mu\text{mol m}^{-2} \text{s}^{-1}$). Plants were manually watered every second day for 5 wk. Active buds from 12 to 20 individuals of each genotype were collected, and leaves were removed to isolate the shoot apex (Fig. 1a). Shoot apices were then cut in half with a scalpel and cleared of all visible differentiated tissues under a binocular magnifying glass. Samples that contained a shoot apical meristem (Fig. 1a) were then frozen in liquid nitrogen and stored at –80°C.

Chromatin extraction and DNase I hypersensitive sites

Shoot apical meristems from 10 individuals (c. 500 mg fresh weight per extraction) of the 101-74 *P. trichocarpa* genotype were ground to a fine powder in liquid N₂ and used for chromatin extraction. Six chromatin extraction protocols (Galbraith *et al.*, 1983; Steinmüller & Apel, 1986; Lechner *et al.*, 2000; Causevic *et al.*, 2006; Loureiro *et al.*, 2007) were tested in parallel (Supporting Information Table S1). Chemicals were purchased from Sigma-Aldrich (Saint-Quentin Fallavier, France). After extraction, the chromatin pellet was solubilized in DNase I (Roche, Meylan, France) digestion buffer composed of 10 mM

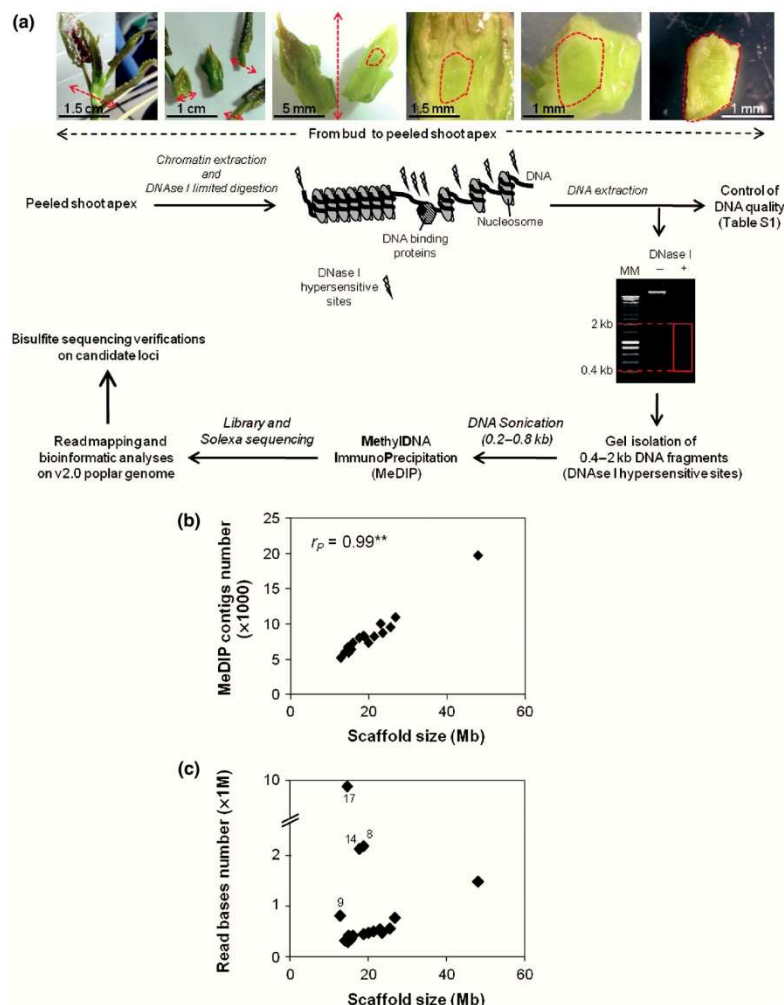


Fig. 1 The DNase I-MethylDNA Immuno-precipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) approach. (a) Description of the DNase I-MeDIP-SEQ approach. The pictures show the shoot apical meristem isolation procedure, from left to right. Arrows with red dotted lines show sections of vegetal material; dotted line surrounds the meristem used for the DNase I-MeDIP-SEQ procedure. White flashes represent DNase I action on hypersensitive sites, that is, nucleosome-free DNA. (b, c) Distribution of sequences from the DNase I-MeDIP-SEQ fraction on *Populus trichocarpa* scaffolds. Number of (b) MeDIP contigs and (c) number of 36-mers reads are shown for each scaffold (v2 genome version) depending on its size in Mb. MeDIP contigs are contiguous loci where one or several 36-mer reads were mapped. A significant Pearson linear correlation ($P < 0.01$) between the number of MeDIP contigs and scaffold size is indicated by **. 8, 9, 14 and 17 are scaffold numbers.

TrisHCl (pH 7.8), 250 mM saccharose, 5 mM $MgCl_2$, 5 mM KCl and 0.1% (v/v) 2-mercaptoethanol (Crawford *et al.*, 2006). Different concentrations of DNase I, ranging from 0.1 to 1 units ml^{-1} , and durations of incubation, from 0 to 10 min, were tested to achieve limited chromatin digestion at hypersensitive sites. The reaction was stopped by adding 0.5 volume of a solution containing 2% (w/v) NaDodSO₄ and 80 mM EDTA equilibrated at pH 8.0 and 0.5 volume of 4 M NaCl. Genomic DNA was isolated after overnight proteinase K incubation at 37°C, phenol/chloroform/isoamyl alcohol (25 : 24 : 1) (v/v/v) extraction and ethanol precipitation. The DNase I hypersensitive fraction, ranging from 0.4 to 2.0 kb, was recovered after electrophoresis on a low melting point 0.8% agarose gel and ethidium bromide staining using the QIAEX II Gel extraction kit (Qiagen, Courtaboeuf, France) according to the manufacturer's recommendations.

Methyl DNA ImmunoPrecipitation (MeDIP)

Purified DNA (0.4–2 kb) extracted from the DNase I hypersensitive chromatin fraction was randomly nebulized using a Nebulizer kit (Invitrogen, Cergy-Pontoise, France) at a pressure of 30 psi for 2 min to obtain fragments ranging from 0.2 to 1 kb. Rabbit polyclonal anti-5mC antibodies (250 µg; Maine Biotechnology Service, Portland, Maine, USA) were cross-linked to 300 µl of a batch of *N*-hydroxysuccinimide (NHS)-activated Sepharose matrix (NHS HP SpinTrap kit; GE HealthCare, Aulnay-sous-Bois, France) in accordance with the manufacturer's instructions (www.gelifesciences.com/trap). Matrix–antibody coupling was estimated using PD-10 desalting columns (GE HealthCare) according to the manufacturer's recommendations and was estimated at >70%. The nebulized DNA (2.5 µg) was solubilized in a 10 mM Na-phosphate buffer (pH 7.0)

containing 140 mM NaCl and 0.05% (v/v) Triton-X (IP buffer 1X) and incubated with the affinity matrix under regular slow mixing for 1 h at room temperature and for a further 3 h at 4°C. The batch was washed three times with the IP buffer 1X before resuspension in a proteinase K digestion solution (50 mM Tris at pH 8.0, 10 mM EDTA, 0.5% (w/v) SDS and 20 mg ml⁻¹ proteinase K) at 50°C for 3 h. DNA was then isolated following the classical procedure of phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. This purification was followed by a second step using a MinElute PCR purification kit (Qiagen). DNA bound (methylated DNA) and unbound (1 : 60 ratio) to the affinity matrix was recovered.

High-throughput sequencing of the MeDIP fraction (MeDIP-SEQ) and controls

The purified poplar methyl-DNA immunoprecipitated fraction was physically fragmented by nebulization into 50–200-bp fragments, end-repaired, ligated with bar-coded adapters, PCR-amplified, purified on an agarose gel and used to generate a DNA colony template library according to the Fasteris procedure (Fasteris, Plan-les-Ouates, Switzerland). Illumina/Solexa sequencing (Bennett, 2004) of the fraction was performed on a Genome Analyzer GAII (Illumina, San Diego, CA, USA) using a Chrysalis 36 cycles v 3.0 sequencing kit and a GAPipeline 1.3.2 data analysis pipeline (Table S2). Illumina quality control was performed by direct capillary sequencing of 30 sequences from the (Illumina, USA) preliminary DNA clone library and comparison with the *P. trichocarpa* genome (Blast). The methylation status of two of these sequences was confirmed by bisulfite sequencing (see primers in Table S3). Semiquantitative PCR was performed on random sonicated DNA and the DNA used for the Illumina sequencing (the DNase I-MeDIP-SEQ fraction) to amplify two sequences: one covered by the DNase I-MeDIP Illumina sequencing reads (*POPTR_0006s21000*; read depth = 3.4) and one not covered at all (*POPTR_0011s13770*). Random Illumina sequencing of *P. trichocarpa* genome was performed with genomic DNA extracted from young leaves of the *P. trichocarpa* 101-74 genotype on the GAII platform (Centre National de Génotypage, Evry, France; P. Faivre-Rampant *et al.*, unpublished). 32 312 796 reads were generated, achieving a mean coverage of 16.44, covering 93.5% of the reference genome by at least one read.

Bioinformatics analyses

Short reads mapping Mapping of the 36-mer reads from Illumina/Solexa sequencing or mapping after *de novo* assembly with the Exact DE Novo Assembler (EDENA, v2.1.1; Hernandez *et al.*, 2008) was performed on v2.0 of the 403-Mb *P. trichocarpa* unmasked genome (<http://www.phytozome.net/poplar.php>; Tuskan *et al.*, 2006) using Mapping and Assembly with Qualities (MAQ, v0.7.1; <http://maq.sourceforge.net/>) software and confirmed using ELAND v2.0 (Efficient Local Alignment of Nucleotide Data, casava v1.6.0, Illumina, USA). Two mismatches per 36-mer read were allowed in the first 24 bases on

the reference sequence for MAQ and in the first 32 bases for ELAND. The sequencing data are available in the NCBI Sequence Read Archive (SRA) database (SRA number: SRA050249) and on the Popgenie website (<http://www.popgenie.org/>).

Sequence annotation Analysis of the content of repeats was performed with RepeatMasker (<http://www.repeatmasker.org/>) using *P. trichocarpa* and *Arabidopsis thaliana* (data not shown) as DNA sources. A few repeated sequences were validated as transposable elements using PlotRep (<http://repeats.abc.hu/cgi-bin/plotrep.pl>) and RepPop (<http://csbl.bmb.uga.edu/~ffzhou/RepPop/>) with default parameters. Identification of transposase was performed using FGENESH software (<http://linux1.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind>) and confirmed by TBlastX analysis (*e*-value < 10⁻³⁰) on a flowering plant transcripts database. Identification of potential small RNAs was carried out with psRNATarget (<http://plantgrn.noble.org/psRNATarget/>) on mRNA sequence with a maximum expectation of 3. Gene annotation from Phytozome poplar genome v2.0 (<http://www.phytozome.net/poplar.php>) was used. Gene ontology class enrichment was produced with AGRIGO software (<http://bioinfo.cau.edu.cn/agriGO/analysis.php>). Nucleosome occupancy was a computational prediction produced by SEGAL LAB software version 3.0 (http://genie.weizmann.ac.il/software/nucleo_prediction.html). 'Low redundancy' and 'high redundancy' coding sequences were determined according to their *e*-values after a blastN comparison of 500 nucleotides centered on each sequence with *P. trichocarpa* v2.0 genome. If there were more than two results with an *e*-value under 10⁻⁵⁰, sequences were considered as highly redundant and otherwise they were classified as lowly redundant.

Expression data Shoot apical meristem microarray expression data were obtained from the Popgenie database, using the tissue comparison experiment data (Street *et al.*, 2008), available in the Umeå Plant Science Centre database (<http://www.upsbase.db.umu.se/>) with the ref number UMA-0020, sample L1 (apical region). The data used were relative expression compared with other tissues. Digital northern heat maps representing the library distribution of expressed sequence tags (ESTs) representing gene models within PopulusDB were produced using the PopGenIE DigitalNorthern tool (<http://popgenie.org/tool/digitalnorthern>).

Bisulfite sequencing and methyl-sensitive PCR amplification

Genomic DNA was extracted from poplar shoot apical meristems using an RNase A digestion, phenol/chloroform extraction and ethanol precipitation procedure (Causevic *et al.*, 2005). Controls for bisulfite conversion efficiency and amplification and the primer design, PCR and sequencing procedures have been recently described (Trap-Gentil *et al.*, 2011). Quantification of cytosine methylation percentages for each cytosine position was obtained from three biological and two technical replicates for each gene and each genotype. Methyl-sensitive PCR (MS-PCR) using *McrBC* (New England BioLabs, Ipswich, UK) was

performed to verify the bisulfite sequencing data according to the detailed procedure of Trap-Gentil *et al.* (2011). Primers used for bisulfite sequencing and MS-PCR are listed in Table S3. Three biological (SD = 3.72%) and three technical (SD = 0.99%) replicates were performed for each sequence.

RNA extraction and semiquantitative RT-PCR

Total RNAs were isolated from poplar shoot apical meristems using Nucleospin[®] RNA Plant (Macherey-Nagel, Hoerd, France). Approximately 500 ng of total RNA was reverse-transcribed using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Courtaboeuf, France). Constitutively expressed genes encoding, respectively, a tubulin and an actin, *POPTR_0001s25410* and *POPTR_0019s02630*, were used as internal standards to normalize the amount of mRNA in the PCR reaction according to Trap-Gentil *et al.* (2011). *POPTR_0001s35160* amplification was also used to confirm that there was no genomic DNA contamination, as primers were designed based on a coding sequence containing a short intron (209 bp). Furthermore, a control without reverse transcriptase during the cDNA synthesis was performed to confirm the absence of genomic DNA in all the total RNA preparations. Primers for PCR were designed using EPRIMER3 software (<http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=eprimer3>) and are listed in Table S3. The number of PCR cycles was adjusted to avoid reaching saturation. PCR products were visualized on a 8% polyacrylamide gel after ethidium bromide staining. Three biological and two technical replicates were performed for each gene and genotype.

Statistical analyses

Statistical analyses were performed using the SPSS statistical software package (SPSS version 11.0.1 PC; SPSS, Chicago, IL, USA). Means are expressed with their standard errors and compared by analysis of variance (one-way ANOVA; general linear model (GLM) procedure). Fraction effects on mapping distribution were evaluated using the χ^2 homogeneity test. Relationships between pairs of continuous variables were analyzed by linear regression analysis (Pearson's correlation coefficients (r)). Statistical tests were considered significant at: *, $P < 0.05$; **, $P < 0.01$; or ***, $P < 0.001$.

Results

MeDIP-sequencing of the DNase I hypersensitive chromatin fraction extracted from shoot apical meristem cells

Shoot apical meristems from *P. trichocarpa* genotype 101-74 were isolated from all differentiated tissues (Fig. 1a). Chromatin extraction was optimized by testing six protocols to achieve efficient DNase I digestibility and high quality (purity, concentration and integrity) of the extracted DNA (Table S1 and the Materials and Methods section). The most efficient and reliable

protocol (Causevic *et al.*, 2006; Table S1) was further used to isolate DNase I hypersensitive sites ranging from 0.4 to 2.0 kb (c. 30% of total DNA according to gel scan analysis), representing 2.5 μ g of DNA, which was then gel-extracted, sonicated (0.2–0.8 kb) and immunoprecipitated with 5-methylcytosine antibodies (MeDIP approach; see the Materials and Methods section for conditions). Purified methylated DNA (10 ng), representing 0.4% of the initial amount, was then sequenced using Illumina/Solexa technology and generated over 8 million reads of 36 bases (MeDIP reads). The sequencing data are available in the NCBI SRA database (SRA number: SRA050249) and on the Popgenie website (<http://www.popgenie.org/>; Fig. 2c).

Mapping of DNase I-MeDIP-SEQ reads on the *P. trichocarpa* genome assembly

After filtration, 180 650 MeDIP contigs (729 378 36-mers total; 8.7% of total reads), composed of 1–15 991 MeDIP reads, with a mean value of four reads, were mapped on the *P. trichocarpa* v2.0 genome assembly (Table S2). MeDIP contigs ranging from 36 to 5078 bp covered 1.9% of the *P. trichocarpa* v2.0 genome assembly (7 834 392 bp). Their read depth, defined as the number of MeDIP reads divided by the length of the covered genomic region, varied from 1 to 459, with a mean value of 3.35.

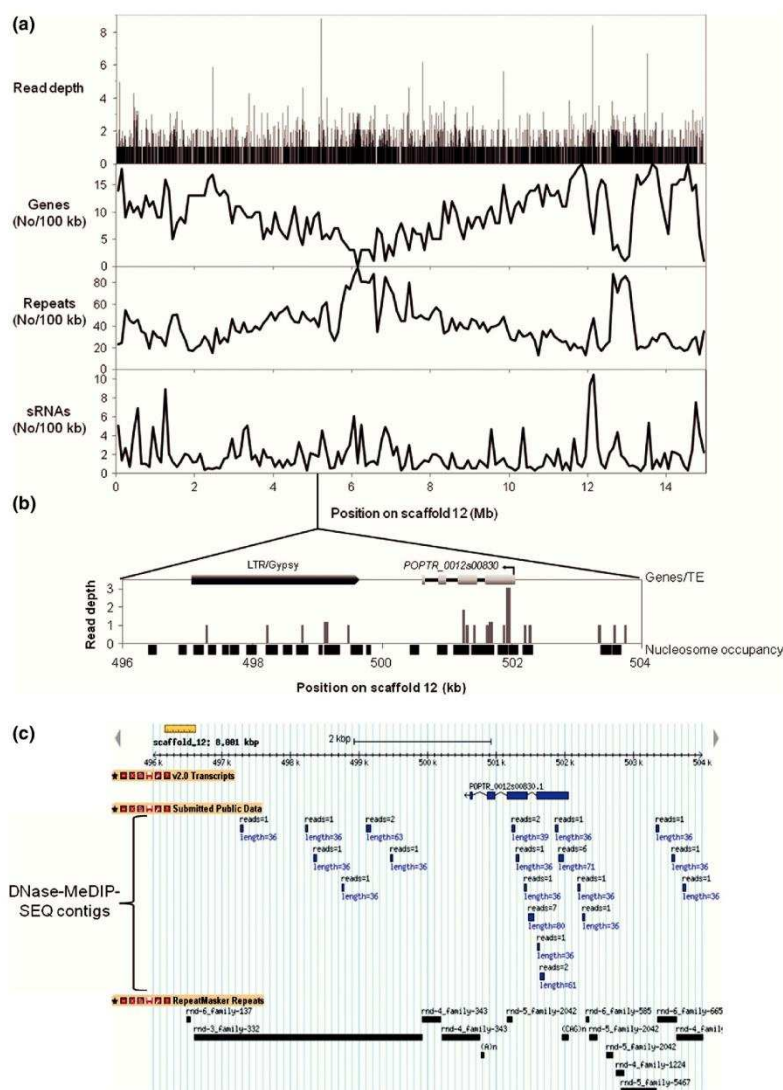
Validation of MeDIP-SEQ and mapping

Immunoprecipitation was assessed by semiquantitative PCR on two loci methylated or not according to prior bisulfite sequencing and showed amplification in the MeDIP fraction only for the methylated sequence (*POPTR_0006s21000*; Fig. S1a) but not for the nonmethylated sequence (*PrTUB2*, *POPTR_0001s25410*; Trap-Gentil *et al.*, 2011; Fig. S2a). Bisulfite sequencing and MS-PCR (Fig. S2b, Table S3) of 34 regions (13 586 bp in total) randomly selected among genic, intergenic and repeated sequences confirmed that all sequences mapped by MeDIP reads were methylated. In addition, a positive linear correlation (Pearson coefficient, $r_p = 0.77$; $P < 0.05$) was established between the density of MeDIP reads and the methylcytosine percentage obtained by bisulfite sequencing. Finally, the mapping of MeDIP reads showed differential distributions on the *P. trichocarpa* genome compared with a random sequencing (Illumina) of the total genome, as evaluated by the χ^2 homogeneity test, for example throughout the first 1 Mb of scaffold 1 ($\chi^2 = 17.78^*$).

Mapping of DNase I-MeDIP-SEQ reads at scaffold space level

A linear positive correlation ($r_p = 0.99$; $P < 0.01$) was established between the number of mapped MeDIP contigs and the scaffold size of *P. trichocarpa* v2.0 (Fig. 1b). However, no correlation could be detected between the number of mapped MeDIP reads by scaffold and their sizes (Fig. 1b). Indeed, scaffolds 8, 9, 14 and 17 showed a high number of MeDIP reads with respect to their sizes (Fig. 1c). These reads mapped on repeated sequences

Fig. 2 Scaffold distribution of the DNase I-MethylDNA Immunoprecipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) fraction. (a) Distribution of MeDIP reads, genes, repeats and sRNA densities along scaffold 12 (v2.0 genome version). Repeats, gene densities and sRNA mapping were obtained from the Popgenie database (<http://www.popgenie.org/popgenie1/>; Klevebring *et al.*, 2009). (b) Locus on scaffold 12 (v2.0 genome) with a Long Terminal Repeat (LTR) and a gene model (POPTR_0012a00830). Exons and introns of the gene model are represented by light gray rectangles and black lines, respectively. The arrow indicates gene orientation. MeDIP contigs are represented by gray rectangles whose width is the covered region and height is read depth (read depth = number of reads/peak surface). Black rectangles represent nucleosome occupancy > 0.5. (c) Popgenie Gbrowser view of the locus from scaffold 12 shown in (b). The first track is v2.0 transcripts, the second is MeDIP contigs mapping with their number of MeDIP reads and length, and the last track is repeated sequences as detected by RepeatMasker.



(Fig. S3a) and more precisely on rDNA sequences (18S/26S and 5S rDNA), which were particularly abundant in these four scaffolds (Fig. S3b).

Heterogeneity of the DNase-MeDIP read distribution was observed all along each scaffold and within them (Fig. 2). A linear correlation was established between the distribution of DNase-MeDIP reads and sRNA along scaffolds ($r_p = 0.501$; $P < 0.001$ on scaffold 12), while no relationships could be established with the densities of repeats or genes (Fig. 2a). The ratio between the densities of DNase I-MeDIP reads and the whole-genome MeDIP reads (Vining *et al.*, 2012; <http://http:poplar.cgrb.oregonstate.edu>) was calculated and compared with the repeat density for scaffold 12 (Fig. S3c). A significant negative correlation was detected ($r_s = -0.191$; $P < 0.05$) in relation to the very

low density of DNase I-MeDIP reads compared with whole-genome MeDIP reads in repeat-rich regions.

Mapping of DNase I-MeDIP-SEQ reads at the coding sequence space level

Mapped MeDIP reads were distributed between intergenic regions (61%), genes (26%) and repeated sequences (13%; Fig. 3a). A depletion of *c.* 30% of repeated sequence was observed compare with the whole genome. Most of the repeats corresponded to retro-elements (58%) and DNA transposons (29%; Fig. 3b). It was noteworthy that 74% of v2.0 model genes (41 377 models) were mapped by MeDIP reads in the body region (74%; Fig. 3c) and/or in the ± 2 kb 5' and 3'

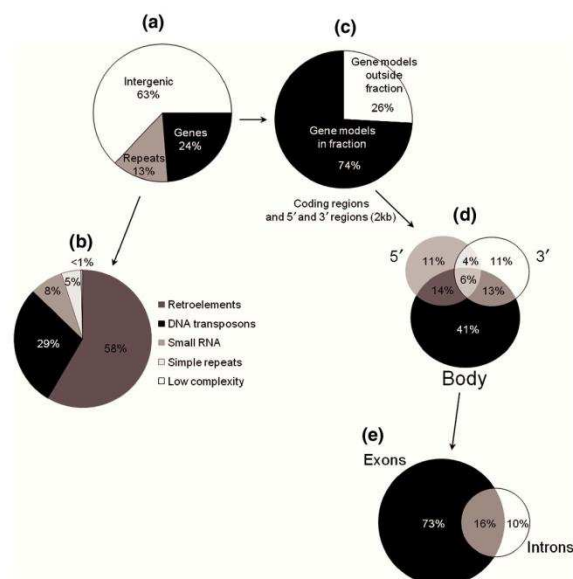


Fig. 3 Nature of sequences in the DNase I-MethylDNA Immunoprecipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) fraction. (a) Pie chart representing proportion of regions covering v2.0 gene models, repeats or intergenic loci in the DNase I-MeDIP-SEQ fraction. (b) Pie chart representing proportion of different repeat classes in the DNase I-MeDIP-SEQ fraction. (c) Pie chart representing proportion of v2.0 gene models contained or not in the DNase I-MeDIP-SEQ fraction. Genes with at least one mapped read in one of their parts are considered as contained in the DNase I-MeDIP-SEQ fraction. Body, 2-kb upstream and 2-kb downstream regions are taken as parts of the gene. (d) Venn diagram with number of v2.0 genes whose body, 2-kb upstream and 2-kb downstream parts were mapped with at least one read. (e) Venn diagram with number of v2.0 genes whose exon (s) or intron(s) was mapped with at least one read.

flanking regions (59%; Fig. 3d). The gene-body regions were mostly mapped on exons (89%; Fig. 3e) close to the beginning or the end of predicted nucleosome sites (see one example in Fig. 2b).

The distribution of MeDIP reads along gene models showed a peak in the body part compared with flanking regions, with strong variations at the borders of the genes (Fig. 4). This distribution was affected by gene size ($\chi^2 = 6.68$, $P < 0.01$; Fig. 4a). Thus, long genes (over 2 or 5 kb), which were found in higher proportions in the DNase I-MeDIP-SEQ fraction than in the whole genome (Fig. S4), exhibited fewer MeDIP reads in the central part of the gene body and more in the flanking regions than small genes (under 2 kb; Fig. 4a).

Concerning rDNA clusters, only some 5S clusters were mapped (data not shown), while all major 18S-5.8S-26S clusters (on scaffolds 14 and 17) and minor dispersed 18S/26S loci (on scaffolds 8 and 9) were covered by DNase I-MeDIP-SEQ reads (Fig. 5a). In addition, more MeDIP reads were mapped between 18S, 5.8S and 26S units than inside them (Fig. 5a). Bisulfite sequencing showed that this distribution is related to the density of methylation of these sequences (Fig. 5b). Variations in the

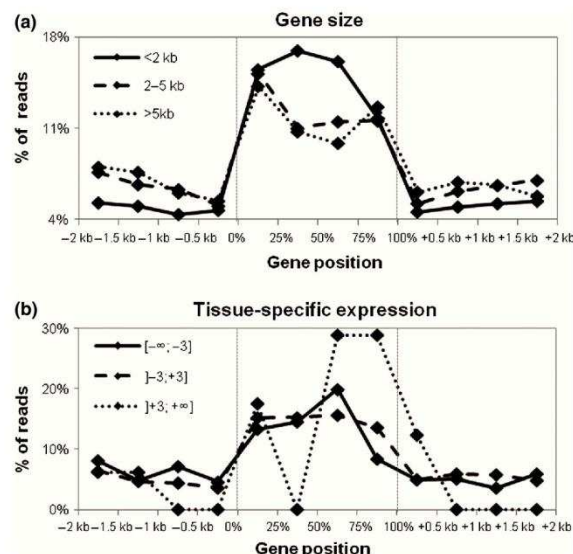


Fig. 4 Sequencing read profile along gene models (from 2 kb upstream to 2 kb downstream). For each gene part, read intensity is shown as a percentage of total reads from 2 kb upstream to 2 kb downstream. (a) Genes are separated according to their size: [0 kb, 2 kb], continuous line; [2 kb, 5 kb], dashed line; [5 kb, +∞], dotted line. The χ^2 value was calculated on the read distribution inside the gene body and was significant at $P < 0.01$. (b) Genes are separated according to their expression log value in the apical region (gene expression microarray experiments on tissue comparison; Street *et al.*, 2008): $[-\infty, -3]$, continuous line; $[-3, +3]$, dashed line; $[+3, +\infty]$, dotted line.

levels of methylation of non-CG sites were also found all along a 18S-5.8S-26S unit and associated with a CHH methylation restricted to CAA sites.

Expression and gene ontology analysis of genes mapped by DNase I-MeDIP-SEQ reads

The relative tissue-specific expression level in SAM compared with other poplar tissues (Street *et al.*, 2008) was shown to affect the MeDIP read distribution along genes (Fig. 4b). Thus, SAM-specific overexpressed genes ($\log_2$ values $> +3$) exhibited two internal peaks at the beginning and at the end (the highest) of the body gene, respectively, and very few mapped MeDIP reads in the flanking regions (Fig. 4b). Furthermore, the enrichment ratio for genes in the DNase I-MeDIP-SEQ fraction compared with the whole genome was also affected by their relative tissue-specific expression level in SAM depending on the region of the gene considered (exon, intron, or 5' or 3' flanking region; Fig. 6). Indeed, for all regions, expressed genes with no tissue specificity ($\log_2$ values ≈ 0) were found in conserved ratios (≈ 1) between the DNase I-MeDIP-SEQ fraction and the whole genome. In contrast, genes with SAM-specific expression (overexpressed with $\log_2 > +4$ or repressed with $\log_2 < -4$) had variable ratios ranging from 0 (not found in the DNase I-MeDIP-SEQ fraction) to 2 (enriched in the DNase I-MeDIP-SEQ fraction; Fig. 6). Among

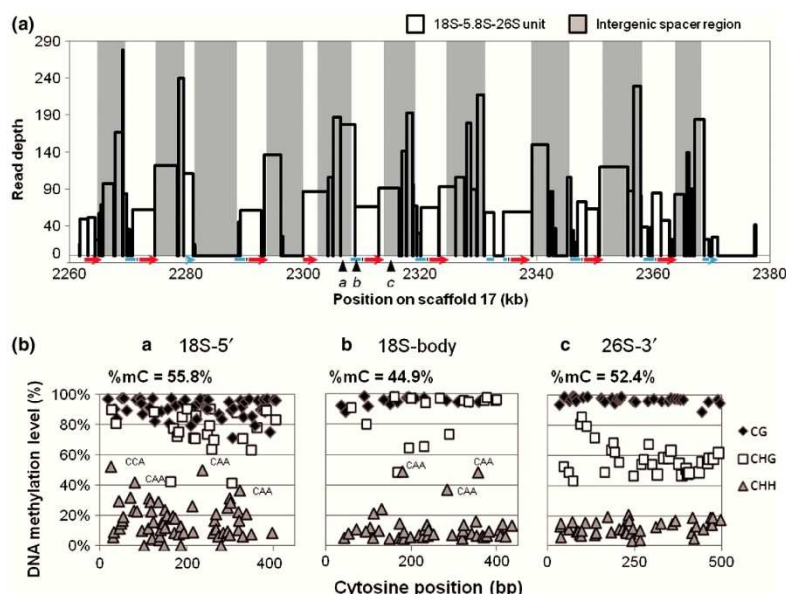


Fig. 5 DNA methylation of rDNA in the DNase I-MethylDNA Immunoprecipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) fraction. (a) Locus on scaffold 17 containing rDNA clusters. MedIP contigs are represented by white rectangles whose width is the covered region and height is the read depth (read depth = number of reads/peak surface). 18S-5.8S-26S ribosomal transcription units were localized by blast of 18S (blue traits), 5.8S (black traits) and 26S (red traits) poplar sequences (AF206999.1, AJ006440.1 and AF479118.1 GenBank accession numbers, respectively) in the v2.0 genome. The arrow shows the orientation of units. Gray regions separate each ribosomal transcription unit. (b) Bisulfite sequencing analysis of DNA methylation in *Populus trichocarpa* shoot apex of loci 'a', 'b' and 'c' corresponding, respectively, to the 5' region of a ribosomal unit, the body of 18S rDNA and the 3' region of a ribosomal unit. CG, CHG and CHH sites are represented by black diamonds, white squares and gray triangles, respectively. DNA methylation mean values are given for each sequence. Sequence of CHH contexts methylated above 35% is shown.

them, genes mapped on intronic regions were mostly under-represented in the DNase I-MeDIP-SEQ fraction. Digital northern analysis showed that genes mapped on intronic regions exhibited a variable level of expression in SAM (number of ESTs in each cDNA library; Fig. S5). To validate our approach, we confirmed the relative expression profiles of 24 genes by semiquantitative RT-PCR (Table S4) and by correlating the public data with personal transcriptomic microarray data on SAM (data not shown).

A gene ontology study (singular enrichment analysis) for mapped genes showed that all biological categories were represented but with variations depending on the gene regions mapped. Interestingly, many expressed genes involved in the determination and the control of SAM activity, such as *LEAFY*, *CLAVATA2* and *WUSCHEL*, were found in the DNase I-MeDIP-SEQ fraction (Table S4) and confirmed by bisulfite sequencing. Mapping in the 5' and 3' flanking regions gave a similar distribution between GO categories ($r_p = 0.77$; $P < 0.01$) but this distribution was different for genes mapped in their body ($r_p = -0.69$; $P < 0.01$). Genes mapped in flanking regions were depleted ($< 50\%$) in transcription regulator activity, biological regulation and regulation of biological process, while genes mapped in their bodies were depleted in electron carrier activity and response to stimulus (Fig. S6). Detailed GO analysis (singular enrichment analysis) also identified over-represented classes of proteins: helicase, ATP-binding and ATP-ase activity for introns and exons and RNA-directed DNA polymerase activity for 5' and 3' regions (Figs S7–S10).

Methylation at the cytosine level: contexts and levels

The three cytosine methylation contexts displayed enrichment ratios ranging from 1.17 to 2.71 between the DNase I-MeDIP-SEQ fraction and the whole genome (Fig. 7a). For non-CG contexts, the highest ratios were found for contexts containing at least two cytosines (Fig. 7b,c).

Bisulfite sequencing of 23 randomly selected sequences found in the DNase I-MeDIP-SEQ fraction (10 312 bp total; Table S3), representing 373 CG, 445 CHG and 1133 CHH, confirmed that CG context was more frequently methylated (334 mCG; mean value: 80%) than CHG (239 mCHG; 40%) and CHH (166 mCHH; 20%) contexts (Fig. 8a). The levels and the context preferences of methylation were dependent on the type of sequences (gene coding protein, rDNA or transposable element (TE)) and the redundancy of genes in the genome (≤ 2 or > 2 ; Fig. 8b,c and Table S5). Indeed, coding sequences with one or two copies exhibited strong CG methylation ($> 70\%$) and low non-CG methylation ($< 10\%$) with preference for C(C/A)G sites but none for CHH contexts (Figs 8, 9a). Coding sequences with more than two copies as well as rDNA and TE exhibited high CG ($> 80\%$) and CHG ($> 60\%$) methylation with moderate methylation ($> 20\%$) of CHH, except for rDNA, which showed a lower CHH methylation level ($< 10\%$). All these sequences showed site-specific methylation on C(A/T)G and CAA sites (Figs 5, 8, S1). Computational analysis did not detect TE inside or within 500 bp around genes, whether they were highly redundant or not.

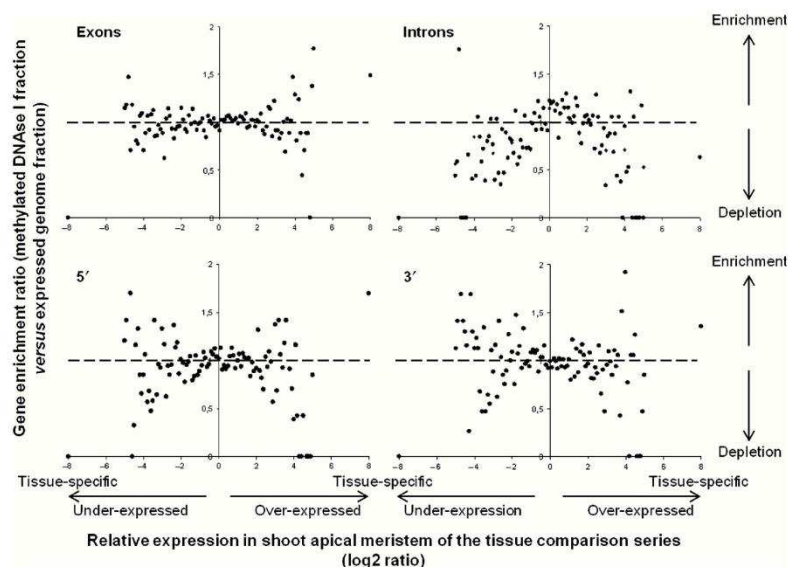


Fig. 6 Relative expression of genes from the DNase I-MethylDNA Immunoprecipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) fraction compared with the whole genome in the apical region of vegetative buds. Relative expression is given as normalized log₂ values. The comparison between the DNase I-MeDIP-SEQ fraction and the whole genome is expressed as the ratio between relative frequencies for each expression class. The comparison was performed for genes whose exons, introns, 5' regions or 3' regions are found in the DNase I-MeDIP-SEQ fraction. The expression data are extracted from the Street *et al.* (2008) tissue comparison study.

Genetic variations: DNA methylation and expression of DNA METHYLTRANSFERASE 1 (*MET1*) genes in *P. trichocarpa* and two *P. × euramericana* hybrid clones

The two *P. trichocarpa* *MET1*-like genes (83.9% nucleotide similarity), *PtMET1a* (POPTR_0019s00240) and *PtMET1b* (POPTR_0004s14140), were mapped by DNase I-MeDIP-SEQ reads, particularly in exons, but not at the same positions (Fig. 9a,b). Bisulfite sequencing confirmed significant methylation only at CG sites between the two genes (Fig. 9a,b). The methylation profiles for these genes were assessed in two *P. × euramericana* hybrid clones, Soligo and Carpaccio, and showed both genetic and epigenetic variations associated with the level of expression. Thus, *PtMET1a* exhibited lower methylation levels in the hybrid clones in relation with mutations (CG to TG transitions) but higher levels of mRNA accumulation (Fig. 9c). By contrast, *PtMET1b* had no mutations among genotypes and showed no methylation or expression in 101-74 and Soligo, while Carpaccio was highly methylated in all cytosine contexts and exhibited expression (Fig. 9c).

Discussion

DNase I-MeDIP-SEQ strategy for methylome assembly studies

Several methods, including DNA methylation microarray analysis, MeDIP-SEQ and direct bisulfite sequencing, have been used in recent years to assess the methylome in several organisms, such as rice (*Oryza sativa*), human and honeybee (Weng *et al.*, 2009;

Zemach *et al.*, 2010). The choice of method is mainly dependent on the quality of genomic resources, the availability of molecular tools, and the objectives in terms of genome coverage and scale resolution, but costs and the need for further bioinformatics analyses must also be considered. In this context, MeDIP-SEQ was the most suitable method for genomic studies in poplar (Vining *et al.*, 2012). Our objective was to establish a draft of the methylome that will highlight gene-body DNA methylation characteristics in poplar, a model tree. In this context, the strategy developed in the present work was to characterize by MeDIP-SEQ the methylome of the DNase I hypersensitive chromatin fraction, for two main reasons. First, this genomic fraction should be strongly enriched in genes and regulatory sequences, allowing focus on gene-body DNA methylation, and depleted in repeated loci, which cause difficulties in bioinformatics analyses. Indeed, the poplar genome is 30% heterochromatin and 33 Mb of unassembled sequences are not incorporated in the actual genome assembly (403 Mb). Secondly, use of this fraction will allow the mapping of epigenetic marks in a specific structural chromatin context: the open chromatin, permissive to transcription. For this purpose, experimental conditions for chromatin extraction, efficient DNase I digestibility and the extraction of high-quality DNA were optimized.

The DNase I-MeDIP-SEQ mapping yielded a proportion of mapped reads, a coverage, and a number and size of contigs in agreement with those reported previously in similar studies (Vining *et al.*, 2012) and provides a useful public database (Sjödin *et al.*, 2009) for other investigations. The procedure was also validated using PCR amplifications of methylated and non-methylated sequences in the immunoprecipitated fraction,

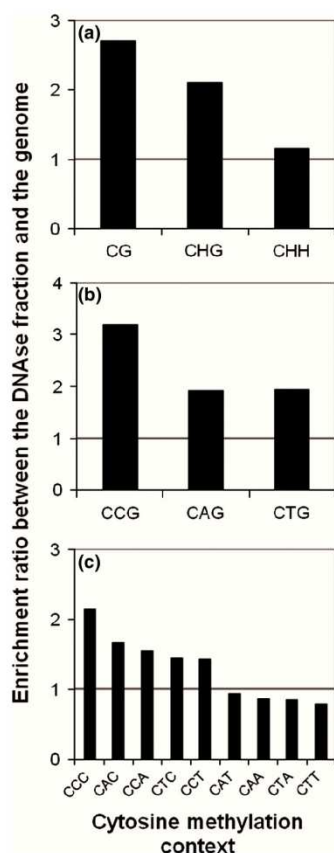


Fig. 7 Enrichment in cytosine methylation contexts in the DNase I-MethylDNA Immunoprecipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) fraction. (a) Enrichment ratio of CG, CHG and CHH sites between the DNase I-MeDIP-SEQ fraction and v2.0 genome. This ratio is given for each (b) CHG and (c) CHH combination. The enrichment ratio is the ratio between each site frequency in the DNase I-MeDIP-SEQ fraction and the same frequency in the v2.0 genome.

comparison with the mapping of a random sequencing of the *P. trichocarpa* genome and bisulfite sequencing of randomly selected regions. A linear distribution was observed for DNase I-MeDIP-SEQ contigs along scaffolds correlated to their size as well as a depletion of DNase I-MeDIP-SEQ contigs in putative highly methylated pericentromeric regions, which were strongly targeted in previous total methylation approaches, such as in poplar (Vining *et al.*, 2012). The low genome coverage (1.9%) is consistent with the low global DNA methylation in poplar apices (*c.* 10%; Gourcilleau *et al.*, 2010), mostly located in heterochromatic regions and nucleosome-bound DNA (Chodavarapu *et al.*, 2010), but also with the use of 30% of extracted DNA after limited DNase I digestion performed according to the procedure of Crawford *et al.* (2006). With 1.9% of the *P. trichocarpa* genome covered, sequences of 30 679 genes (including 2 kb of upstream and downstream regions) were mapped, representing 74% of poplar v2.0 gene models (26% promoter-methylated genes and

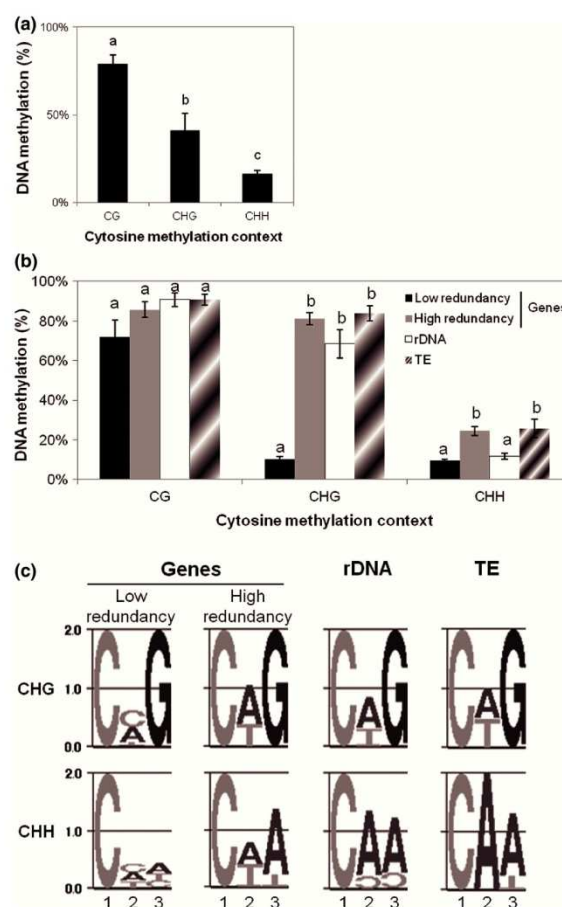


Fig. 8 DNA methylation of CG and non-CG contexts in the DNase I-MethylDNA Immunoprecipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) fraction. (a) Average DNA methylation values (bisulfite sequencing) of CG, CHG and CHH contexts in 23 sequences found in the DNase I-MeDIP-SEQ fraction and randomly chosen. Means are given with $\pm$ SE. Significant DNA methylation differences ($P < 0.05$) among the three nucleotide contexts, as determined by one-way ANOVA, are shown by letters 'a', 'b' and 'c'. (b) Average DNA methylation values of CG, CHG and CHH in lowly (black) and highly (gray) redundant coding sequences, rDNA (white) and TE (hatched). Means are given with their standard errors. 'Low redundancy' coding sequences: $n = 8$; 'high redundancy' coding sequences: $n = 6$; rDNA: $n = 4$; TE: $n = 3$. For each cytosine context, significant DNA methylation differences ($P < 0.05$) among the four sequence types, as determined by one-way ANOVA, are shown by letters 'a' and 'b'. (c) Sequence logos showing preferentially methylated CHG and CHH sites in coding with 'low redundancy', coding with 'high redundancy', rDNA and TE sequences, from left to right. CHG and CHH sequence logos were obtained using GENIO/LOGO software (<http://www.biogenio.com/logo/>). Sequence logos were made with the same number of each CHG or CHH combination to eliminate site frequency-dependent bias. The number of analyzed sites in lowly redundant coding sequences: $n = 155$ CHG and 463 CHH; in highly redundant coding sequences: $n = 176$ CHG and 501 CHH; in rDNA: $n = 95$ CHG and 221 CHH; in TE: $n = 55$ CHG and 98 CHH.

55% body-methylated genes, including genes that were methylated at both features) and confirming the ability of our strategy to

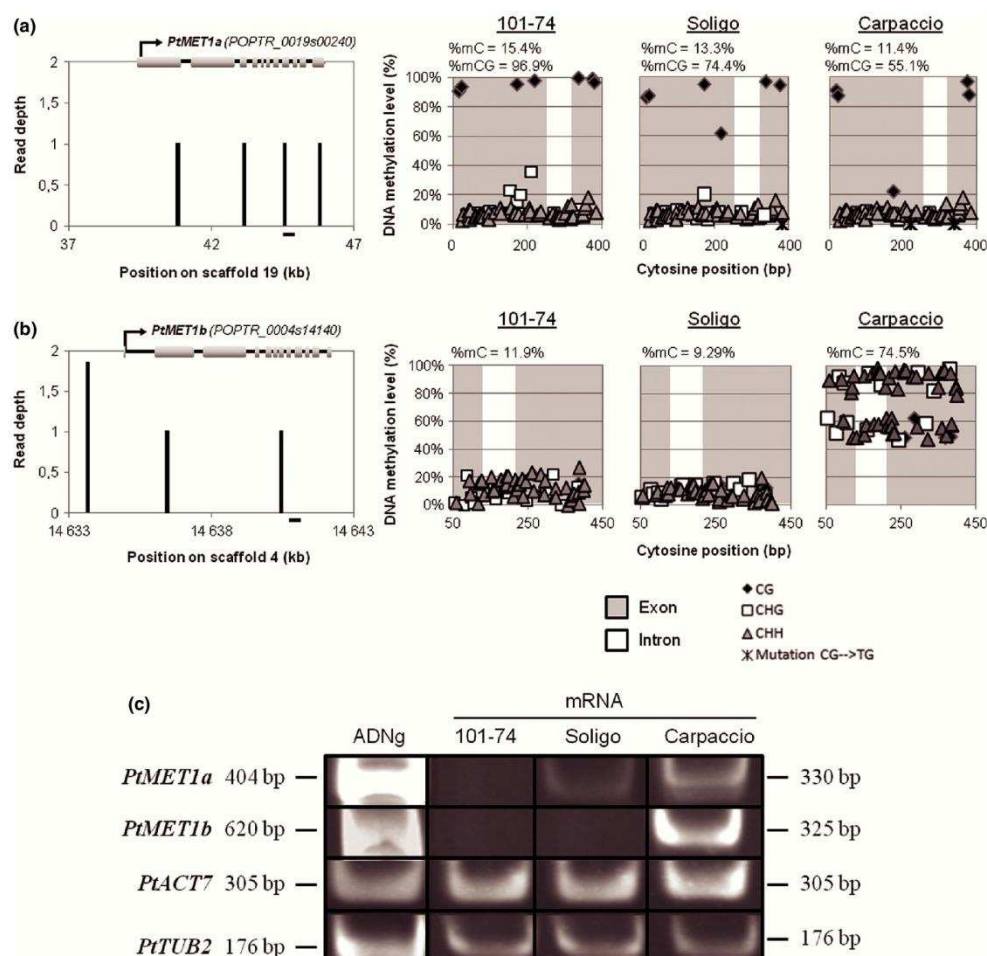


Fig. 9 DNA methylation and mRNA relative abundance of two poplar MET1 genes in different genetic backgrounds. (a) Genomic context of POPTR_0019s00240 (*PtMET1a*, left) and partial DNA methylation profile in the shoot apex of three genetic backgrounds (right). (b) Genomic context of POPTR_0004s14140 (*PtMET1b*, left) and partial DNA methylation profile in the shoot apex of three genetic backgrounds (right). The two genes are shown with their MeDIP contigs from the DNase I-MethylDNA Immunoprecipitation followed by Illumina/Solexa sequencing (DNase I-MeDIP-SEQ) fraction. The black bar above the gene represents the locus analyzed by bisulfite sequencing. The three genetic backgrounds are *Populus trichocarpa* (the 101-74 genotype) and *P. x euramericana* (Soligo and Carpaccio genotypes). CG, CHG and CHH sites are represented by black diamonds, white squares and gray triangles, respectively. CG sites mutated to TG in *P. x euramericana* compared with *P. trichocarpa* are shown by asterisks. Exon and intron regions are gray and white, respectively. DNA methylation mean values are given for each sequence. (c) mRNA relative abundance of POPTR_0019s00240 (*PtMET1a*) and POPTR_0004s14140 (*PtMET1b*) in *P. trichocarpa* (the 101-74 genotype) and *P. x euramericana* (Soligo and Carpaccio genotypes) shoot apex assessed by RT-PCR analyses. *PtACTIN7* (POPTR_0019s02630) and *PtTUBULIN-β2* (POPTR_0001s25410) cDNA amplification was used for normalization between samples. Amplification on genomic DNA was used as a positive control.

focus on gene-body DNA methylation vs repeated sequences. These percentages contrast with percentages of 5% obtained for promoter-methylated genes and 33% for body-methylated genes in *A. thaliana* (Zhang *et al.*, 2006) and 17% for promoter-methylated genes and 15% for body-methylated genes in poplar using the whole-genome approach (Vining *et al.*, 2012). However, quantitative differences in comparisons with these data should be interpreted with caution because procedures, calculations and statistical methods were different. Moreover, previous data on poplar relate to a tissue comparison and were obtained using a whole-genome MeDIP-SEQ approach (Vining *et al.*, 2012). One

other explanation could be that our mapping was concentrated on a very small fraction of the genome (1.9%) rich in genes, increasing the sensibility of the detection. It is also possible that these data are related to the undifferentiated or highly dividing characteristics of meristematic cells, already observed at the transcriptome and proteome level in maize (*Zea mays*) (Dembinsky *et al.*, 2007). Further analyses will be needed to clarify this point. This kind of approach allows a low-cost and simplified methylome bioinformatic analysis that could be useful for organisms with a large content of repeated sequences (Morse *et al.*, 2009; Schnable *et al.*, 2009) or whose assembly is incomplete.

DNA methylation and chromatin structure

Methylated DNase I hypersensitive sites were interspersed across the genome, with a depletion in putative pericentromeric regions, a strong enrichment in genes and a large proportion of intergenic regions that might be distal and proximal to regulatory sequences (Crawford *et al.*, 2006). Some of these interspersed hypersensitive sites were mapped on repeated sequences methylated in the three cytosine contexts, suggesting that, as in *A. thaliana*, the *P. trichocarpa* epigenome is a mosaic of domains with different chromatin states (Roudier *et al.*, 2011). However, in some scaffolds, long adjacent loci corresponding to rDNA were mapped with strong density, raising questions about ribosomal chromatin organization. It must be admitted that only a few rDNA copies among thousands are in an open chromatin state and transcriptionally active (Tucker *et al.*, 2010). This is contradictory with the finding of two 18S-5.8S-26S clusters (on scaffolds 14 and 17) and two additional dispersed 18S-5.8S-26S loci (on scaffolds 8 and 9) that were almost entirely covered by the DNase I-MeDIP-SEQ mapping. This could be explained by the incomplete genome assembly in rDNA loci with <100 copies of 18S-5.8S-26S units found by blast localization, whereas this number has been reckoned at several thousand by restriction analyses (Faire-Rampant *et al.*, 1992). This inaccurate assembly might also explain why we found four rDNA-containing scaffolds while only three chromosomes were found using cytological approaches (Tuskan *et al.*, 2006; Islam-Faridi *et al.*, 2009). In addition, rDNA active copies are much more numerous in meristematic cells than in differentiated cells (Qian *et al.*, 2006). In this context, our most interesting finding was that rDNA loci were more frequently mapped between 18S, 5.8S and 26S units than inside them by DNase I-MeDIP-SEQ reads, which is consistent with previous localization of DNase hypersensitive sites in barley (*Hordeum vulgare*) rDNA supposed to co-localize with *Poll* binding sites (Dimitrova *et al.*, 2009). In addition, CHH contexts were poorly methylated, which is consistent with 5S rDNA low CHH methylation as a consequence of demethylation targeting by DEMETER and DEMETER-LIKE enzymes in *A. thaliana* (Woo *et al.*, 2008). Finally, these data show that high CG and CHG methylation can co-localize with low nucleosome occupancy (Mathieu *et al.*, 2002).

Taken together, these data confirmed the relevance of methylome studies focused on a specific chromatin state to propose new functional hypothesis and improve our knowledge of epigenetic control.

Insights regarding gene-body DNA methylation

The DNase I-MeDIP-SEQ mapping on the gene body for half of the poplar model genes, preferentially in exons, showed that gene-body DNA methylation is more extensive than previously reported (Vining *et al.*, 2012) and is associated with DNase I hypersensitive sites and the open chromatin state. In *A. thaliana* and rice, DNase I hypersensitive sites were detected mainly in the promoter but also in the gene body (Zhang *et al.*, 2012a,b), potentially corresponding to origins of replication (ORIs) which

are contained in the gene body (Costas *et al.*, 2011). In active ORI sites, DNA is methylated and wrapped in nucleosomes with H2A.Z, a histone variant that compromises nucleosome stability and becomes hypersensitive to DNase (Jin *et al.*, 2009). This phenomenon could be particularly common in poplar meristematic dividing cells, and further investigations are needed to determine the role of gene-body DNA methylation in replication.

In *A. thaliana*, gene-body methylation was detected in highly expressed and constitutively active genes, whereas promoter-methylated genes show a greater degree of tissue-specific expression (Zhang *et al.*, 2006). By contrast, gene-body methylation in poplar was correlated to tissue-specific expression and had a more negative effect on transcription than promoter-gene methylation (Zilberman *et al.*, 2007; Vining *et al.*, 2012). The proportion of genes with reported expression in microarray analysis was slightly higher (38%) for genes mapped by DNase I-MeDIP-SEQ reads than for unmapped genes (30%), but no relationship with gene expression level was found. This could be explained by the different environment and *Populus* species used to generate microarray expression data (Street *et al.*, 2008). To address this issue, a methylome/transcriptome cross study has been performed on the same biological material (C. Lafon-Placette, unpublished). However, variations in relation to their size or tissue-specific expression were observed. As long genes were previously shown to be more methylated in *A. thaliana* (Zhang *et al.*, 2008), the depletion of reads in the central part of long poplar genes might reflect low chromatin accessibility for DNase I preventing binding of transcription factors to a cryptic initiation site (Zilberman *et al.*, 2007). Methylation of SAM-specific overexpressed genes was low in the proximal promoter (<1 kb), consistent with previous findings (Zhang *et al.*, 2006), and high at the end of the gene body, in agreement with protection against cryptic initiation or a role in termination. Interestingly, genes with SAM-specific expression (under- or over-expressed) exhibited variations (enrichment or depletion) of methylation and/or DNase accessibility in their gene body and 5'/3' flanking regions compared with ubiquitous genes. This analysis confirmed the possible role of DNA methylation in directing or maintaining tissue-specific gene expression, as proposed by Vining *et al.* (2012). However, genes mapped in intronic regions displaying SAM-specific expression patterns were strongly under-represented in our methylome. The co-localization of DNA methylation and the nucleosome on exons (Chodavarapu *et al.*, 2010) could not alone explain the depletion in introns only for SAM-specific expressed genes. Recently, the role of intragenic DNA methylation in identification of coding regions was assessed in the rat, and suggested an important role for DNA methylation in the regulation of splicing events and the final constitution of the protein sequence (Sati *et al.*, 2012). Our data suggest that this process will be affected by the gene tissue-specific expression profile.

Single- and two-copy genes were mostly methylated in CG contexts, as previously shown in many eukaryotes (Feng *et al.*, 2010a), possibly through protection from non-CG methylation by the jmjC domain-containing protein INCREASE IN BONSAI METHYLATION 1, as described in *A. thaliana* (Miura *et al.*, 2009). By contrast, redundant genes were methylated in

the three contexts, probably as a result of small RNA targeting (Teixeira & Colot, 2010). Methylated contexts in redundant genes were different from those in lowly redundant genes and mostly similar to those in TEs and rDNA, suggesting a specific recognition of repeated genes by the DNA methylation machinery. Considering the recent genome duplication event that occurred in poplar (Tuskan *et al.*, 2006), these data are consistent with theories involving epigenetics in rapid extinction of new duplicated genes to avoid expression redundancy, which could lead to phenotypic abnormalities (Abrouk *et al.*, 2010). However, as only one copy has been studied for each gene, the investigation of DNA methylation in all copies will be necessary to reveal whether this pattern is specific to one copy or affects all copies.

In conclusion, although most genes in poplar SAM cells are partly in a DNA-methylated accessible chromatin state, this conformation can be modified depending on gene tissue-specific expression, gene size or redundancy and can be involved in the control of transcription, splicing or even replication (Zilberman *et al.*, 2007).

Poplar, a model tree for epigenetic studies

Our data are consistent with recent data obtained in poplar, showing a high level of CHG methylation, an interaction between gene-body and promoter methylation, and tissue-specific expression (Feng *et al.*, 2010a; Vining *et al.*, 2012). We also describe new data: functional mapping of methylation on accessible chromatin, higher proportion of methylated genes, variations of gene-body DNA methylation depending on gene size, poplar genotypes and specific cytosine methylation context.

In particular, the investigation of two genes (*PtMET1a* and *PtMET1b*) showed a difference in DNA methylation patterns among poplar genotypes as a result of both genetic (CG→TG transitions) and epigenetic variations (C/mC) on exons at CG sites, paralleling expression differences. In poplar, epigenetic variations have previously been reported among hybrids (Gourcilleau *et al.*, 2010) as well as among clone histories (Raj *et al.*, 2011). The present study was performed on three genotypes (obtained from distinct nurseries) and two genes; genome-wide studies are now needed in several genotypes with characterized clone histories to confirm that epigenetic variations between genotypes or clone histories are extensive and affect related gene expression in poplar.

The investigation of variations of DNA methylation in response to environmental changes and between hybrids is a future challenge, and poplar is an appropriate model tree for such epigenetic studies. In addition, studies of SAM cells may answer questions about the impact of epigenetics on development, plasticity, seasonality, memory and inheritance, providing insights that may be of value for tree breeding in the context of global climate change.

Acknowledgements

The authors are grateful to D. Auguin and C. Hébrard (University Orléans, France) for helpful discussions. We thank A. Guichard (University Orléans), Patrick Poursat and M. C. Lesage-Descauses

(INRA, Orléans), FASTERIS (Plan-les-Ouates, Switzerland), V. Vidal and B. Jesson from Imaxio (Clermont-Ferrand) and K. Vining (Oregon University) for technical assistance.

References

- Abrouk M, Murat F, Pont C, Messing J, Jackson S, Faraut T, Tannier E, Plomion C, Cooke R, Feuillet C. 2010. Palaeogenomics of plants: syntenic-based modelling of extinct ancestors. *Trends in Plant Science* 15: 479–487.
- Beck S, Rakyen VK. 2008. The methylome: approaches for global DNA methylation profiling. *Trends in Genetics* 24: 231–237.
- Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5: 433–438.
- Causevic A, Delaunay A, Ounnar S, Righetta M, Delmotte F, Brignolas F, Hagege D, Maury S. 2005. DNA methylating and demethylating treatments modify phenotype and cell wall differentiation state in sugarbeet cell lines. *Plant Physiology and Biochemistry: PPB/Société Française De Physiologie Végétale* 43: 681–691.
- Causevic A, Gentil M-V, Delaunay A, El-Soud WA, Garcia Z, Pannetier C, Brignolas F, Hagege D, Maury S. 2006. Relationship between DNA methylation and histone acetylation levels, cell redox and cell differentiation states in sugarbeet lines. *Planta* 224: 812–827.
- Chodavarapu RK, Feng S, Bernatavichute YV, Chen P-Y, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ *et al.* 2010. Relationship between nucleosome positioning and DNA methylation. *Nature* 466: 388–392.
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452: 215–219.
- Costas C, de la Paz Sanchez M, Stroud H, Yu Y, Oliveros JC, Feng S, Benguria A, López-Vidriero I, Zhang X, Solano R *et al.* 2011. Genome-wide mapping of *Arabidopsis thaliana* origins of DNA replication and their associated epigenetic marks. *Nature Structural & Molecular Biology* 18: 395–400.
- Crawford GE, Davis S, Scacheri PC, Renaud G, Halawi MJ, Erdos MR, Green R, Meltzer PS, Wolfsberg TG, Collins FS. 2006. DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. *Nature Methods* 3: 503–509.
- Crawford GE, Holt IE, Mullikin JC, Tai D, Blakesley R, Bouffard G, Young A, Masiello C, Green ED, Wolfsberg TG *et al.* 2004. Identifying gene regulatory elements by genome-wide recovery of DNase hypersensitive sites. *Proceedings of the National Academy of Sciences, USA* 101: 992–997.
- Dembinsky D, Woll K, Saleem M, Liu Y, Fu Y, Borsuk LA, Lamkemeyer T, Fladerer C, Madlung J, Barbazuk B *et al.* 2007. Transcriptomic and proteomic analyses of pericycle cells of the maize primary root. *Plant physiology* 145: 575–588.
- Dimitrova AD, Ananiev ED, Gecheff KI. 2009. DNase I hypersensitive sites within the intergenic spacer of ribosomal RNA genes in reconstructed barley karyotypes. *Biotechnology & Biotechnological Equipment* 23: 1039–1043.
- Faivre-Rampant P, Jeandroz S, Lefevre F, Lemoine M, Villar M, Berville A. 1992. Ribosomal DNA studies in poplars: *Populus deltoides*, *P. nigra*, *P. trichocarpa*, *P. maximowiczii*, and *P. alba*. *Genome* 35: 733–740.
- Feng S, Cokus SJ, Zhang X, Chen P-Y, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME *et al.* 2010a. Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences, USA* 107: 8689–8694.
- Feng S, Jacobsen SE, Reik W. 2010b. Epigenetic reprogramming in plant and animal development. *Science* 330: 622–627.
- Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Fireozabady E. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science (New York, N.Y.)* 220: 1049–1051.
- Goll MG, Bestor TH. 2005. Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry* 74: 481–514.
- Gourcilleau D, Bogeat-Triboulet M-B, Le Thiec D, Lafon-Placette C, Delaunay A, El-Soud WA, Brignolas F, Maury S. 2010. DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity. *Annals of Forest Science* 67: 208.

- Hamanishi ET, Campbell MM. 2011. Genome-wide responses to drought in forest trees. *Forestry* 84: 273–283.
- Hernandez D, François P, Farinelli L, Osterås M, Schrenzel J. 2008. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome research* 18: 802–809.
- Islam-Faridi MN, Nelson CD, DiFazio SP, Gunter LE, Tuskan GA. 2009. Cytogenetic analysis of *Populus trichocarpa* – ribosomal DNA, telomere repeat sequence, and marker-selected BACs. *Cytogenetic and Genome Research* 125: 74–80.
- Jansson S, Douglas CJ. 2007. *Populus*: a model system for plant biology. *Annual Review of Plant Biology* 58: 435–458.
- Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark ‘nucleosome-free regions’ of active promoters and other regulatory regions. *Nature Genetics* 41: 941–945.
- Klevebring D, Street NR, Fahlgren N, Kasschau KD, Carrington JC, Lundeberg J, Jansson S. 2009. Genome-wide profiling of *Populus* small RNAs. *BMC Genomics* 10: 620.
- Lechner T, Lusser A, Pipal A, Brosch G, Loidl A, Goralik-Schramel M, Sendra R, Wegener S, Walton JD, Loidl P. 2000. RPD3-type histone deacetylases in maize embryos. *Biochemistry* 39: 1683–1692.
- Lister R, Pelizzola M, Dowen HV, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M *et al.* 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462: 315–322.
- Loureiro J, Rodriguez E, Dolezel J, Santos C. 2007. Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Annals of Botany* 100: 875–888.
- Mathieu O, Yukawa Y, Sugiura M, Picard G, Tourmente S. 2002. 5S rRNA genes expression is not inhibited by DNA methylation in Arabidopsis. *Plant Journal: For Cell and Molecular Biology* 29: 313–323.
- Miura A, Nakamura M, Inagaki S, Kobayashi A, Saze H, Kakutani T. 2009. An Arabidopsis jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. *The EMBO Journal* 28: 1078–1086.
- Morse AM, Peterson DG, Islam-Faridi MN, Smith KE, Magbanua Z, Garcia SA, Kubisiak TL, Amerson HV, Carlson JE, Nelson CD *et al.* 2009. Evolution of genome size and complexity in *Pinus*. *PLoS ONE* 4: e4332.
- Pecinka A, Dinh HQ, Baubec T, Rosa M, Lettner N, Mittelsten Scheid O. 2010. Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis. *The Plant Cell* 22: 3118–3129.
- Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. 2007. DNA demethylation in the Arabidopsis genome. *Proceedings of the National Academy of Sciences, USA* 104: 6752–6757.
- Qian J, Lavker RM, Tseng H. 2006. Mapping ribosomal RNA transcription activity in the mouse eye. *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 235: 1984–1993.
- Raj S, Brautigam K, Hamanishi ET, Wilkins O, Thomas BR, Schroeder W, Mansfield SD, Plant AL, Campbell MM. 2011. Clone history shapes *Populus* drought responses. *Proceedings of the National Academy of Sciences, USA* 108: 12521–12526.
- Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L *et al.* 2011. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *The EMBO Journal* 30: 1928–1938.
- Sati S, Tanwar VS, Kumar KA, Patowary A, Jain V, Ghosh S, Ahmad S, Singh M, Reddy SU, Chandak GR *et al.* 2012. High resolution methylome map of rat indicates role of intragenic DNA methylation in identification of coding region. *PLoS ONE* 7: e31621.
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA *et al.* 2009. The B73 Maize genome: complexity, diversity, and dynamics. *Science* 326: 1112–1115.
- Sjödin A, Street NR, Sandberg G, Gustafsson P, Jansson S. 2009. The *Populus* genome integrative explorer (PopGenIE): a new resource for exploring the *Populus* genome. *New Phytologist* 182: 1013–1025.
- Steinmuller K, Apel K. 1986. A simple and efficient procedure for isolating plant chromatin which is suitable for studies of DNase I-sensitive domains and hypersensitive sites. *Plant Molecular Biology* 7: 87–94.
- Street N, Sjödin A, Bylesjö M, Gustafsson P, Trygg J, Jansson S. 2008. A cross-species transcriptomics approach to identify genes involved in leaf development. *BMC Genomics* 9: 589.
- Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics* 9: 465–476.
- Teixeira FK, Colot V. 2010. Repeat elements and the Arabidopsis DNA methylation landscape. *Heredity* 105: 14–23.
- Trap-Gentil M-V, Hébrard C, Lafon-Placette C, Delaunay A, Hagege D, Joseph C, Brignolas F, Lefebvre M, Barnes S, Maury S. 2011. Time course and amplitude of DNA methylation in the shoot apical meristem are critical points for bolting induction in sugar beet and bolting tolerance between genotypes. *Journal of Experimental Botany* 62: 2585–2597.
- Tucker S, Vitins A, Pikaard CS. 2010. Nucleolar dominance and ribosomal RNA gene silencing. *Current Opinion in Cell Biology* 22: 351–356.
- Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al.* 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604.
- Vining KJ, Pomraning KR, Wilhelm LJ, Priest HD, Pellegrini M, Mockler TC, Freitag M, Strauss S. 2012. Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* 13: 27.
- Weng Y-I, Huang TH-M, Yan PS. 2009. Methylated DNA immunoprecipitation and microarray-based analysis: detection of DNA methylation in breast cancer cell lines. *Methods in Molecular Biology (Clifton, N. J.)* 590: 165–176.
- Woo HR, Dittmer TA, Richards EJ. 2008. Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in Arabidopsis (T. Kakutani, Ed.). *PLoS Genetics* 4: e1000156.
- Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328: 916–919.
- Zhang X, Shiu S, Cal A, Borevitz JO. 2008. Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays (GP Copenhaver, Ed.). *PLoS Genetics* 4: e1000032.
- Zhang W, Wu Y, Schnable JC, Zeng Z, Freeling M, Crawford GE, Jiang J. 2012a. High-resolution mapping of open chromatin in the rice genome. *Genome research* 22: 151–162.
- Zhang W, Zhang T, Wu Y, Jiang J. 2012b. Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. *The Plant Cell* 24: 2719–2731.
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW-L, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE *et al.* 2006. Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell* 126: 1189–1201.
- Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. 2007. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics* 39: 61–69.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 DNA methylation of a transposable element and a highly redundant gene in the DNase I-MeDIP-SEQ fraction.

Fig. S2 Control PCR for sequence enrichment of the DNase I-MeDIP-SEQ fraction.

Fig. S3 Repeat content of the DNase I-MeDIP-SEQ fraction.

430 Research

New
Phytologist

Fig. S4 Enrichment of long genes in the DNase I-MeDIP-SEQ fraction.

Fig. S5 Digital northern heat map representation of genes whose introns are in the DNase I-MeDIP-SEQ fraction.

Fig. S6 General view of enriched GO molecular functions in genes whose (a) upstream or (b) downstream region or (c) body is contained in the DNase I-MeDIP-SEQ fraction.

Fig. S7 Enriched GO molecular functions in genes whose introns are contained in the DNase I-MeDIP-SEQ fraction.

Fig. S8 Enriched GO molecular functions in genes whose exons are contained in the DNase I-MeDIP-SEQ fraction.

Fig. S9 Enriched GO molecular functions in genes whose 5' region is contained in the DNase I-MeDIP-SEQ fraction.

Fig. S10 Enriched GO molecular functions in genes whose 3' region is contained in the DNase I-MeDIP-SEQ fraction.

Table S1 Efficiency estimation of six chromatin extraction protocols

Table S2 Coordinates of MeDIP reads mapped on v2.0 *P. trichocarpa* genome

Table S3 List of primers used for semi-quantitative PCR or bisulfite sequencing

Table S4 List of genes from the DNase I-MeDIP-SEQ fraction whose expression was studied by RT-PCR

Table S5 Criteria chosen for redundancy features of genes

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



About New Phytologist

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <25 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@ornl.gov)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**

Supporting Information

Fig. S1. DNA methylation of a transposable element and a highly redundant gene in the DNase I-MeDIP-SEQ fraction. (a) Two LTR elements on scaffold 11 (left) and their DNA methylation profile analyzed by bisulfite sequencing (right). (b) *POPTR_0006s21000*, a highly redundant v2.0 gene model on scaffold 6 (left) and its DNA methylation profile (right). The LTR elements were identified with both RepeatMasker and PlotRep (see Material and Methods). Exons and introns of the gene model are represented by light grey rectangles and black lines respectively. The arrow indicates gene orientation. The black bar represents the locus analyzed by bisulfite sequencing. MeDIP contigs are represented by grey rectangles whose width is the covered region and height is read depth (read depth = number of reads/peak surface). Bisulfite sequencing analysis was performed on *P. trichocarpa* shoot apex DNA. CG, CHG and CHH sites are represented by black lozenges, white squares and grey triangles respectively. DNA methylation mean value is given.

Fig. S2. Control PCR for sequence methylation of the DNase I-MeDIP-SEQ fraction. (a) Amplification of 1 methylated and 1 unmethylated sequence on control DNA (random sonicated DNA) and DNase I-digested immunoprecipitated methylated DNA (bound). Sequences were then analysed by bisulfite sequencing to know their DNA methylation state. The unmethylated locus is a part of a tubulin gene (*POPTR_0001s25410*) whereas the methylated locus is a part of a highly redundant gene (*POPTR_0006s21000*, Fig. S1). The unmethylated sequence was not found in the DNase I-MeDIP-SEQ fraction whereas the methylated one was. (b) Confirmation of DNA methylation state with methyl-sensitive PCR of 1 methylated and 1 unmethylated sequence using McrBC enzyme. Sequences were then analysed by bisulfite sequencing to know their DNA methylation state. The unmethylated locus is a part of a ubiquitin-like gene (*POPTR_0001s04700*) whereas the methylated locus is a part of a highly redundant gene (*POPTR_0013s03400*). The unmethylated sequence was not found in the DNase I-MeDIP-SEQ fraction whereas the methylated one was.

Fig. S3. Repeat content of the DNase I-MeDIP-SEQ fraction. (a) For each scaffold, ratio of repeat content between v2.0 genome and the DNase I-MeDIP-SEQ fraction. Genome repeat content corresponds to the masked bases in *P. trichocarpa* v2.0 masked genome. Repeat content in the DNase I-MeDIP-SEQ fraction was estimated with RepeatMasker (see Material and Methods). (b) For each scaffold, proportion of repeated sequences in the DNase I-MeDIP-SEQ fraction (in intergenic regions only). Repeats were classified with RepeatMasker according to their nature: retroelements, DNA transposons, ribosomal DNA, simple repeats and low complexity repeats. Arrows indicate the four scaffolds with abnormal reads number shown in Fig. 2. (c) Repeats density, expressed as percentage of covered bases every 100 kb, and reads per kilobase (RPKM) ratio between DNase – MeDIP –SEQ and MeDIP –SEQ (Vining *et al.*, 2012) data.

Fig. S4. Enrichment of long genes in the DNase I-MeDIP-SEQ fraction. For each gene size class, ratio between frequency in DNase I-MeDIP-SEQ fraction (black bars) or in highly enriched genes (peaks with read depth > 2, white bars) and frequency in complete v2.0 genome are shown. Gene size classes are [0; 2 kb[, [2 kb; 5 kb[, > 5 kb.

Fig. S5. Digital northern heat map representation of genes whose introns are in the DNase I-MeDIP-SEQ fraction. Heat map representation of a digital northern to examine the library distribution of the genes whose introns are in the DNase I-MeDIP-SEQ fraction. Genes are represented vertically and their names are not indicated for the sake of clarity. Different tissues are shown. The expression intensity corresponds to the number of counted EST. The results are based on the data contained in PopulusDB.

Fig. S6. General view of enriched GO molecular functions in genes whose (a) upstream, (b) downstream region or (c) body is contained in the DNase I-MeDIP-SEQ fraction. Singular Enrichment Analysis (SEA) was performed for genes whose each of the three regions is in the DNase I-MeDIP-SEQ fraction and for complete v2.0 genome on agriGO website. The values shown here are a ratio between GO terms abundance of each

fraction described above and total v2.0 genome. Thus a value higher than 1 means enrichment towards genome and vice versa. The GO terms selected here are children terms of root one (or called secondary level terms) or significant terms of secondary level terms. Thus, the bar chart gives a brief portray since the GO terms are relatively general description.

Fig. S7. Enriched GO molecular functions in genes whose introns are contained in the DNase I-MeDIP-SEQ fraction. Singular Enrichment Analysis (SEA) for genes whose body is in the DNase I-MeDIP-SEQ fraction was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Fig. S8. Enriched GO molecular functions in genes whose exons are contained in the DNase I-MeDIP-SEQ fraction. Singular Enrichment Analysis (SEA) for genes whose body is in the DNase I-MeDIP-SEQ fraction was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Fig. S9. Enriched GO molecular functions in genes whose 5' region is contained in the DNase I-MeDIP-SEQ fraction. Singular Enrichment Analysis (SEA) for genes whose 5' region is in the DNase I-MeDIP-SEQ fraction was performed towards complete v2.0 genome

on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Fig. S10. Enriched GO molecular functions in genes whose 3' region is contained in the DNase I-MeDIP-SEQ fraction. Singular Enrichment Analysis (SEA) for genes whose 3' region is in the DNase I-MeDIP-SEQ fraction was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Table S1. Efficiency estimation of six chromatin extraction protocols. DNA yield and purity were estimated with DO measurement after DNA extraction. DNA random degradation and DNase digestibility were defined by the quantity of DNA fragments under 2 kb without DNase I digestion or with an incubation of 5 minutes with 0.1 units of DNase I respectively.

Table S2. Coordinates of MeDIP reads mapped on v2.0 *P. trichocarpa* genome. The first column is the scaffold number, the second and third ones are the start and end of MeDIP contigs respectively, the fourth one is the length of each MeDIP contig and the last column is the number of reads forming each MeDIP contig.

Table S3. List of primers used for semi-quantitative PCR or bisulfite sequencing.

Table S4. List of genes from the DNase I-MeDIP-SEQ fraction whose expression was studied by RT-PCR. The two first columns are their v2.0 names and annotation respectively.

The two next columns indicate the number of MeDIP contigs in these genes and where they were mapped (gene body or 2 kb 5'/3' regions). The sixth column shows microarray expression values extracted from (Sjodin *et al.*, 2008). Last column exhibits relative expression values according to semi-quantitative RT-PCR. Expression level were divided arbitrarily in 'not expressed' (0), 'lowly expressed' (+), 'moderately expressed' (++) and 'strongly expressed' (+++) compared with *PtTUB2* (*POPTR_0001s25410*) expression level.

Table S5. Criteria chosen for redundancy features of genes. Two different redundancy thresholds based on blast parameters, are given. An e-value of 10⁻⁵⁰ has been chosen as the most relevant threshold. Related CG, CHG and CHH DNA methylation levels, assessed by bisulfite sequencing, are given.

Fig. S1

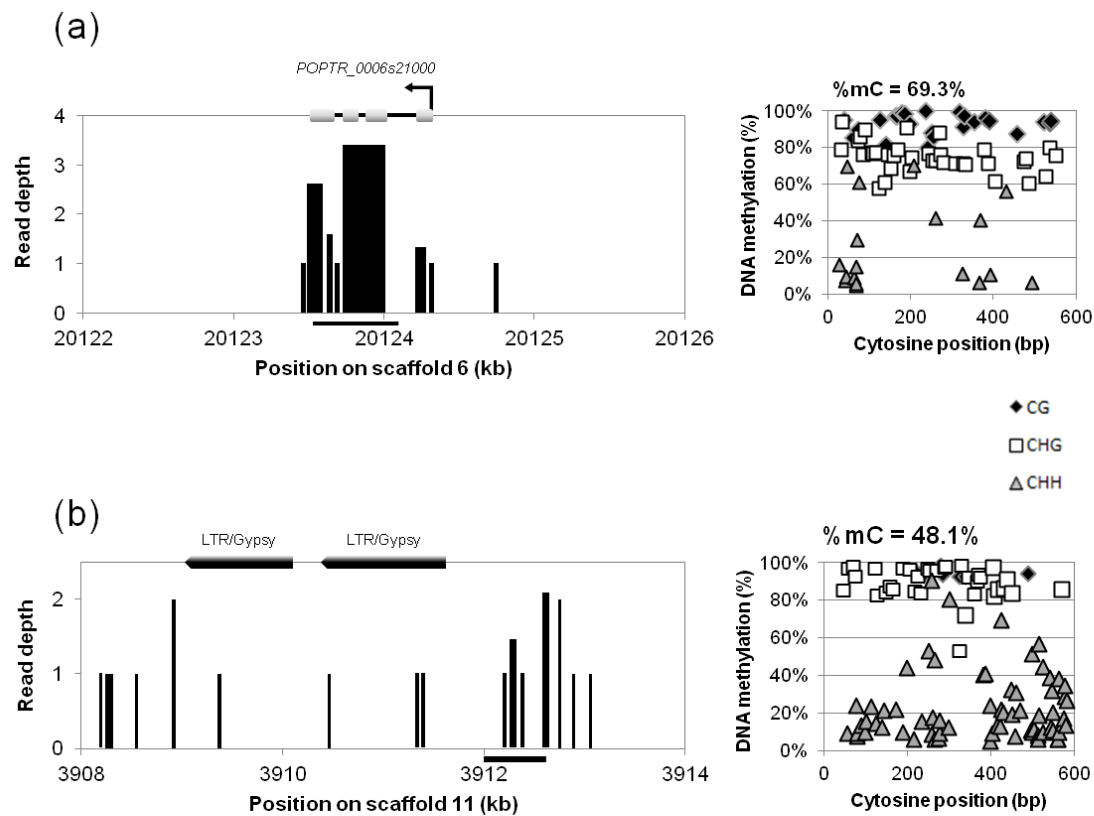
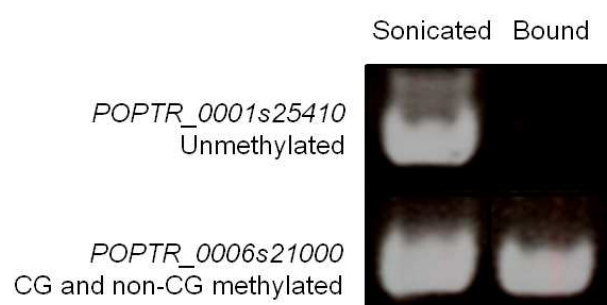


Fig. S2

(a)



(b)

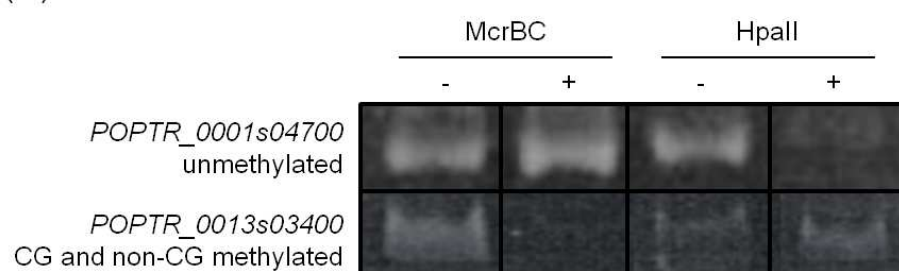


Fig. S3

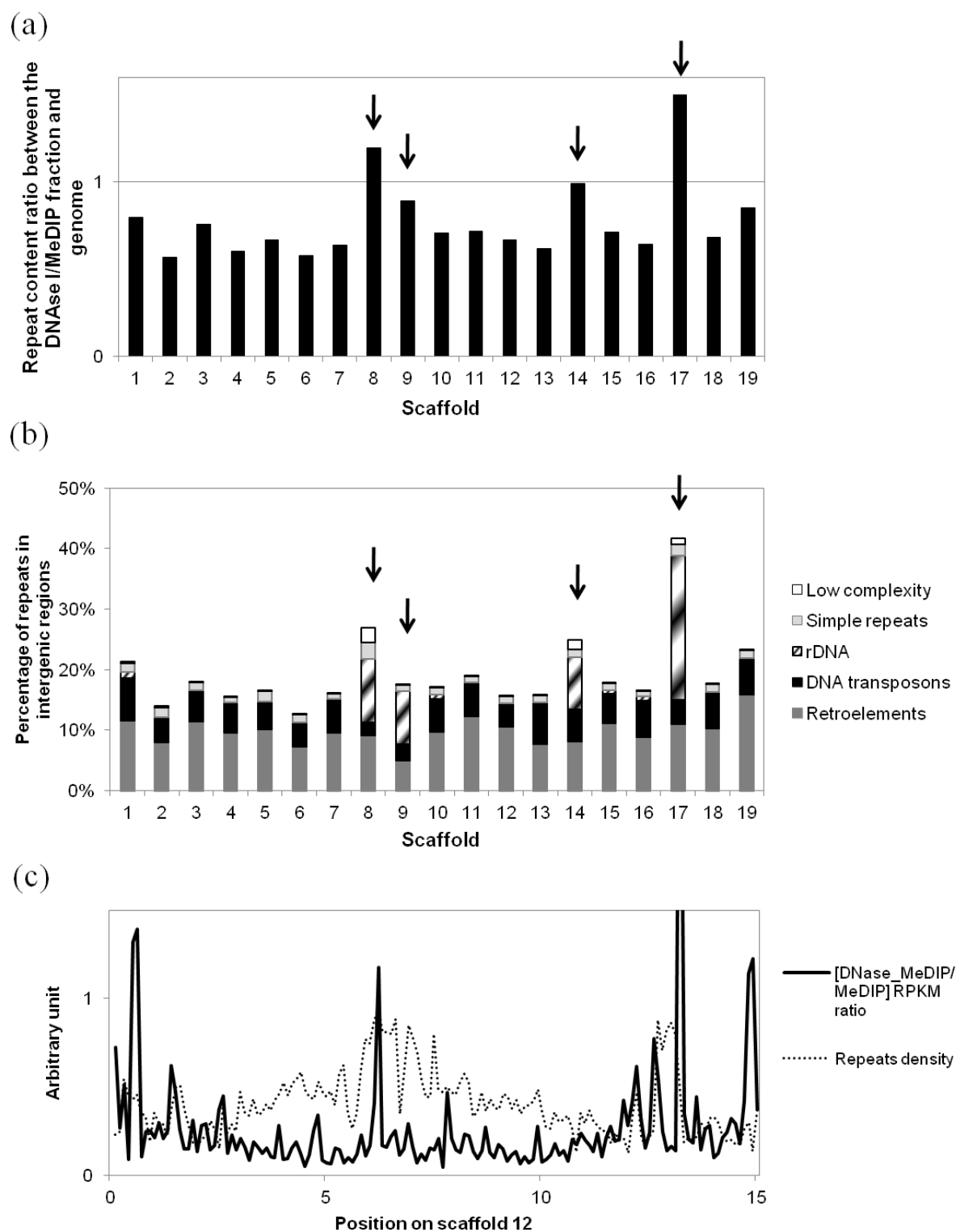


Fig. S4

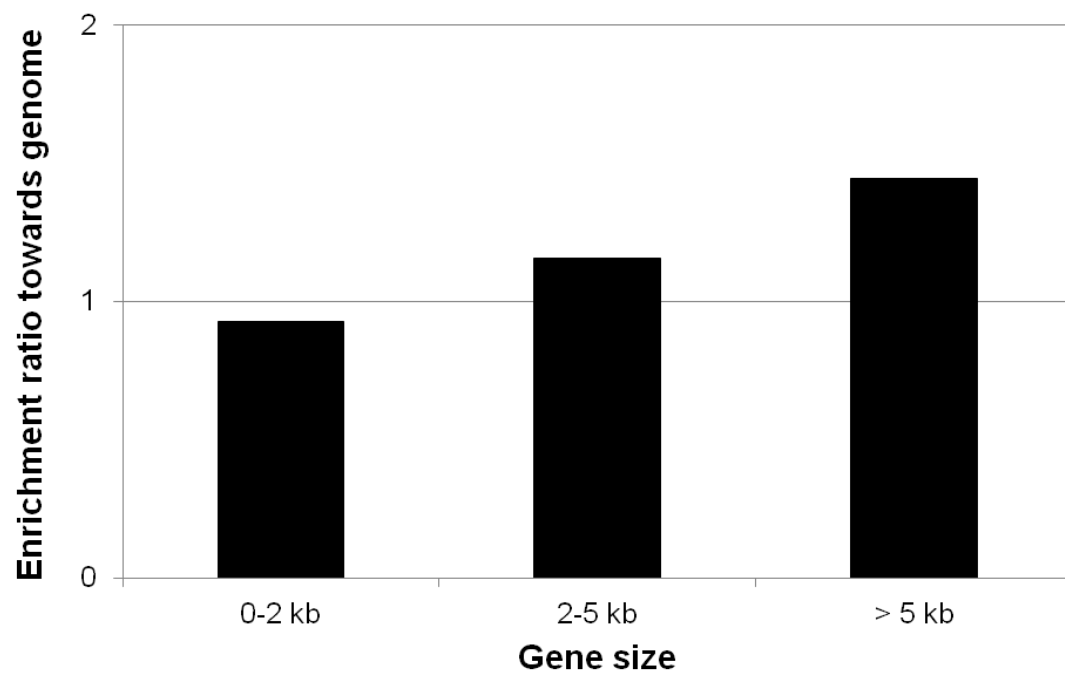


Fig. S5

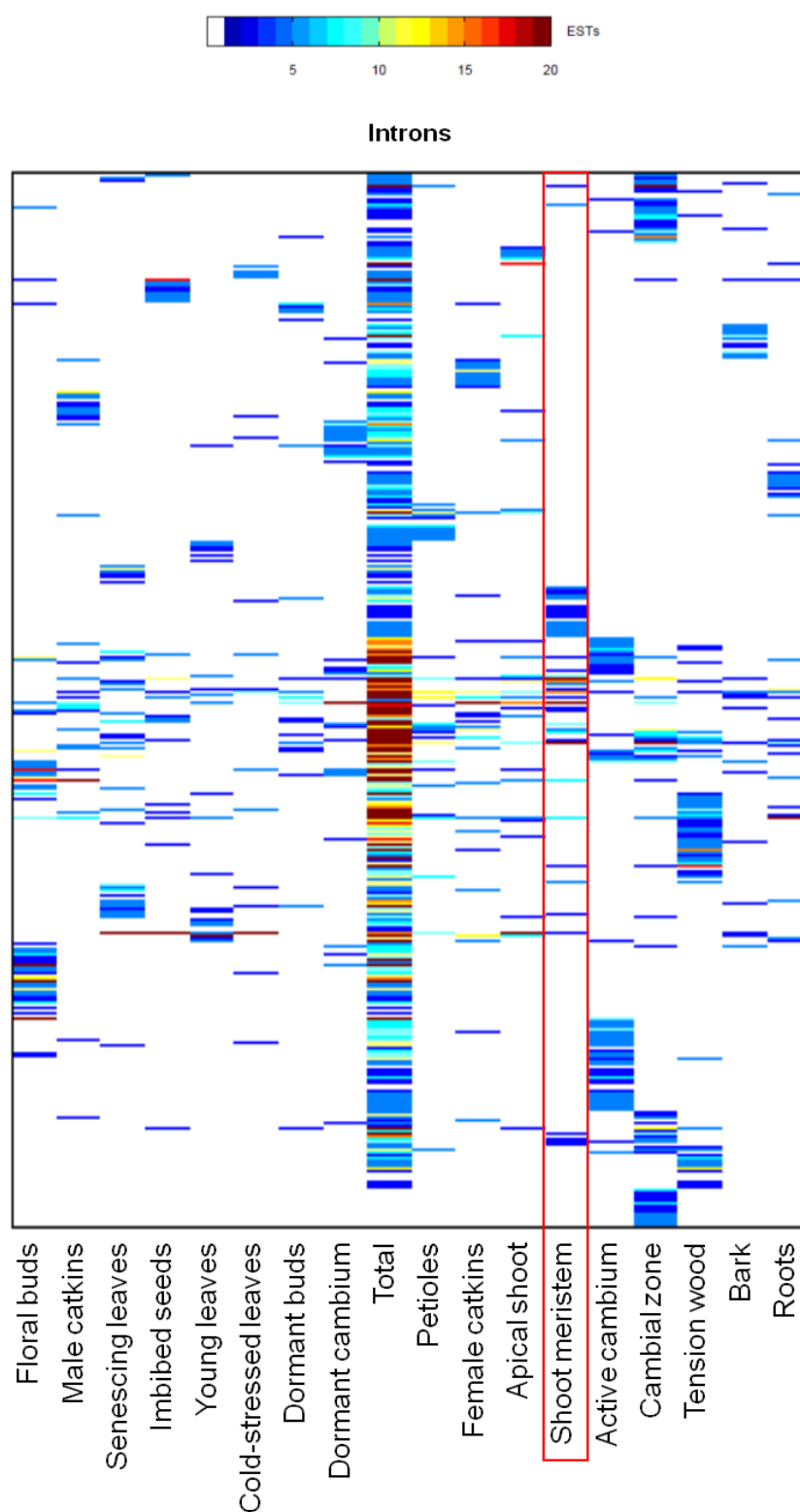


Fig. S6

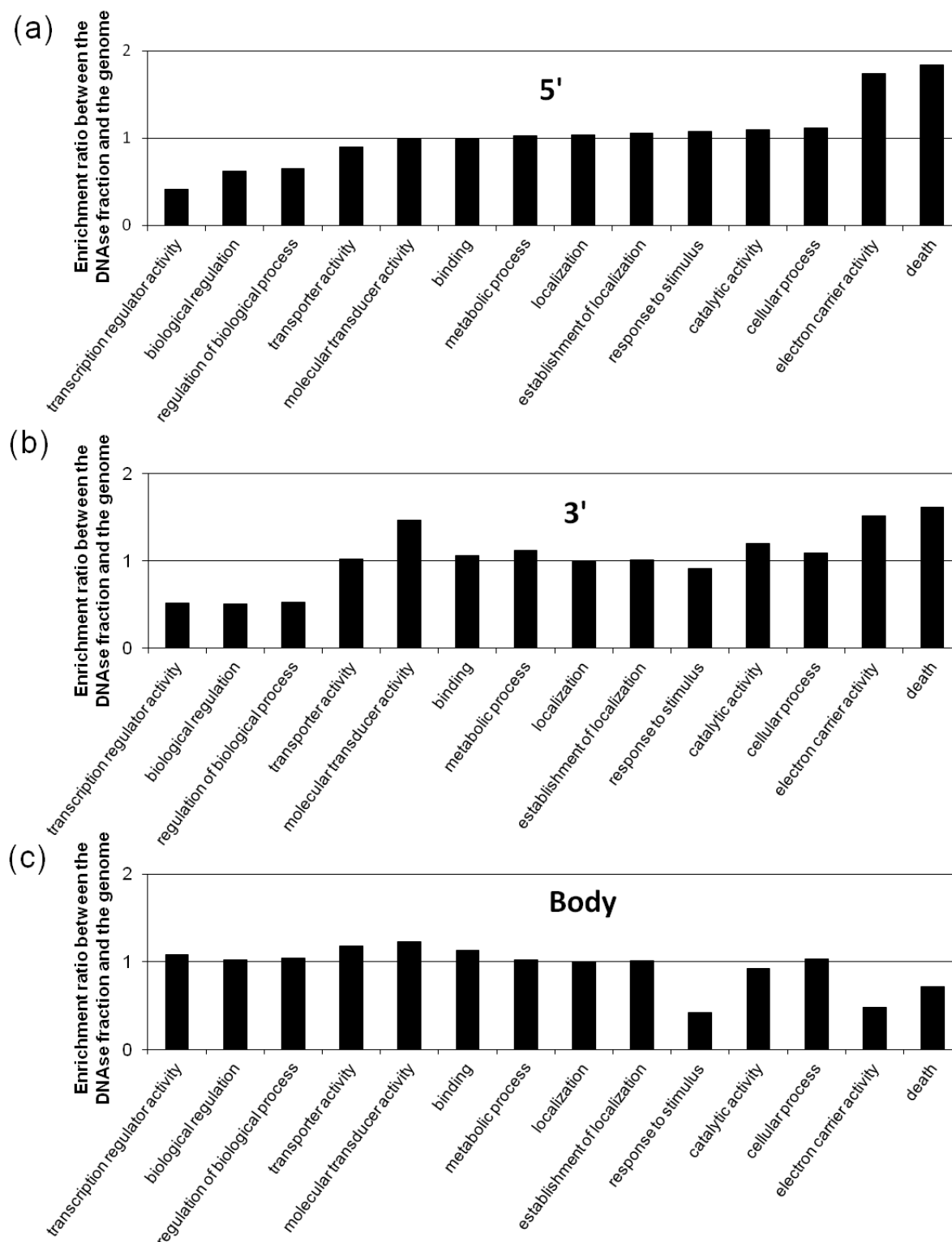


Table S1. Efficiency estimation of six chromatin extraction protocols.

Concentration ng of DNA/mg of fresh weight	Purity		Degradation	DNaseI digestibility	Reference
	260/280	260/230			
170	2.0	1.9	low	high	Causevic <i>et al.</i> (2006)
92	1.8	1.3	moderate	high	Loureiro <i>et al.</i> (2007); GPB buffer
53	1.8	1.8	low	moderate	Galbraith <i>et al.</i> (1983)
28	1.9	2.1	high	moderate	Loureiro <i>et al.</i> (2007); WPB buffer
20	1.8	1.9	moderate	moderate	Lechner <i>et al.</i> (1999)
19	1.8	1.8	high	nd	Steinmuller <i>et al.</i> (1986)

Table S5. Criteria chosen for redundancy features of genes.

Poplar v2.0 gene	No of loci with eval<10 ⁻¹⁰⁰ >75% identity on 500 bp	No of loci with eval<10 ⁻⁵⁰ 500 bp	CG	CHG	CHH
POPTR_0006s21000	4	9	87,2%	70,4%	30,3%
POPTR_0012s12740	3	6	93,8%	85,7%	16,1%
POPTR_0013s03400	5	9	66,4%	72,4%	29,3%
POPTR_0010s01240	5	13	86,4%	83,9%	22,8%
POPTR_0016s11070	13	19	91,9%	89,4%	20,3%
POPTR_0019s05300	>20	>20	88,7%	84,8%	28,0%
POPTR_0007s07830	2	2	83,3%	18,2%	12,9%
POPTR_0008s14350	1	2	84,2%	8,2%	9,9%
POPTR_0019s00240	2	2	97,6%	13,1%	8,3%
POPTR_0006s11720	1	2	85,6%	8,9%	8,5%
POPTR_0010s16200	1	1	68,0%	7,7%	9,3%
POPTR_0010s13870	2	2	57,6%	9,1%	8,6%
POPTR_0014s10610	2	2	87,0%	15,0%	13,6%
POPTR_0003s17720	1	1	23,3%	8,1%	8,6%

3. Analyses complémentaires

Des analyses *in silico* complémentaires ont été menées afin d'expliquer d'une part, les raisons d'une méthylation en CHG plus importante chez le peuplier comparé à d'autres plantes comme *Arabidopsis* et le riz (Feng *et al.*, 2010a) et d'autre part, les raisons d'une préférence de méthylation en CHH (Figure IV.2) plus proche de l'Homme que d'*Arabidopsis* (Cokus *et al.*, 2008; Lister *et al.*, 2009). Nous avons pour cela cherché des particularités sur les enzymes de méthylation du peuplier. Le peuplier possède des gènes codant pour les trois types de DNMT : MET, CMT et DRM (voir introduction ; Goll & Bestor, 2005).

Chez *Arabidopsis*, quatre gènes *MET* sont retrouvées, *MET1* dont le rôle est bien connu et les trois autres dont le rôle a été peu étudié. Une étude suppose, d'après le profil d'expression des gènes codant pour ces dernières, qu'elles joueraient un rôle dans l'empreinte parentale dans l'albumen (Hsieh *et al.*, 2011), mais aucune étude enzymatique ou par mutants n'a été faite à ce jour. En ce qui concerne les trois CMT, CMT3 joue le rôle de méthylation de maintenance des sites CHG, CMT2 semble catalytiquement inactive puisque ne pouvant compléter les mutants *cmt3* et CMT1 est une protéine avortée du fait de la coupure du gène en deux par un rétroélément (Bartee *et al.*, 2001). Enfin, pour les DRM, DRM1 et DRM2 ont une fonction redondante tandis que DRM3, bien que possédant un site catalytique non fonctionnel, joue également un rôle dans la méthylation *de novo* (Henderson *et al.*, 2010).

L'annotation du génome du peuplier n'étant pas complète, il a été nécessaire de l'effectuer pour les DNMT. Par analyse blastX entre génome du peuplier (v2.2) et d'*Arabidopsis* (TAIR9), des gènes candidats ont été identifiés pour les familles d'enzymes :

- *POPTR_0004s14140* et *POPTR_0019s00240* pour la famille des **MET**
- *POPTR_0003s21510*, *POPTR_0003s21520*, *POPTR_0003s09920*, *POPTR_0003s09930* et *POPTR_0001s04700* pour la famille des **CMT**
- *POPTR_0001s35160*, *POPTR_0014s04840* et *POPTR_0010s16200* pour la famille des **DRM**

Les différents motifs composant le domaine méthyltransférase chez les DNMT de peuplier ont été délimités d'après les séquences consensus définies dans des travaux précédents (Pavlopoulou & Kossida, 2007; Figure IV.3). Des analyses par alignement (ClustalW) ont été réalisées pour déterminer *in silico* les potentiels homologues des DNMT d'*Arabidopsis* chez le peuplier. Le motif IV, site actif, a été choisi en priorité pour ces alignements. La même analyse a été menée chez le riz et a donné des résultats similaires, mais dans un souci de clarté ne sera pas montrée ici.

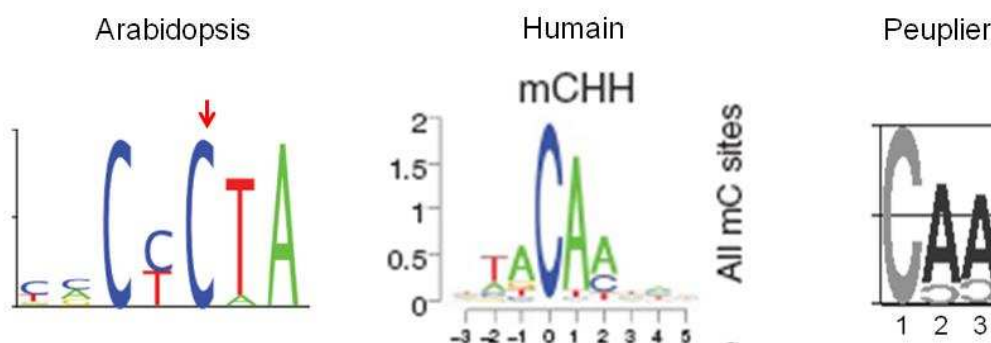


Figure IV.2. Contexte CHH préférentiellement méthylés chez Arabidopsis, l'humain et le peuplier (Cokus *et al.*, 2008; Lister *et al.*, 2009). La flèche rouge indique la cytosine méthylée pour Arabidopsis pour pallier à l'absence de numérotation des cytosines.

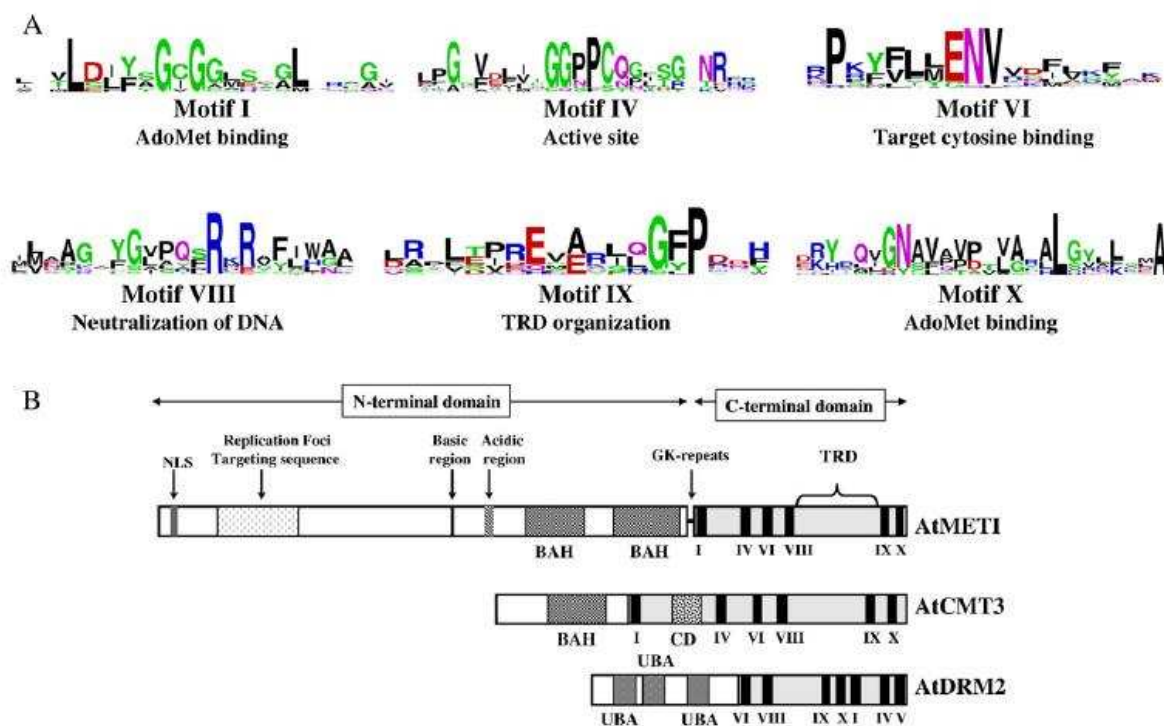


Figure IV.3. A) Logo des séquences consensus des motifs importants pour le domaine méthyltransférase. Motif I et X : fixation du cofacteur S adénosyl-L-méthionine (AdoMet) ; Motif IV : site actif ; Motif VI : fixation à la cytosine cible ; Motif VIII : Neutralisation de la charge négative de l'ADN ; Motif IX : organisation du motif TRD ; TRD : *Target Recognition Domain*, situé entre les motifs VIII et IX, définit la spécificité de méthylation de séquence et de base. B) Représentation schématique des trois DNMT principales chez Arabidopsis, MET1, CMT3 et DRM2. BAH : *bromo-adjacent homology domain* ; CD : *chromodomain* ; UBA : *ubiquitin-associated domain*. D'après (Pavlopoulou & Kossida, 2007).

Concernant les DRM (Figure IV.4A-B), *POPTR_0001s35160* (*PtDRM1*) est le candidat le plus probable comme homologue d'*AtDRM1*, *POPTR_0014s04840* (*PtDRM2*) pour *AtDRM2* et *POTR_0010s16200* (*PtDRM3*) pour le gène codant la DNMT catalytiquement inactive *AtDRM3*. L'alignement en utilisant la séquence des autres motifs a donné des résultats similaires (données non montrées). De manière intéressante, *PtDRM2* est tronquée (Figure IV.4C), ne possédant pas de domaine N-terminal et lui manquant les motifs VIII et IX du domaine méthyltransférase. De ce fait, le *Target Recognition Domain* (TRD), localisé entre les motifs VII et IX (Figure IV.3), a été difficilement délimité et basé sur l'alignement avec *PtDRM1*. Le TRD putatif de *PtDRM2* a une taille identique à celle du TRD des DNMT3 (57 aa). D'après quelques analyses de RT-PCR dans deux fonds génétiques (hybrides *P. × euramericana*) et deux conditions de régime hydrique, *PtDRM1* est dans l'ensemble exprimé avec un niveau moyen tandis que même avec un nombre important de cycles PCR (40), l'ADNc de *PtDRM2* reste faiblement amplifié (Figure IV.4D). *DRM1* et 3 semblent conservées entre peuplier et *Arabidopsis* mais *DRM2* est altérée chez *P. trichocarpa*. On peut se demander si cette dernière est toujours active compte tenu de ses nombreux motifs manquant et du faible niveau d'expression du gène *PtDRM2*. Même si elle est catalytiquement mutée, il est possible qu'elle joue toujours un rôle dans la méthylation de l'ADN *de novo*, à l'instar de *DRM3* chez *Arabidopsis* (Henderson et al., 2010). Dans le cas où elle serait active, la taille de son motif TRD en fait un potentiel candidat pour expliquer la préférence de méthylation des sites CHH chez le peuplier. En effet, la taille du TRD est un des déterminants de la spécificité de méthylation de séquence, autrement dit de contexte de méthylation de l'enzyme (Pavlopoulou & Kossida, 2007). *PtDRM2* ayant un TRD de taille identique à DNMT3 chez l'Homme, ceci pourrait expliquer pourquoi les sites CAA sont préférentiellement méthylés chez le peuplier tout comme chez l'humain, contrairement à *Arabidopsis* où les sites CTA sont préférentiels (Cokus et al., 2008; Lister et al., 2009). La complémentation de mutants d'*Arabidopsis* *drm* permettrait à la fois de savoir si *PtDRM2* est active et par un séquençage génomique bisulfite de savoir si cette enzyme est responsable de la spécificité de méthylation des sites CAA.

Pour la famille des CMT, il s'avère que *POTR_0003s21510* et *POPTR_0003s21520*, deux gènes éloignés de moins de 3 kb, correspondent probablement à un seul gène puisque l'un code pour des motifs du site catalytique que l'autre ne possède pas et *vice versa*. Il est probable que l'outil de détection de gènes utilisé pour l'annotation du génome ait un seuil de taille maximum d'intron relativement faible. Le cas de figure est similaire pour *POPTR_0003s09920* et *POPTR_0003s09930*, deux gènes proches (moins de 2 kb). La séquence du gène fusionné a été retrouvée à l'aide de prédiction de séquence codante par

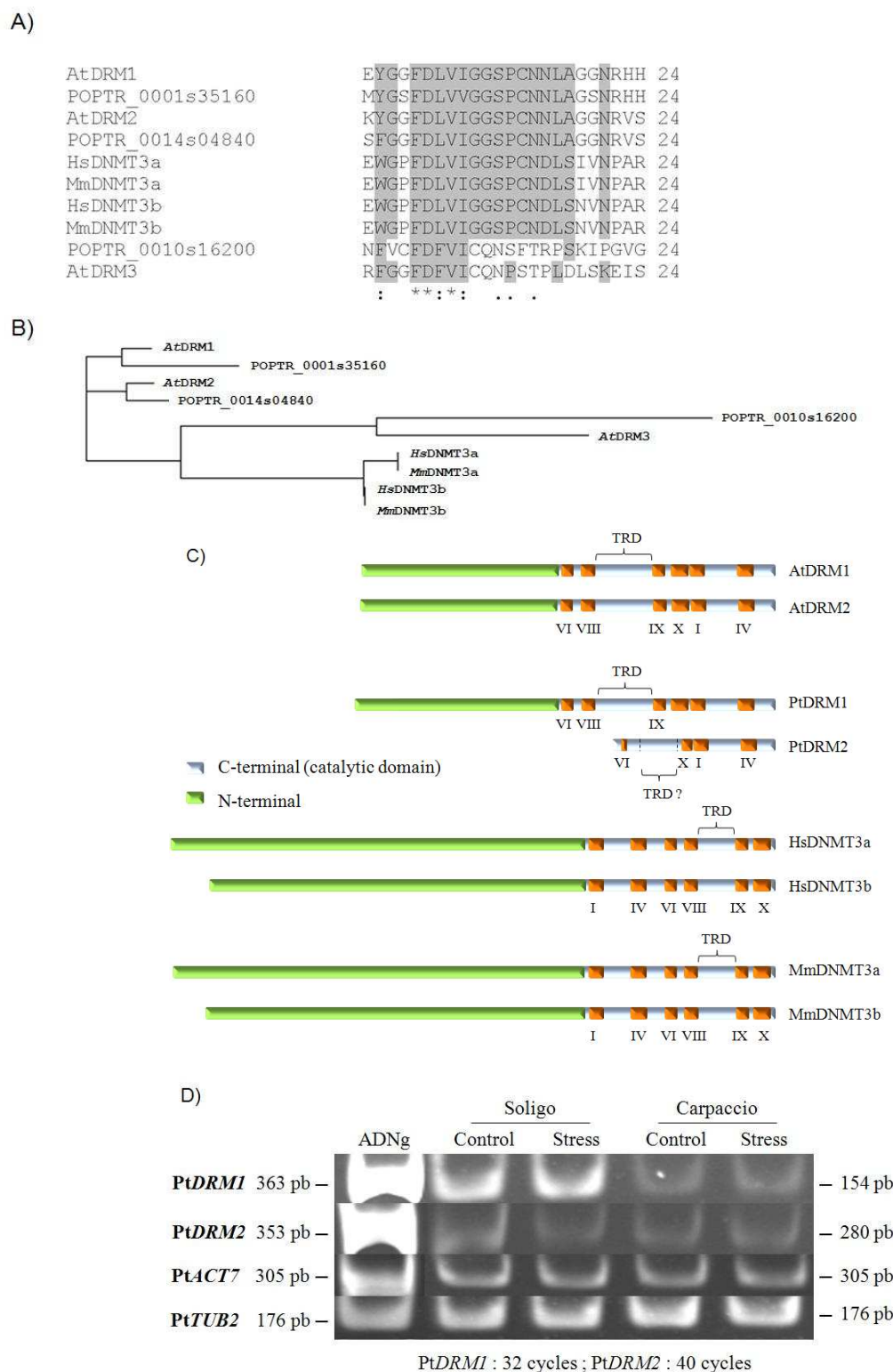
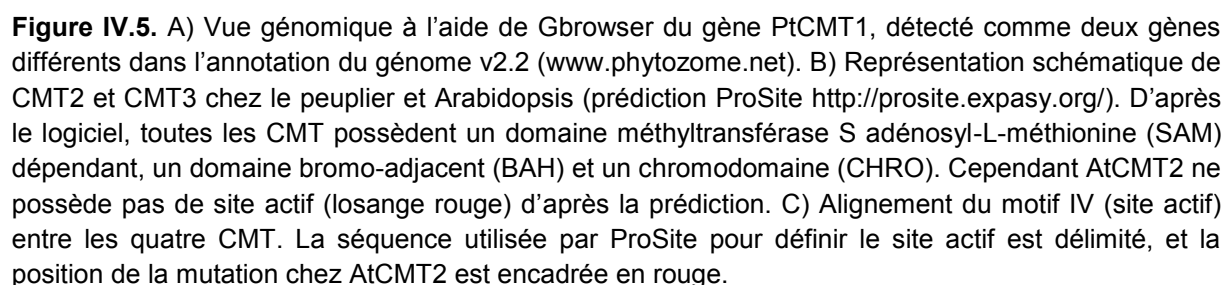


Figure IV.4. A) Alignement du motif IV du site catalytique de DRM1, DRM2, DRM3 d'Arabidopsis, des deux formes DNMT3a et DNMT3b chez l'homme (*Hs*) et la souris (*Mm*) et des protéines POPTR_0001s35160, POPTR_0014s04840 et POPTR_0010s16200 de peuplier. B) Cladogramme correspond à l'alignement du motif IV des DNMT *de novo*. C) Représentation schématisée des enzymes avec le détail des motifs du domaine catalytique. D) Expression des gènes *PtDRM1* et *PtDRM2* dans l'apex mesurée par RT-PCR semi-quantitative, chez deux hybrides *P. × euramericana* 'Soligo' et 'Carpaccio' en condition témoin et déficit hydrique, dispositif expérimental d'après (Gourcilleau *et al.*, 2010).

FGENESH sur la région génomique englobant les deux gènes. Un exon supplémentaire a ainsi été trouvé entre les deux gènes. Par alignement du motif IV, *POTR_0003s21510* et *POPTR_0003s21520* (PtCMT1), *POPTR_0003s09920* et *POPTR_0003s09930* (PtCMT2) et *POPTR_0001s04700* (PtCMT3) ont été définis comme homologues probables de AtCMT1, AtCMT2 et AtCMT3 respectivement. Il est intéressant de constater que PtCMT1 possède, comme AtCMT1, un rétrotransposon le coupant (Figure IV.5A). Une prédiction de domaine protéique effectuée avec ProSite semble confirmer que AtCMT2 est inactive, possédant un site actif muté et non reconnu par le logiciel (Figure IV.5B). Ce logiciel détecte un site actif pour AtCMT3, PtCMT3 mais aussi pour PtCMT2. Un alignement effectué avec le logiciel de prédiction de domaine PFAM (Figure IV.5C) confirme qu'un acide aminé consensus au sein des sites actifs de DNMT, l'aspartate (D), est remplacé par une glycine (G) dans la séquence de AtCMT2 alors que les autres enzymes possèdent bien un aspartate ou glutamate (remplacement semi-conservatif). Par ailleurs, l'interrogation des bases de données d'expression chez Arabidopsis et le peuplier (eFP browser) révèle que AtCMT2 n'a pas de transcrit trouvé dans les nombreuses conditions proposées (développement, stress...) contrairement à AtCMT3, alors que PtCMT2 et PtCMT3 sont notamment exprimés lors du développement de la feuille avec des profils différents (Figure IV.6). Aussi, les CMT, et plus particulièrement CMT2, présente des caractéristiques intéressantes chez le peuplier. Un site actif prédit et une expression dans différents tissus penchent en faveur d'une enzyme effective, contrairement à Arabidopsis, où CMT2 est inactive puisque ne complétant pas les mutants *cmt3* (Bartee *et al.*, 2001) et son gène étant très peu exprimé. Ces indices pourraient expliquer pourquoi les sites CHG sont méthylés à un niveau plus important chez le peuplier (Feng *et al.*, 2010a), CMT2 et CMT3 agissant de manière synergétique ou tout du moins redondante. Ces hypothèses restent néanmoins à être confirmées, par des mesures d'activité enzymatique des différentes CMT par exemple, ou par des approches de mutants RNAi chez le peuplier.

Enfin, pour les MET, les deux enzymes chez le peuplier sont étonnamment proches (89,5% d'identité sur la totalité de la séquence protéique, contre 69,3% entre AtMET1 et AtMET2 par exemple). Du fait de cette forte identité, l'alignement du motif IV et également du motif TRD place les deux MET de peuplier comme homologues du même gène, AtMET1 (Figure IV.7). De manière générale, il est intéressant de noter que malgré un et deux événements de duplication du génome ayant eu lieu respectivement dans les genres *Populus* et *Arabidopsis* (Tuskan *et al.*, 2006; Abrouk *et al.*, 2010), un nombre similaire de DNMT est retrouvé chez Arabidopsis et le peuplier. Après les événements de duplication de génome, les gènes et particulièrement ceux ayant des rôles centraux seraient rapidement éteints voire délétés, évitant une redondance transcriptionnelle pouvant aboutir à des



problèmes développementaux (Abrouk *et al.*, 2010). Cependant, une autre possibilité est la diversification de fonctions ou de patterns d'expression des gènes dupliqués, ce qui semble être le cas des quatre *MET* retrouvées chez *Arabidopsis* (Hsieh *et al.*, 2011). Il est possible que ces quatre gènes proviennent de la duplication d'un *MET* ancestral, suite aux deux évènements de duplication de génome ayant eu lieu dans le genre *Arabidopsis* (Abrouk *et al.*, 2010). Si les deux *PtMET* sont issus de la duplication d'un seul gène, ceci confirme l'hypothèse d'un seul gène *MET* ancestral commun aux genres *Arabidopsis* et *Populus*.

En conclusion, nos analyses *in silico* suggèrent que la méthylation CHG élevée chez le peuplier pourrait s'expliquer par l'activité enzymatique CMT2, redondante ou synergique avec celle de CMT3. Par ailleurs, l'enzyme DRM2 du peuplier présente des caractéristiques communes aux enzymes de méthylation *de novo* chez l'Homme, expliquant potentiellement une préférence de méthylation au sein des sites CHH plus proche de l'humain que d'*Arabidopsis*.

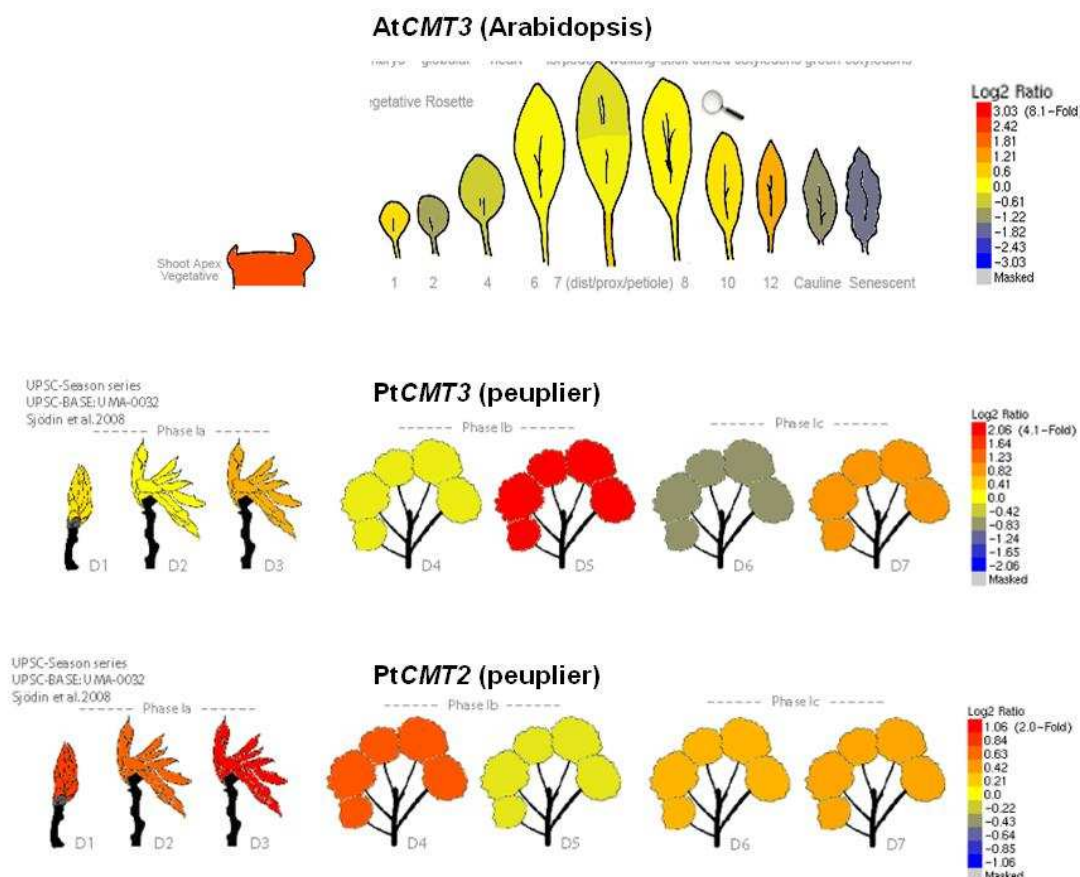


Figure IV.6. Données d'expression relative d'après eFP browser (Winter *et al.*, 2007; Sjodin *et al.*, 2008) des gènes AtCMT3, PtCMT2 et PtCMT3 lors du développement de la feuille (<http://bar.utoronto.ca/> et <http://popgenie.org/>) chez Arabidopsis et le peuplier.

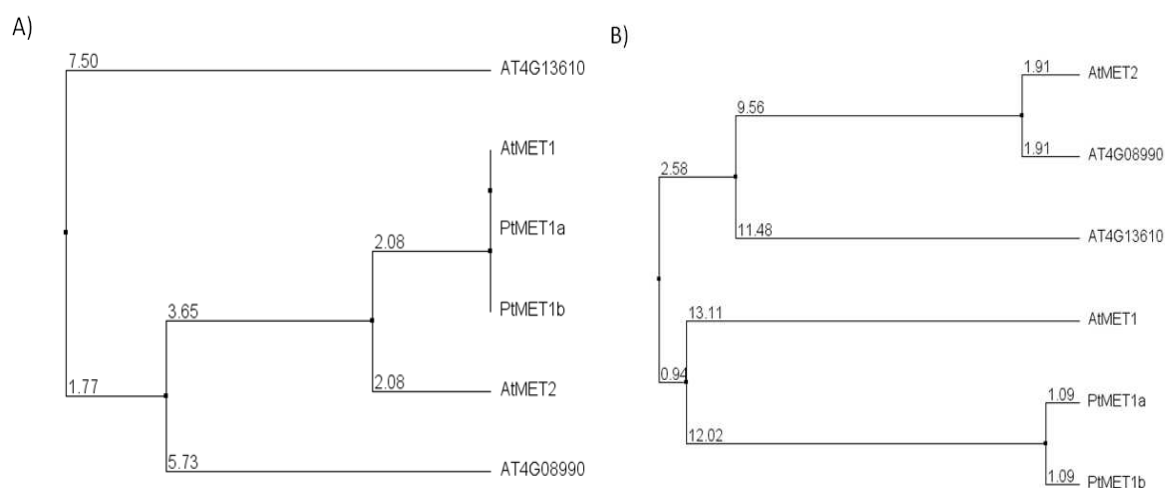


Figure IV.7. A) Arbre phylogénétique entre les MET de peuplier et d'Arabidopsis en se basant sur la séquence du motif IV. B) Même arbre en se basant sur la séquence du TRD.

4. Limites et perspectives

Notre approche a permis d'établir un méthylome du méristème apical du peuplier à partir d'une analyse simplifiée, permettant de cibler les gènes et renseignant sur les liens entre compaction de la chromatine et méthylation de l'ADN. Cette approche serait certainement pertinente pour l'étude d'organismes dont le génome est très riche en séquences répétées comme le pin ou le maïs (Morse *et al.*, 2009; Schnable *et al.*, 2009).

Une première limite de ce travail est liée au génotype étudié de l'espèce *P. trichocarpa*, le clone INRA '101-74', qui est différent du clone 'Nisqually-1' dont le génome était la référence (Tuskan *et al.*, 2006). Ce génotype a été choisi car il est le père d'une population en ségrégation de plus de 2100 individus [croisement *P. deltoides* ('73038-62') × *P. trichocarpa* ('101-74')] créée par l'INRA d'Orléans et pour laquelle des caractères agronomiques comme la tolérance à la rouille (*M. larici-populina*) ont été analysés. Le séquençage a généré des lectures courtes (36-mères) et le mapping n'autorisait qu'un maximum de 2 mésappariements. Il est donc probable qu'un certain nombre de lectures, (difficilement chiffrable), ait été éliminé du fait de la distance génétique entre '101-74' et 'Nisqually-1'. Une solution envisagée à ce problème avait été un assemblage *de novo* mais les zones mappées étant dans l'ensemble courtes et isolées, cette approche ne s'est pas révélée possible excepté pour les loci d'ADNr, fortement couverts par notre séquençage.

Une autre limite de ce travail relève de la séparation de la chromatine non condensée combinée à l'immunoprécipitation de l'ADN méthylé. En effet, la distribution génomique de cette fraction est le produit de l'interaction de ces deux variables, ne permettant pas de distinguer l'effet de l'une ou de l'autre. Comme dans l'étude présente, d'autres données sur le méthylome du peuplier confirment le rôle de la méthylation de l'ADN dans le contrôle de l'expression tissu-spécifique des gènes (Figure IV.8 ; Vining *et al.*, 2012) et une perspective majeure à ce travail serait de mener une étude de l'accessibilité de la chromatine à la DNase I chez le peuplier, permettant de définir plus précisément le rôle de ce facteur dans la régulation des gènes, notamment tissu-spécifique. Les deux premières études majeures chez les plantes de la chromatine non condensée à l'échelle du génome (approche DNase I) sont très récentes (Zhang *et al.*, 2012a,b) et montrent que l'interaction entre accessibilité de la chromatine et méthylation de l'ADN participe au contrôle de l'expression tissu-spécifique des gènes (Figure IV.9 ; Zhang *et al.*, 2012a). Ces résultats témoignent de la pertinence de ce type de travail pour la compréhension du rôle de la structure de la chromatine dans l'expression des gènes. La superposition de données relatives à la méthylation de l'ADN rajoute une information supplémentaire décrivant cette structure.

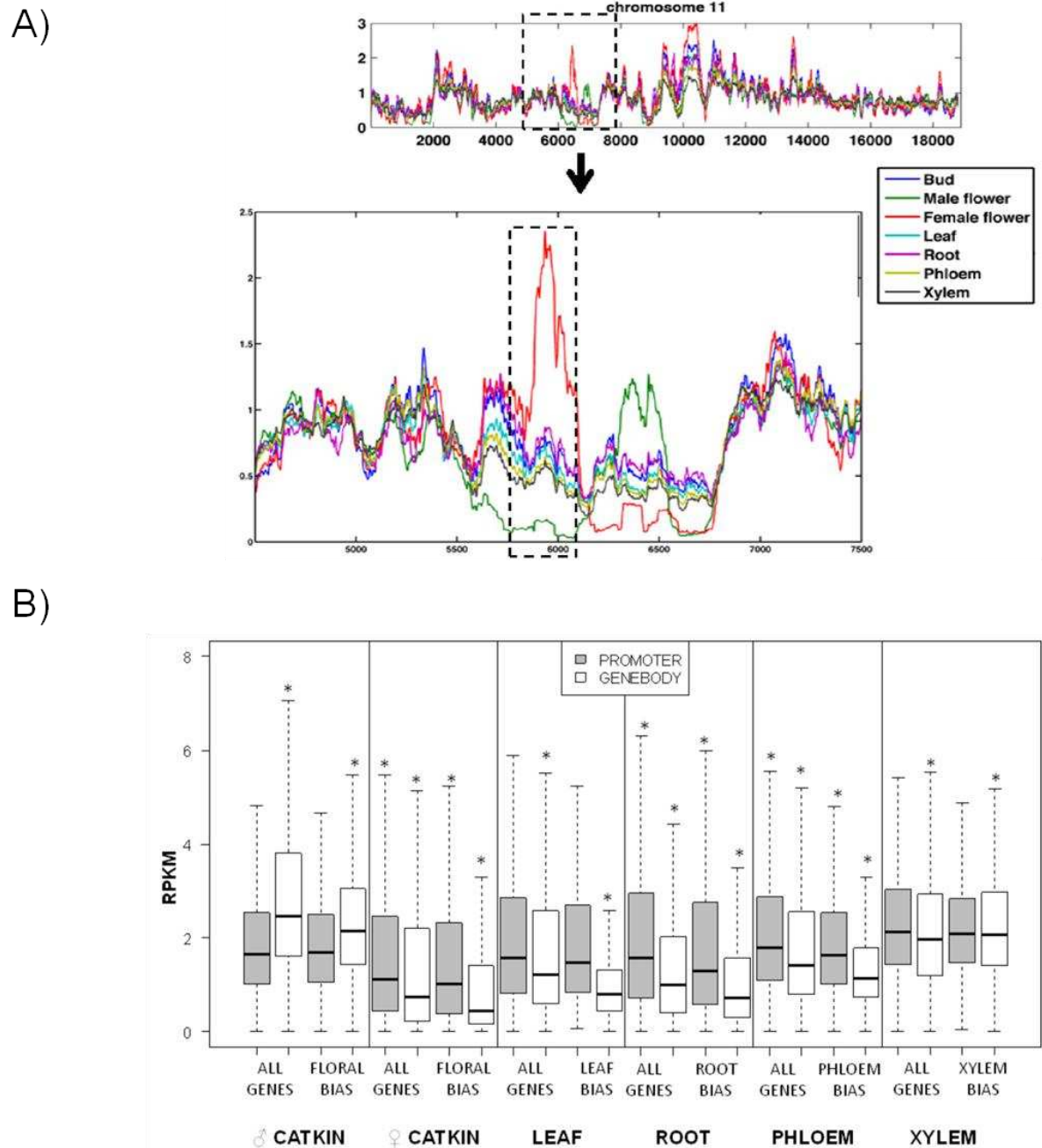


Figure IV.8. Différences de méthylation de l'ADN et d'expression entre tissus chez le peuplier (Vining *et al.*, 2012). A) Cartographie de la méthylation de l'ADN sur le chromosome 11 pour les bourgeons (*Bud*), les fleurs mâles (*Male flower*) et femelles (*Female flower*), les feuilles (*Leaf*), les racines (*Root*), le phloème (*Phloem*) et le xylème (*Xylem*) et agrandissement sur une région différemment méthylée entre tissus. B) Représentation en boîte à moustache de la distribution de la méthylation de l'ADN pour l'ensemble des gènes (*all genes*) et les gènes particulièrement exprimés dans un tissu donné (*[name of tissue] bias*). La méthylation de l'ADN dans le promoteur est représenté en gris et celle du corps du gène en blanc. L'astérisque indique une différence de moyenne de méthylation de l'ADN significative entre l'ensemble des gènes et les gènes particulièrement exprimés dans un tissu donné.

Ainsi, les données obtenues dans ce travail, au vu de la bibliographie existante, permettent une compréhension intégrée du rôle de la structure de la chromatine, de la méthylation de l'ADN et de leurs interactions dans l'expression des gènes. Notre travail reste d'actualité puisque notamment le rôle précis de la méthylation de l'ADN dans le corps des gènes reste toujours à clarifier. Enfin, nos données publiques peuvent alimenter des études dynamiques des changements de méthylation de l'ADN liés à des variations environnementales ou génotypiques chez le peuplier ou en comparaison avec d'autres espèces.

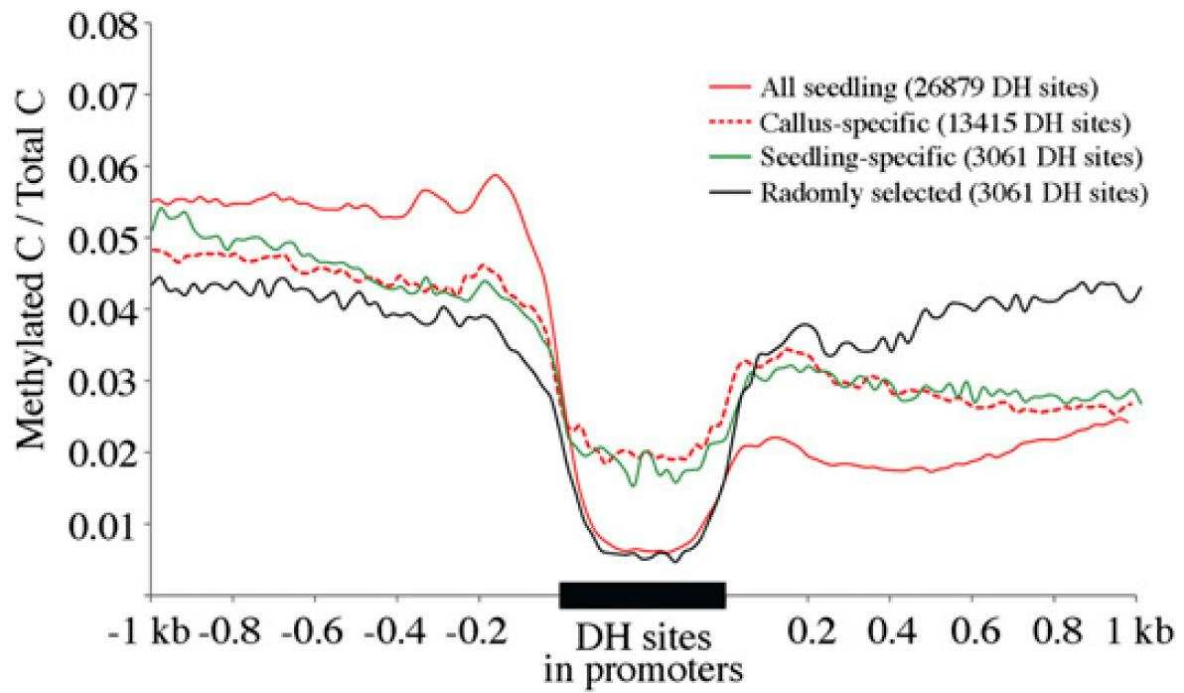


Figure IV.9. Niveau de méthylation des sites hypersensibles à la DNase I (DH) dans les promoteurs chez le riz. Les différentes courbes correspondent à différents profils d'expression de gènes : ubiquistes, exprimés spécifiquement dans les cals ou les plantules (Zhang *et al.*, 2012a).

Chapitre 2 : Impact d'une sécheresse sur la méthylation de l'ADN chez des hybrides de peuplier (*P. deltoides* × *P. nigra*)

1. Présentation synthétique des travaux

Au démarrage de ce travail, quelques études avaient déjà démontré l'existence d'une variabilité naturelle de la méthylation de l'ADN chez des plantes annuelles (Riddle & Richards, 2002; Vaughn *et al.*, 2007; Zhang *et al.*, 2008). De plus, quelques publications avaient mis en évidence des variations de méthylation de l'ADN en réponse à un stress osmotique ou une déshydratation dans des organes ou des cultures cellulaires (Kovarík *et al.*; Sabbah *et al.*, 1995; Labra *et al.*, 2002). En revanche, aucune donnée concernant les variations génotypiques de marques épigénétiques ou leur possible rôle dans la plasticité phénotypique en réponse à la sécheresse n'était disponible chez les plantes pérennes comme le peuplier.

L'objectif de ce travail était d'évaluer le rôle de marques épigénétiques au niveau de l'apex caulinaire dans la plasticité phénotypique des hybrides de peuplier de type *Populus deltoides* × *P. nigra* en réponse à un déficit hydrique. Pour cela, nous avons (i) estimé la variabilité génotypique de la méthylation de l'ADN et de l'acétylation des histones parmi six cultivars *Populus deltoides* × *P. nigra*, (ii) déterminé la plasticité de ces marques en réponse à un déficit hydrique et (iii) mis en évidence des relations entre ces variables épigénétiques et des caractères liés à la production de biomasse.

A cette fin, des boutures des six génotypes ont été soumises en serre à une sécheresse courte et modérée (pas d'altération significative du potentiel hydrique foliaire de base sur 17 jours). Des caractères phénotypiques indicateurs de leur potentiel de croissance ont été mesurés (biomasse, hauteur, surface foliaire). Différentes marques épigénétiques dans le méristème apical ont été étudiées à l'échelle globale (méthylation de l'ADN par HPLC, acétylation des histones H3 et H4 par western blot) ainsi que l'activité et l'accumulation d'enzymes responsables de l'établissement de ces marques (DNMT et HISTONE DESACETYLASE, HDAC). La détermination du niveau global de méthylation de l'ADN a été effectuée par hydrolyse enzymatique de l'ADN en nucléosides, ensuite séparés puis quantifiés par HPLC (Figure V.1 ; Gentil & Maury, 2007). Des corrélations statistiques entre les caractères liés à la production de biomasse et les marques épigénétiques ont ensuite été testées.

Les résultats ont montré des variations génotypiques et en réponse au déficit hydrique pour les variables de croissance et la méthylation globale de l'ADN mais pas pour l'acétylation des histones. Plus précisément, le déficit hydrique a induit une diminution globale de croissance, alors que la méthylation globale de l'ADN présentait des variations entre génotypes, en réponse au déficit hydrique et ceci de façon génotype-dépendante. Une

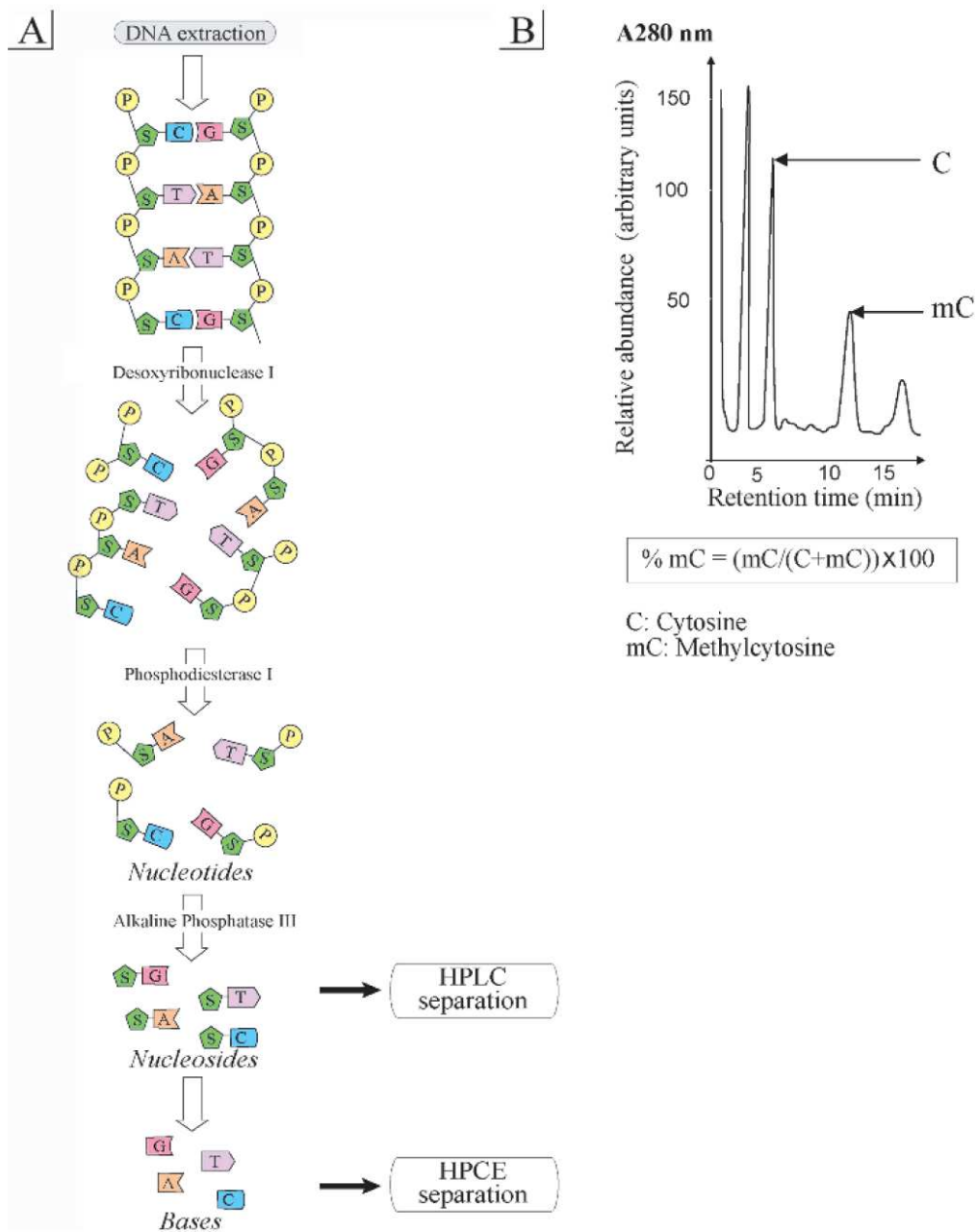


Figure V.1. Détermination du niveau global de méthylation de l'ADN. A) Hydrolyse enzymatique de l'ADN. B) Chromatogramme d'HPLC pour la détermination du pourcentage de méthylcytosine. P : groupement phosphate ; S : sucre ; A, T, C, G : bases (Gentil & Maury, 2007).

corrélation positive a été mise en évidence entre la méthylation de l'ADN et des caractères liés à la production de biomasse (hauteur, biomasse aérienne) en condition hydrique non limitante. En revanche, l'activité DNMT et l'accumulation de ces enzymes n'expliquaient pas seules les variations de méthylation de l'ADN.

En conclusion, Les variations de la méthylation de l'ADN génotype-dépendantes suggèrent la possibilité d'une réponse très sensible et d'une stratégie épigénétique différente selon les hybrides, ce qui est cohérent avec leur différence de plasticité dans des conditions de déficit hydrique plus marqué. Par ailleurs, la mise en évidence d'une corrélation entre la méthylation de l'ADN et des caractères liés à la production de biomasse ouvre une perspective comme un potentiel marqueur de la production de biomasse chez le peuplier et pose la question du rôle de cette marque épigénétique dans la plasticité phénotypique de la tige feuillée en réponse à des variations de disponibilité en eau du sol.

2. Variations de la méthylation globale de l'ADN de l'apex caulinaire au sein de différents hybrides de peuplier et en réponse à la sécheresse (Gourcilleau et al. 2010, *Annals of Forest Science* : article 3)

Ann. For. Sci. 67 (2010) 208
© INRA, EDP Sciences, 2010
DOI: 10.1051/forest/2009101

Available online at:
www.afs-journal.org

Original article

DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity

Delphine GOURCILLEAU¹, Marie-Béatrice BOGEAT-TRIBOULOT², Didier LE THIEC²,
Clément LAFON-PLACETTE^{1,3}, Alain DELAUNAY^{1,3}, Walid Abu EL-SOUD⁴, Franck BRIGNOLAS^{1,3},
Stéphane MAURY^{1,3*}

¹ Laboratoire de Biologie des Ligneux et des Grandes Cultures, UPRES EA 1207, Faculté des Sciences, Université d'Orléans, rue de Chartres, BP 6759, 45067 Orléans Cedex 2, France

² INRA, UMR1137 INRA-UHP Écologie et Écophysiologie Forestières, IFR 110 Génomique, EFABA, 54280 Champenoux, France

³ INRA, USC2030 Arbres et Réponses aux Contraintes Hydrique et Environnementales (ARCHE), 45067 Orléans, France

⁴ Division of Molecular Biology, Biocenter, Innsbruck Medical University, Fritz Pregl-Str. 3, 6020 Innsbruck, Austria

(Received 20 April 2009; accepted 19 August 2009)

Keywords:

methylcytosine /
shoot apex /
vegetative development /
water deficit /
productivity

Mots-clés :

apex caulinaire /
déficit hydrique /
développement végétatif /
méthylcytosine /
productivité

Abstract

• Several reports on annual plants have already shown the involvement of epigenetic modifiers such as DNA methylation in their adaptation to abiotic stresses.
• Nevertheless, the genotypic variations of epigenetic modifiers, their possible correlations with morphological traits and the impact of water deficit have not been described for perennial plants.
• Six genotypes of *Populus deltoides* × *P. nigra* were subjected or not to a moderate water deficit treatment. Various morphological traits such as the height of the plants, their biomass and the total leaf area were measured to characterize the productivity in both conditions. Levels of DNA methylation, histone acetylation and the activities and isoform accumulation of the corresponding enzymes were measured at the shoot apex, the site of morphogenesis. Genotypic variation was observed for the morphological traits and the epigenetic variables and correlations were established among them. Genotypic variation for DNA methylation was detected in hybrid poplars. A positive correlation was demonstrated between DNA methylation percentage and productivity under well watered conditions.
• While there was a general decrease of growth for all genotypes in response to a moderate water deficit, genotypic dependant variations of DNA methylation were found suggesting different strategies among hybrids.

Résumé – Méthylation de l'ADN et acétylation des histones : variations génotypiques chez des peupliers hybrides, impact d'un déficit hydrique et relations avec la productivité.

• Plusieurs études sur des plantes annuelles ont déjà montré l'implication des modifications épigénétiques telles que la méthylation de l'ADN dans la plasticité de leurs réponses aux contraintes abiotiques.
• Néanmoins, les variations génotypiques de ces modifications épigénétiques, leur possible corrélation avec des variables de croissance et l'impact d'un déficit hydrique n'ont pas été décrits sur une plante pérenne.
• Six génotypes de *Populus deltoides* × *P. nigra* ont été soumis ou non à un déficit hydrique modéré et plusieurs variables de croissance ont été mesurées afin de caractériser leur productivité. Les niveaux de méthylation de l'ADN, d'acétylation des histones, les activités enzymatiques et l'accumulation des isoformes correspondantes ont été mesurés sur des apex caulinaires, site de la morphogenèse. Des variations génotypiques ont été observées pour les variables de croissance et épigénétiques. Une corrélation positive a été mise en évidence entre la méthylation de l'ADN et la productivité en condition hydrique favorable.
• Bien qu'il y ait une diminution générale de la croissance de tous les génotypes en réponse à un déficit hydrique modéré, des variations génotype-dépendant de la méthylation de l'ADN ont été trouvées suggérant différentes stratégies entre hybrides.

* Corresponding author: stephane.maury@univ-orleans.fr

1. INTRODUCTION

Poplars are among the fastest growing trees in temperate latitudes but their high productivity is associated with large water requirements (Tschaplinski et al., 1994) and they are one of the most sensitive to water deficit. For this reason, the natural poplar growing area is restricted to river banks and its cultivation to alluvial plains. Nevertheless, significant genotypic variability in the response to water deficit has been reported (Monclus et al., 2006). In order to limit water losses and to optimize water absorption, poplar possesses several mechanisms such as stomatal closure, as well as leaf surface reduction and root surface maintenance via changes of carbon allocation among organs (Braatne et al., 1992; Roden et al., 1990). These mechanisms to adjust to environmental variation define the high developmental plasticity of plants. Studies on annual plants, organs or cell cultures have already demonstrated the role of epigenetic mechanisms in developmental plasticity (Boyko and Kovalchuk, 2007; Chen and Tian, 2007). Indeed, epigenetic modifiers play an important role in genome organization and stability, the control of gene expression and inheritance via chromatin structural reworking (Zilberman et al., 2007).

Chromatin is defined as the association of DNA with histone proteins (H₂A, H₂B, H₃ and H₄) with reversible condensation states (Allis et al., 2007). DNA methylation (Gehring and Henikoff, 2007) or histone acetylation (Chen and Tian, 2007; Loidl, 2004) are epigenetic modifications that lead towards either heterochromatin or euchromatin formation. Indeed, these modifications may ensure the recruitment of multi-enzymatic complexes responsible for dynamic chromatin changes (Allis et al., 2007). Heterochromatin is associated with a negative transcriptional state, due to the incapacity of transcriptional activators or general transcription machinery to gain entry to the DNA sequence.

DNA methylation corresponds to the addition of a methyl group (-CH₃) from *S*-adenosyl-L-methionine on a cytosine, and is catalyzed by DNA methyltransferases (DNMT; EC 2.1.37) (Allis et al., 2007; Pavlopoulou and Kossida, 2007) at the level of CG, CNG (where N is any base) or non-symmetric CHH genomic sequences (where H is any base except G). Histone acetylation or deacetylation is the addition or the removal of an acetyl group (-CO-CH₃) on various lysine residues of the N-terminal extremity of histones, and is catalyzed by histone acetyltransferases (HAT; EC 2.3.1) and histones deacetylases (HDAC; EC 3.5.1) respectively (Allis et al., 2007; Loidl, 2004).

While a few studies have demonstrated natural variation of DNA methylation in annual plants (Riddle and Richards, 2002; Vaughn et al., 2007; Zhang et al., 2008) or in organ and cell cultures under osmotic or dehydration stresses (Kovarik et al., 1997; Labra et al., 2002; Sabbah et al., 1995), no data are available in perennial species such as poplar concerning genotypic variation of various epigenetic modifiers, or of their possible correlation with growth and the impact of water deficit. Therefore, the aim of this study is to test a possible role of epigenetic mechanisms in the plasticity of hybrid poplars in response to water deficit. Elements of answers will be provided

to the following questions: (1) Is there any genotypic variation for epigenetic modifiers such as DNA methylation and histone acetylation in hybrid poplars? (2) If so, could this variability be correlated to the genotypic variability observed in morphological traits? and (3) What is the impact of water deficit on these variables and their eventual correlations?

Six genotypes of *Populus deltoides* × *P. nigra*, which is the best-selling hybrid in Europe and a model of woody perennial plants, were submitted to a moderate water deficit. At the end of the experiment, growth was assessed through morphological traits (stem and root biomasses, plant height, total leaf area and leaf number) and leaf structure (specific leaf area), which were measured for each genotype. In parallel, shoot apices containing meristematic cells and being the site of shoot morphogenesis were collected from each plant. DNA methylation percentages, histones H₃ and H₄ acetylation ratios, and their associated enzymatic activities (DNMT and HDAC) and isoform accumulations were measured. Our results show that genotypic variation of epigenetic modifiers in poplar shoot apex is correlated to productivity. DNA methylation is affected in a genotype dependent way by a moderate water deficit suggesting different strategies among hybrids.

2. MATERIALS AND METHODS

2.1. Plant material, growth conditions and water deficit induction

The experiment was conducted on six genotypes of *Populus euramericana* (*Populus deltoides* × *P. nigra*): Carpaccio (female and diploid), I45-51 (male and possible triploid), NL3972 (female and diploid), Triplo (male and triploid), Lambro (male and diploid) and Soligo (male and diploid). Twelve cuttings per genotype were planted in 10L pots containing a mixture of peat and sand (50/50, v/v) amended with magnesian chalk (60 g/100 L) and fertilized with a slow releasing fertilizer (Nutricote T100, 13/13/13/2, N/P/K/Mg + trace elements). Plants were placed in a greenhouse exposed to natural daylight (ranging from 350 to 900 μmol m⁻² s⁻¹) and where temperature was maintained in the range 15–27 °C with humidity oscillating between 50 and 75%. Plants were watered to field capacity three times a day for 5 weeks, until plant height was between 0.6 and 0.9 m, depending on the genotype. A moderate water deficit was then applied for 17 days to the half of the cuttings, the other half remaining fully watered (controls). While soil volumetric water content (SWC) at field capacity was 32%, it was maintained in the range 7–13% (minimum–maximum) by controlled watering every morning. A correlation between weight of pots (Wt) and SWC measured by time-domain reflectometry was established previously and allowed a control of SWC by weighing the pots.

2.2. Shoot apex harvest and morphological traits measurement

At the end of the experiment, the shoot apex was harvested, frozen in liquid nitrogen and conserved at –80 °C until epigenetic analyses. Then, the height of plants (m) and the length of each individual leaf (m) were measured; the number of leaves was consequently obtained.

The surface area of each individual leaf (m^2) was calculated from abacs previously established on a sample of leaves covering the whole size range (one for each genotype). The total leaf area (m^2) was then calculated for each plant. Roots were washed to remove substrate and then detached from the cutting. Dry weight of leaves, stems and roots (g) were determined after oven-drying for 48 h at 60 °C. Specific leaf area ($cm^2 g^{-1}$) was calculated as the ratio of total leaf area and leaf number.

2.3. Determination of the genomic DNA methylation percentage

Shoot apices were ground to a fine powder. Genomic DNA was extracted, enzymatically hydrolysed into nucleosides and analysed by High Performance Liquid Chromatography (HPLC) according to Causevic et al. (2005). A new hydrophobic column GeminiTM (150 × 4.6 mm, 5 µm, Phenomenex, Le Pecq, France) was used. Isocratic mobile phase is composed of 0.5% methanol (v/v) and 5 mM acetic acid in water. A flow rate 1.5 mL min⁻¹ was used. Identification of cytosine (C) and methylcytosine (mC) were assessed by comigration under the same HPLC conditions with commercial standards (Sigma-Aldrich, Saint-Quentin Fallavier, France). The methylcytosine percentages were calculated using the following formula: %mC = (mC/(C + mC)) × 100, where "C" represents 2'-deoxycytidine content and "mC" 5-methyl-2'-deoxycytidine content.

2.4. Protein extraction, enzymatic activities and immunodetection

Total soluble proteins and chromatin bound histones were extracted separately from shoot apex according to Causevic et al. (2006). The protein concentration was determined using Protein Assay reagent (Bio-Rad, Marnes-la-Coquette, France) and a standard curve established with different solutions of bovine serum albumin (from 0 to 33 µg µL⁻¹).

Similar amounts of protein extracts and a molecular weight marker (Prestained SDS-PAGE Standards, Low Range, Bio-Rad) were separated by SDS-PAGE and electrotransferred on a nitrocellulose membrane (Causevic et al., 2005; 2006). Blotted membranes with soluble proteins were hybridized with different antibodies: carrot anti-DNA methyltransferase 1 (DNMT1; Bernacchia et al., 1998; Causevic et al., 2005) antibodies at 1:300 dilution (v/v) and three maize anti-Histone deacetylase antibodies (HDAC; Pipal et al., 2003; Causevic et al., 2006) raised against the three classes of HDAC (anti-HD1A at 1:2000 dilution (v/v), anti-HD1B at 1:1500 dilution (v/v) and anti-HD2 at 1:1000 dilution (v/v)). Blotted membranes with histone extracts were hybridized with anti-acetylated lysine at 1:1000 dilution (v/v) or anti-non acetylated H₃ histone at 1:200 dilution (v/v) (Santa Cruz Biotechnology, California, and Cell Signalling Technology, Danvers, USA). Cross-reacting bands were identified using anti-rabbit or anti-goat immunoglobulin conjugates labelled with alkaline phosphatase (Sigma-Aldrich) at 1:2000 dilutions (v/v) and stained with BCIP/NBT liquid substrate system (Sigma-Aldrich).

Blots were scanned and the intensity of the immunodetected bands quantified in arbitrary units using imaging software (ImageTool for Windows version 3.00) providing quantitative data. Three independent extractions were analysed by western blot in duplicate for all genotypes, treatments and antibodies. Fixed amounts of a histone H₃

enriched preparation from calf thymus (arginine-rich subgroup f3; Sigma-Aldrich), a maize-HD1A recombinant protein and a carrot-DNMT1 recombinant protein were used as positive controls as already described (Causevic et al., 2005; 2006). These controls served as internal standards for quantification after immunodetection and comparison between membranes. A mean value was calculated for each immunodetected band using the corresponding internal standard and after the subtraction of the aspecific background. Then, mean amounts of the 100 kDa DNMT and 80 kDa HDAC bands were expressed as percentages of the total signal detected for a given protein sample and antibody. This total signal corresponds to the additional mean signal of bands at 100, 70, 50, 30 and 15 kDa for DNMT and 80, 75, 65, 45 and 30 kDa for HDAC. Acetylation ratios of histones H₃ or H₄ were calculated according to the following formula: acetylation ratio of H₃ or H₄ histone (AcH₃ or AcH₄) = amount of acetylated histone H₃ or H₄ / amount of non acetylated histone H₃. Anti-acetylated histone H₃ antibodies were also tested to confirm the data as previously published (Causevic et al., 2006).

Using total soluble protein extracts, DNMT and HDAC activities were measured as described by Causevic et al. (2005; 2006) respectively. DNMT activity was measured using non methylated genomic DNA from a *dam⁻dcn⁻ E. coli* strain as substrate, and ³H-SAM as a methyl donor and is expressed in pkatal per µg of protein. HDAC activity was assayed using an artificial substrate, N-(4-methyl-7-coumarinyl)-N-(t-butyloxycarbonyl)-N-acetyllysine (MAL) (Sigma-Aldrich). HDAC activity is expressed in percentage of deacetylated MAL min⁻¹ per µg protein after HPLC separation and detection.

2.5. Statistical analysis

Statistical analyses were carried out using the SPSS statistical software package (SPSS version 11.0.1 PC, Chicago, IL, USA). Genotype and water deficit effects were evaluated by two-way ANOVA (GLM procedure) using the following model: $Y_{ijk} = \mu + G_i + T_j + (G_i \times T_j) + \varepsilon_{ijk}$; where Y_{ijk} are individual values, μ the general mean, G_i the effect of the genotype i , T_j the effect of the treatment j , $G_i \times T_j$ the genotype by treatment interactions and ε_{ijk} the error. Statistical tests were considered significant at * $P \leq 0.05$, ** $P \leq 0.01$ or *** $P \leq 0.001$.

Genotypic variation was described using principal component analysis (PCA) from genotypic means of morphological traits (stem biomass, root biomass, height, total leaf area, leaf number and specific leaf area) measured in well-watered or water-limited conditions. Variables were standardized and orthogonal factors (scores F_1 and F_2) were built successively as linear combinations of these variables, maximizing the percentage of the variability explained by these factors. The epigenetic variables were projected as supplementary variables in this $F_1 \times F_2$ plane. Their coordinates corresponded to their linear correlation coefficients with scores F_1 and F_2 axis of the PCA.

3. RESULTS

3.1. Genotypic variation of morphological traits and water deficit effects

Significant genotype and treatment effects were detected for all morphological traits (Fig. 1), except for specific leaf area

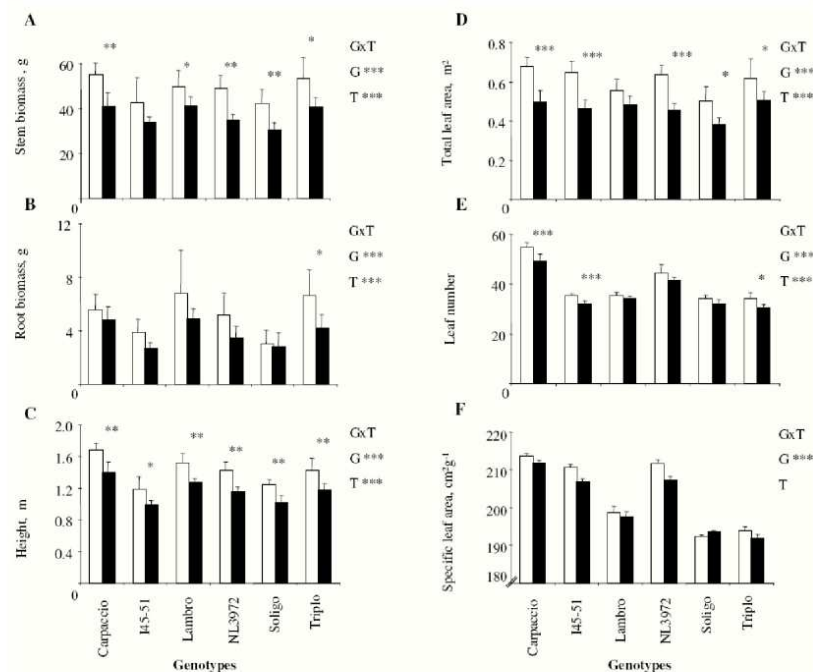


Figure 1. Genotypic variation and water deficit effect on morphological traits related to productivity in six *Populus deltoides* × *P. nigra* hybrids growing in glasshouse in well-watered (white bar) or in water deficit (black bar) conditions. Studied traits encompassed: (A) stem biomass, (B) root biomass, (C) height of plants, (D) total leaf area, (E) leaf number and (F) specific leaf area. For each graph, G indicates the genotype effect, T the treatment effect and (GxT) genotype by treatment effect. Means are accompanied by their standard errors SE ($n = 6$). Significant differences between well-watered and water deficit conditions are indicated by asterisk: * $P \leq 0.05$, ** $P \leq 0.01$ and *** $P \leq 0.001$.

for which only a genotype effect was observed. No genotype by treatment interaction effect was observed.

Under well-watered conditions, genotypes differed for height, leaf number, total leaf area and specific leaf area (Fig. 1). For most of the traits, Carpaccio and Soligo were the extreme genotypes with height ranging from 1.23 ± 0.06 to 1.68 ± 0.08 m, total leaf area from 0.50 ± 0.07 to 0.67 ± 0.04 m², leaf number from 34.2 ± 1.3 to 54.8 ± 1.9 leaves and specific leaf area from 190 ± 5 to 210 ± 5 cm² g⁻¹.

The applied water deficit was moderate and leaf predawn water potential of stressed plants did not differ from controls (data not shown). However, height, stem biomass, leaf number and total leaf area decreased significantly in response to water deficit for most of the genotypes. The range of variation depended on the genotype. For example, stem biomass decreased from about 17% for Lambro to 28% for NL3972. Root biomass only decreased significantly for Triplo.

3.2. Genotypic variations of DNA methylation and water deficit effects

A significant genotype by treatment interaction was observed for percentage methylcytosine (%mC) of shoot apices, indicating that the effect of the treatment depended on the

genotype (Fig. 2A). In the well-watered samples, %mC ranged from $4.1 \pm 1.6\%$ for I45-51 to $11.5 \pm 3.4\%$ for Carpaccio. In response to water deficit, %mC decreased significantly for Carpaccio, Lambro and Triplo, while it strongly increased for Soligo and remained unchanged for I45-51 and NL3972 (Fig. 2A).

Neither genotype nor treatment effects were observed for DNA methyltransferase (DNMT) activity except for Carpaccio and Lambro, for which DNMT activity decreased by about 40% after a water deficit treatment (Fig. 2B). Only a tendency is seen between %mC and DNMT activity ($r = 0.75$ at $P = 0.08$).

Immunodetection of the isoforms of DNMT class 1 revealed five bands: one over 100 kDa and smaller ones with sizes of 70, 50, 30 and 15 kDa, respectively (Fig. 2C). Distinct patterns of accumulation for these bands were observed between genotypes. These patterns were not profoundly affected by water deficit except for Carpaccio, which showed a strong decrease for all bands. A positive correlation was established between DNMT activity and the amount of the 100 kDa DNMT isoform (expressed as a percentage of total signal) ($r = 0.61^*$; supplementary¹ figure S1A).

¹ Supplementary material is available on line only at www.afs-journal.org.

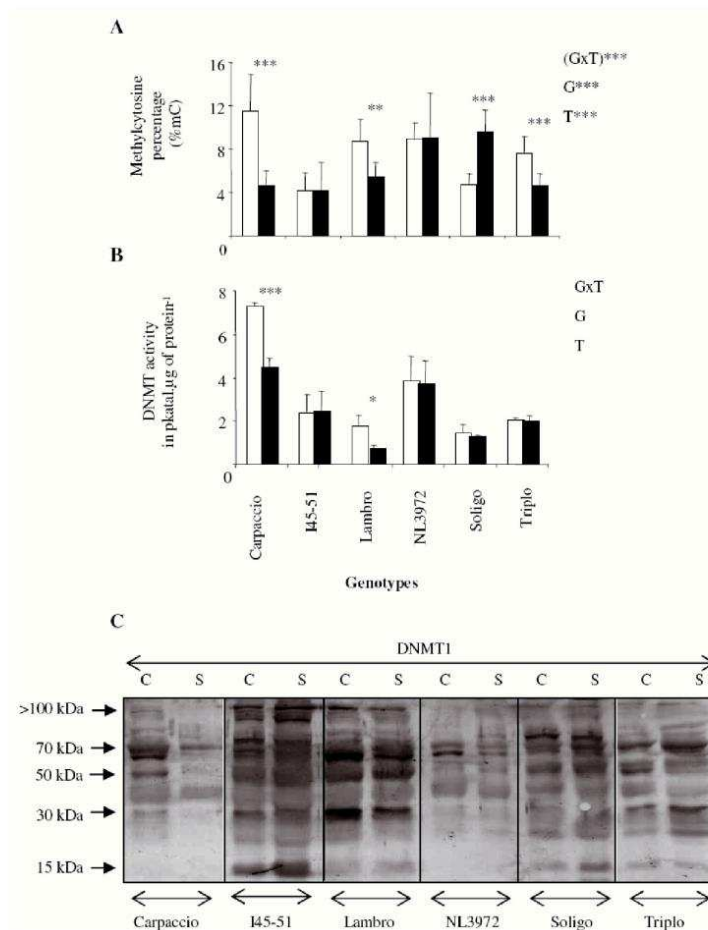


Figure 2. Genotypic variation and water deficit effect on traits related to DNA methylation in six *Populus deltoides* × *P. nigra* hybrids growing in glasshouse in well-watered (white bar) or in water deficit (black bar) conditions. Studied traits encompassed: (A) percentages of methylation of cytosine residues (%mC) and (B) total DNMT activity (see Materials and Methods section for details). For each graph, G indicates the genotype effect, T the treatment effect and (GxT) genotype by treatment effect. Means are accompanied by their standard errors SE ($n = 6$). Significant differences between well-watered and water deficit conditions are indicated by asterisk: * $P \leq 0.05$, ** $P \leq 0.01$ and *** $P \leq 0.001$ (C) immunodetection in the six genotypes of isoforms of DNMT class I. C (Control) represents plants in well-watered condition and S (Stressed) plants in water deficit condition. Sizes of the different isoforms are indicated in kiloDalton.

3.3. Genotypic variation of histone acetylation and water deficit effects

Using antibodies raised against acetylated forms of histones H₃ and H₄, it was possible to calculate the corresponding acetylation ratios (AcH₃ and AcH₄) by normalising with the amount of non-acetylated histone H₃ (Figs. 3A and 3B; see material and methods for details). Neither genotype nor treatment effects were observed for the ratios of AcH₃ or AcH₄; these were around 0.3, indicating a general hypoacetylation of these histones in poplar.

For Histone deacetylase (HDAC) activity, only a genotype effect was detected. Values ranged from 0.08% MAL deacety-

lated substrate $\text{min}^{-1} \mu\text{g}^{-1}$ of proteins for I45-51 to 11.10% for Carpaccio (Fig. 3C). While no treatment effect was detected by ANOVA for HDAC activity, significant differences were observed for five genotypes in response to water deficit. Thus, an increase in HDAC activity was observed for I45-51 and Triplo while Lambro, NL3972 and Soligo showed a decrease in HDAC activity.

Immunodetection using antibodies raised against each HDAC class described in plants (HD1A, HD1B and HD2, respectively) revealed several bands for each of them ranging from 80 kDa to 30 kDa (Fig. S3D). HD1A antibodies detected isoforms at 80, 75, 65, 45 and 30 kDa, while HD1B antibodies

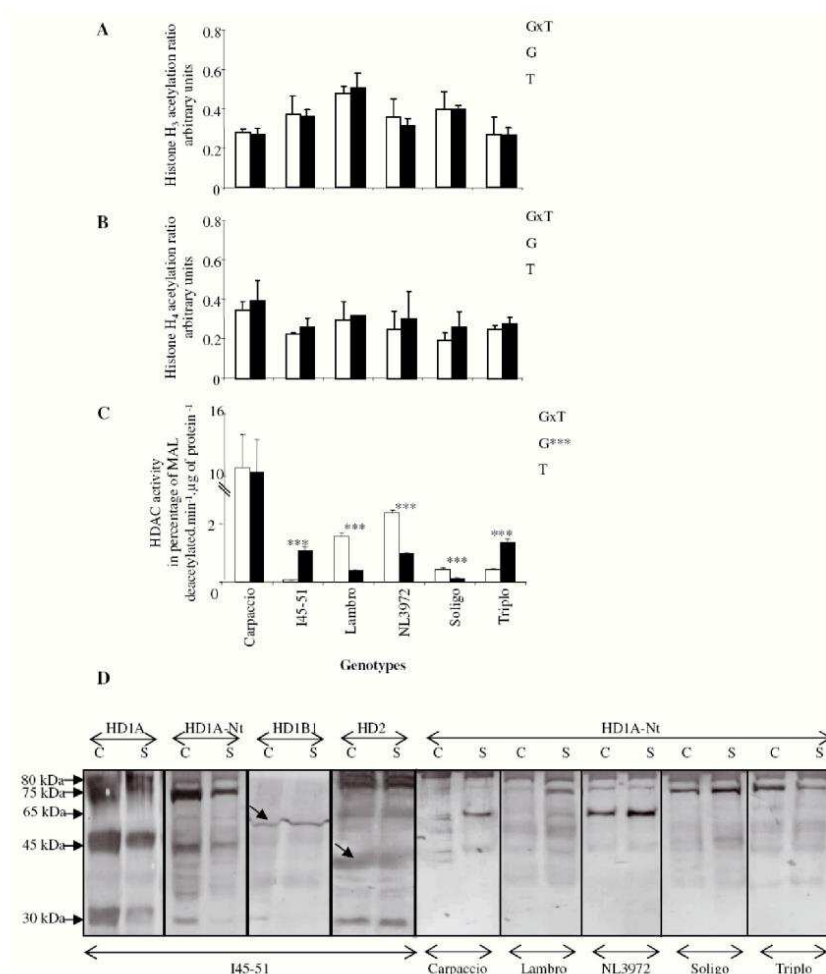


Figure 3. Genotypic variation and water deficit effect on traits related to histone acetylation in six *Populus deltoides* × *P. nigra* hybrids growing in glasshouse in well-watered (white bar) or in water deficit (black bar) conditions. Studied traits encompassed: (A) histone H₃ acetylation ratio (AcH₃). (B) histone H₄ acetylation ratio (AcH₄). (C) HDAC activity. For each graph, G indicates the genotype effect, T the treatment effect and (GxT) genotype by treatment effect. Means are accompanied by their standard errors SE ($n = 6$). Significant differences between well-watered and water deficit conditions are indicated by asterisk: * $P \leq 0.05$, ** $P \leq 0.01$ and *** $P \leq 0.001$. (D) immunodetection of the three classes of HDAC using anti-HD1A antibody, truncated anti-HD1A antibody (HD1A-Nterminal part of the protein), anti-HD1B1 antibody and anti-HD2 antibody. C (Control) represents plants in well-watered condition and S (Stressed) plants in water deficit condition. Sizes of the different isoforms are indicated in kiloDalton. Two arrows indicate specific bands mentioned in the result section.

revealed an additional specific band of about 50 kDa and HD2 one at 40 kDa. Blots with antibodies raised against the N-terminal part of HD1A (HD1A-Nt) presented a different profile, with additional bands of weak intensity, from those with HD1A antibodies. The most significant change between genotypes or treatments was in the relative amounts of the HD1A bands smaller than 80 kDa (Fig. 3D). A positive correlation was established between HDAC activity and the amount of the 80 kDa HDAC isoform expressed in percentage of total signal ($r = 0.63^*$; supplementary¹ figure S1B).

3.4. Correlation between morphological traits and epigenetic modifiers

Principal component analyses (PCA) were performed using genotypic means of the morphological traits measured in well-watered or water deficit conditions to characterize the variability of morphological traits between hybrid poplars (Fig. 4). Under the control conditions, the main plane of PCA explained 88.0% of the variability, with 60.7% for F_1 alone (Fig. 4A). The F_3 axis did not differentiate the variables (data

Table I. Linear correlations (Pearson coefficient, r) computed between morphological traits, $F1$ and $F2$ axes of PCA (see Fig. 4) and epigenetic variables. Only significant values are indicated by asterisk: * $P \leq 0.05$, ** $P \leq 0.01$ and *** $P \leq 0.001$. "Control" and "Stressed" represent plants in well-watered or in water deficit conditions, respectively.

	%mC		DNMT		AcH ₃		AcH ₄		HDAC	
	C	S	C	S	C	S	C	S	C	S
SB	0.89*									
RB										
Hgh	0.98***						0.83*	0.85*	0.81*	
TLA										
LN			0.98***	0.85*					0.94**	0.85*
SLA										
$F1$ axis	0.92**		0.86*						0.82*	
$F2$ axis					-0.90*	-0.83*				

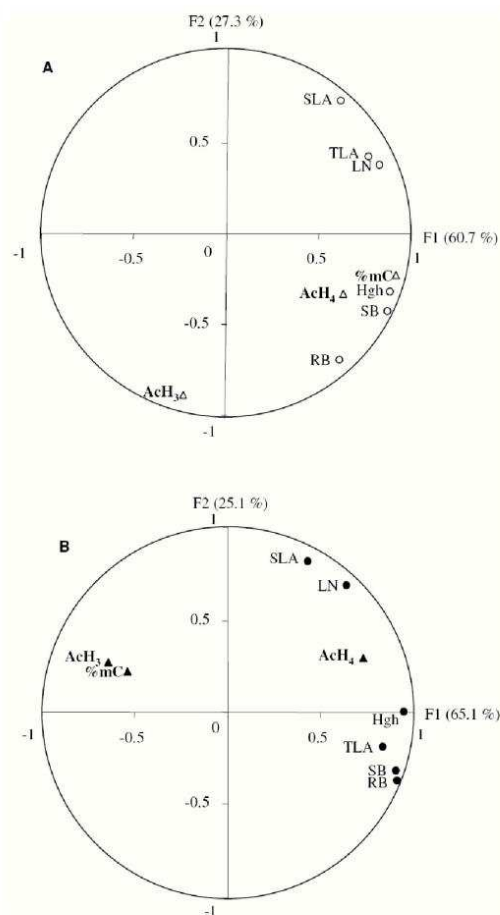


Figure 4. Principal Component Analyses (PCA) built using morphological traits (circles): stem biomass (SB), root biomass (RB), height (Hgh), total leaf area (TLA), leaf number (LN) and specific leaf area (SLA). Epigenetic variables (triangles): percentages of methylation of cytosine residues (%mC), histone H₃ and H₄ acetylation ratios (respectively AcH₃ and AcH₄) have been projected as supplementary variables using their Pearson coefficients with scores of PCA as coordinates. (A) in well-watered condition, (B) in water deficit condition.

not shown). The $F1$ axis was mainly defined by height, leaf number and stem biomass. In this group, most variables scaled positively. The $F2$ axis was mainly defined by specific leaf area. Linear correlations between methylcytosine percentage or histone acetylation ratios and scores of $F1$ and $F2$ axes were computed and used to project them as supplementary variables in these planes. Positive correlations were detected between methylcytosine percentage and stem biomass, height and scores of $F1$ axis from PCA (Tab. I and Fig. 5). AcH₃ was only negatively correlated to scores of $F2$ axis, while AcH₄ scaled positively with height (Tab. I; supplementary¹ figures S2A and S3A).

In the water deficit conditions, the main plane of PCA explained 90.2% of the variability, with 65.1% for $F1$ alone (Fig. 4B). The $F3$ axis did not differentiate the variables (data not shown). The $F1$ axis was defined by total leaf area, height, stem biomass and root biomass and most of these variables scaled positively (data not shown); the $F2$ axis was mainly defined by specific leaf area. The negative correlation between scores of $F2$ axis and AcH₃ was conserved (Tab. I; supplementary¹ figure S2B). AcH₄ still scaled positively with height but also with root biomass (Tab. I; supplementary¹ figure S3B and S3C).

4. DISCUSSION

4.1. Morphological traits distinguish hybrid poplars and are decreased by a moderate water deficit

Genotypes differed for morphological traits as already reported in previous studies (Bonhomme et al., 2008; Marron et al., 2005; Monclus et al., 2006). In response to water deficit a general decrease of most of these morphological traits was recorded. The leaf variables adjustment ensures a reduction of water losses but affects the photosynthesis activity and consequently the biomass production (Monclus et al., 2006; Roden et al., 1990). The absence of a genotype specific response may be explained by the fact that the constraint applied was moderate and short. Indeed, previous results show that in response to longer and stronger droughts poplars (and particularly those of this study) are differentially sensitive to water deficit (Marron et al., 2005; Monclus et al., 2006). This highlights that water deficit tolerance is often the result of a temporal integration of

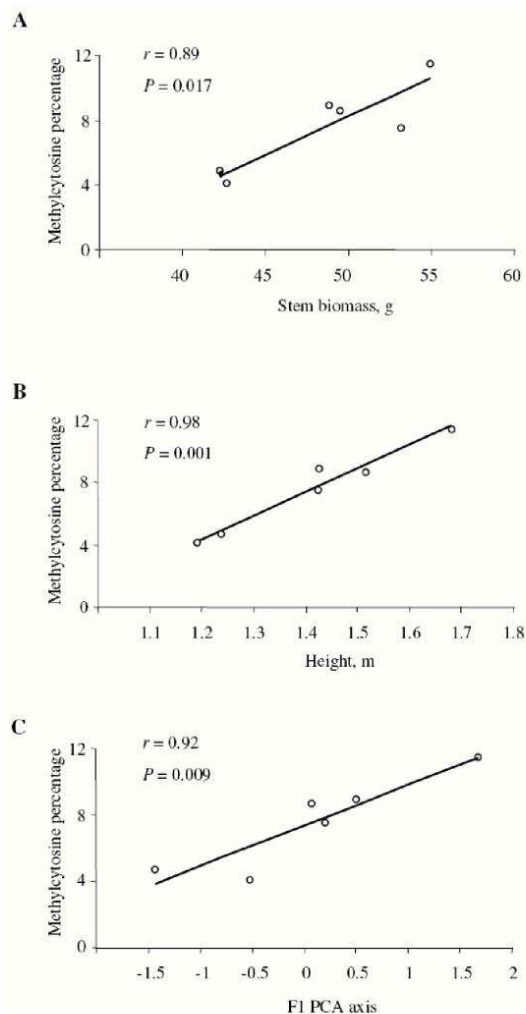


Figure 5. Relationship (linear correlation) between methylation percentages (%mC) and (A) stem biomass, (B) height and (C) F1 axis of PCA, measured on the apex of six poplar genotypes in well-watered condition.

fine physiological differences, not easily revealed in short term experiments. Possibly for similar reasons specific leaf area was not affected, either (Marron et al., 2005; Monclus et al., 2006).

The water deficit was moderate and the predawn leaf water potential (Ψ_{pd}) was unaffected (data not shown). Thus, Ψ_{pd} is one of the last parameters to be affected during an increasing water deficit (Bogeat-Triboulot et al., 2007). As the growth of aerial parts is the most sensitive parameter to water deficit, we assessed the impact of the moderate water deficit on morphological traits.

4.2. Differences in DNA methylation levels between genotypes

Because epigenetic patterns participate in the control of gene expression (Zhang et al., 2006; Zilberman et al., 2007) and are transmitted by mitosis, changes that arise in dividing cells in the shoot apex during environmental stress have the possibility of being propagated to daughter cells (Boyko and Kovalchuk, 2007; Chen and Tian, 2007; Gehring and Henikoff, 2007). Thus, epigenetic changes in shoot apex would explain, in part, variation for morphological traits.

Genotypic variation was observed for DNA methylation percentages that ranged from 3.9% for Soligo to 11.1% for Carpaccio, with no relationship with the characteristics of sex or ploidy of the six genotypes. Nevertheless, these percentages are relatively low compared with other plant species where the values ranged from 5% in *Arabidopsis* to more than 37% in *Helianthus* (Causevic et al., 2005). This is usually attributable to differences in repetitive DNA content among species, which is mostly associated with heterochromatin and is the major target of DNA methylation (Gehring and Henikoff, 2007).

Variations of DNA methylation have already been reported on *Arabidopsis* ecotypes for selected loci (Riddle and Richards, 2002; Vaughn et al., 2007). A similar degree of methylation polymorphism was observed, and gene expression was not generally affected by differences in DNA methylation. In contrast, in two *Arabidopsis* accessions and their reciprocal F1 hybrids Zhang et al. (2008) found a significant negative correlation between the degree of methylation variation within immediate upstream/downstream coding regions and the degree of expression variation. Furthermore, methylation polymorphism within genic regions showed a weak positive correlation with expression variation. These data suggest a possible relationship between natural CG methylation variation and gene expression variation that should be tested in poplars. Indeed, data on the *Arabidopsis* methylome have revealed that transposons and other repeats are heavily methylated, more than one-third of all genes contain methylation whereas fewer than 5% of expressed genes were shown to have methylated promoters (Zhang et al., 2006). Interestingly, genes methylated in transcribed regions are highly expressed and constitutively active, whereas promoter-methylated genes show a greater degree of tissue specific expression.

In addition to genetic variations among genotypes, the variations of global DNA methylation percentages observed in this study between poplar genotypes could have two main origins: (i) the heterogeneity of the plant material. Indeed, shoot apex of these genotypes could differ for their ratio of differentiated/undifferentiated cells. Two results seem to render this possibility less likely (but do not exclude it): the direct observation of shoot apex organization revealed no differences between genotypes (data not shown) and the histone H₃ and H₄ acetylation ratios were not significantly different between genotypes; (ii) the activity of DNA methyltransferases (Pavlopoulou and Kossida, 2007). Several DNMT isoforms were assessed in hybrid poplars in accordance with databases (Tuskan et al., 2006) and a process of limited proteolysis at the C-terminal part (Bernacchia et al., 1998;

Causevic et al., 2005). The relative amount of a 100 kDa DNMT class 1 isoform was positively correlated with the total DNMT enzymatic activity, in agreement with a previous report (Causevic et al., 2005). Nevertheless, no significant correlation was found between the total DNMT activity and the global methylation percentage. This suggests the involvement of distinct mitotic index between poplar genotypes, cell replication without methylation maintenance and/or variable DNA glycosylase-lyase activities.

4.3. Genotype specific changes in DNA methylation levels in response to water deficit

The water deficit also affected the levels of global DNA methylation percentages, with a significant interaction between genotype and treatment. Effects of environmental variations such as temperature, diseases, light, salinity, ABA and water stress on epigenetic parameters have already been shown in plants (Boyko and Kovalchuk, 2007; Chen and Tian, 2007). The few studies on water stress have analysed overall levels of DNA hypermethylation during osmotic stress in tobacco and potato cell cultures (Kovarik et al., 1997; Sabbah et al., 1995) or during dehydration in pea root tip cells (Labra et al., 2002). The two fold variation in DNA methylation levels in our system are in agreement with the previous reports. While poplar genotypes showed reduced growth in water deficit conditions, a significant genotype effect was observed for DNA methylation variations. This suggests that DNA methylation could participate to the fine-tuning of the reported control of gene expression in poplar during water stress (Plomion et al., 2006; Bogeat-Triboulot et al., 2007) including genotype-specific differences (Bonhomme et al., 2009). A complementary study should be done on the same plant material to confirm this hypothesis.

No correlation between DNA methylation and morphological traits was observed in response to a moderate water deficit certainly in relation to the genotype specific variations of DNA methylation. Nevertheless, morphological traits have already been shown to be genotype specific under intense water deficit (Monclus et al., 2006). It is then possible that epigenetic variations are more sensitive or precocious. Furthermore, genotype specific variations of molecular markers such as DNA methylation, mRNA and proteins reflect different strategies even if the morphological consequences are similar. Indeed, different kinetics and amplitude of the response between genotypes reflect their variable phenotypic plasticity (Monclus et al., 2006).

4.4. Histones H₃ and H₄ acetylation change neither between hybrids nor in response to moderate water deficit

Using polyclonal antibodies raised against acetylated lysine, we showed that acetylation ratios of histones H₃ and H₄ were consistent with previously-reported values (Causevic et al., 2006). No significant difference could be observed between genotypes or in response to a moderate water deficit.

This result apparently contrasts with the dynamic and reversible changes of histone H₃ acetylation and methylation reported during submergence of rice plants (Tsuji et al., 2006). Nevertheless, a detailed kinetic study of the variation of histone acetylation ratios on various lysine positions during water deficit as well as a gene specific analysis are needed to reach firm conclusions. Several HDAC isoforms were assessed in hybrid poplars in accordance with poplar databases (Tuskan et al., 2006) and a limited proteolysis of the C-terminal part previously reported in other plants (Causevic et al., 2006; Loidl, 2004; Pipal et al., 2003). The relative amount of a 80 kDa HDAC isoform was positively correlated to the total HDAC enzymatic activity. This is in agreement with previous reports showing that immunoprecipitation of the corresponding isoform is followed by a strong decrease of HDAC activity in the residual supernatant (Causevic, 2005; Causevic et al., 2006). This isoform and HDAC activity were affected by water deficit in poplar in accordance with a report in *Arabidopsis* about the role of this enzyme in the abscisic acid response (Sridha and Wu, 2006).

4.5. Epigenetic traits scaled with morphological traits

Linear correlations were established between DNA methylation percentages, histones H₃ or H₄ acetylation ratios and morphological traits. Significant linear correlations were observed between epigenetic variables and the scores of F1 or F2 axes of PCAs representing the variability of morphological traits. Furthermore, some traits were highly correlated such as %mC with height or stem biomass. The occurrence of correlations do not mean a possible causal link between these variables and the biological significance, if any, remains to be established. Nevertheless, considering the role of DNA methylation in the fine-tuning control of gene expression, we hypothesise that epigenetic traits may be determinants of poplar productivity. The identification of the genomic sequences that show differential methylation should now be envisaged using recently published approaches (Gentil and Maury, 2007; Zhang et al., 2006). This will permit the discovery of biomarkers that could represent valuable tools for poplar breeders.

Acknowledgements: The authors thank G. Moreau from University of Orléans (France) for technical assistance. We also thank N. Ningre, B. Clerc and J.M. Gioria for their contribution in managing the experiments. We are grateful to Professor P. Loidl from Innsbruck Medical University (Austria) for providing HDAC antibodies. Professors J.-L. Julien from University of Clermont-Ferrand (France), F.M. Delmotte from University of Orléans (France) and Dr S. Barnes from SES-VanderHave (Belgium) are gratefully acknowledged for their careful reading of the manuscript. We also thank C. Bastien from INRA Orléans (France) for helpful discussions.

REFERENCES

- Allis C.D., Jenuwein T., Reinberg D., and Caparros M.-L., 2007. Epigenetics, Cold Spring Harbor Laboratory Press, New York, 502 p.
- Bernacchia G., Primo A., Giorgetti L., Pitto L., and Cella R., 1998. Carrot DNA methyltransferase is encoded by two classes of genes with differing patterns of expression. *Plant J.* 13: 317–329.

Ann. For. Sci. 67 (2010) 208

D. Gourcilleau et al.

- Bogeat-Triboulot M.-B., Brosche M., Renaut J., Jouve L., Le Thiec D., Fayyaz P., Vinocur B., Witters E., Laukens K., Teichmann T., Altman A., Hausman J.-F., Polle A., Kangasjärvi J., and Dreyer E., 2007. Gradual soil water depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in *Populus euphratica*, a poplar growing in arid regions. *Plant Physiol.* 143: 876–892.
- Bonhomme L., Barbaroux C., Monclus R., Morabito D., Berthelot A., Villar M., Dreyer E., and Brignolas F., 2008. Genetic variation in productivity, leaf traits and carbon isotope discrimination in hybrid poplars cultivated on contrasting sites. *Ann. For. Sci.* 65: 503.
- Bonhomme L., Monclus R., Vincent D., Carpin S., Lomenech A.-M., Plomion C., Brignolas F., and Morabito D., 2009. Leaf proteome analysis of eight *Populus* × *euramericana* genotypes: genetic variation in drought response and in water-use efficiency involves photosynthesis-related proteins. *Proteomics* 9: 4121–4142.
- Boyko A. and Kovalchuk I., 2007. Epigenetic control of plant stress response. *Environ. Mol. Mutagen.* 49: 61–72.
- Braatne J.H., Hinckley T.M., and Stettler R.F., 1992. Influence of soil water on the physiological and morphological components of plant water balance in *Populus trichocarpa*, *Populus deltoides* and their F1 hybrids. *Tree Physiol.* 11: 325–339.
- Causevic A., 2005. Étude comparative de quelques paramètres du contrôle épigénétique en relation avec le développement de lignées de betterave sucrière. Thèse de doctorat. Faculté des Sciences, Université Orléans, France, 170 p.
- Causevic A., Delaunay A., Ounnar S., Righezza M., Delmotte F.M., Brignolas F., Hagège D., and Maury S., 2005. DNA methylating and demethylating treatments modify phenotype and cell wall differentiation state in sugarbeet cell lines. *Plant Physiol. Biochem.* 43: 681–691.
- Causevic A., Gentil M.-V., Delaunay A., El Soud W.A., Garcia Z., Pannetier C., Brignolas F., Hagège D., and Maury S., 2006. Relationship between DNA methylation and histone acetylation levels, cell redox and cell differentiation states in sugarbeet. *Planta* 224: 812–827.
- Chen Z.J. and Tian L., 2007. Roles of dynamic and reversible histone acetylation in plant development and polyploidy. *Biochim. Biophys. Acta* 1769: 295–307.
- Gehring M. and Henikoff S., 2007. DNA methylation dynamics in plant genomes. *Biochim. Biophys. Acta* 1769: 276–286.
- Gentil M.-V. and Maury S., 2007. Characterization of epigenetic biomarkers using new molecular approaches. In: Varshney R. and Tuberosa R. (Eds.), *Genomics-Assisted Crop Improvement: Genomics Approaches and Platforms*, Springer, Dordrecht, pp. 351–377.
- Kovarik A., Koukalová B., Bezdík M., and Opatrný Z., 1997. Hypermethylation of tobacco heterochromatic loci in response to osmotic stress. *Theor. Appl. Genet.* 95: 301–306.
- Labra M., Ghiani A., Citterio S., Sgorbati S., Sala F., Vannini C., Ruffini-Castiglione M., and Bracale M., 2002. Analysis of cytosine methylation pattern in response to water deficit in pea root tips. *Plant Biol. (Stuttg.)* 4: 694–699.
- Loidl P., 2004. A plant dialect of the histone language. *Trends Plant Sci.* 9: 84–90.
- Marron N., Villar M., Dreyer E., Delay D., Boudouresque E., Petit J.M., Delmotte F.M., Guehl J., and Brignolas F., 2005. Diversity of leaf traits related to productivity in 31 *Populus deltoides* × *Populus nigra* clones. *Tree Physiol.* 25: 425–435.
- Monclus R., Dreyer E., Villar M., Delmotte F.M., Delay D., Petit J.M., Barbaroux C., Le Thiec D., Bréchet C., and Brignolas F., 2006. Impact of drought on productivity and water use efficiency in 29 genotypes of *Populus deltoides* × *Populus nigra*. *New Phytol.* 169: 765–777.
- Pavlopoulou A. and Kossida S., 2007. Plant cytosine-5 DNA methyltransferases: structure, function, and molecular evolution. *Genomics* 90: 530–541.
- Pipal A., Goralik-Schramel M., Lusser A., Lanzanova C., Sarg B., Loidl A., Lindner H., Rossi V., and Loidl P., 2003. Regulation and processing of maize histone deacetylase Hda1 by limited proteolysis. *Plant Cell.* 15: 1904–1917.
- Plomion C., Lalanne C., Claverol S., Meddour H., Kohler A., Bogeat-Triboulot M.-B., Barre A., Le Provost G., Dumazet H., Jacob D., Bastien C., Dreyer E., de Daruvar A., Guehl J.-M., Schmitter J.-M., Martin F., and Bonneau M., 2006. Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics* 6: 6509–6527.
- Riddle N.C. and Richards E.J., 2002. The control of natural variation in cytosine methylation in *Arabidopsis*. *Genetics* 162: 355–363.
- Roden J., van Volkenburgh E., and Hinckley T.M., 1990. Cellular basis for limitation of poplar leaf growth by water deficit. *Tree Physiol.* 6: 211–219.
- Sabbah S., Raise M., and Tal M., 1995. Methylation of DNA in NaCl-adapted cells of potato. *Plant Cell Rep.* 14: 467–470.
- Sridha S. and Wu K., 2006. Identification of AtHD2C as a novel regulator of abscisic acid responses in *Arabidopsis*. *Plant J.* 46: 124–133.
- Tschaplinski T.J., Tuskan G.A., and Gunderson C.A., 1994. Water-stress tolerance of black and eastern cottonwood clone and four hybrid progeny. I. Growth, water relations and gas exchange. *Can. J. For. Res.* 24: 364–371.
- Tsuji H., Saika H., Tsutsumi N., Hirai A., and Nakazono M., 2006. Dynamic and reversible changes in histone H3-Lys4 methylation and H3 acetylation occurring at submergence-inducible genes in rice. *Plant Cell Physiol.* 47: 995–1003.
- Tuskan G.A. et al., 2006. The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604.
- Vaughn M.W., Tanurdzic M., Lippman Z., Jiang H., Carrasquillo R., Rabinowicz P.D., Dedhia N., McCombie W.R., Agier N., Bulski A., Colot V., Doerge R.W., and Martienssen R.A., 2007. Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biology* 5: e174.
- Zhang X., Yazaki J., Sundaresan A., Cokus S., Chan S.W.-L., Chen H., Henderson I.R., Shinn P., Pellegrini M., Jacobsen S.E., and Ecker J.R., 2006. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 126: 1189–1201.
- Zhang X., Shiu S., Cal A., and Borevitz J.O., 2008. Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. *PLoS Genetics* 4: e1000032. Doi:10.1371/journal.pgen.1000032.
- Zilberman D., Gehring M., Tran R.K., Ballinger T., and Henikoff S., 2007. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics* 39: 61–69.

Ann. For. Sci. 67 (2010) 208

Online Material

2

D. Gourcilleau et al., *Online Material*

Supplementary figures

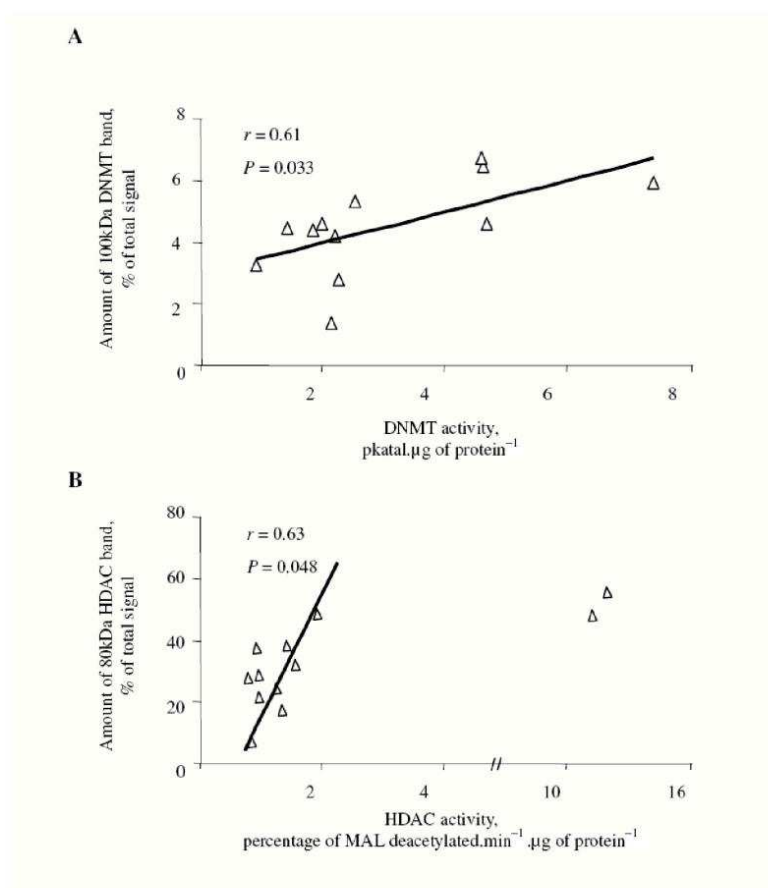


Figure S1. Relationship (linear correlation) measured on the apex of the six poplar genotypes in well-watered and in water deficit conditions between (A) DNMT activity and amount of DNMT band at 100 kDa (expressed in percentage of total signal, see materials and methods section for details); (B) HDAC activity and the ratio between the amount of HDAC band at 80 kDa (expressed in percentage of total signal, see materials and methods section for details). Regarding the out of range values for Carpaccio, linear correlation was computed only for the five other genotypes.

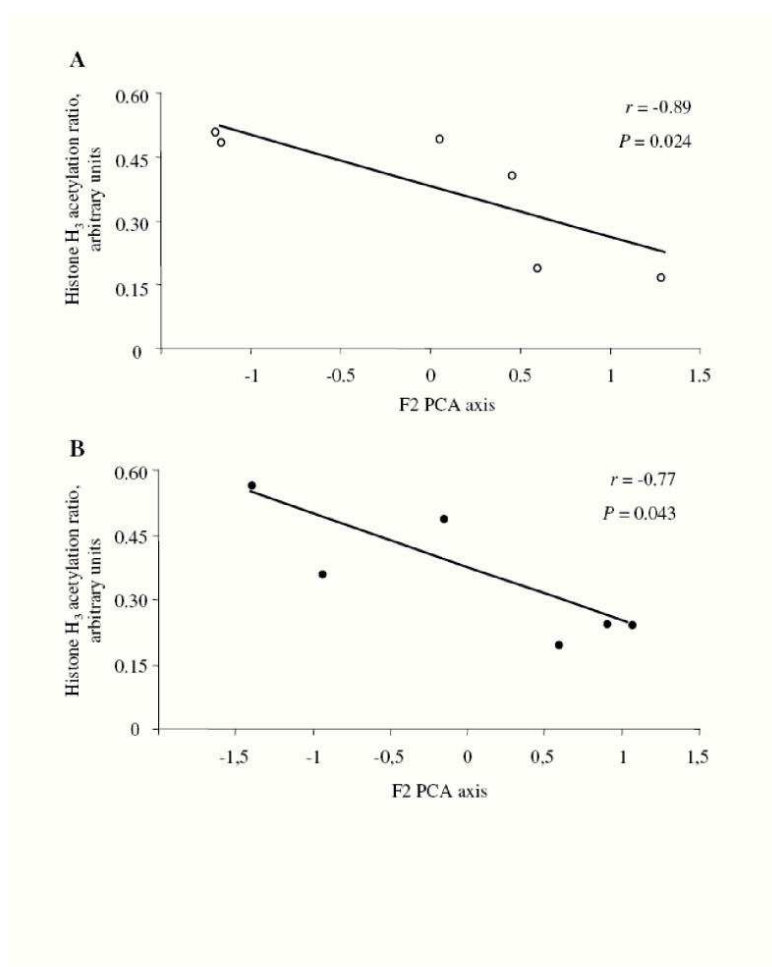


Figure S2. Relationship (linear correlation) between ratio of acetylated histone H₃ (AcH₃) and F2 axis of PCA measured on the apex of the six poplar genotypes (A) in well-watered condition (white circle) and (B) water deficit condition (black circle).

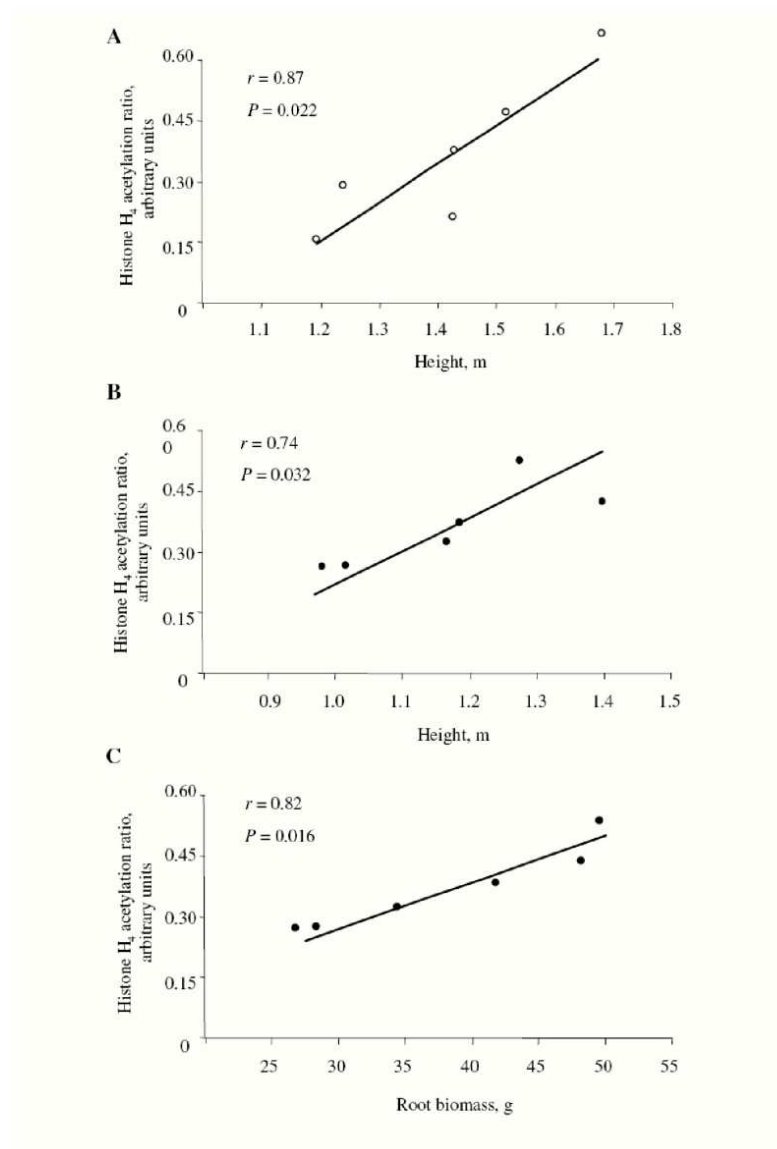


Figure S3. Relationship (linear correlation) between ratio of acetylated histone H₄ and Height measured on the apex of the six poplar genotypes (A) in well-watered condition (white circle), (B) in water deficit condition (black circle) and (C) Root biomass in water deficit condition (black circle).

3. Travaux complémentaires

Nos précédents résultats (Gourcilleau *et al.*, 2010) ont été obtenus sur le méristème apical caulinaire qui est le site de la morphogenèse de la tige feuillée. Cette zone présente d'intenses divisions cellulaires au cours desquelles une transmission mitotique de marques chromatinienne serait possible ce qui correspondrait alors à un phénomène épigénétique (Roudier *et al.*, 2009 ; voir introduction). De plus, le méristème organise la formation des futurs organes aériens et ce sous contrôle environnemental comme par exemple la morphogenèse des feuilles lors d'une sécheresse (Marron *et al.*, 2002, 2003). Plusieurs questions se posent alors, telles que :

1. Les primordia foliaires et les feuilles qui émergent de l'apex au cours d'un déficit hydrique sont-ils aussi affectés par des remaniements épigénétiques ?
2. Les remaniements épigénétiques suite à un déficit hydrique sont-ils stables dans le temps chez le peuplier, plante pérenne ?
3. La corrélation entre la méthylation de l'ADN dans l'apex et la production de biomasse est-elle retrouvée chez d'autres hybrides de peuplier ?

Afin de tester ces hypothèses et pour compléter nos données obtenues sur l'apex caulinaire, nous avons entrepris d'étudier les variations de méthylation de l'ADN au cours de l'ontogenèse de la feuille lors d'un déficit hydrique mais également de suivre la marque de méthylation plusieurs mois après une sécheresse chez de nouveaux hybrides. Ces travaux ont pu profiter de dispositifs expérimentaux disponibles à l'INRA de Nancy et d'Orléans et qui sont déjà publiés (Cohen *et al.*, 2010 ; Fichot *et al.*, 2010).

3.1. Ontogenèse de la feuille

La méthylation globale de l'ADN a été mesurée sur les feuilles du rang 0 (défini comme la première feuille en partant de l'apex d'au moins 20 mm de longueur) au rang 10, en condition hydrique non limitante ou en déficit hydrique modéré (comparable à Gourcilleau *et al.*, 2010). Deux hybrides *P. × euramericana*, 'Carpaccio' et 'Soligo', présentant une tolérance à la sécheresse ainsi qu'un comportement épigénétique très différents dans les expériences publiées dans Gourcilleau *et al.* (2010), ont été étudiés. Le dispositif expérimental (Cohen *et al.*, 2010) incluait entre autre un déficit hydrique de 12 jours avec $10\% < REW < 20\%$ (severe stress dans la Figure V.2A). Le stress n'a eu aucun effet sur le potentiel hydrique foliaire de base. En revanche, une diminution de la vitesse de croissance, de la conductance stomatique et de l'assimilation nette de CO₂ a été observée (Cohen *et al.*, 2010; Figure V.2B-C).

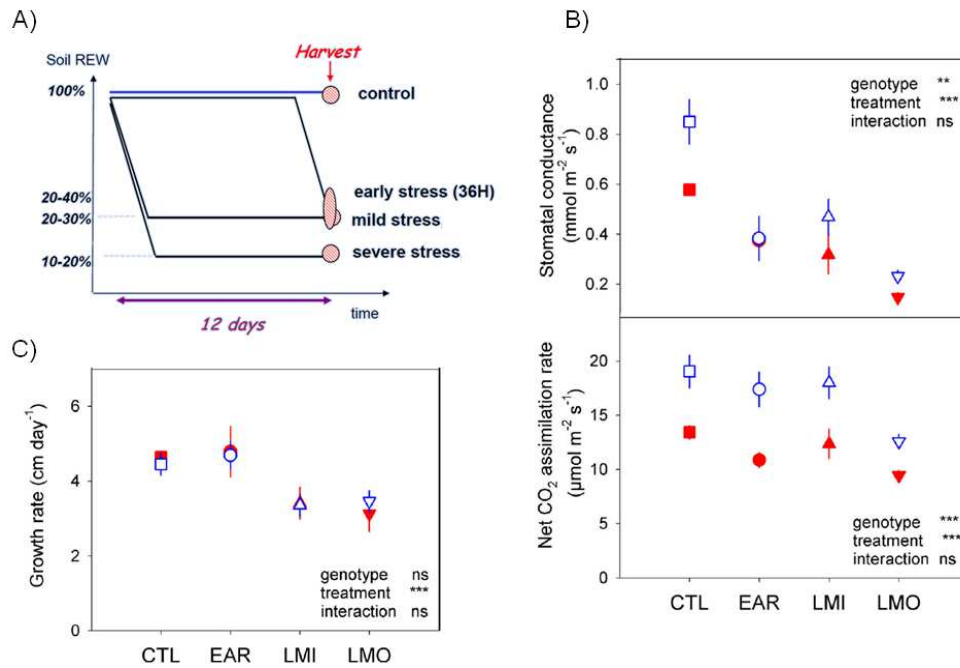


Figure V.2. Réponse écophysiolgique de deux hybrides *P. × euramericana* à des déficits hydriques d'intensité et durée variables. A) Suivi du Soil REW : *Soil Relative Extractable Water* (réserve relative en eau du sol) lors de l'expérience. B) Conductance stomatique (haut) et assimilation nette de CO₂ (bas). C) Vitesse de croissance. Les points rouges représentent le clone 'Carpaccio' et les bleus 'Soligo'. CTL : *Control*, EAR : *Early stress*, LMI : *Mild stress*, LMO : *Severe stress*. Les moyennes sont exprimées ± erreur standard, $n = 6$. La significativité des effets principaux de l'ANOVA 2 facteurs est donnée (génotype, traitement, interaction génotype × traitement). Les niveaux de significativité sont : * $P < 0,05$, ** $P < 0,01$, *** $P < 0,001$. ns, non significatif. D'après Cohen *et al.*, 2010.

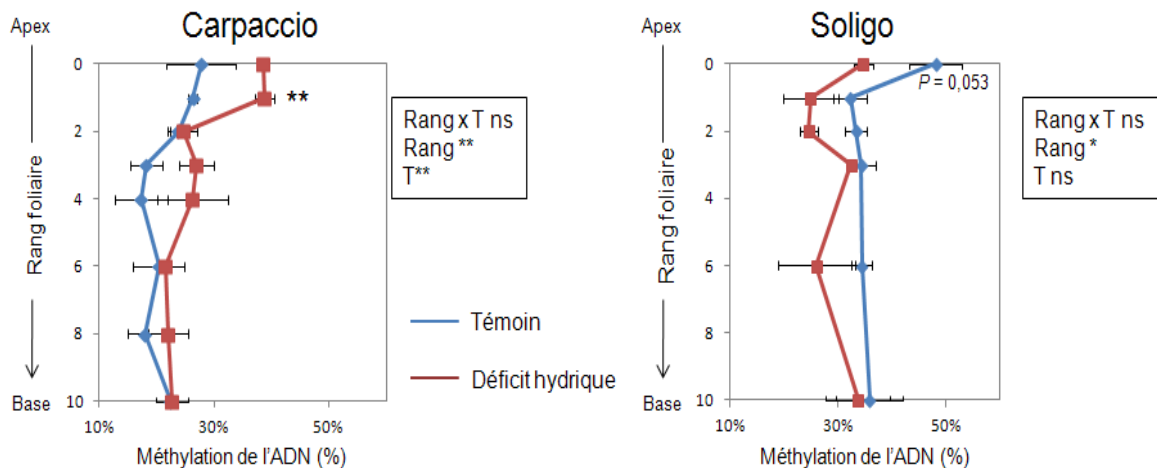


Figure V.3. Niveau de méthylation globale de l'ADN dans les cellules des feuilles selon le rang foliaire chez les hybrides *P. × euramericana* 'Carpaccio' et 'Soligo'. Le témoin correspond à la condition *Control* (CTL) et le déficit hydrique à la condition *Severe stress* (LMO, voir Figure V.2). Les moyennes sont exprimées ± erreur standard, $n = 6$. La significativité des effets principaux de l'ANOVA 2 facteurs est donnée (rang ; T : traitement ; RangxT : interaction rang × traitement). Les niveaux de significativité sont : * $P < 0,05$, ** $P < 0,01$. ns, non significatif.

Chez le génotype 'Carpaccio', un effet du rang foliaire et du déficit hydrique sur le niveau de méthylation globale de l'ADN de la feuille a été observé (Figure V.3) avec une hyperméthylation des jeunes feuilles et plus particulièrement de la feuille de rang 1 en condition de déficit hydrique (40 % contre environ 25 % en témoin). Le pourcentage de méthylation globale dans les feuilles plus âgées était similaire quelle que soit la condition hydrique (environ 25%). Pour le génotype 'Soligo', un effet lié au rang foliaire uniquement a été détecté (valeurs autour de 35 %) même si pour la feuille de rang 0, la valeur de *P* était proche du seuil de significativité et que la contrainte induisait une hypométhylation globale (Figure V.3). Le comportement antagoniste des 2 génotypes pour la méthylation avait déjà été observé au niveau des apex (Gourcilleau *et al.*, 2010), ; il est ici retrouvé mais dans le sens opposé au niveau des feuilles. Néanmoins, ces données semblent indiquer que les modifications de la méthylation globale de l'ADN dans l'apex en réponse à un déficit hydrique affectent également les plus jeunes feuilles en formation, et ce de manière génotype dépendante, alors que les feuilles déjà formées ne sont pas affectées. Une étude récente a d'ailleurs montré que les feuilles formées lors d'un déficit hydrique présentaient une diminution de la densité de stomates impliquant la régulation de certains gènes lors de l'ontogenèse foliaire alors que les feuilles déjà mûres avant ce stress ne subissaient pas un tel changement morphologique (Hamanishi *et al.*, 2012).

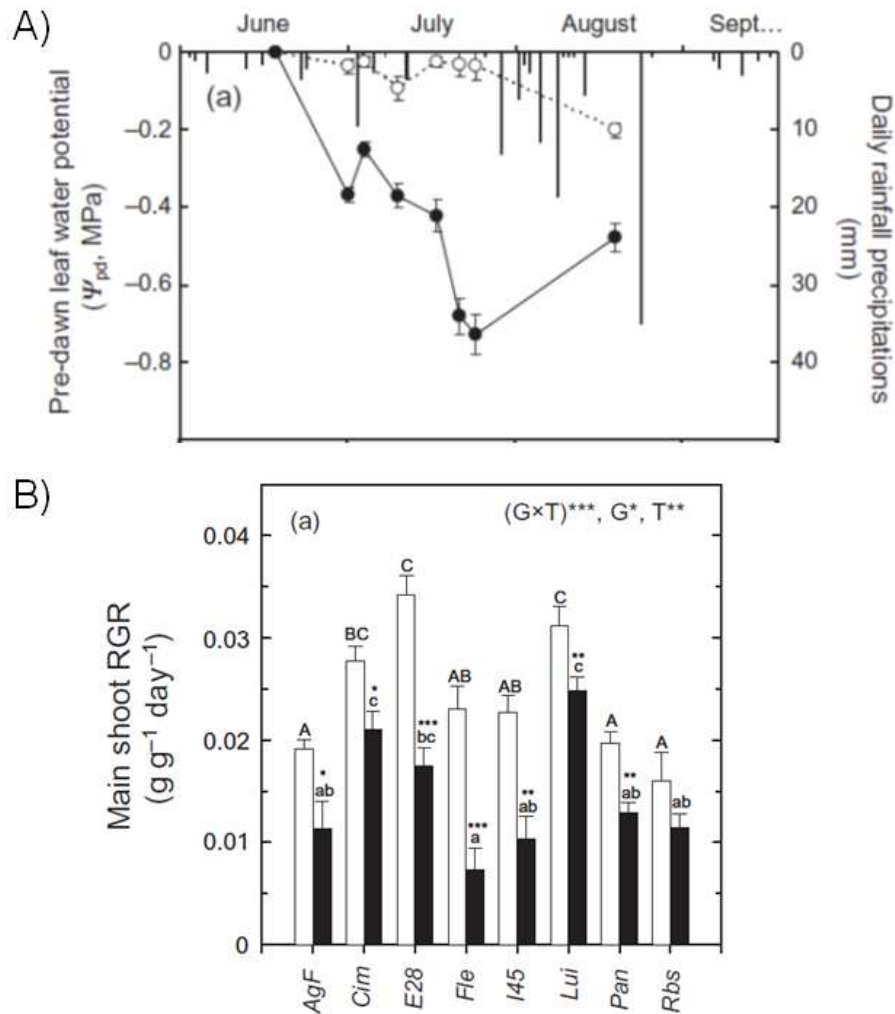


Figure V.4. Réponse écophysologique de huit cultivars hybrides *P. × euramericana* à un déficit hydrique estival. A) Suivi estival du potentiel hydrique de base (Ψ_{pd} , cercles) et des précipitations journalières (mm, barres verticales). Les moyennes de Ψ_{pd} sont exprimées $\pm$ erreur standard, $n = 15$. B) Vitesse de croissance relative (RGR) de la tige principale en condition hydrique non limitante (barres blanches) et en déficit hydrique (barres noires). Les moyennes sont exprimées $\pm$ erreur standard, $n = 5$. La significativité des effets principaux de l'ANOVA 2 facteurs est donnée (G : effet génotype ; T : traitement ; (G×T) : interaction génotype × traitement). Les lettres indiquent les différences significatives (test post hoc de Tukey), les majuscules étant pour les arbres bien irrigués et les minuscules pour les arbres exposés au déficit hydrique. Les niveaux de significativité sont : * $P < 0,05$, ** $P < 0,01$, *** $P < 0,001$. ns, non significatif. Abréviations des génotypes : AgF, 'Agathe_F' ; Cim, 'Cima' ; E28, 'Eco28' ; Fle, 'Flevo' ; I45, 'I45-51' ; Lui, 'Luisa_Avanzo' ; Pan, 'Pannonia' ; Rbs, 'Robusta'. D'après (Fichot et al., 2010).

3.2. Stabilité dans le temps de la marque épigénétique chez des hybrides *P. × euramericana*

La méthylation globale a été mesurée dans les apex caulinaires de huit cultivars hybrides *P. × euramericana* cultivés en pépinière (Figure V.4A) ayant subi ou non un déficit hydrique estival six mois avant la récolte des apex caulinaires (en janvier, apex dormants). Le déficit hydrique s'est traduit par une diminution progressive du potentiel hydrique foliaire de base au cours de la saison estivale ainsi qu'une diminution significative de la croissance des arbres, cette dernière étant toutefois génotype-dépendante (Fichot *et al.*, 2010; Figure V.4B).

Six mois après la sécheresse, un effet significatif, dépendant du génotype, a pu être retrouvé sur la méthylation globale de l'ADN de l'apex caulinaire (Figure V.5A). Plus particulièrement, une hypométhylation significative a été observée pour les génotypes 'Cima' et 'Eco28' qui ont de fortes vitesses relatives de croissance. En condition hydrique non limitante, une corrélation linéaire positive entre vitesse relative de croissance (RGR) et méthylation globale de l'ADN de l'apex caulinaire six mois plus tard a été détectée ($r_P = 0,71^*$; Figure V.5B). Chez les individus ayant subi un déficit hydrique, une corrélation négative, cette fois, a été observée ($r_P = -0,76^*$; Figure V.5C).

Ainsi, six mois après le déficit hydrique, en période de dormance, des différences de méthylation globale de l'ADN de l'apex caulinaire (avec des valeurs plus grandes que dans les apex en activité ; Gourcilleau *et al.*, 2010) entre les 2 conditions étaient observables chez certains génotypes. Ce résultat suggère la possibilité d'une mémoire épigénétique de la contrainte mais qui est fortement génotype-dépendante. Par ailleurs, comme précédemment montré dans Gourcilleau *et al.*, 2010, une corrélation positive entre méthylation globale de l'ADN de l'apex caulinaire et des caractères liés à la production de biomasse a été retrouvée en condition hydrique non limitante. Enfin, une corrélation négative entre méthylation globale de l'ADN de l'apex caulinaire et la vitesse de croissance en déficit hydrique a été trouvée certainement car les effets du déficit hydrique étaient plus forts et génotype-dépendants contrairement à Gourcilleau *et al.* (2010). Cette corrélation complexifie le rôle potentiel de la méthylation de l'ADN dans la morphogenèse de la tige feuillée chez le peuplier. Une raison qui pourrait expliquer le sens inverse des deux corrélations est que les variations de méthylation de l'ADN n'affectent pas les mêmes loci selon la condition hydrique étudiée.

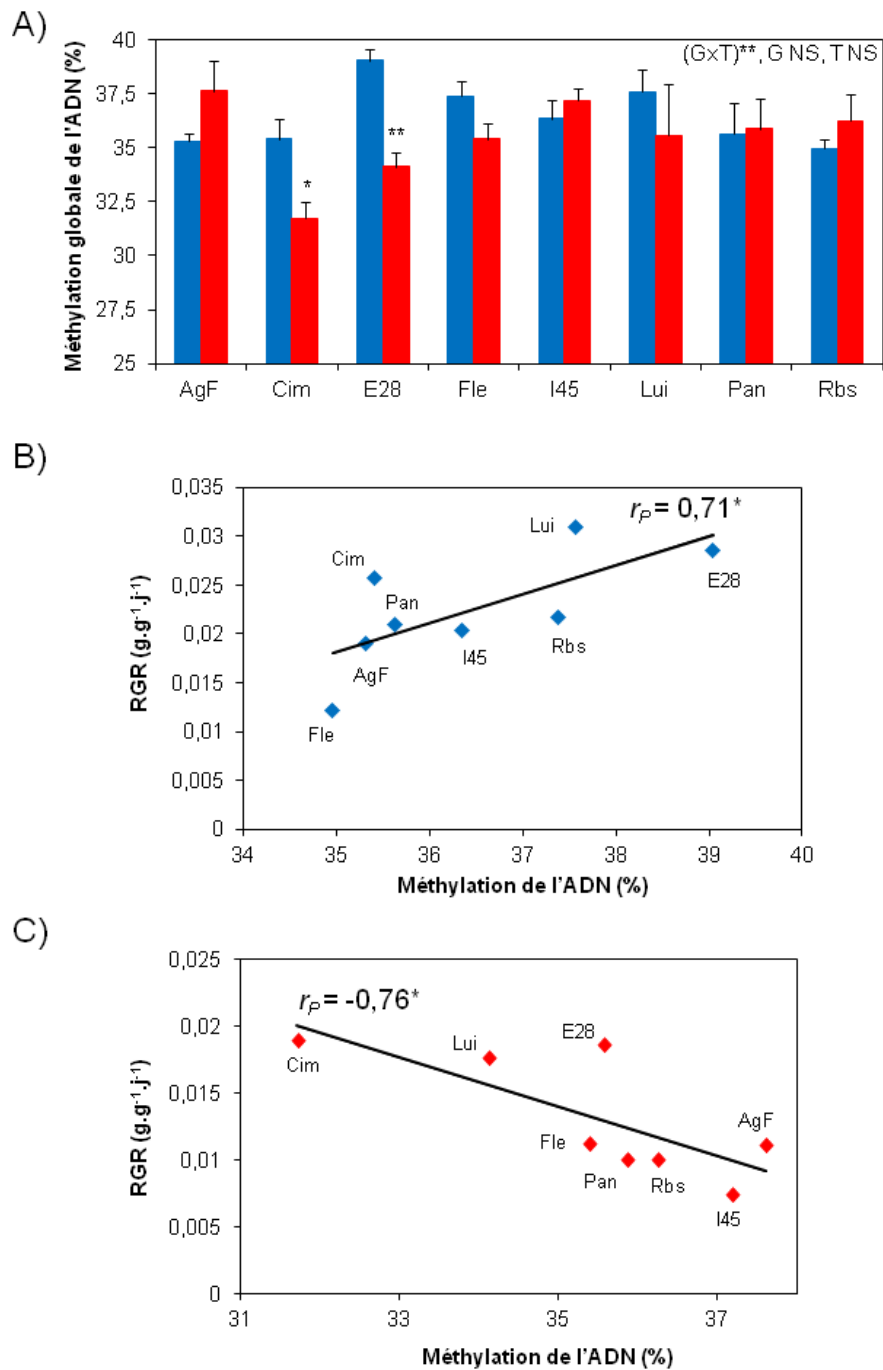


Figure V.5. Corrélation entre méthylation globale de l'ADN dans le méristème apical et production de biomasse. A) Niveau de méthylation globale de l'ADN dans le méristème apical de la tige principale chez huit hybrides *P. × euramericana* en condition hydrique non limitante (barres bleues) et en déficit hydrique (barres rouges). Les moyennes sont exprimées $\pm$ erreur standard, $n = 6$. La significativité des effets principaux de l'ANOVA 2 facteurs est donnée (G : effet génotype ; T : traitement ; (GxT) : interaction génotype × traitement). Les niveaux de significativité sont : * $P < 0,05$, ** $P < 0,01$. NS, non significatif. Les corrélations de Pearson entre vitesse de croissance relative (RGR) et méthylation globale de l'ADN sont données en condition hydrique non limitante (B) et en déficit hydrique (C), avec leurs coefficients (r_P) et leur significativité (* $P < 0,05$). Pour l'abréviation des génotypes, voir Figure V.4. Données de RGR issues de (Fichot *et al.*, 2010).

4. Perspectives et limites

Les travaux présentés dans ce chapitre avaient pour but d'étudier l'impact d'une sécheresse sur la méthylation globale de l'ADN chez différents cultivars hybrides de peuplier (*Populus deltoides* × *P. nigra*). Plus précisément, les objectifs de ce chapitre étaient d'évaluer les variations de la méthylation globale de l'ADN (i) dans les apex et les feuilles, (ii) en réponse à des variations de la disponibilité en eau, (iii) au sein de différents génotypes de peuplier (iv) immédiatement au cours de la contrainte ou plusieurs mois après et (v) de mettre en évidence une possible relation statistique entre la méthylation globale de l'ADN et des variables liées à la production de biomasse.

Ce travail a permis la première publication en épigénétique sur le peuplier (Gourcilleau *et al.*, 2010) et a révélé une variation de la méthylation globale de l'ADN dans l'apex caulinaire selon la disponibilité en eau (condition hydrique optimale vs déficit hydrique, $\Delta\%mC = 13\%$) et au sein des hybrides étudiés ($\Delta\%mC = 8\%$), selon l'organe et son âge (jeunes feuilles vs feuilles mûres, $\Delta\%mC = 15\%$) et plusieurs mois après un déficit hydrique. En revanche, les variations observées dans des organes mûres comme les feuilles sont bien plus faibles quelle que soit la condition environnementale ou le génotype.

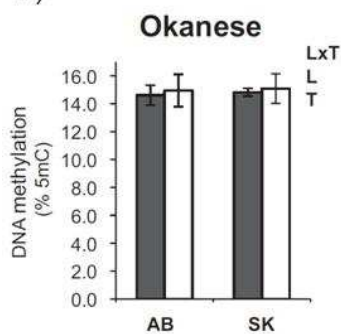
Récemment, une autre étude s'est intéressée chez le peuplier à la variabilité entre génotypes issus de différentes localisations géographiques au Canada et en fonction d'une contrainte hydrique (Raj *et al.*, 2011). Une faible gamme de variation dans les feuilles mûres a été retrouvée au sein de différents génotypes étudiés ($\Delta\%mC < 3\%$). Un effet de la contrainte génotype-dépendant a également été rapporté dans cette étude (Figure V.6). Enfin, cette étude a montré que suivant l'origine géographique de la plante ayant servi à faire les boutures, la méthylation de l'ADN était différente. Les auteurs concluent que les différences de méthylation de l'ADN sont parallèles à celles au niveau transcriptomique dans les feuilles lors d'une sécheresse suggérant une base épigénomique pour les divergences de transcriptomique qui sont génotype- et origine- ('history') dépendantes.

Nos résultats ouvrent deux perspectives principales. Tout d'abord, il s'agit de mieux caractériser la corrélation trouvée entre la méthylation de l'ADN globale dans l'apex caulinaire et des variables liées à la production de biomasse. Cette perspective est abordée dans le cadre du projet ANR Sylvabiom (http://www4.nancy.inra.fr/eef_eng/News/SYLVABIOM) coordonné par Jean-Charles Bastien de l'INRA d'Orléans (tâche 1.5 sous la responsabilité de S. Maury) et comprenant une soixantaine de génotypes de peuplier (*Populus deltoides* × *P. nigra*) plantés sur trois sites pédoclimatiques distincts et cultivés en taillis à courte ou très courte rotation (TCR et TTCR,

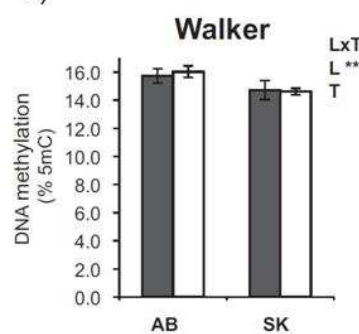
A)



B)



C)



D)

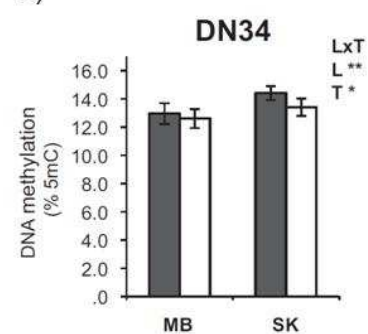


Figure V.6. Effet provenance sur la méthylation globale de l'ADN dans les feuilles de génotypes de peuplier hybrides (Raj *et al.*, 2011). A) Localisation géographique des trois pépinières d'origine (AB, SK et MB) et du *common garden* (Toronto) où a eu lieu l'expérience. Le niveau global de méthylation de l'ADN est montré pour les plants des génotypes Okanese (A), Walker (B) et DN34 (C) cultivés en condition hydrique non limitante (barres grises) ou en déficit hydrique (barres blanches). AB, SK et MB indique les pépinières d'où proviennent les boutures. Pour chaque graphique, L indique l'effet provenance, T l'effet traitement et L×T l'interaction entre effets provenance et traitement (* $P < 0,05$, ** $P < 0,01$). Les valeurs moyennes et leurs écart-types sont représentés ($n = 6$; trois réplicats biologiques × deux réplicats techniques).

respectivement). Ensuite, il paraît pertinent de caractériser à l'échelle génomique les loci affectés par les remaniements de la méthylation de l'ADN au cours d'un déficit hydrique. En effet, certaines études ont montré que les variations épigénétiques lors de stress affectent préférentiellement les séquences répétées de type transposon, cibles privilégiées de la méthylation de l'ADN et donc responsables en majorité du niveau global de méthylation d'un génome (Pecinka & Mittelsten Scheid, 2012; Bucher *et al.*, 2012). De ce fait, bien que modifiant celui-ci de manière significative, les variations globales de la méthylation de l'ADN observées lors d'un déficit hydrique n'auraient qu'un rôle mineur à jouer dans la plasticité développementale du peuplier. De plus, les corrélations observées entre production de biomasse et méthylation de l'ADN pourraient très bien être expliquées par des mécanismes cellulaires autres que la transcription des gènes comme la réplication de l'ADN ou la stabilité des télomères, impliquant également la méthylation de l'ADN (Moore *et al.*, 2011; Vanyushin & Ashapkin, 2011). Aussi, l'identification des loci affectés par les changements de méthylation lors de variations de la disponibilité en eau ainsi que l'étude de leur expression semblent indispensables pour préciser le rôle de la méthylation de l'ADN dans la plasticité phénotypique de la tige feuillée chez le peuplier. Cette perspective est abordée dans le chapitre suivant de ce manuscrit.

**Chapitre 3 : Caractérisation de loci affectés pour leur expression et leur
méthylation lors de variations de la disponibilité en eau du sol pour le
génotype 'Carpaccio' (*P. deltoides* × *P. nigra*)**

1. Présentation synthétique des travaux

Les changements climatiques prédits pour les prochaines décennies, comprenant des épisodes de sécheresse, sont une menace pour la santé des forêts et la production de bois (Intergovernmental Panel on Climate Change, 2007; Allen et al., 2010; Hamanishi & Campbell, 2011). Dans ce contexte, comprendre les mécanismes moléculaires de la plasticité phénotypique des arbres en réponse à des fluctuations de la disponibilité en eau du sol est primordial. Les travaux présentés dans le chapitre précédent ont montré des variations de la méthylation globale de l'ADN chez différents génotypes de peuplier hybrides (*Populus deltoides* × *P. nigra*) en réponse à un déficit hydrique, suggérant une possible implication de cette marque épigénétique dans la plasticité phénotypique en réponse à la sécheresse chez le peuplier (Gourcilleau et al., 2010). Cependant, afin de préciser ce rôle éventuel, il était nécessaire de caractériser les loci affectés par les variations de méthylation de l'ADN lors de fluctuations de la disponibilité en eau du sol.

Dans ce contexte, les objectifs de ce travail était de caractériser (i) les changements transcriptomiques et épigénomiques (méthylation de l'ADN) dans le méristème apical caulinaire en réponse à des variations de la disponibilité en eau du sol, (ii) les relations entre ces deux mécanismes et plus particulièrement en se focalisant sur le corps des gènes et (iii) les gènes dont les profils d'expression et/ou de méthylation changent en réponse à une variation de disponibilité en eau suggérant une implication dans la plasticité phénotypique.

A cette fin, le génotype 'Carpaccio' (*P. deltoides* × *P. nigra*) précédemment étudié dans Gourcilleau et al. (2010) a été soumis à un cycle de déficit hydrique modéré/réarrosage et sa réponse écophysologique a été étudiée. En parallèle, les variations épigénomiques (méthylation de l'ADN) et transcriptomiques dans le méristème apical caulinaire ont été suivies par des analyses de type microarray.

Selon leur expression dans les trois conditions hydriques (témoin, déficit hydrique et déficit hydrique suivi d'un réarrosage), les gènes ont été regroupés en clusters, certains étant spécifiquement induits lors du réarrosage ou d'autres ayant un profil spécifique du méristème apical en comparaison avec des données publiques de transcriptomique dans la feuille. Les changements de méthylation de l'ADN ont été observés dans l'ensemble du génome, le déficit hydrique étant associé à une hypométhylation modérée alors que la réirrigation a provoqué une forte hyperméthylation comparé à la condition témoin. Ces modifications épigénétiques ont particulièrement affecté le corps des gènes exprimés et les éléments *cis* régulateurs associés à la réponse à la sécheresse. Enfin, des gènes impliqués dans la signalisation des phytohormones étaient différenciellement exprimés et méthylés en

fonction de la disponibilité en eau, tels qu'un cluster de gènes codant pour des sulfotransférases (SOT). Cette famille d'enzymes affecte notamment des brassinostéroïdes (BR) pour leur inactivation ou intervient dans la biosynthèse de glucosinolates. Ces résultats suggèrent un contrôle épigénétique et transcriptomique de gènes impliqués dans la signalisation hormonale avec notamment l'inactivation du signal brassinostéroïdien lors du réarrosage et avant la reprise de croissance chez le peuplier.

2. Réponses du méthylome et du transcriptome à des variations de disponibilité en eau du sol (Lafon-Placette et al., en préparation pour *Plant Cell* : article 4)

Epigenomic and transcriptomic changes in the shoot apical meristem mediate developmental plasticity of *Populus euramericana* hybrid in response to variations in water availability through phytohormone signaling control

Clément Lafon-Placette^{1,2}, Alain Delaunay^{1,2}, Irène Hummel³, Béline Jesson⁴, Marie-Béatrice Bogeat-Triboulot³, Cécile Montgiraud^{1,2}, Franck Brignolas^{1,2}, Stéphane Maury^{1,2,*}

¹ Université d'Orléans, Faculté des Sciences, Laboratoire de Biologie des Ligneux et des Grandes Cultures (LBLGC), UPRES EA 1207, Orléans, 45067, France

² INRA, USC1328 Arbres et Réponses aux Contraintes Hydriques et Environnementales (ARCHE), Orléans, 45067, France

³ INRA, UMR1137 INRA-UHP Écologie et Écophysiologie Forestières, IFR 110 Génomique, EFABA, Champenoux, 54280, France

⁴ IMAXIO, Biopôle Clermont-Limagne, 63360 Saint-Beauzire, France

* To whom correspondence should be addressed. Tel: +33 (0)2 38 41 70 22; Fax: +33 (0)2 38 49 49 11; Email: stephane.maury@univ-orleans.fr

INTRODUCTION

Drought is a significant threat to forest health and agro-ecosystem productivity (Hamanishi & Campbell, 2011). Indeed, even if trees, as sessile organisms, evolved mechanisms to contend with variable water availability, this ability has limits that could be overtaken by predicted climate changes, including unprecedented long and severe periods of drought (Intergovernmental Panel on Climate Change, 2007; Allen *et al.*, 2010). To face successive drought episodes, the response to limited water supply alone might not be sufficient for trees and the ability to recover growth after rewatering, i.e. reversibility of drought effects, might also be important, depending in particular on drought intensity and extent of cell damage that occurred during stress (Gallé *et al.*, 2007; Xu *et al.*, 2010). In other words, phenotypic plasticity of trees, defined as the range of phenotypes that can express a genotype in response to environmental variations (Lande, 2009; Nicotra *et al.*, 2010), might have a great influence on their survival to predicted climate changes.

With the availability of its genome and its important natural genetic and phenotypic variations, *Populus* became a model tree (Tuskan *et al.*, 2006; Jansson & Douglas, 2007). Poplars (*Populus* spp.) are among the fastest growing trees in temperate latitudes. Their high productivity is associated with large water requirements. As a consequence, poplars are among the trees with the highest sensitivity to water deficit. The concept of water deficit tolerance, when applied to cultivated tree species such as poplars, has been defined as the ability to limit the decrease of biomass production in response to a moderate water deficit (Passioura, 2002). This tolerance might be mediated by phenotypic adjustments to limited water supply, allowing the maintenance of CO₂ assimilation. Poplars phenotypic plasticity in response to variations in water availability concerns several physiological and morphological traits, including stomatal conductance and formation (Liu & Dickmann, 1992; Blake *et al.*, 1996; Harvey & van den Driessche, 1997; Marron *et al.*, 2002; Hamanashi *et al.*, 2012), osmotic adjustment (Gebre *et al.*, 1994; Marron *et al.*, 2002), leaf expansion and morphogenesis, leaf abscission and adjustment of the root : shoot ratio (Liu & Dickmann, 1992; Chen *et al.*, 1997; Ibrahim *et al.*, 1997; Tschaplinski *et al.*, 1998; Marron *et al.*, 2003).

In a molecular point of view, plants drought transcriptional response is mediated through two regulatory pathways, one involving the phytohormone abscissic acid (ABA), called ABA-dependent, and the other one being ABA-independent (Seki *et al.*, 2007). The ABA-dependent regulatory network involves transcription factors activated by ABA that bind on ABA-Responsive Elements (ABRE) promoter motives, thus called ABRE binding factors (ABF). The ABA-independent pathway is similar, with promoters containing Dehydration-Responsive Elements (DRE) bound by DRE binding factors (DREB) (Des Marais & Juenger, 2010). Genes shown to be regulated through these two networks belong to a wide range of functions, from gene transcription regulation (like MYB transcription factors) to cellular component stabilization such as the Late Embryogenesis Abundant (LEA) genes (Abe *et al.*, 2003; Bies-Ethève *et al.*, 2008). In addition, other phytohormones like brassinosteroids act synergistically with ABA, by inducing both ABA biosynthesis and expression of ABA-responsive genes (Choudhary *et al.*, 2012). Concerning molecular pathways that allow growth recovery after a drought stress, few is known except antioxydative enzymes and aquaporins that are thought respectively to

decrease ROS accumulated during the stress and to restore water transport in the plant (Berta *et al.*, 2010; Xu *et al.*, 2010; Filippou *et al.*, 2011). Transcriptomic and proteomic responses to drought are well documented for roots and leaves, organs responsible for water uptake and loss (Plomion *et al.*, 2006; Bonhomme *et al.*, 2009b; Cohen *et al.*, 2010; Raj *et al.*, 2011), but site of shoot morphogenesis, the shoot apical meristem (SAM), has been less studied whereas it might control shoot developmental response to variations in water availability.

In plants, epigenetic reprogramming occurs at developmental key steps and in response to environmental stimuli (Feng *et al.*, 2010b). These epigenomic events play important roles in genome protection, control of gene expression and inheritance via chromatin structural reworking (Teixeira & Colot, 2010). Among epigenetic marks, DNA methylation shows high stability through mitosis and meiosis and has been well investigated. In plants, methylated cytosines are found mainly in CG dinucleotides and to a lower extent in CHG and asymmetric CHH contexts (where H is A, T, or C) and are propagated through DNA replication respectively by DNA METHYLTRANSFERASE 1 (MET1), the plant-specific CHROMOMETHYLASE 3 (CMT3) enzyme and small RNAs pathway-mediated DOMAINS REARRANGED METHYLTRANSFERASE (DRM) action (Goll & Bestor, 2005; Teixeira & Colot, 2010). Unmethylated CG and non-CG sites can also be *de novo* methylated by the DRM enzymes and conversely methylated cytosines can be replaced by unmodified bases through the action of DEMETER (DME) and DEMETER-LIKE (DML) proteins (Penterman *et al.*, 2007). With novel technologies allowing the survey of DNA methylation at genome scale, studies in *Arabidopsis* revealed that DNA methylation is concentrated in heterochromatin and repeats, but is also present in 33 % of genes (Zhang *et al.*, 2006; Cokus *et al.*, 2008). This pattern was found in other Eukaryotes but despite its high conservation, the functional role of gene body methylation still needs to be clarified (Feng *et al.*, 2010; Lafon-Placette *et al.* in press; Zemach *et al.*, 2010).

Recent studies of *P. trichocarpa* methylome revealed particularities of this model such as a higher CHG methylation level compared to other plants (Feng *et al.*, 2010a) and showed that gene body DNA methylation is extensive in open chromatin state, linked to structural gene characteristics and correlated to tissue specific gene expression (Lafon-Placette *et al.*, in press; Vining *et al.*, 2012). Interestingly, global DNA methylation in shoot apex has previously been shown to vary between *P. × euramericana* hybrids and in response to a water deficit in a genotype-dependent way (Gourcilleau *et al.*, 2010). Furthermore, differences in transcript abundance patterns in response to drought in leaves were based on differences in geographic origin of clones and paralleled the differences in global DNA methylation, suggesting an epigenomic basis for the clone history-dependent transcriptome divergence (Raj *et al.*, 2011). However, the nature and extent of loci affected by DNA methylation and expression changes depending on water availability in the shoot apical meristem remain unknown. In addition, few is known about the molecular pathways allowing growth recovery after a drought stress and such an investigation could help to elucidate the role of DNA methylation in developmental plasticity.

In this context, our objectives were to characterize i) SAM transcriptomics and DNA methylation changes at the genomic scale in response to variations of water availability, ii)

relationships between these two processes particularly by focusing on gene-body DNA methylation iii) molecular actors with differential expression and/or methylation patterns involved in the phenotypic plasticity towards water availability. For this purpose, we studied the physiological effects of a moderate water deficit/rewatering cycle on 'Carpaccio', a *P. × euramericana* hybrid, and used a microarray approach to assess variations both in SAM transcriptome and methylome. Clusters of gene expression were identified for each water status, some being specifically up-regulated during re-irrigation or being specific to the SAM compared to available leaf transcriptomic data. DNA methylation changes were genome-wide, water deficit being associated to a moderate hypo-methylation while re-irrigation exhibited large hyper-methylation compared to control conditions and particularly affected gene body of expressed or promoters containing water stress boxes genes. Finally, several genes involved in phytohormone signaling were differentially methylated and expressed such as a cluster of *SULFOTRANSFERASE* (SOT) encoding genes containing brassinosteroid SOT playing a potential role in the inactivation of brassinosteroids. This cluster was up-regulated and targeted by gene-body DNA methylation after rewatering, suggesting a role in the inactivation of phytohormone water stress signal before growth recovery.

MATERIAL AND METHODS

Plant material, growth conditions and treatments

Unrooted dormant stem cuttings of *Populus* × *euramericana* (*P. deltoides* × *P. nigra*) 'Carpaccio' genotype were obtained from INRA nursery (Nancy, France in 2008). Ten-years old mother plants are cut off each spring. Cuttings were 22 cm long and 0.5-2 cm wide and were harvested in February from plants grown in INRA nursery. Unrooted cuttings (18 for each hybrid) were then stocked 1 month in 2°C cold dark room. Before potting, cuttings soaked in water in vertical position to favor rooting. Rooted cuttings of *Populus* × *euramericana* (*P. deltoides* × *P. nigra*) 'Carpaccio' genotype (6 cuttings per treatment: 'Control', 'Water deficit' and 'Re-irrigation') were planted in 20L pots containing a mixture of peat and sand (50/50, v/v) amended with magnesian chalk (60 g/100 L) and fertilized with a low releasing fertilizer (Nutricote T100, 13/13/13/2, N/P/K/Mg + trace elements). Plants were placed in a greenhouse (Nancy, France in April 2008) exposed to natural daylight (ranging from 350 to 900 μmol m⁻² s⁻¹) and where temperature was maintained in the range 19–26 °C with humidity oscillating between 50 and 75%. Plants were watered to field capacity three times a day for 5 weeks, until plant height was between 0.6 and 0.9 m. Soil water availability was determined following relative extractable water (REW), which was maintained in the range of 80-86% (minimum-maximum) for the fully watered treatment and 17–23% for the water deficit by controlled watering every morning. A moderate water deficit was applied for 14 days to the cuttings ('water deficit') while cuttings in 'control' condition remained fully watered. For the 'Re-irrigation' treatment, a moderate water deficit was applied for 9 days and for the 5 following days cuttings were fully watered. For each treatment, the six cuttings were assigned to physiological monitoring. Leaf predawn water potential, leaf full turgor osmotic pressure, height growth rate and gas exchange were measured as previously described (Bogeat-Triboulot *et al.*, 2007). At the end of the 14 days treatment, active buds from each individual of each treatment were collected, cleared from leaves to isolate shoot apex. Shoot apices were then cut in their half with a scalpel and cleared from all visible differentiated tissues under binocular magnifying glass as described in Lafon-Placette *et al* (*in press*). After this step, samples that contain shoot apical meristem were frozen in liquid nitrogen and conserved at –80°C.

Microarray probes design

Custom microarray probes were designed with eArray software (<https://earray.chem.agilent.com/earray/>; Agilent technologies, Massy, France). Both transcriptomics and CH3 microarray probes were 60-mer oligomers.

For transcriptomics analyses, at least one probe was designed for each v2.0 *P. trichocarpa* transcript with a positive bias in 3' of mRNA, using best probe distribution method that limits GC content of oligomers. When probes could not differentiate between transcripts (more than 70% identity on 50 residues), a second probe was design in 5' of the first one. With this design, 45907 probes each matching with one v2.0 *P. trichocarpa* transcript were generated while 5770 additional probes allowed the differentiation between transcripts. In addition, 50 probes were duplicated 10 times to be used as

reproducibility controls and internal Agilent control probes were also used. For hybridization, 8 × 60K custom microarray slides were used (Agilent technologies, France).

CH3 microarray probes design was different depending on four defined biological groups. In the first group, composed of total v2.0 *P. trichocarpa* genes, 5 probes were designed, 1 probe in 0.5 kb upstream and 1 probe in 0.5 kb downstream region of genes, and 3 probes were located within gene body. Genes involved in morphogenesis or drought response according to bibliography, constituting the second group, were entirely covered from 0.5 kb upstream to 0.5 downstream with overlapping microarray probes designed every 30 bp. Loci from a previous methylome study (Lafon-Placette et al., in press) were the third and fourth group, corresponding to genes and intergenic regions, with probes designed every 140 bp or 780 bp respectively between 0.5 kb upstream and 0.5 downstream of loci. In addition, 50 probes were duplicated 10 times to be used as reproducibility controls and internal Agilent control probes were also used. For hybridization, 1 × 1 M custom microarray slides were used (Agilent technologies, France).

RNA preparation and microarray hybridization

Total RNAs were independently isolated from 3 individual shoot apical meristems (50 mg fresh weight) per treatment using Nucleospin® RNA Plant (Macherey-Nagel, Hoerd, France). A control PCR was performed on RNA to confirm the absence of genomic DNA in all the total RNAs preparations. Bioanalyzer 2100 (Agilent Technologies, Massy, France) was used to estimate DNA contamination and RNA integrity with RNA Integrity Number (RIN) measurement (RIN threshold of 8). Fluorescent cRNA (complementary RNA) was generated from 600 ng of total RNA according to Agilent Technologies recommendations. Briefly, RNA spike mix, which is a mix of 10 polyadenylated RNAs synthesized *in vitro* from *E1A* adenovirus gene, was prepared with One-Color RNA Spike-In Kit (Agilent Technologies, France) to be used as positive control for labeling and data analysis. For each sample, 80 ng of total RNAs were reverse transcribed, followed by T7 RNA polymerase-mediated simultaneous cRNA synthesis and incorporation of cyanine 3-labeled CTP, using Input Quick Amp Labeling Kit, One Color (Agilent Technologies, France). Fluorescent cRNA was purified with RNeasy Mini Kit (Qiagen, France). Quality control of synthesized fluorescent cRNA was based on cRNA and Cy3 concentration measurement as well as cRNA fragments size. For each sample, a yield of 825 ng of cRNA, a specific activity of 6 pmol_{Cy3}/g_{cRNA} and a 200-2000 nucleotides fragments size range were chosen as quality thresholds.

RNA hybridization was conducted according to Agilent Technologies instructions. Briefly, the labeled cRNAs were fragmented in Fragmentation mix (Agilent Technologies) during 30 min at 60°C, before addition of 2X GEx Hybridization Buffer HI-RPM (Agilent Technologies) and loading onto the arrays. The arrays were incubated at 65°C for 17 hours in Agilent hybridization chambers and then wash according to the protocol at room temperature.

Methyl DNA ImmunoPrecipitation (MeDIP) and microarray hybridization

Shoot apical meristem from 3 individuals (about 150 mg fresh weight per extraction) of the 'Carpaccio' *P. × euramericana* genotype for each treatment ('Control', 'Water deficit' and 'Re-irrigation') were

grounded into a fine powder in liquid N₂ and used for DNA extraction. CTAB DNA extraction protocol was used (Doyle & Doyle, 1987). Chemicals were purchased from Sigma-Aldrich (Saint-Quentin Fallavier, France). Briefly, for each water treatment, 150 mg of shoot apical meristem powder were solubilised in 1.5 mL of a pre-heated lysis buffer composed of 100 mM Tris-HCl (pH 8.0), 20 mM EDTA (pH 8.0), 2% (w/v) CTAB, 1.4 M NaCl, 0.2% (v/v) β-mercaptoethanol, 4% (w/v) PVP, 0.1 mg/mL Proteinase K. Samples were incubated at 60°C for 1h on a rotator. DNA was isolated after 2 chloroform/isoamyl alcohol (24:1) (v/v) extractions, RNase A digestion and isopropanol precipitation. DNA pellets were solubilised in ultrapure water and concentrations were estimated with Nanodrop quantification.

DNA was then sonicated using a VC 505 Vibra-Cell sonicator (Fisher Scientific, Illkirch, France) to obtain fragments ranging from 0.2 to 0.8 kb. Methylated DNA was immunoprecipitated according to (Weng *et al.*, 2009). Briefly, 11 µg of sonicated DNA were diluted in 250 µL of IP buffer 1X composed of 10 mM Na₃PO₄ (pH 7.0), 140 mM NaCl, 0.05% Triton X-100 and denaturated during 10 min in boiling water. After 10 min of cooling in ice, 4 µg of DNA were kept at -20°C to be used as input control for microarray analysis. The 7 µg of DNA left were incubated overnight with 10 µg of mouse monoclonal anti-5mC antibodies (1 µg/µL; Eurogentec, France) on a rotator in a 4°C cold room. The DNA-antibodies mixture was coupled with Dynabeads Protein G (Invitrogen, Cergy-Pontoise, France) during 2 hours at 4°C. Beads were washed 4 times with IP Buffer 1X using a magnetic rack (Invitrogen, France). Methylated DNA was eluted by digesting antibodies overnight with 5 µL of Proteinase K (20 mg/mL; Sigma-Aldrich, France) in 250 µL of a digestion buffer composed of 50 mM Tris (pH 8.0), 10 mM EDTA pH 8.0, 0.5% SDS at 50°C on a rotator. DNA was isolated with 1 phenol/chloroform/isoamyl alcohol (25:24:1) (v/v/v) and 2 chloroform/isoamyl alcohol (24:1) (v/v) extractions using phase lock gel heavy tubes (Eppendorf, Le Pecq, France) and ethanol precipitation. DNA pellets were solubilised in ultrapure water and concentrations were estimated with Nanodrop quantification.

For each sample, labeled DNA was generated according to Agilent Technologies recommendations. Briefly, near 500 ng of input control and immunoprecipitated (IP) methylated DNA were labeled using Genomic DNA Enzymatic Labeling Kit (Agilent Technologies, France). The use of random primers and the exo-Klenow fragment allowed the differential labeling of input control and IP methylated DNA with cyanine 3-labeled dUTP and cyanine 5-labeled dUTP respectively. Labeled DNA was purified using 30 kDa Amicon columns (Millipore, Molsheim, France). Quality control of synthesized labeled DNA was based on DNA and Cy5/Cy3 concentration measurement. For each sample, a yield of 2.5 µg of labeled DNA and a specific activity of 7-20 pmol_{Cy5}/g_{DNA} and 18-25 pmol_{Cy3}/g_{DNA} were chosen as quality minimum thresholds.

DNA hybridization was conducted in Imaxio (Clermont-Ferrand, France) according to Agilent Technologies instructions. Briefly, Cy5- and Cy3-labeled DNAs were pooled in ddH₂O before addition of Hybridization mix (Agilent Technologies), incubation at 95°C for 3 min and then at 37°C for 30 min. Samples were loaded onto the arrays and hybridized at 67°C for 40 hours at 20 rpm in Agilent hybridization chambers. Arrays were then washed according to the protocol at room temperature.

Data collection, normalization and cluster analysis

Microarray slides were scanned with G2505C scanner (Agilent Technologies, France) and raw signal data were extracted using Feature Extraction 10.7 software. Reproducibility and reliability of each single microarray was assessed using Quality Control report data (Feature Extraction, Agilent Technologies).

One-color transcriptomics data were normalized per chip to the 75th percentile and per probe to medians and were transformed to log values. Data filtering was based on Feature Extraction flag criteria: non outlier population; not saturated, uniform and significant probe signal; probe signal above background. A probe signal was considered as "well above background" if gBGSubSignal was greater than 2.6 gBG_SD. To pass the filtering, a probe needed to pass all flag criteria for all replicates of at least one experimental condition (Control, Water Deficit or Re-irrigation). After this data filtering, 33005 probes were available. To determine if there were genes differentially expressed between experimental conditions, one-way ANOVA and Welch's *t*-test were performed with MeV software (Saeed *et al.*, 2003). Statistical tests were considered significant at ** $P < 0.01$. Benjamini-Hochberg's correction was used for false positive discovery. Average linkage hierarchical clustering analysis was applied using Pearson correlation.

For CH3 microarray, data normalization, performed by Feature Extraction software, was used to correct cyanine bias with the linear method: the mean signal of probes selected for quality criteria (signal above background, non-outlier and whose ranks are similar between the two dyes) allowed the normalization of all probes values. For each probe, Cy5/Cy3 log ratio was then calculated. Data filtering was based on Feature Extraction flag criteria: non outlier population; not saturated, uniform and significant Cy3 and Cy5 probe signal; Cy3 and Cy5 probe signal above background. A probe signal was considered as "well above background" if g/rBGSubSignal was greater than 2.6 g/rBG_SD. To pass the filtering, a probe needed to pass all flag criteria for all replicates of at least one experimental condition (Control, Water Deficit or Re-irrigation). After this data filtering, 441214 probes were available. To determine if there were sequences differentially methylated, a probe signal fold change of 2 between at least 2 experimental conditions was chosen as a threshold. To reduce the number of false positive, additional filtering criteria were used: probes with mean signal between experimental conditions below 5% mean signal of total probes were eliminated; per condition and per dye, probes with signal below 10% mean signal of total probes were also eliminated. Probes needed a GC content comprised between 20% and 80% and a T_m comprised between 60°C and 85°C. Finally, when focusing on genes and 0.5 kb upstream/downstream regions, only unique probes (matching with one locus on v2.0 *P. trichocarpa* genome with an identity threshold of 50 nt out of 60) were selected.

Bioinformatic analyses

Differentially expressed or methylated genes were annotated using the Gene Ontology class enrichment produced with AgriGO software (<http://bioinfo.cau.edu.cn/agriGO/analysis.php>). Analysis of repeats content was performed with the RepeatMasker database using *A. thaliana* as DNA sources (<http://www.repeatmasker.org/>). Poplar genomic data were used from Phytozome (<http://www.phytozome.net/poplar.php>; (Tuskan *et al.*, 2006)) for genomic data and from PopGenIE

<http://www.popgenie.org/>) for methylome data (Lafon-Placette et al. in press). For all these softwares, default parameters were applied.

Quantitative PCR (*En cours avec IMAXIO*)

Approximately 500 ng of total RNA were reverse transcribed using High capacity cDNA reverse transcription kit (Applied Biosystems, Courtaboeuf, France). Constitutively expressed genes encoding respectively a tubulin and an actin, *POPTR_0001s25410* and *POPTR_0019s02630*, were used as internal standards to normalize the amount of mRNA in PCR reaction according to (Trap-Gentil *et al.*, 2011). *POPTR_0001s35160* amplification was also used to confirm that there was no genomic DNA contamination since primers were designed on a coding sequence containing a short intron (209 bp). Furthermore, a control without reverse transcriptase during the cDNA synthesis was performed to confirm the absence of genomic DNA in all the total RNAs preparations. Primers for PCR were designed with eprimer3 software (<http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=eprimer3>) and are listed in Table S1. Three biological and two technical replicates will be performed for each gene and genotype.

Bisulfite sequencing (*En cours*)

Bisulfite treatment, which results in the conversion of unmethylated cytosines to uracils but does not affect methylated cytosines, was used to determine the cytosine methylation status in CG, CHG, or asymmetric CHH (H could be A, T, or C) contexts. Conversion was performed using Epitect bisulfite kit (Qiagen, France) and 500 ng of genomic DNA according to manufacturer's recommendations. Controls for bisulfite conversion efficiency and amplification as well as primers design, PCR and sequencing procedures have been recently detailed (Trap-Gentil *et al.*, 2011; Lafon-Placette et al. (in press)). PCR products will be directly sequenced without any cloning step according to (Leakey *et al.*, 2008) and (Jiang *et al.*, 2010). Direct quantification of cytosine methylation percentages for each cytosine position (Leakey *et al.*, 2008; Jiang *et al.*, 2010) will be obtained from three biological and two technical replicates for each gene and each genotype. Primers used for bisulfite sequencing are listed in Table S1. Three biological and three technical replicates will be performed for each sequence.

Determination of DNA methylation percentages by HPLC

The global percentage of methylation of genomic DNA was determined by high-performance liquid chromatography (HPLC) using a Gemini™ column (150 × 4.6 mm, 5 µm; Phenomenex, Le Pecq, France) and according to Gourcilleau et al. (2010). Controls for this procedure (RNA contamination and confirmation by enzyme restriction analysis) were performed as described in Trap-Gentil et al. (2011). The methylcytosine percentages were calculated using the formula $\%mC = [mC/(C + mC)] \times 100$, where 'C' represents 2'-deoxycytidine content and 'mC' 5-methyl-2'-deoxycytidine content. Three independent analyses and three replicates were performed for each measurement

Statistical analyses

Statistical analyses were performed using the SPSS statistical software package (SPSS version 11.0.1 PC, Chicago, Illinois, USA). Means are expressed with their standard error and compared by analysis of variance [one-way ANOVA; general linear model (GLM) procedure]. Fraction effects on mapping distribution were evaluated using the Chi square (χ^2) homogeneity test. Relationships between gene expression and DNA methylation were analyzed by linear regression analysis [Pearson's correlation coefficients (r_P)]. Statistical tests were considered significant at * $P < 0.05$, ** $P < 0.01$ or *** $P < 0.001$.

RESULTS

Physiological states in response to variations in water availability

Three 'Carpaccio' *P. × euramericana* hybrid populations of six clones were each submitted to different water availability statuses during 14 days (Fig. 1a). The first population (control; Ctrl) was continuously watered to field capacity. The second population (water deficit; WD) was submitted to a short (14 days) and moderate water deficit, with relative extractable water not going down under 20%. The third population (water deficit and re-irrigation; WD-RI) was also placed under the same water deficit during 9 days and then re-irrigated, being fully watered for the 5 remaining days. Physiological monitoring of these 3 populations showed that the water deficit induced a significant decrease in stomatal conductance (Fig. 1b) and height growth rate (Fig. 1c) while full turgor osmotic potential as well as leaf mass per area were significantly increased (Fig. 1d and e). In addition, neither leaf predawn water potential nor net CO₂ assimilation of stressed plants differed from controls (data not shown). Stomatal conductance and full turgor osmotic potential came back to the control level when plants were re-irrigated after water deficit, while height growth rate and leaf mass per area did not. At the end of the treatment, shoot apical meristems were isolated from the shoot apex of each plant and used for further analyses. These data indicated that the three populations are in distinct physiological states and that the water stress effects were moderate but significant.

Shoot Apical Meristem (SAM) transcriptomics in response to variations in water availability

SAM transcriptomic response to water availability was assessed with custom arrays (8x60K) designed using the 45 033 v2.0 *P. trichocarpa* model transcripts (see Material and Methods). Probe design (51 677; Table S2) and transcriptomics data are available in GEO database (*dépôt dans GEO en cours*). Microarray transcript abundance data among the three experimental conditions were independently validated by quantitative RT-PCR (qRT-PCR) for a sample of genes (Fig. S1 *analyses en cours*). After data filtration and normalization (Fig. S2), 33 005 filtered probes (63.8% of designed probes) were available for further study (Table S2). A hierarchical clustering performed on these filtered probes gathered together the 3 biological replicates of the 'water deficit' condition while 'Control' and 'Re-irrigation' conditions were mixed (Fig. 2a).

Using a fold change threshold of two, 1 488 probes were found among the 3 conditions and most of them (1 260 probes with 675 probes specific to this conditions) was significantly different between 'water deficit' and 're-irrigation' (Fig. 2b). According to one-way ANOVA analysis ($P < 0.01$), 1 499 significant probes corresponding to 1 386 differentially expressed genes were identified (about 3% of the *P. trichocarpa* model transcripts; Table S3). Among these differentially expressed genes, 6 clusters could be defined according to gene expression pattern between the three conditions (Pearson hierarchical clustering, Fig. 3): one cluster was specifically up-regulated in 'control' (n°5), two in 'water stress' (n°2 and 3), one in 're-irrigation' (n°1), one in 'control' and 'water deficit' (n°4) and one in 'control' and 're-irrigation' (n°6). Over-represented functions were identified for 4 clusters (n°1, 3, 5 and 6) using GO annotation terms (Singular Enrichment Analysis; Fig. S3-6 and Fig. 3). Cluster 5 was enriched in transcription factor encoding genes that were down-regulated in both water deficit and re-

irrigation and played role in cell multiplication, leaf fate or flowering/meristem functioning (Fig. 3, Fig.S7a and Table S4). In cluster 3, genes were upregulated during the water deficit and downregulated after re-irrigation, exhibiting enrichment in protein tyrosine kinase or membrane transport protein functions (Fig. S4 and Fig. 3). Cluster 2 with up-regulated genes in 'water deficit' condition displayed functions in transport and response to biotic/abiotic stimuli (Fig. 3 and Fig. S7b). Some of them still exhibited up-regulation in 're-irrigation' condition. Cluster 1 with genes overexpressed during re-irrigation preferentially encoded sulfotransferases and oxydoreductases (Fig. S3 and Fig. 3).

Genes with expression specifically affected by the water deficit (identified by Welch's *t*-test at $P < 0.01$ between control and water deficit condition) were compared for their transcriptional pattern in mature leaves from the same experimental system (Bogeat-Triboulot et al., unpublished data; GEO accession number: GSE17226; Fig. S8). Hierarchical clustering gathered samples by organ and in a second time by environmental condition. Only 10 differentially expressed genes between control and water deficit were common to shoot apex and mature leaves experiments (Welch's *t*-test at $P < 0.01$) but with well-known drought-responsive genes like DREB or Aquaporin encoding genes (Table S5). Altogether, our data showed that shoot apical meristem displayed tissue-specific patterns of gene expression under each water availability condition, the 're-irrigation' condition being not only an intermediate state between control and water deficit but rather a third transcriptomic situation.

Genomic features of SAM DNA methylation changes in response to variations in water availability

Methylome response to water availability was assessed with custom arrays (1M) using a MeDIP microarray approach. Arrays were designed based on v2.0 *P. trichocarpa* genomic dataset (see Material and Methods). Probe design (988 544 probes, Table S2) and methylome data are available in GEO database (*dépôt dans GEO en cours*). Microarray methylation enrichment data among the three experimental conditions were independently validated by bisulfite sequencing for a sample of genes (Fig. S9 *analyses en cours*). After data filtration and normalization, 441 214 filtered probes (44.6% of designed probes) were available for further study (Table S2). A hierarchical clustering performed on filtered probes gathered together 'Control' and 'Water Deficit' conditions while 'Re-irrigation' remained as an out-group (Fig. 4a). Using a fold change threshold of two, 270 064 differential probes were identified (27.3% of total probes; Table S2), with most of them (268 389) involving the 're-irrigation' condition (Fig. 4b). Microarray DNA methylation data were mapped on v2.0 *P. trichocarpa* genome scaffolds for the 3 experimental conditions, as well as genes, repeats and sRNAs densities (scaffold 19 as an example in Fig. 5a). 'Control' and 'water deficit' conditions were strongly methylated in pericentromeric repeat-rich regions as well as in sRNA-rich loci and displayed strong DNA methylation variations in gene-rich regions. Globally, water deficit condition was associated with a lower methylation level mostly in gene-rich regions while hypermethylation affected few regions, mostly repeat-rich loci compared to control condition. The 're-irrigation' condition had a distinct pattern of methylation with weak DNA methylation variations within scaffold and globally higher level of methylation compared the two other conditions. HPLC global DNA methylation measurement

confirmed these global mappings and particularly the hypermethylated state of the 're-irrigation' condition (40% for Ctrl, 32% for WD and 51% for WD-RI; Fig. 5b). Moreover, mean probes methylation signal (-1.11 for Ctrl, -1.51 for WD and -0.11 for WD-RI) paralleled global DNA methylation levels between water treatments. Methylome changes between the 3 conditions affected both genic and intergenic regions with near 28% of designed probes being differentially methylated (Fig. 6a). Among intergenic regions, repeated sequences being the most reworked by DNA methylation were DNA transposons (Fig. 6b). Genic regions mapped by differentially methylated probes corresponded to 16 330 genes (2.5% of poplar gene models) belonging to a broad range of functions (Table S6, Fig. S10). Within gene-related loci, a higher percentage of significant probes was found in gene body (3 times more) than the 0.5 kb neighboring 5' and 3' parts (Fig. 6a and 6c). The internal part of the gene-body was associated to the maximum of significant methylation changes. Nevertheless, promoters containing water deficit related element (DRE and ABRE) as well as pathogen responsive element (GCC box) were preferentially affected by DNA methylation changes (Fig. 6d). ABRE elements were the most modified motif during water deficit. Gene body and flanking regions were strongly targeted by DNA methylation reworking after re-irrigation, resulting in hypermethylation for the majority of genes (Fig. 6c,d and 7). Among these differentially methylated gene regions, several clusters could be defined according to the patterns between the three conditions (Pearson hierarchical clustering, Fig. 7). One large cluster showed hypermethylation in 're-irrigation' and 'control' (with a lower level however). Genes exhibiting this pattern in 5' flanking part were involved in gene expression while those showing this profile in their body and 3' parts had function in ubiquitin related protein catabolism. One small cluster was hypermethylated in 'water stress' and 're-irrigation' and one hyper-methylated in 're-irrigation' only. In 5' and 3' parts, an additional cluster was defined with hypermethylation in the three conditions.

Altogether, our data showed that shoot apical meristem displayed, under each water availability condition, specific patterns of DNA methylation particularly on gene-body and promoters containing water deficit related elements, the 're-irrigation' condition being hyper-methylated.

DNA methylation reworking and gene expression

Genes expressed in at least one experimental condition were compared to non expressed genes for their DNA methylation changes depending on the gene parts (5', 3' and gene body; Fig. 8). The most significant difference was found for gene body with 60% of expressed genes being differentially methylated versus 41% for non-expressed genes (Fig. 8). Genes differentially expressed between the three experimental conditions were affected by DNA methylation changes in similar proportions than expressed genes (data not shown).

Using the transcriptomic clustering for differentially expressed genes, no clear clustering could be observed for the DNA methylation patterns in upstream and downstream regions as well as in gene body (Fig. S11). Genic regions were globally hypermethylated in the 're-irrigation' condition, hypomethylated in the water deficit one and intermediate in the control condition. However, the transcriptomic clusters 4 and 5 (Fig. 3), defined by up-regulated in control condition displayed slightly higher gene-body DNA methylation compared with genes from other clusters in this condition (Fig.

S11). Correlations between transcriptomic and methylation data were tested in all conditions (Table S7-8). Water deficit transcriptome was negatively correlated to control and re-irrigation one ($r_p = -0.251$ at $P = 5.10^{-232}$ and $r_p = -0.514$ at $P = 0$) while DNA methylation levels were strongly correlated between conditions within each gene part (5', body or 3'; Table S7). Each genic region was reworked by DNA methylation independently from other parts according to low correlations observed (Table S8). A positive significant one was detected in gene body between methylome and transcriptome levels for control condition ($r_p = 0.154$ at $P = 3.10^{-73}$) while a negative correlation was observed for re-irrigation ($r_p = -0.094$ at $P = 8.10^{-28}$). Very low correlations were found between methylome and transcriptome variations (Table S8).

Finally, 839 differentially expressed and methylated genes representing 60.5% of differentially expressed genes and 5.1% of differentially methylated genes were identified and showed over representing of transcription factor function (Fig. S12). Using more stringent thresholds, a set of 36 genes was identified and 21 of them displayed annotation in Arabidopsis (Tables I). Several functions were represented, and particularly 8 genes (38%) playing roles in phytohormone signaling during environmental stimuli: *DIR1* (lipid signaling in Systemic Acquired Resistance), *RDR1* (control of jasmonate and other phytohormones during insect attack), *3-oxo-5-alpha-steroid 4-dehydrogenase family protein* (metabolism of steroid hormones), *PRT1* (protein degradation in signaling of phytohormones response or environmental adaptation), *EXL5* (brassinosteroid responsive genes and growth regulation), *ABA DEFICIENT 2* (ABA metabolism), *ERF023* (ethylene responsive transcription factor) and *HYDROXYJASMONATE SULFOTRANSFERASE 2A* (*Jasmonate metabolism*) (Table I). Altogether, our data showed that DNA methylation reworking in response to variations of water availability could not be related alone to the transcriptomic changes (quantitative as well as qualitative) and *vice et versa* even if some significant correlations could be found between the two variables. Nevertheless, most of the differentially expressed genes were modified for their DNA methylation state, particularly genes involved in phytohormone signaling during environmental stimuli and thus possibly participating to the phenotypic plasticity.

Epigenetic control of a *SULFOTRANSFERASE* (SOT) gene cluster involved in brassinosteroid metabolism in response to variations in water availability

Sulfotransferase encoding genes were over-represented in up-regulated genes after re-irrigation and were found among the 36 differentially expressed and methylated genes (cluster 1; Fig. 3; Table I). Identification of sulfotransferase encoding genes (blast analysis) in *P. trichocarpa* v2.0 genome revealed three main clusters: one localized on scaffold 3 with 13 genes, one on scaffold 11 with 9 genes and a minor cluster was found on scaffold 1 with 3 genes (Fig. S13a). Most of the genes from scaffolds 1 and 3 had significant and similar expression patterns, being up-regulated after re-irrigation, while differentially expressed genes from scaffold 11 had opposite behavior with up-regulation after water deficit and down-regulation after re-irrigation (Fig. 9a-c). All the differentially expressed genes from scaffolds 1 and 3 clusters, and to a lower extend from scaffold 11 cluster, were also differentially methylated in their body while unaffected genes for their expression were differentially methylated in their promoter or not affected by DNA methylation changes (Fig. 9d-f). According to a phylogenetic

analysis based on the active site conserved protein region and using all *P. trichocarpa* and *A. thaliana* sulfotransferase sequences, strong similarities were found between *P. trichocarpa* scaffolds 1 and 3 cluster genes and characterized (*AT2G03760* and *AT2G14920*) or putative *A. thaliana* brassinosteroid sulfotransferases, while scaffold 11 cluster genes were related to *AT1G74090*, *AT1G74100* and *AT1G18590*, annotated as desulfoglucosinolate sulfotransferases encoding genes (Fig. S14b,c). In addition, the expression of other genes involved in the brassinosteroid homeostasis and biosynthesis displayed opposite patterns of expression between the water deficit and the re-irrigation conditions (Fig. S15). Altogether, our data proposed that distinct clusters of SOT are epigenetically and transcriptionally controlled during changes in water availability and depending on their substrates.

REFERENCES

- Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003.** Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant cell* **15**: 63–78.
- Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears DD, Hogg EH (Ted), et al. 2010.** A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* **259**: 660–684.
- Berta M, Giovannelli A, Sebastiani F, Camussi A, Racchi ML. 2010.** Transcriptome changes in the cambial region of poplar (*Populus alba* L.) in response to water deficit. *Plant Biology* **12**: 341–354.
- Bies-Ethève N, Gaubier-Comella P, Debures A, Lasserre E, Jobet E, Raynal M, Cooke R, Delseny M. 2008.** Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. *Plant molecular biology* **67**: 107–124.
- Bogeat-Triboulot M-B, Brosché M, Renaut J, Jouve L, Le Thiec D, Fayyaz P, Vinocur B, Witters E, Laukens K, Teichmann T, et al. 2007.** Gradual soil water depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in *Populus euphratica*, a poplar growing in arid regions. *Plant physiology* **143**: 876–892.
- Bonhomme L, Monclus R, Vincent D, Carpin S, Lomenech A-M, Plomion C, Brignolas F, Morabito D. 2009.** Leaf proteome analysis of eight *Populus xeuramericana* genotypes: genetic variation in drought response and in water-use efficiency involves photosynthesis-related proteins. *Proteomics* **9**: 4121–4142.
- Choi H, Hong J, Ha J, Kang J, Kim SY. 2000.** ABFs, a family of ABA-responsive element binding factors. *The Journal of biological chemistry* **275**: 1723–1730.
- Choudhary SP, Yu J-Q, Yamaguchi-Shinozaki K, Shinozaki K, Tran L-SP. 2012.** Benefits of brassinosteroid crosstalk. *Trends in plant science*.
- Cohen D, Bogeat-Triboulot M-B, Tisserant E, Balzergue S, Martin-Magniette M-L, Lelandais G, Ningre N, Renou J-P, Tamby J-P, Le Thiec D, et al. 2010.** Comparative transcriptomics of drought responses in *Populus*: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. *BMC genomics* **11**: 630.
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschield CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. 2008.** Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**: 215–219.
- Doyle J, Doyle J. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.
- Feng S, Cokus SJ, Zhang X, Chen P-Y, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH,**

- Halpern ME, et al. 2010a.** Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences* **107**: 8689–8694.
- Feng S, Jacobsen SE, Reik W. 2010b.** Epigenetic Reprogramming in Plant and Animal Development. *Science* **330**: 622–627.
- Filippou P, Antoniou C, Fotopoulos V. 2011.** Effect of drought and rewatering on the cellular status and antioxidant response of *Medicago truncatula* plants. *Plant Signaling & Behavior* **6**: 270–277.
- Gallé A, Haldimann P, Feller U. 2007.** Photosynthetic performance and water relations in young pubescent oak (*Quercus pubescens*) trees during drought stress and recovery. *The New phytologist* **174**: 799–810.
- Goll MG, Bestor TH. 2005.** Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry* **74**: 481–514.
- Gourcilleau D, Bogeat-Triboulot M-B, Le Thiec D, Lafon-Placette C, Delaunay A, El-Soud WA, Brignolas F, Maury S. 2010.** DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity. *Annals of Forest Science* **67**: 208.
- Grob U, Stüber K. 1987.** Discrimination of phytochrome dependent light inducible from non-light inducible plant genes. Prediction of a common light-responsive element (LRE) in phytochrome dependent light inducible plant genes. *Nucleic acids research* **15**: 9957–9973.
- Hamanishi ET, Campbell MM. 2011.** Genome-wide responses to drought in forest trees. *Forestry* **84**: 273–283.
- Intergovernmental Panel on Climate Change. 2007.** IPCC Fourth Assessment Report: Climate Change 2007 (AR4). Geneva, Switzerland: IPCC.
- Jansson S, Douglas CJ. 2007.** *Populus*: A Model System for Plant Biology. *Annual Review of Plant Biology* **58**: 435–458.
- Jiang M, Zhang Y, Fei J, Chang X, Fan W, Qian X, Zhang T, Lu D. 2010.** Rapid quantification of DNA methylation by measuring relative peak heights in direct bisulfite-PCR sequencing traces. *Laboratory Investigation; a Journal of Technical Methods and Pathology* **90**: 282–290.
- Klevebring D, Street NR, Fahlgren N, Kasschau KD, Carrington JC, Lundeberg J, Jansson S. 2009.** Genome-wide profiling of *Populus* small RNAs. *BMC Genomics* **10**: 620.
- Kültz D, Csonka L. 1999.** What sets the TonE during osmotic stress? *Proceedings of the National Academy of Sciences* **96**: 1814–1816.
- Lande R. 2009.** Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of evolutionary biology* **22**: 1435–1446.
- Leakey TI, Zielinski J, Siegfried RN, Siegel ER, Fan C-Y, Cooney CA. 2008.** A simple algorithm for quantifying DNA methylation levels on multiple independent CpG sites in bisulfite genomic sequencing electropherograms. *Nucleic Acids Research* **36**: e64–e64.

Des Marais DL, Juenger TE. 2010. Pleiotropy, plasticity, and the evolution of plant abiotic stress tolerance. *Annals of the New York Academy of Sciences* **1206**: 56–79.

Marron N, Delay D, Petit J-M, Dreyer E, Kahlem G, Delmotte FM, Brignolas F. 2002. Physiological traits of two *Populus* × *euramericana* clones, Luisa Avanzo and Dorskamp, during a water stress and re-watering cycle. *Tree Physiology* **22**: 849–858.

Marron N, Dreyer E, Boudouresque E, Delay D, Petit J-M, Delmotte FM, Brignolas F. 2003. Impact of successive drought and re-watering cycles on growth and specific leaf area of two *Populus* × *canadensis* (Moench) clones, 'Dorskamp' and 'Luisa_Avanzo'. *Tree Physiology* **23**: 1225–1235.

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, et al. 2010. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**: 684–692.

Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. 2007. DNA demethylation in the *Arabidopsis* genome. *Proceedings of the National Academy of Sciences* **104**: 6752–6757.

Plomion C, Lalanne C, Claverol S, Meddour H, Kohler A, Bogeat-Triboulot M-B, Barre A, Le Provost G, Dumazet H, Jacob D, et al. 2006. Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics* **6**: 6509–6527.

Qi X, Zhang Y, Chai T. 2007. Characterization of a Novel Plant Promoter Specifically Induced by Heavy Metal and Identification of the Promoter Regions Conferring Heavy Metal Responsiveness. *Plant Physiology* **143**: 50–59.

Raj S, Brautigam K, Hamanishi ET, Wilkins O, Thomas BR, Schroeder W, Mansfield SD, Plant AL, Campbell MM. 2011. Clone history shapes *Populus* drought responses. *Proceedings of the National Academy of Sciences* **108**: 12521–12526.

Rushton PJ, Reinstädler A, Lipka V, Lippok B, Somssich IE. 2002. Synthetic Plant Promoters Containing Defined Regulatory Elements Provide Novel Insights into Pathogen- and Wound-Induced Signaling. *The Plant Cell* **14**: 749–762.

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, et al. 2003. TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* **34**: 374–378.

Seki M, Umezawa T, Urano K, Shinozaki K. 2007. Regulatory metabolic networks in drought stress responses. *Current opinion in plant biology* **10**: 296–302.

Teixeira FK, Colot V. 2010. Repeat elements and the *Arabidopsis* DNA methylation landscape. *Heredity* **105**: 14–23.

Trap-Gentil M-V, Hébrard C, Lafon-Placette C, Delaunay A, Hagège D, Joseph C, Brignolas F, Lefebvre M, Barnes S, Maury S. 2011. Time course and amplitude of DNA methylation in the shoot apical meristem are critical points for bolting induction in sugar beet and bolting tolerance between genotypes. *Journal of Experimental Botany* **62**: 2585–2597.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al. 2006. The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Vining KJ, Pomraning KR, Wilhelm LJ, Priest HD, Pellegrini M, Mockler TC, Freitag M, Strauss S. 2012. Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* **13**: 27.

Weng Y-I, Huang TH-M, Yan PS. 2009. Methylated DNA immunoprecipitation and microarray-based analysis: detection of DNA methylation in breast cancer cell lines. *Methods in Molecular Biology* (Clifton, N.J.) **590**: 165–176.

Xu Z, Zhou G, Shimizu H. 2010. Plant responses to drought and rewatering. *Plant Signaling & Behavior* **5**: 649–654.

Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. *Science* **328**: 916–919.

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW-L, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, et al. 2006. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in *Arabidopsis*. *Cell* **126**: 1189–1201.

FIGURES LEGENDS

Figure 1. Water deficit/rewatering cycle characterization. (a) Monitoring of the soil relative extractable water (REW) during the 14 days experiment for control (Ctrl), water deficit (WD) and rewatering after water deficit (WD-RI) conditions. For details see Material and Methods. Stomatal conductance (b), height growth rate (c), full turgor osmotic potential (d) and leaf mass per area (e) for *Populus deltoides* × *P. nigra* 'Carpaccio' genotype were measured for each water condition at the end of the experiment. Values are means ± SE ($n = 6$). Values marked with different letters are significantly different between water treatments ($P < 0.05$) as determined by one-way ANOVA.

Figure 2. Shoot apical meristem transcriptomic characterization of samples. (a) Hierarchical clustering of samples based on transcriptomics values of all probes. For control (Ctrl), water deficit (WD) and rewatering after water deficit (WD-RI) conditions, three biological replicates were studied and numbered from 1 to 3. (b) Venn diagram showing the number of differential probes (determined by Fold Change > |2|) between each condition.

Figure 3. Heatmap representation of differentially expressed genes between control, water deficit and rewatering conditions. Criterion for differential expression is a significant difference among the three conditions as determined by one-way ANOVA analysis with $P < 0.01$. Pearson hierarchical clustering was performed with TigerMeV.4.8.1 software. GO terms that are significantly enriched within each gene cluster are displayed (SEA analysis with $P < 0.01$ except for the fourth cluster where $P < 0.05$ was chosen).

Figure 4. Shoot apical meristem methylome characterization of samples. (a) Hierarchical clustering of samples based on total methylome values of all probes. Control (Ctrl), water deficit (WD) and rewatering after water deficit (WD-RI) samples are each a pool of three biological replicates. (b) Venn diagram showing the number of differential probes (determined by Fold Change > |2|) between each condition.

Figure 5. Genomic features of DNA methylation changes in shoot apical meristem. (a) Distribution of DNA methylation level in control (Ctrl), water deficit (WD) and rewatering after water deficit (WD-RI) samples in v2.0 *P. trichocarpa* genome scaffold 19 as well as mapping of repeats, genes and sRNA densities, obtained from PopGenIE database (<http://www.popgenie.org/>; Klevebring *et al.*, 2009). (b) Global DNA methylation level in SAM for the three conditions, as determined by HPLC analyses. Values are means ± SE ($n = 4$). Values marked with different letters are significantly different between water treatments ($P < 0.05$) as determined by one-way ANOVA.

Figure 6. Features of loci affected by DNA methylation reworking in shoot apical meristem. Except for (b), values are expressed as a percentage of significantly different probes between water treatments (Fold Change > |2|) compared with total number of designed probes for each locus category. (a) Values are given for genic (gene body + 500 bp flanking regions) and intergenic regions, gene body and 500 bp flanking regions. (b) Values are a ratio between repeat content in significantly different probes and repeat content in total designed probes. Repeat detection was performed with

RepeatMasker on all probes sequence gathered in one. (c) Distribution of DNA methylation changes within "genic regions" from (a) is detailed. Changes are given for water deficit (WD) vs. control (Ctrl) comparison, rewatering after water deficit (WD-RI) vs. WD and WD-RI vs. Ctrl. (d) Values are given for all promoters that contain probes ("genome") or promoters with ABA Responsive Element (ABRE), GCC box, Dehydration Responsive Element (DRE), box S, Metal Responsive Element (MRE) or Light Responsive Element (LRE) motives. ABRE sequence: C/TACGTGGC (Choi *et al.*, 2000); GCC box sequence: AGCCGCC (Rushton *et al.*, 2002); DRE sequence: CCGAC (Kültz & Csonka, 1999); box S sequence: AGCCACC (Rushton *et al.*, 2002); MRE sequence: TGCAGGC (Qi *et al.*, 2007); LRE sequence: ATGATAAGG (Grob & Stüber, 1987).

Figure 7. Heatmap representation of differentially methylated gene parts between control (Ctrl), water deficit (WD) and rewatering (WD-RI) conditions. Gene parts are (a) 0.5 kb upstream region, (b) gene body, (c) 0.5 kb downstream region. The number of concerned genes is indicated for each gene part. Probes were differentially methylated according to fold change value between at least two conditions (Fold Change > |10|). Pearson hierarchical clustering was performed with TigerMeV.4.8.1 software. GO terms that are significantly enriched within each gene cluster are displayed (SEA analysis with $P < 0.01$).

Figure 8. Difference in DNA methylation reworking between expressed and non expressed genes. "Expressed genes" correspond to filtered probes with a signal above background for at least one condition while "non expressed genes" probes had signal below background for the three conditions. Percentage of genes with DNA methylation changes (Fold Change > |2| between at least two conditions) in each part are displayed. "Upstream" and "downstream" are 0.5 kb flanking regions in 5' and 3' respectively.

Figure 9. DNA methylation and expression changes in response to variations in water availability for three SULFOTRANSFERASE genes clusters. Normalized expression log values are shown in control (Ctrl), water deficit (WD) and water deficit followed by rewatering (WD-RI) conditions for differentially expressed genes (one-way ANOVA, Figure 2) in scaffold 1 cluster (a), scaffold 3 cluster (b) and scaffold 11 cluster (c). Means are given with their standard errors ($n = 3$; biological triplicates). Genomic view of DNA methylation variations between rewatering and water deficit conditions (RIvsWD) or between water deficit and control conditions (WDvsCtrl) are shown for scaffold 1 cluster (d), scaffold 3 cluster (e) and scaffold 11 cluster (f). Arrows represent 60-mers microarray probes and are given with their fold change value. Only significantly different probes are shown (Fold Change > |2|). No significantly different probes between water deficit and control conditions could be found in scaffold 1 and 3 gene clusters. When existing, probes were separated in hypomethylated or hypermethylated probes according to their Fold Change values.

SUPPLEMENTARY FIGURES

Figure S1. qPCR verification results. *En cours*

Figure S2. Box plot distribution of normalized expression values for each biological sample.

Ctrl, Control; WD, water deficit; WD-RI, water deficit/rewatering condition. The biological triplicates are arbitrary number from 1 to 3.

Figure S3. Enriched GO molecular functions in differentially expressed genes between 'control', 'water deficit' and 'water deficit/rewatering' conditions (expression cluster 1).

Singular Enrichment Analysis (SEA) for differentially expressed genes from expression cluster 1 (see Figure 3) was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Figure S4. Enriched GO molecular functions in differentially expressed genes between 'control', 'water deficit' and 'water deficit/rewatering' conditions (expression cluster 3).

Singular Enrichment Analysis (SEA) for differentially expressed genes from expression cluster 3 (see Figure 3) was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Figure S5. Enriched GO molecular functions in differentially expressed genes between 'control', 'water deficit' and 'water deficit/rewatering' conditions (expression cluster 5).

Singular Enrichment Analysis (SEA) for differentially expressed genes from expression cluster 5 (see Figure 3) was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Figure S6. Enriched GO molecular functions in differentially expressed genes between 'control', 'water deficit' and 'water deficit/rewatering' conditions (expression cluster 6).

Singular Enrichment Analysis (SEA) for differentially expressed genes from expression cluster 6 (see Figure 3) was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hieratical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information

includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Figure S7. Functions of genes whose expression was impacted during water deficit and remained altered after rewatering. Upregulated and downregulated genes corresponds to expression clusters 2 and 5 as shown in Figure 3. (a) Transcription factors encoding genes downregulated during water deficit and rewatering. The role of these genes was classified according to homologous Arabidopsis gene function (detailed in Table S4). (b) Enriched functions in genes upregulated during water deficit and rewatering. GO terms determination was performed with Classification SuperViewer tool (<http://bar.utoronto.ca/>) using Arabidopsis homologous genes according to v2.0 *P. trichocarpa* annotation. The values shown here are a ratio between GO terms abundance in upregulated genes from expression cluster 2 (Figure 3) and total *P. trichocarpa* v2.0 genome. Thus, a value higher than 1 means enrichment towards genome and vice versa.

Figure S8. Heatmap representation of differentially expressed genes between control and water deficit in shoot apical meristem and mature leaves. Genes showed are those differentially expressed between control and water deficit in shoot apical meristem ('Carpaccio' *P. × euramericana* genotype). Mature leaves transcriptomic data were generated from the same experimental system in 'Carpaccio' and 'Soligo' *P. × euramericana* genotypes (Bogeat-Triboulot et al., unpublished data; GEO accession number: GSE17226). Criterion for differential expression is a significant difference between control and water deficit conditions as determined by Welsch's *t*-test ($P < 0.01$). Pearson hierarchical clustering was performed with TigerMeV.4.8.1 software.

Figure S9. Bisulfite sequencing verification results. *En cours*

Figure S10. Enriched GO biological processes in differentially methylated genes between 'control', 'water deficit' and 'water deficit/rewatering' conditions. Singular Enrichment Analysis (SEA) for differentially methylated genes in 5', 3' and/or body parts between at least two conditions was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hierarchical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Figure S11. Heatmap representation of differentially expressed and methylated genes between control, water deficit and rewatering conditions. Criterion for differential expression is a significant difference among the three conditions as determined by one-way ANOVA analysis with $P < 0.01$. For DNA methylation, gene parts are 0.5 kb upstream region, gene body, 0.5 kb downstream region. Probes were differentially methylated according to fold change value between at least two conditions (Fold Change > |2|). For each gene part, gene order was the one defined by Pearson hierarchical clustering based on gene expression pattern in Figure 3.

Figure S12. Enriched GO molecular functions in differentially methylated and expressed genes between 'control', 'water deficit' and 'water deficit/rewatering' conditions. Singular Enrichment Analysis (SEA) for differentially methylated and expressed genes between at least two conditions was performed towards complete v2.0 genome on agriGO website. The graphic is a GO hierarchical image containing all statistically significant GO molecular function terms. Significance is classified into ten levels which are associated with corresponding specific colors. The smaller of the term's adjusted p-value, the more significant statistically. Inside the box of the significant terms, the information includes: GO term, adjusted p-value, GO description, item number mapping the GO in the query list and v2.0 genome, and total number of query list and v2.0 genome.

Figure S13. *In silico* characterization of the three genomic clusters of SULFOTRANSFERASE genes in *P. trichocarpa*. (a) Localization of the three main clusters in v2.0 *P. trichocarpa* scaffolds. (b) ClustalW alignment using conserved active site of all *A. thaliana* and *P. trichocarpa* sulfotransferase proteins, performed with Jalview (www.jalview.org). Proteins missing this active site were not selected for this alignment. (c) Cladogram of the same proteins based on the conserved active site protein sequence.

Figure S14. Upregulated and downregulated genes involved in brassinosteroid (BR) biosynthesis and homeostasis in response to variations in water availability. Model for BR biosynthesis and homeostasis in Arabidopsis proposed by Choudhary *et al.* (2012). Poplar genes homologous to Arabidopsis ones were defined according to *P. trichocarpa* v2.0 annotation. Differentially expressed genes were determined by Welch's *t*-test between two conditions ($P < 0.01$). Ctrl, control; WD, water deficit; WD-RI, water deficit followed by rewatering. DWF4 (DWARF4) catalyzes the conversion of CN to 6-oxocampestanol (6-Oxo-CN); DWF4 brings conversion of campestanol (CN) and 6-Oxo-CN to 6-deoxocathasterone (dCT) and cathasterone (CT). *DWF4* is under positive regulation of the TCP1 (TEOSINTE BRANCHED 1, CYCLOIDEA and PCF1). Auxin (Aux)-induced transport of BRX (BREVIS RADIX) to the nucleus also promotes *CPD* and *DWF4* expression. Sulfonation of brassinolide (BL), castasterone (CS) and 24-epibrassinosteroids (24-epiBR) by Arabidopsis sulfotransferases (AtST), namely AtST4a and AtST1, renders BRs nonfunctional. *bri1-5 ENHANCED 1 (BEN1)* encodes a protein reducing the synthesis of active BRs by competing with BL and CS precursors for the active site on BR biosynthesis enzymes.

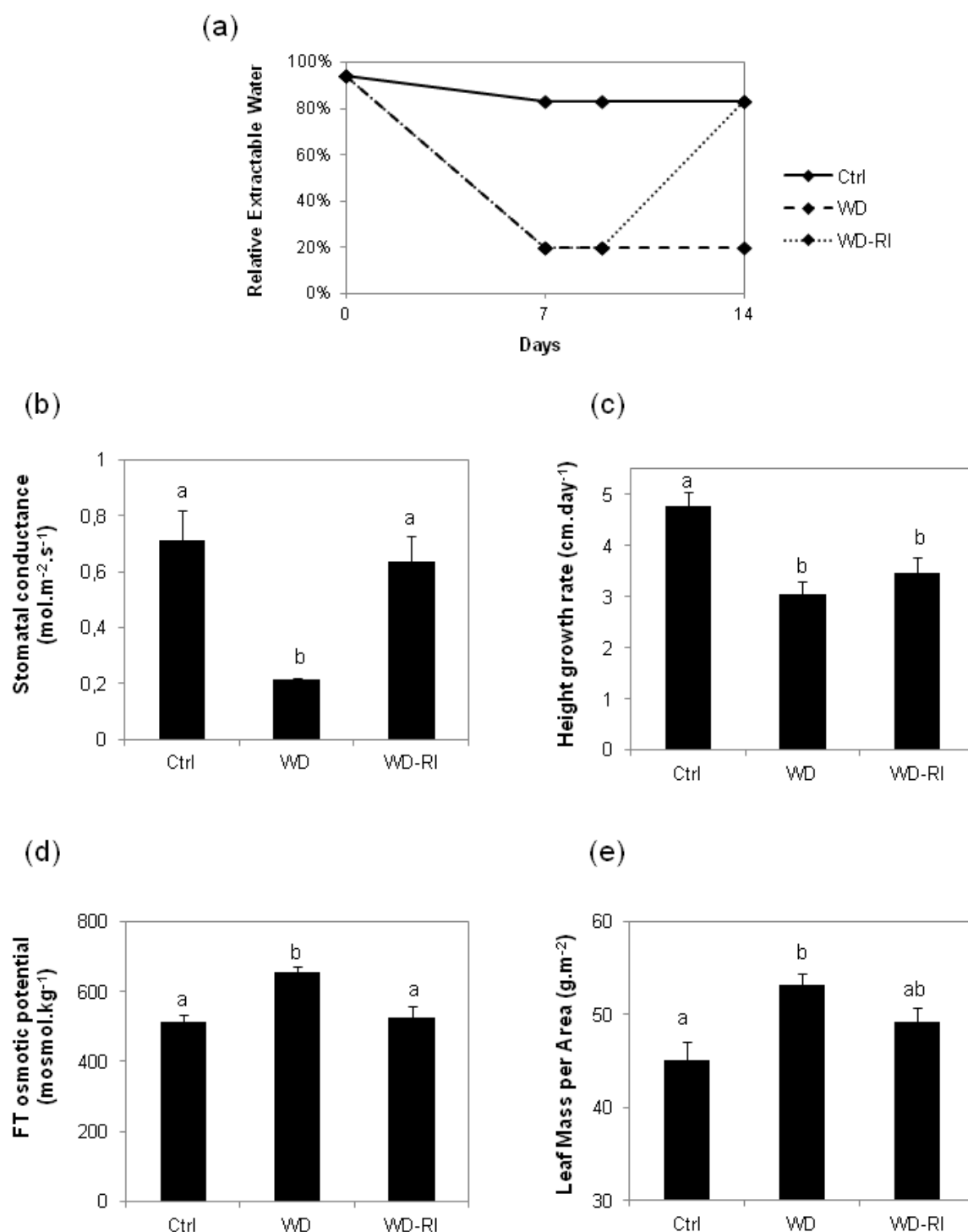


Figure 1

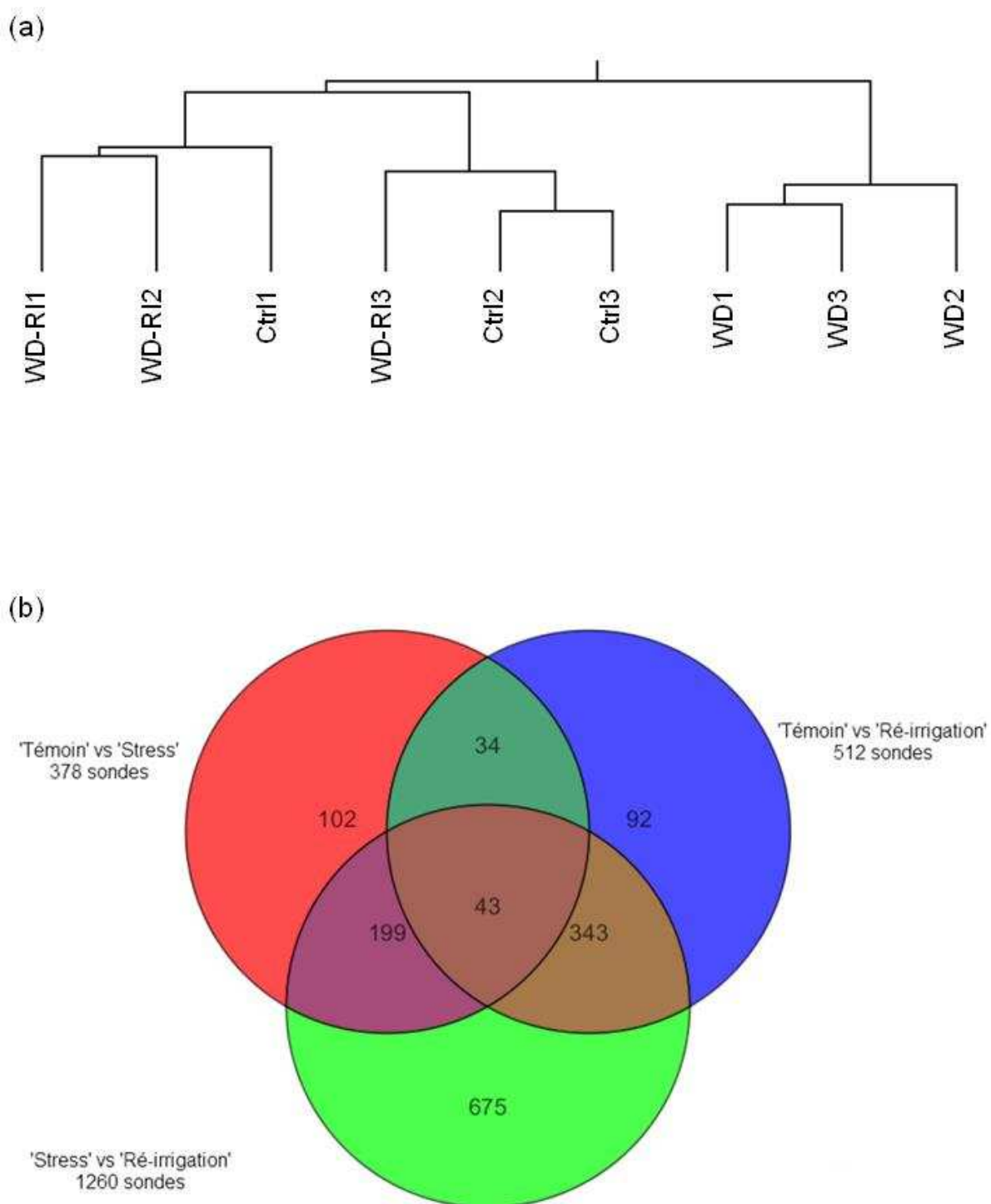


Figure 2

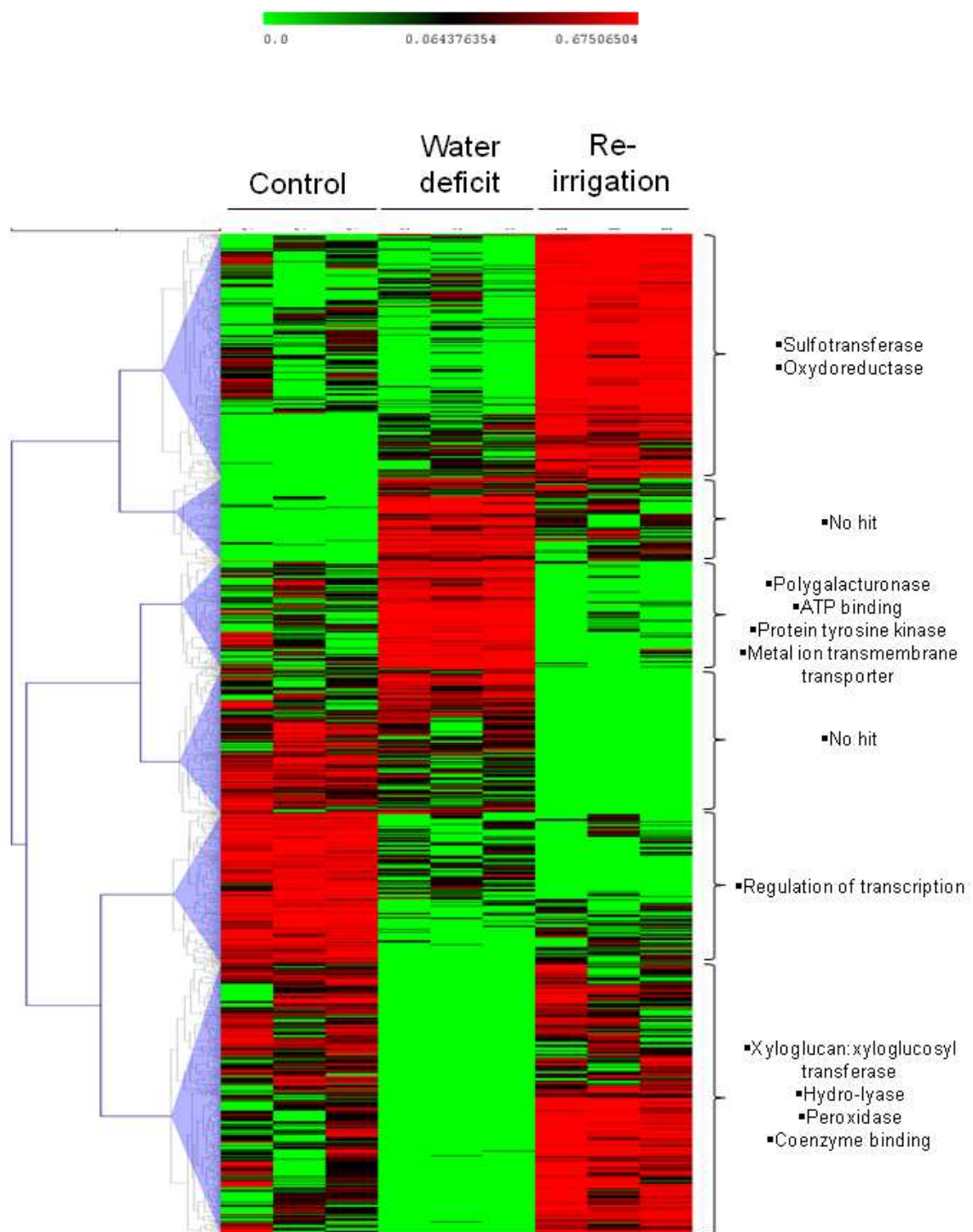
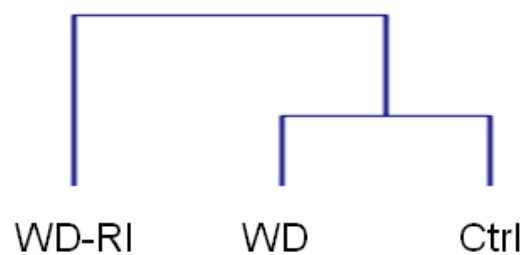


Figure 3

(a)



(b)

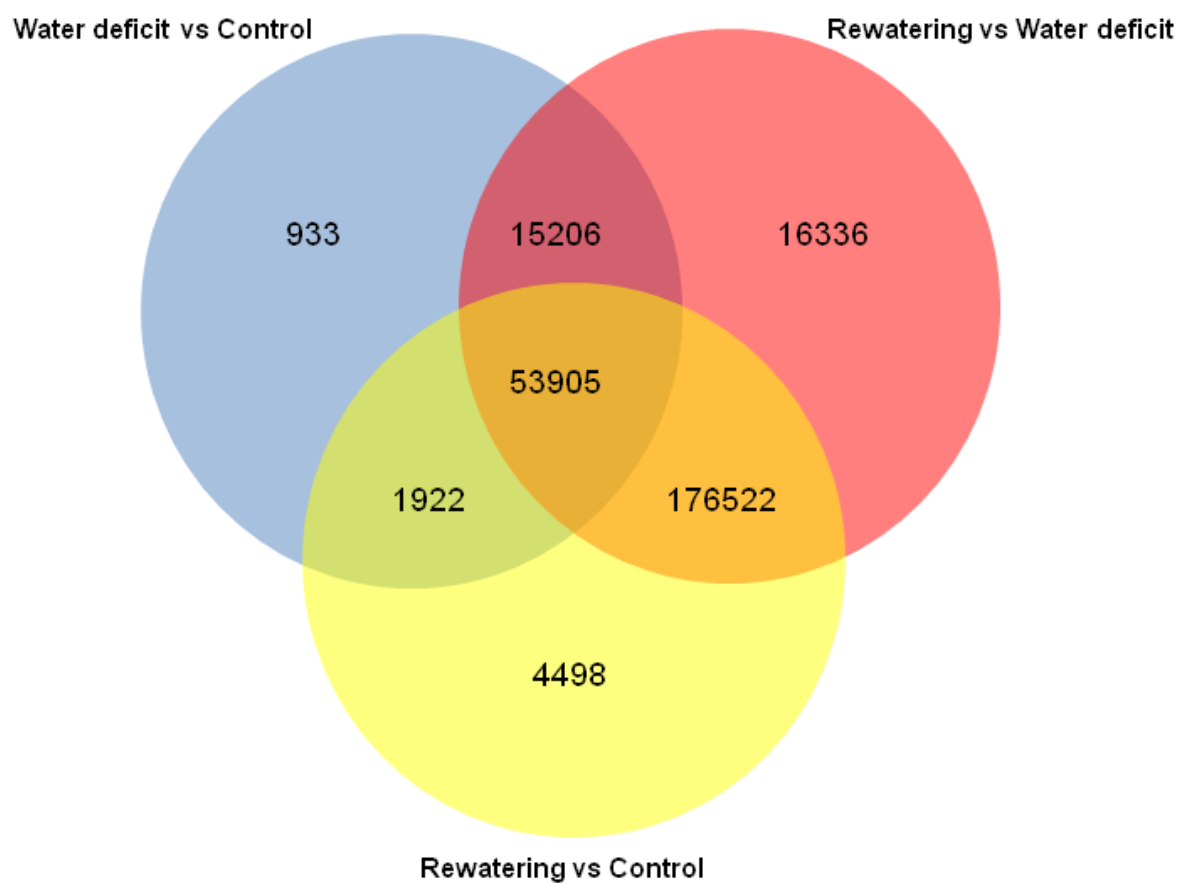


Figure 4

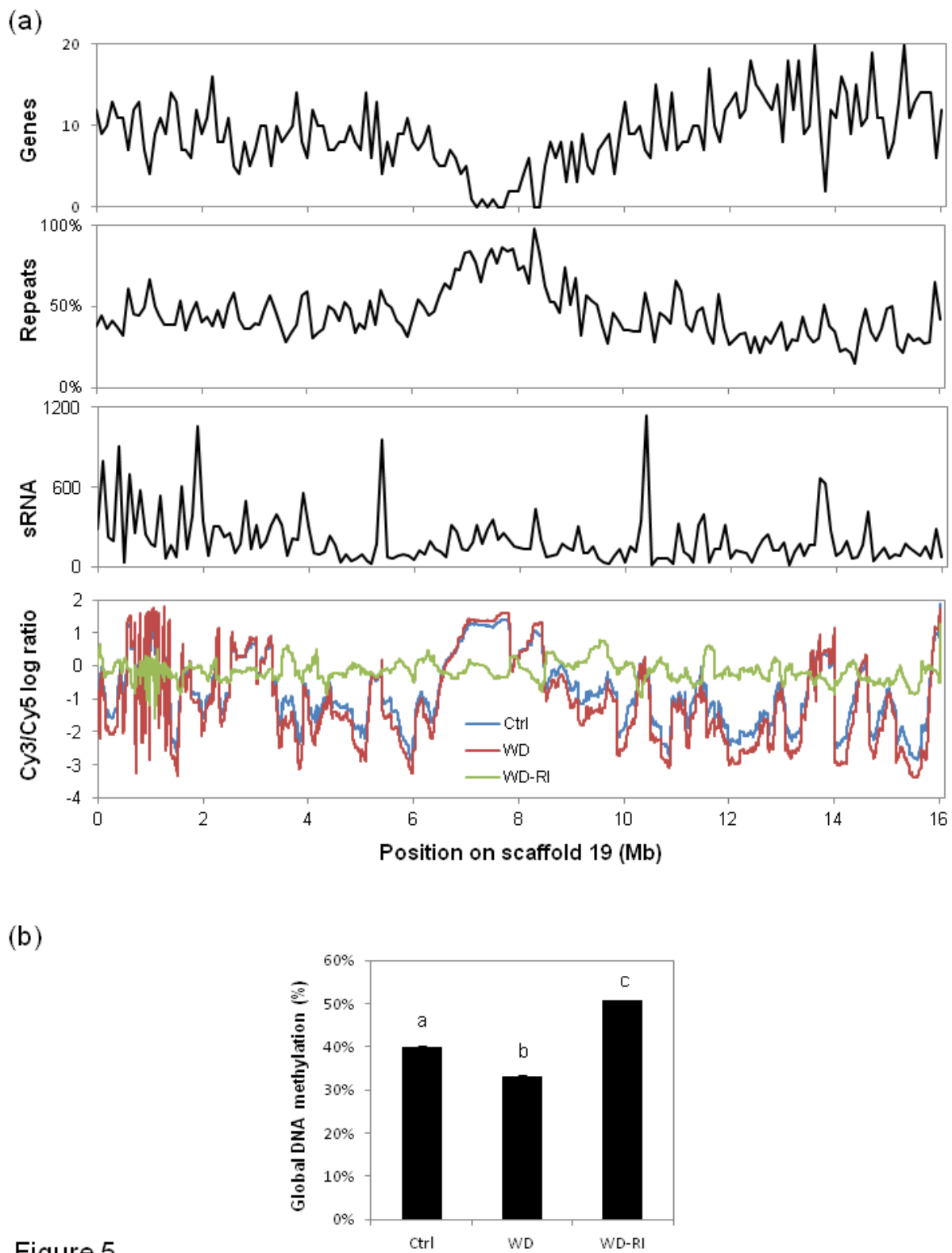


Figure 5

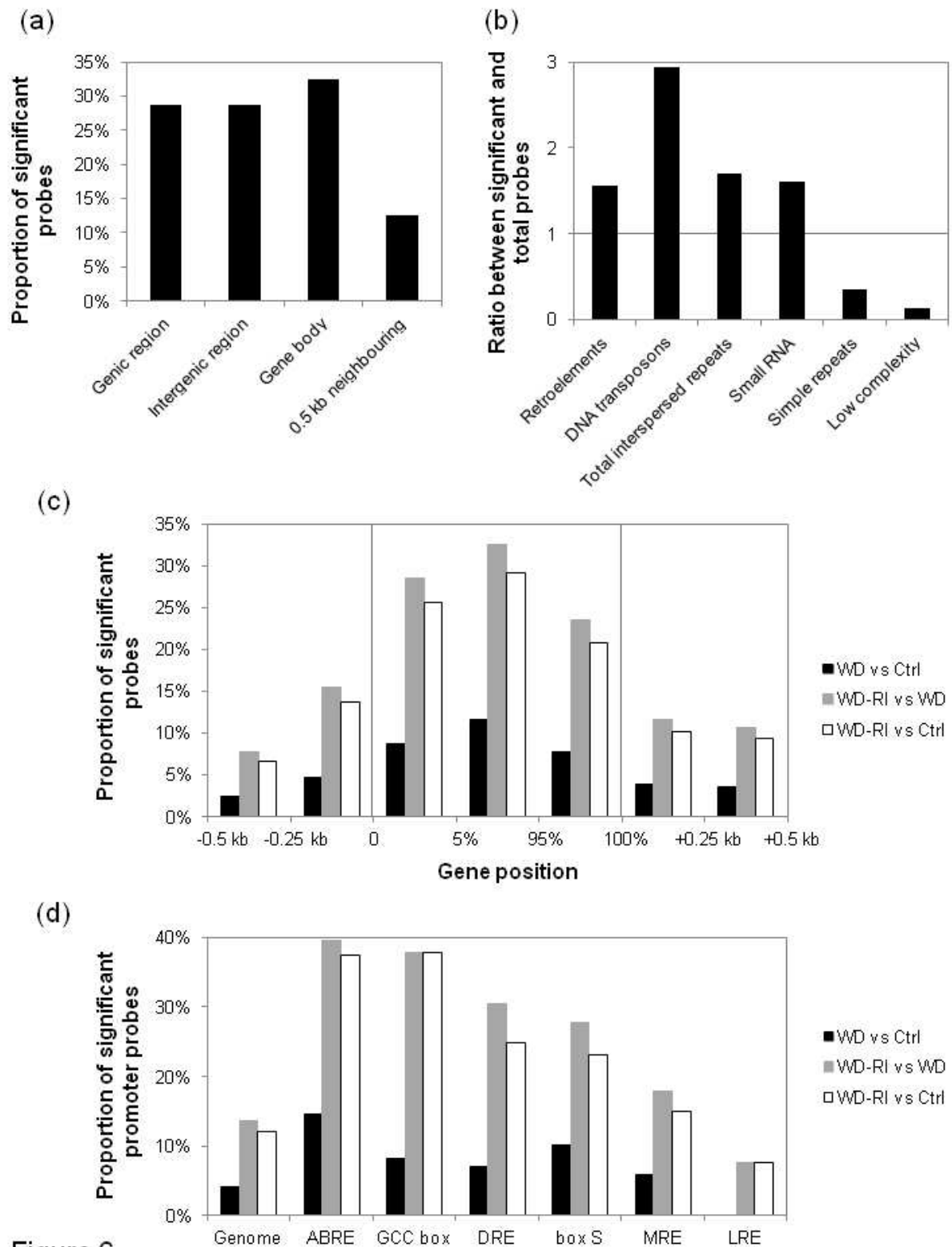


Figure 6

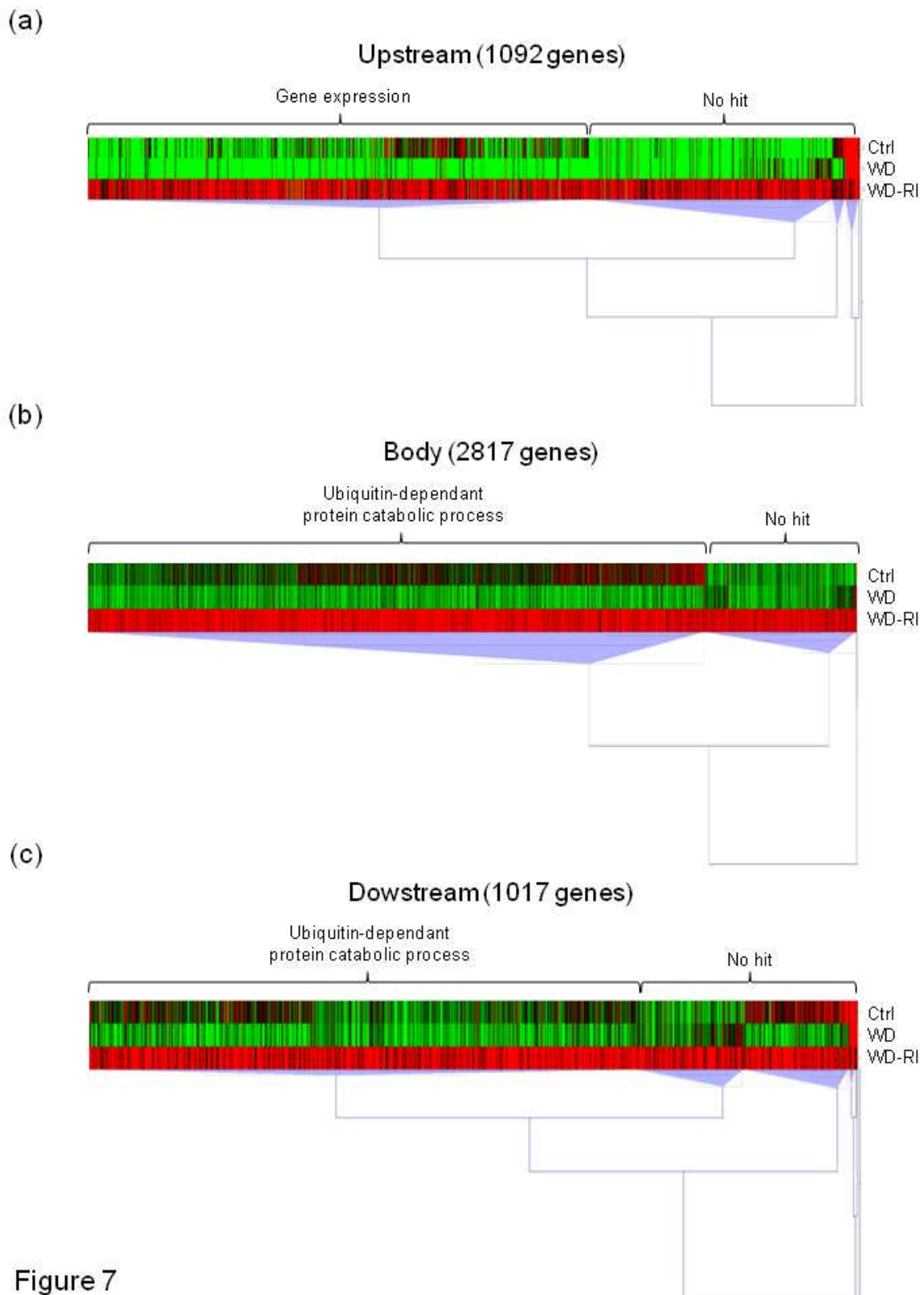


Figure 7

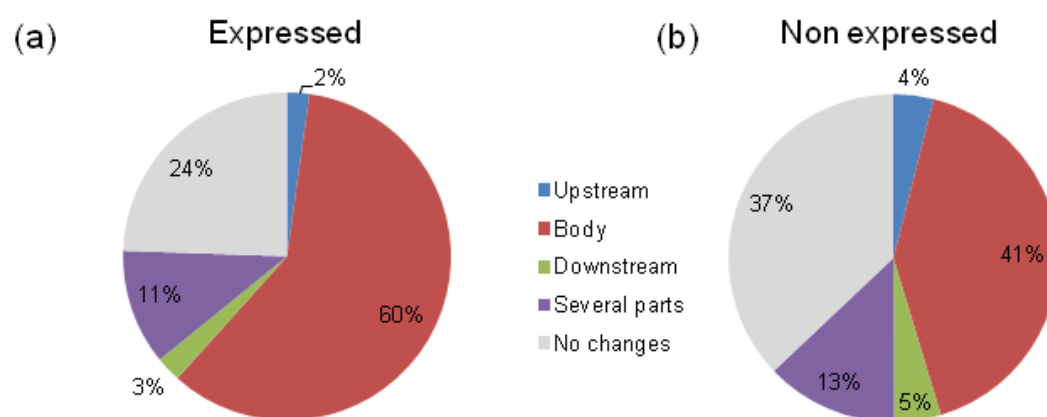


Figure 8

Table 1. Differentially expressed and methylated genes. Given values are fold changes between control (Ctrl), water deficit (WD) or re-irrigation (WD-Ri) conditions. Criteria for differential expression are a significant difference among the three conditions (one-way ANOVA analysis with $P < 0.01$) and a fold change value $> |2|$. A gene was defined as differentially methylated if at least one of its part (5' or 3' 500 bp flanking regions, body) displayed a fold change $> |6|$ between two conditions.

<i>P. trichocarpa</i> v2.0 gene	Arabidopsis homolog	Arabidopsis homolog annotation	Upstream region			Gene body			Downstream region			Gene expression		
			WD-RWD	WD-RiCH	WD/CH	WD-RWD	WD-RiCH	WD/CH	WD-RWD	WD-RiCH	WD/CH	WD-RWD	WD-RiCH	WD/CH
POPTR_000041550	AT5G4465.1	DIR1 DEFECTIVE IN INDUCED RESISTANCE 1; lipid binding / lipid transporter	1.54	2.09	1.37	NA	NA	NA	NA	NA	NA	2.13	1.96	-1.28
POPTR_000422380	AT1G08060.1	ACAT ALPHA CARBOXY-AMIDOHYDROLASE 7; carboxylic dehydratase; zinc ion binding	-1.97	-2.36	-1.23	NA	NA	NA	NA	NA	NA	-2.46	-2.17	-1.09
POPTR_000302900	AT3G1190.1	cysteine proteinase inhibitor-related	-1.63	NA	NA	NA	NA	NA	NA	NA	NA	-2.64	1.36	-1.48
POPTR_000342060	AT4G03730.1	unknown protein	1.23	1.91	1.56	NA	NA	NA	NA	NA	NA	2.46	1.44	-1.87
POPTR_001041610	AT1G04100.1	glycoside hydrolase family 28 protein; polygalacturonase (pectinase) family protein	1.43	2.17	1.60	1.15	1.46	-1.29	NA	NA	NA	-2.37	-1.28	-1.22
POPTR_0003416510	AT4G1940.1	unknown protein	-1.87	NA	-1.01	NA	NA	-2.12	NA	NA	NA	-2.86	1.46	1.07
POPTR_000540980	AT1G07140.1	serine-type endopeptidase; serine-type peptidase	2.15	2.44	1.13	2.35	2.32	-2.06	NA	NA	NA	-3.57	-1.82	-2.25
POPTR_000341320	AT1G14790.1	RCM1 RNA-DEPENDENT RNA POLYMERASE 1; RNA-directed RNA polymerase; nucleic acid binding	4.49	4.23	-1.15	3.26	2.99	-2.22	NA	NA	NA	-3.45	-2.00	-1.73
POPTR_001341480	AT3G2870.1	cell wall protein precursor; pectinase	6.25	4.28	-1.40	NA	NA	NA	NA	NA	NA	1.99	2.12	-1.02
POPTR_001042150	AT3G1960.1	3-oxo-5-oxo-steroid 4-oxohydrolase family protein / steroid 5-oxo-reductase family protein	6.58	3.99	-1.64	NA	NA	NA	NA	NA	NA	1.44	2.46	-1.71
POPTR_000142930	AT3G1960.1	unknown protein	6.64	5.27	1.29	NA	NA	NA	NA	NA	NA	2.42	1.16	-2.06
POPTR_000540980	AT1G08032.1	unknown protein	5.24	5.20	-0.24	NA	NA	NA	NA	NA	NA	2.46	1.26	-1.81
POPTR_000143170	AT3G14005.1	unknown protein	6.88	5.29	-0.02	NA	NA	1.13	NA	NA	NA	3.00	-1.36	-4.15
POPTR_000041460	AT1G08120.1	unknown protein	6.58	2.08	-2.13	3.21	3.48	-1.61	11.14	5.25	-1.20	1.68	-1.19	-2.26
POPTR_000154850	no hit	NA	NA	NA	NA	14.88	3.64	-2.80	2.25	1.24	-1.89	2.98	1.54	-1.46
POPTR_000306570	AT1G07810.1	GER1 (GERMINALINE PROTEIN 1); oxalate oxidase	NA	NA	NA	6.37	4.87	-1.67	NA	NA	NA	2.82	1.19	-1.46
POPTR_000341280	AT1G05390.1	protease-associated (PA) domain-containing protein	NA	NA	NA	45.98	4.81	-1.39	NA	NA	NA	2.05	1.76	-1.16
POPTR_000341710	AT3G24800.1	PR1 (PROTEOLYSIS 1); dehydrin-protein ligase	NA	NA	NA	7.42	6.82	-1.15	NA	NA	NA	2.23	-1.86	1.19
POPTR_000341290	AT5G07010.1	ST3A (SULFOTRANSFERASE 3A); hydroxymethyl sulfotransferase; sulfotransferase	NA	NA	NA	NA	NA	-1.60	NA	NA	NA	1.42	2.54	-1.72
POPTR_000421230	AT3G2740.1	unknown protein	NA	NA	NA	6.86	1.20	-6.09	NA	NA	NA	1.64	2.94	1.41
POPTR_000421600	AT3G17200.1	EXL5 (EXORDIUM LINE 5)	NA	NA	NA	12.13	7.86	-4.63	NA	NA	NA	1.44	2.94	-1.72
POPTR_000422290	AT1G05240.1	ABA2 (ABA DEFICIENT 2); alcohol dehydrogenase; oxidoreductase; xanthin dehydrogenase	NA	NA	NA	7.42	6.10	-0.24	NA	NA	NA	2.21	1.21	-1.83
POPTR_000420530	AT5G05630.1	AT1G01; calcium-activated potassium channel; ion channel; outward rectifier potassium channel	NA	NA	NA	10.85	5.29	-5.25	11.12	6.18	1.45	-4.50	2.25	-2.84
POPTR_000341600	AT3G2740.1	unknown protein	NA	NA	NA	12.13	6.79	-6.69	NA	NA	NA	1.38	2.25	-2.84
POPTR_001041050	AT1G10522.1	unknown protein	NA	NA	NA	3.88	5.14	-1.26	NA	NA	NA	2.20	1.03	-1.86
POPTR_001041910	AT3G08110.1	unknown protein	NA	NA	NA	9.46	4.29	-5.09	NA	NA	NA	1.46	1.96	-1.27
POPTR_001240310	AT3G1790.1	PAF7; acid phosphatase; phosphatase; protein serine/threonine phosphatase	NA	NA	NA	10.77	2.29	-8.68	NA	NA	NA	2.29	2.29	1.02
POPTR_001241310	AT5G05400.1	CYP715A1; electron carrier; heme binding / iron ion binding / monooxygenase; oxygen binding	NA	NA	NA	6.13	7.11	1.16	NA	NA	NA	-4.73	1.30	-3.98
POPTR_001340240	AT1G05470.1	binding / catalytic oxidoreductase	NA	NA	NA	16.19	5.49	-2.42	1.70	4.94	-2.82	2.54	2.17	0.09
POPTR_001341340	AT5G04130.1	FLA13 (FASCICULINE ABABINOLACTAM PROTEIN 13 PRECURSOR)	NA	NA	NA	16.22	5.29	-5.09	NA	NA	NA	2.38	1.73	-1.96
POPTR_001400940	AT1G01200.1	AP2 domain-containing transcription factor; ERF103	NA	NA	NA	7.46	4.14	-3.46	NA	NA	NA	3.10	1.80	-1.27
POPTR_001641330	AT3G08110.1	unknown protein	NA	NA	NA	13.27	2.29	-1.66	NA	NA	NA	2.20	1.86	-1.41
POPTR_001641360	AT1G10870.1	ZPR4 (ZINC TRANSPORTER 4 PRECURSOR)	NA	NA	NA	13.99	5.46	-0.06	NA	NA	NA	-2.46	-1.25	-1.77
POPTR_001641620	AT1G15190.1	unknown protein	NA	NA	NA	11.13	6.79	-4.66	NA	NA	NA	1.98	-1.11	-2.94
POPTR_001640930	AT3G05090.1	unknown protein	NA	NA	NA	6.85	4.25	-1.21	NA	NA	NA	1.15	-2.26	-2.89
POPTR_001640920	AT3G05390.1	unknown protein	NA	NA	NA	6.88	2.20	-2.01	NA	NA	NA	-1.85	1.74	-2.17

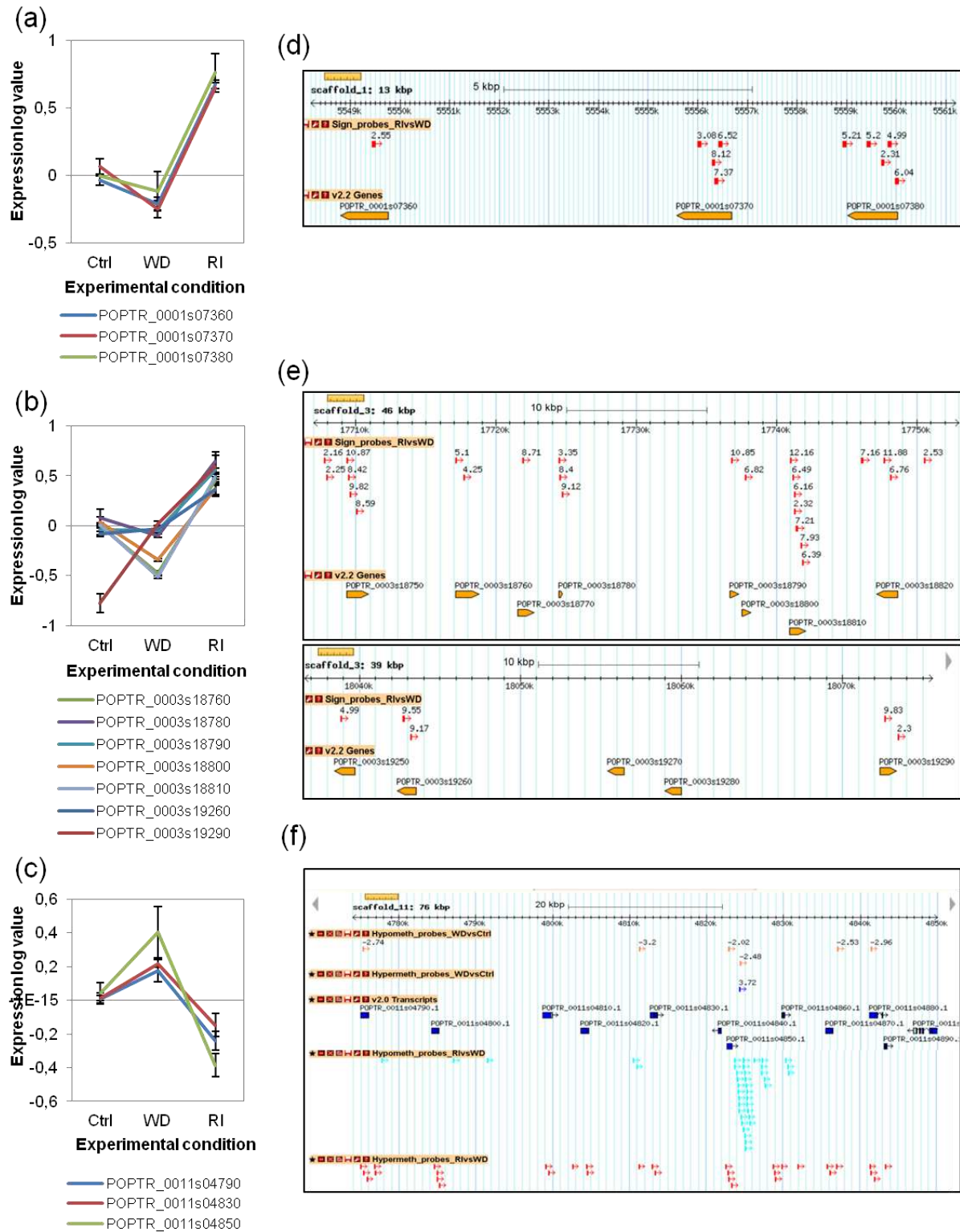


Figure 9

Sup data

qPCR: results

Figure S1

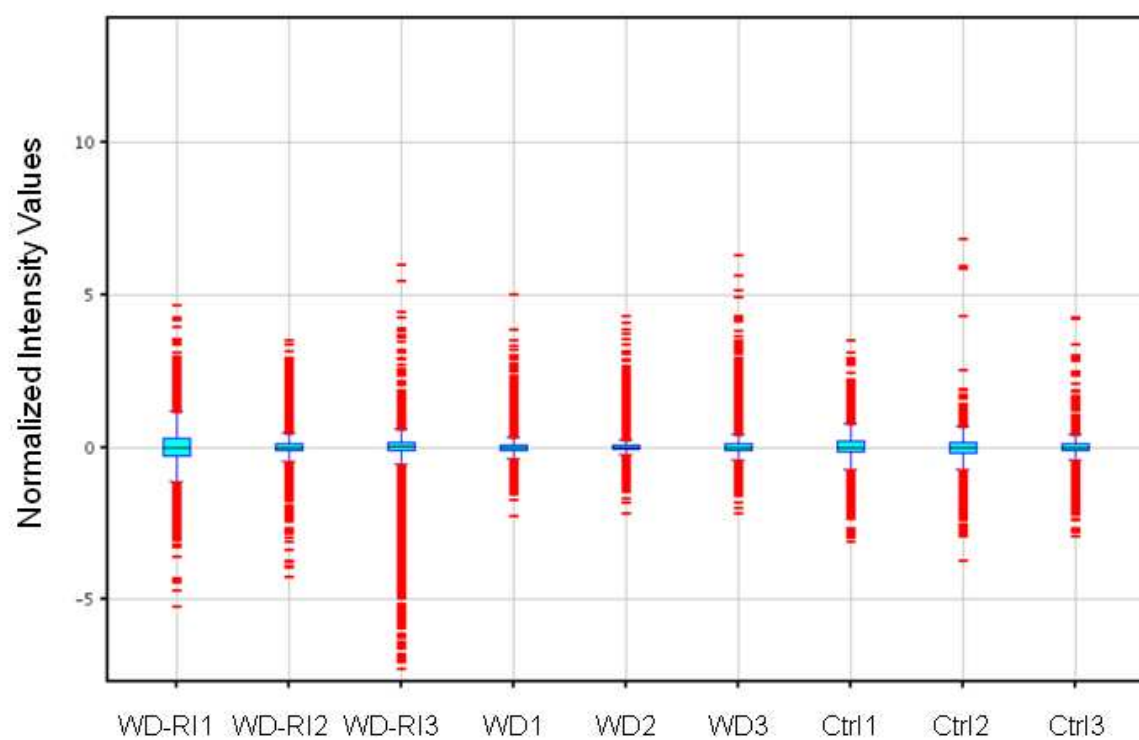


Figure S2

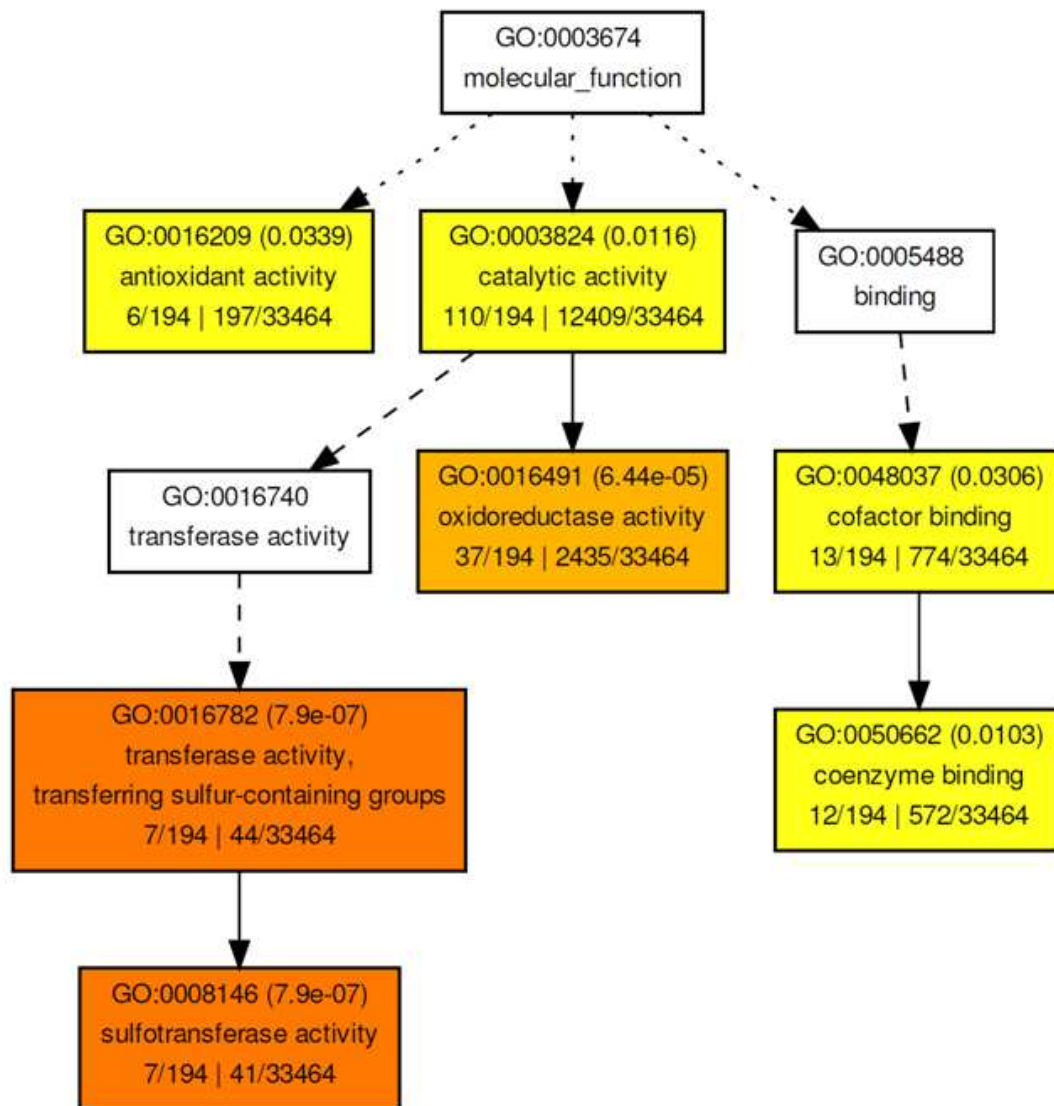


Fig S3. Sign expr changes, cluster 1

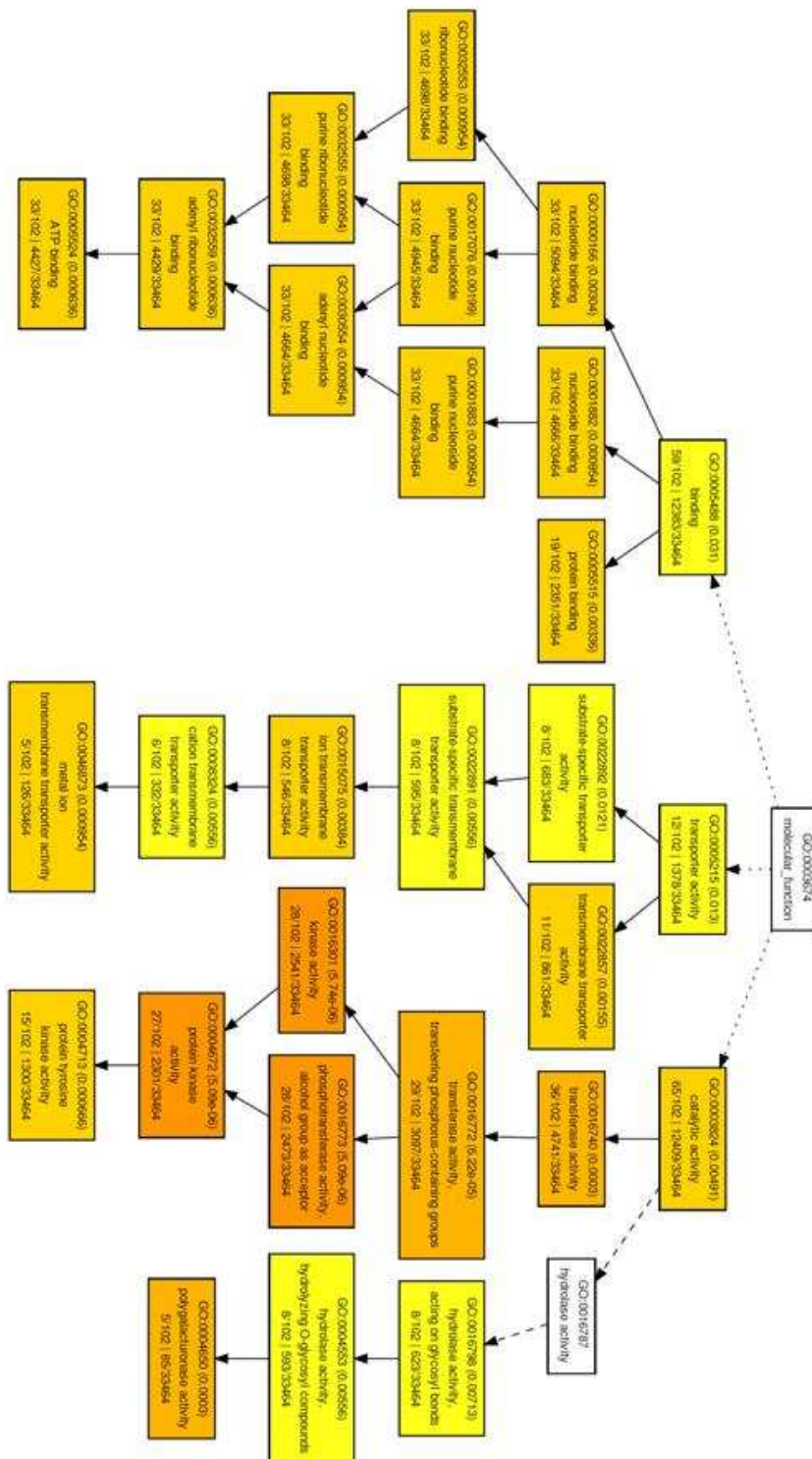


Fig S4. Sign expr changes, cluster 3

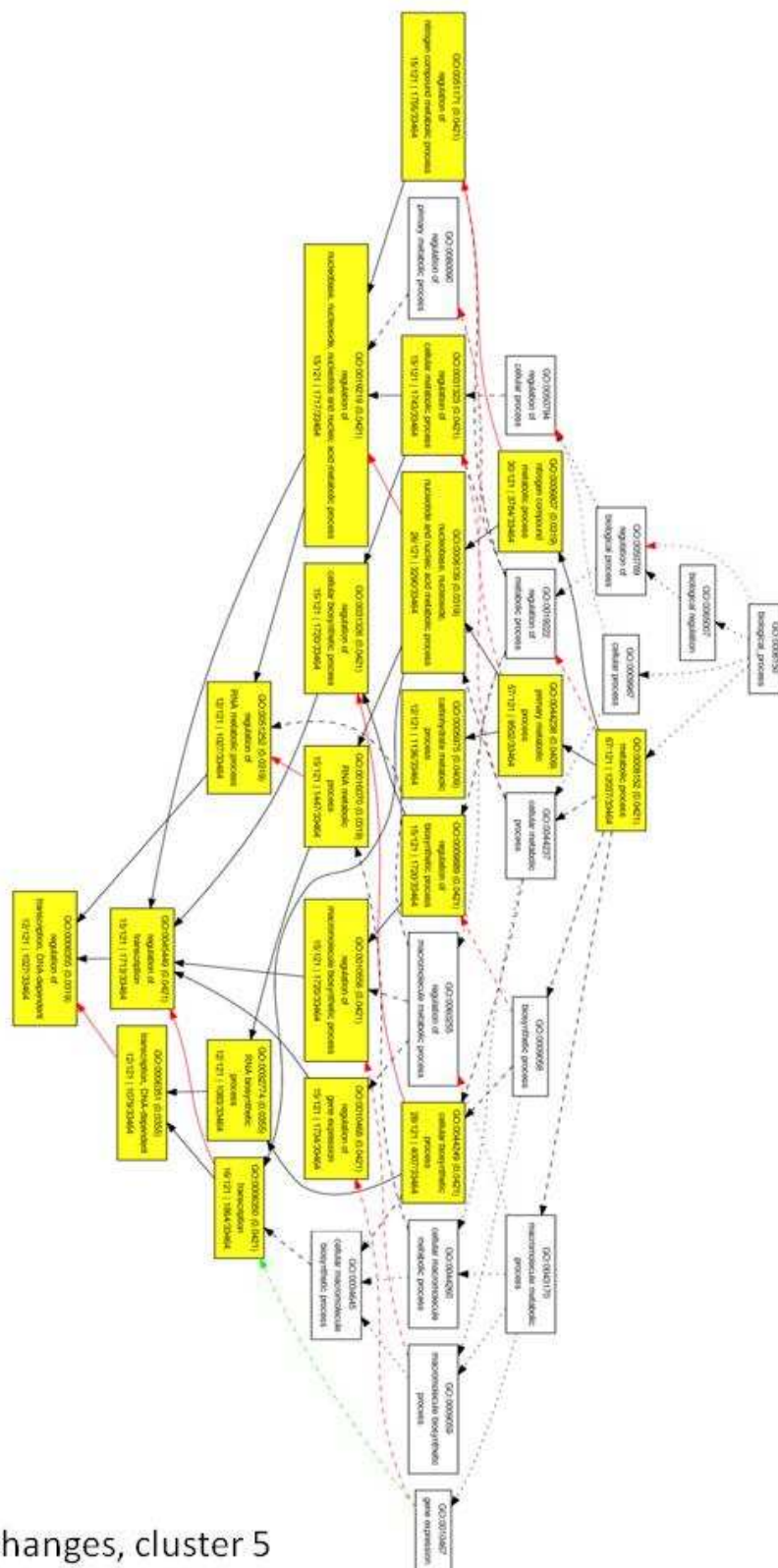


Fig S5. Sign expr changes, cluster 5

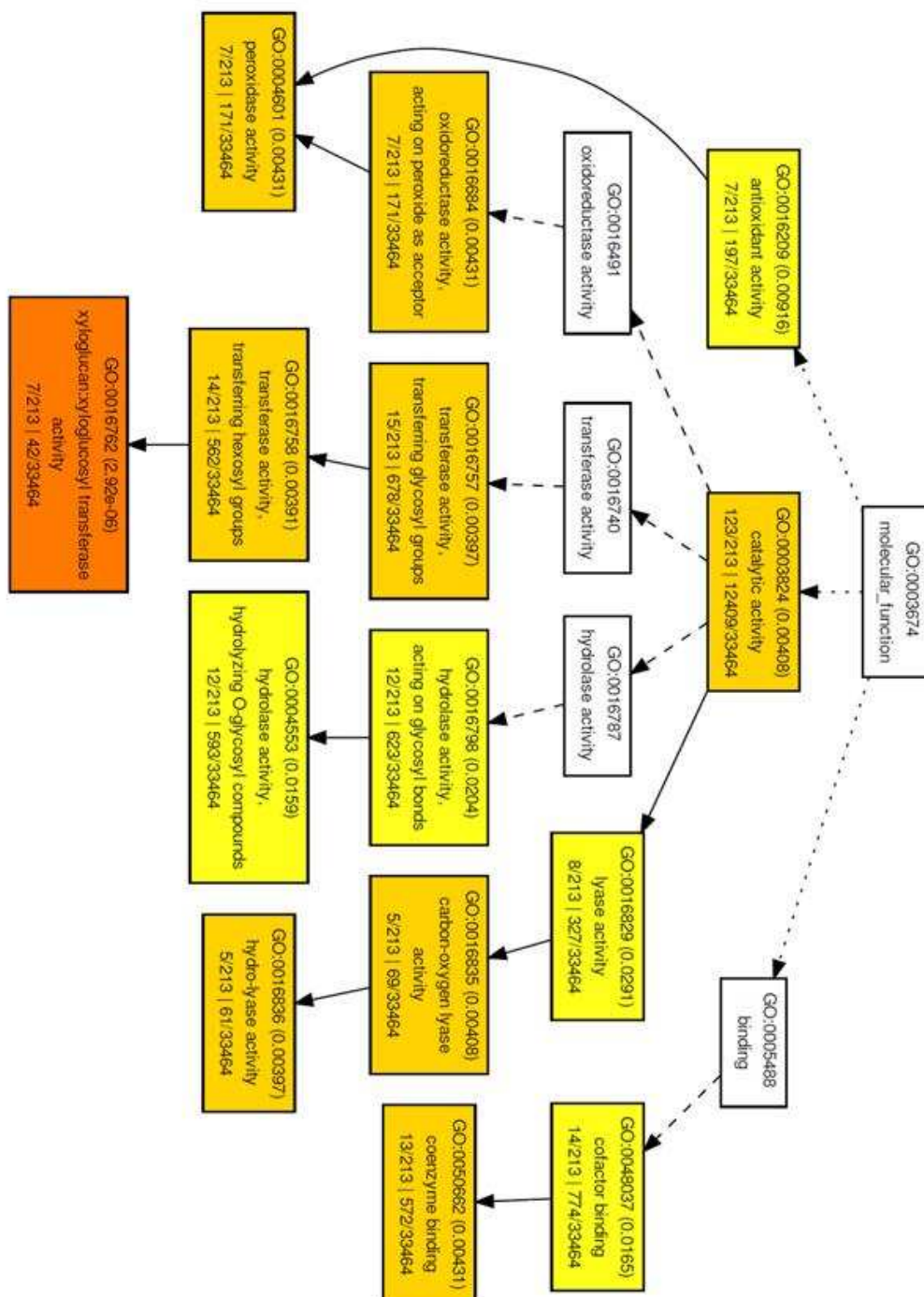
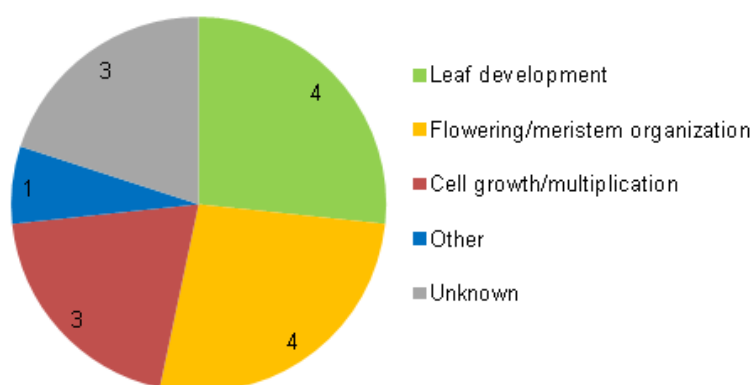


Fig S6. Sign expr changes, cluster 6

(a)



(b)

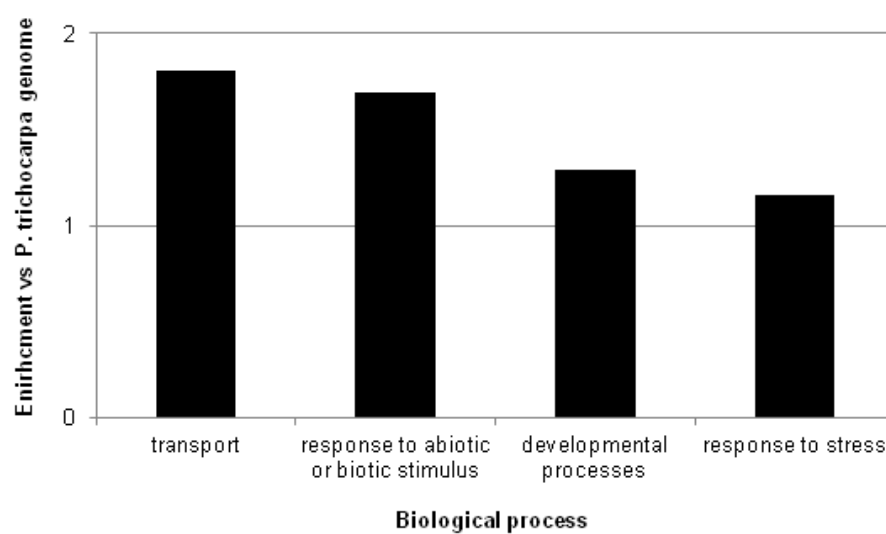


Figure S7

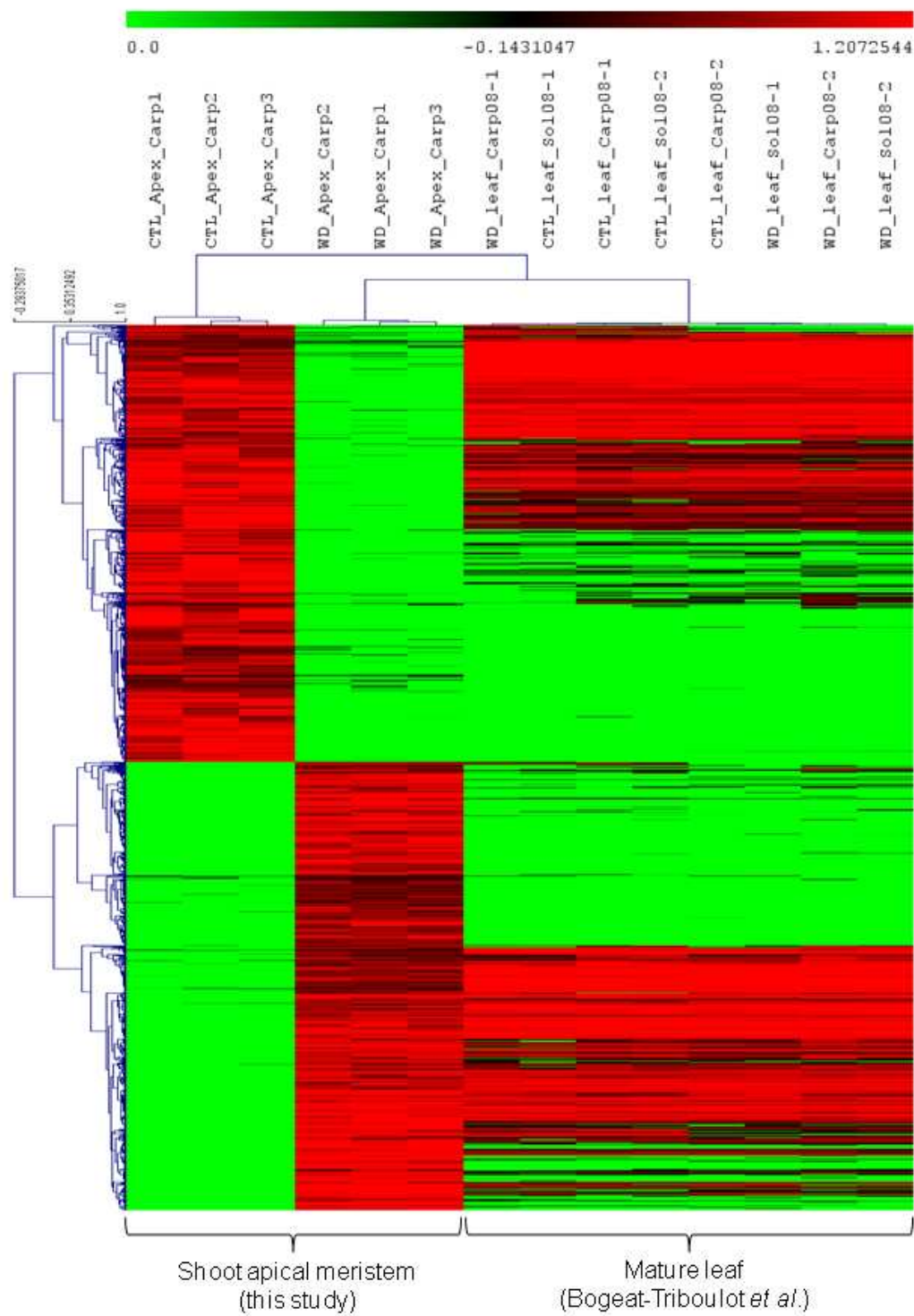


Figure S8

Bisulfite:
DNA methylation levels

Figure S9

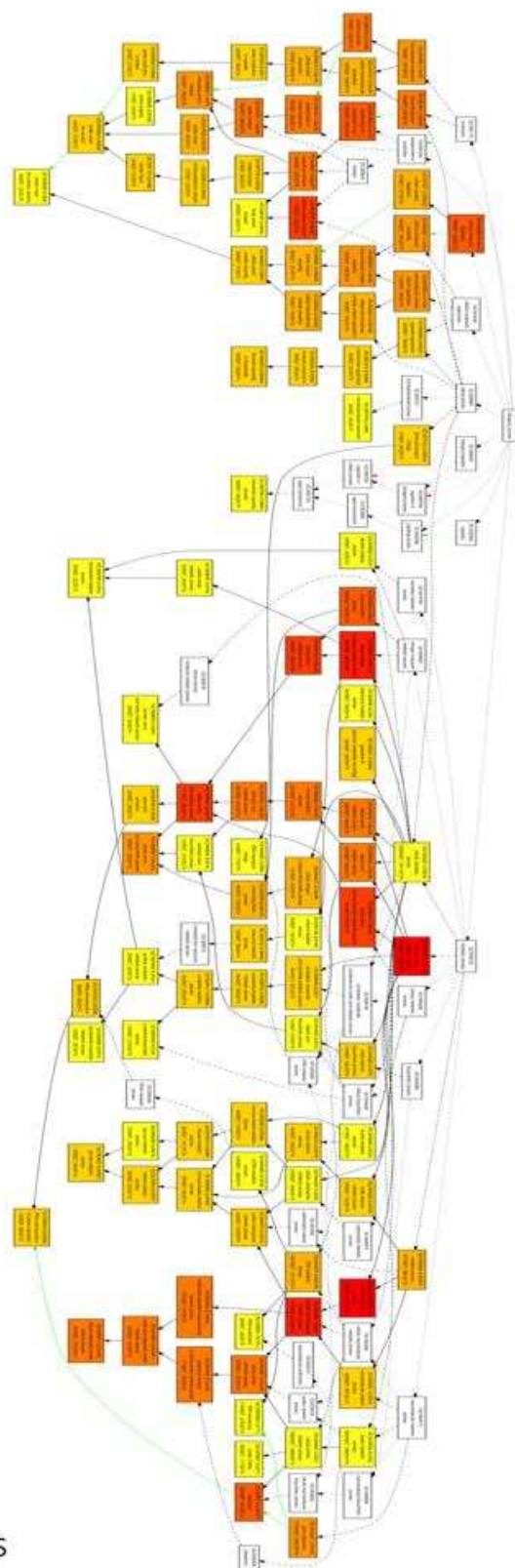


Fig S10. Sign DNA meth changes

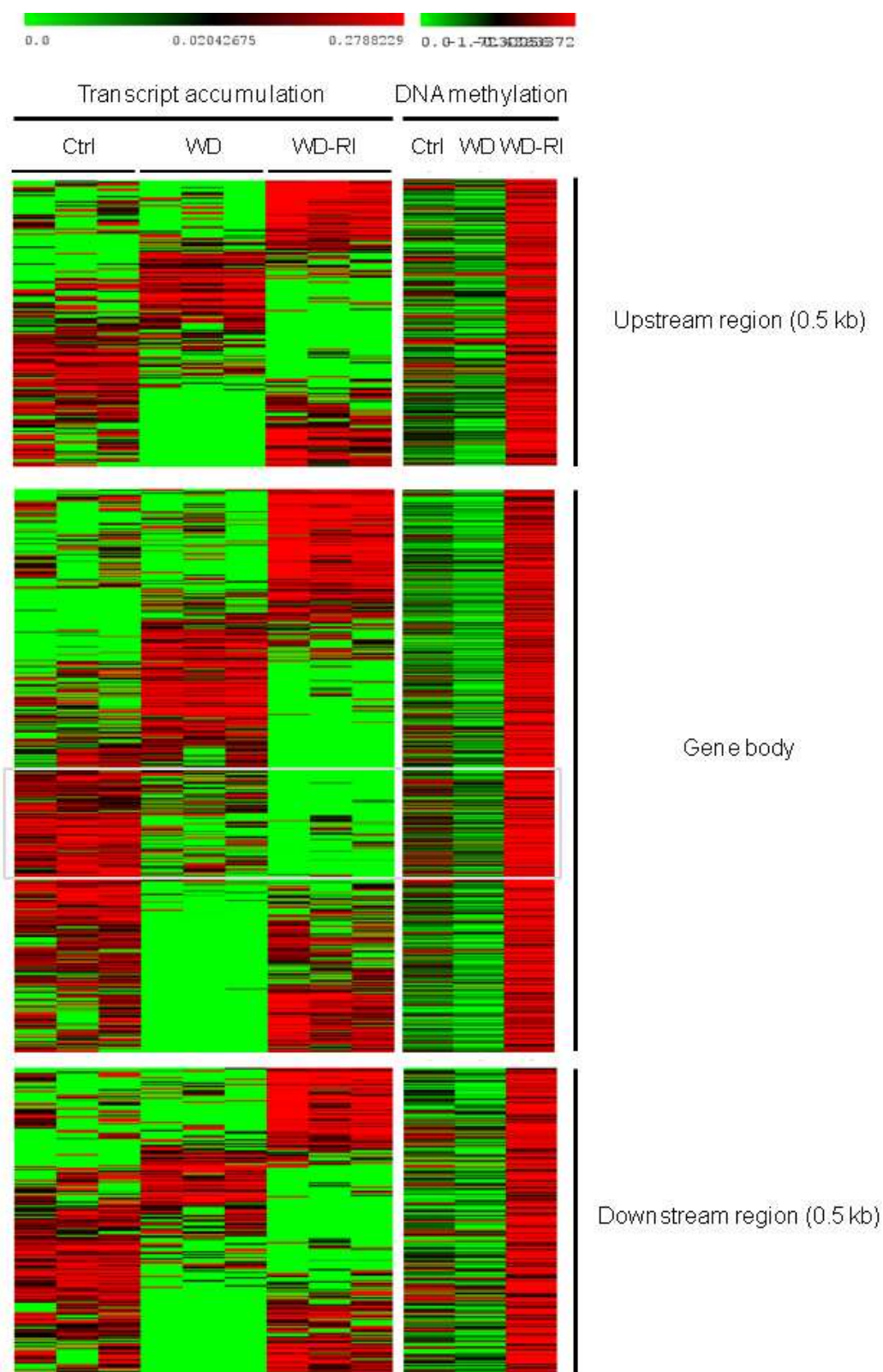


Figure S11

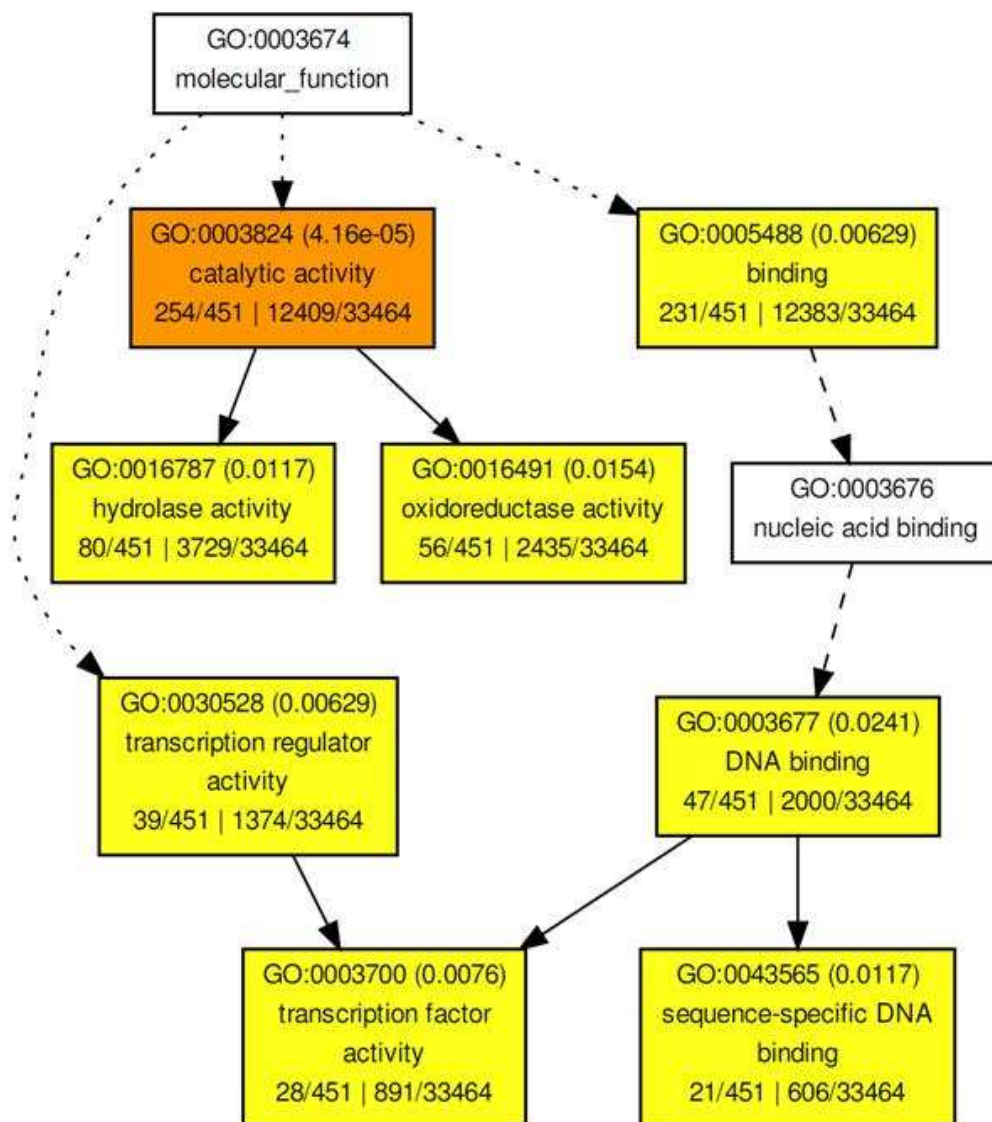


Fig S12. Sign DNA meth and expr changes

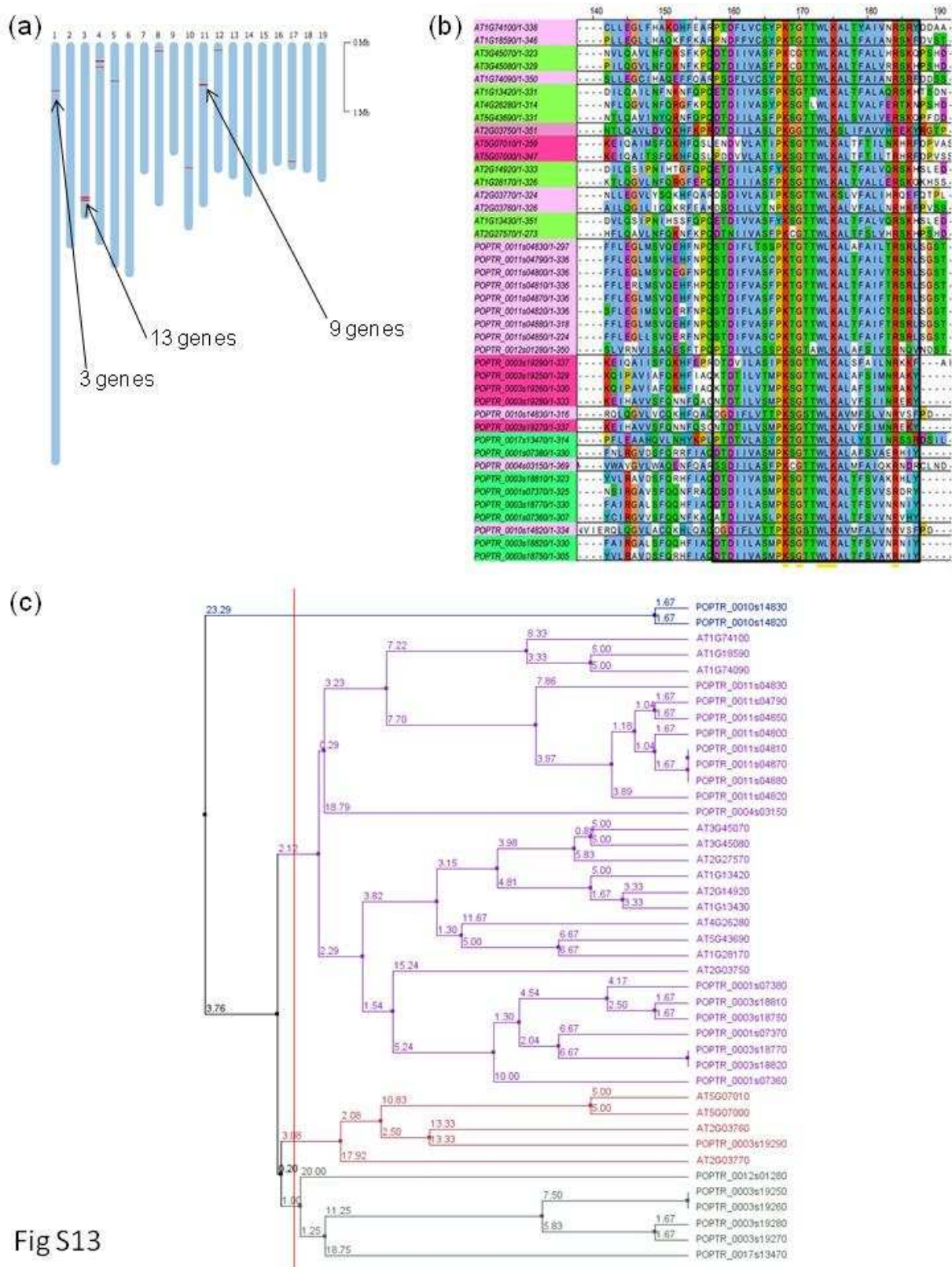


Fig S13

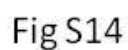


Table S1. qPCR and bisulfite primers (*Excel file*)

Table S2. Microarray probes summary. Agilent custom array design was used both for expression and DNA methylation detection (see Material and Methods). Two criteria were used to detect differentially expressed genes (one-way ANOVA analysis with $P < 0.01$ and Fold Change threshold of 2). For some genes, several probes were designed to allow a distinction between highly homologous sequences.

Probetype	Designed probes	Filtered probes	Significant probes		Significant genes
Expression	51677	33005	1499 (ANOVA)	1488 (Fold change)	1386 (ANOVA)
Methylation	988544	441214	270064		16330

Table S2

Table S3. Differentially expressed genes (*Excel file*). Criterion for differential expression is a significant difference among the three conditions as determined by one-way ANOVA analysis with $P < 0.01$.

Table S4. Downregulated transcription factors encoding genes both during water deficit and after rewatering

P. trichocarpav2.0 gene	Phytozome P. trichocarpa annotation	Arabidopsis homolog	Arabidopsis annotation
POPTR_0001s38120	Pt-PHB.2; similar to PHABULOSA	AT2G34710	Dominant PHB mutations cause transformation of abaxial leaf fates into adaxial leaf fates. Encodes a member of HD-Zip family which contains homeodomain-leucine zipper domains and domain similar to a mammalian sterol binding domain. Has overlapping functions with PHAVOLUTA, REVOLUTA and CORONA.
POPTR_0002s15430	PtpRR1; Myb-like DNA-binding domain	AT4G16110	Encodes a pollen-specific transcription factor involved in the expression of nuclear genes for components of mitochondrial complex I in Arabidopsis. Acts in concert with other type-B ARRs in the cytokinin signaling pathway. AHK3 mediates cytokinin-induced phosphorylation of ARR2 on the Asp-80 residue. This phosphorylation plays a positive role of ARR2 in cytokinin-mediated control of leaf longevity.
POPTR_0005s12750	bZIP transcription factor	AT4G36730	member of a gene family encoding basic leucine zipper proteins (GBFs) which bind the G-box
POPTR_0005s25240	NA	AT1G27660	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
POPTR_0006s11630	similar to transcription activator GRL4	AT2G36400	Growth regulating factor encoding transcription activator. One of the nine members of a GRF gene family, containing nuclear targeting domain. Mutants result in smaller leaves indicating the role of the gene in leaf development. Expressed in root, shoot and flower.
POPTR_0007s01330	DDT domain	AT5G44180	Homeodomain-like transcriptional regulator
POPTR_0009s02020	Pt-ARF10.1; similar to Auxin response factor 10	AT2G28350	Involved in root cap cell differentiation. Mutant plants exhibit developmental defects in leaves, flowers and siliques. the serrated and narrow-leaf mutant phenotypes are associated with the transgene insert. Plants display yellowish cauline leaves and sepals. Flowers have contorted and elongated petals and produce twisted siliques. During sensu stricto germination and post-germination, seeds and plants are hypersensitive to ABA in a dose-dependent manner.
POPTR_0010s24900	NA	AT5G04840	bZIP protein
POPTR_0013s15330	CW-type Zinc Finger	AT2G30470	HSI2 is a member of a novel family of B3 domain proteins with a sequence similar to the ERF-associated amphiphilic repression (EAR) motif. It functions as an active repressor of the Spo minimal promoter (derived from a gene for sweet potato sporamin A1) through the EAR motif. It contains a plant-specific B3 DNA-binding domain. The Arabidopsis genome contains 42 genes with B3 domains which could be classified into three families that are represented by ABI3, ARF1 and RAV1. HSI2 belongs to the ABI3 family. It is expressed at similar levels in all organs. Treatment with 6% sucrose showed a slight increase in transcript levels after 24 h. No changes were observed after treatment with 50 μM ABA. It is localized in the nucleus via a nuclear localization sequence located in the fourth conserved region of the C-terminal B3 domain.
POPTR_0014s06750	QLQ	AT2G45480	Growth regulating factor encoding transcription activator. One of the nine members of a GRF gene family, containing nuclear targeting domain. Involved in leaf development.
POPTR_0014s08020	Transcription mediator complex subunit Med12	AT4G00450	Encodes the Arabidopsis homolog of the transcriptional regulator MED12, is dynamically expressed during embryogenesis and regulates both developmental timing and the radial pattern formation. Involved in flowering time. The mutant enhances the expression of the flowering time (FT) gene. A knockout mutant of this gene showed late-flowering phenotype.
POPTR_0014s12730	Pt-LD.2; similar to LUMINIDEPENDENS	AT4G02560	Encodes a nuclear localized protein with similarity to transcriptional regulators. Recessive mutants are late flowering. Expression of LFY is reduced in LD mutants.
POPTR_0016s06860	CCAAT-binding transcription factor (CBF-B/NF-YA) subunit B	AT5G06510	"nuclear factor Y, subunit A10" (NF-YA10)
POPTR_0018s08110	Pt-HB1.5;	AT4G32880	member of homeodomain-leucine zipper family, acting as a differentiation-promoting transcription factor of the vascular meristems.
POPTR_0010s19680	similar to Paxneb protein-related	AT3G11220	NA

Table S5. Genes differentially expressed between control and water deficit conditions both in apex and leaves of *P. x euramericana* 'Carpaccio' hybrid. Welsch *t*-test was used to identify differentially expressed genes with $P < 0.01$. Common genes to comparison in apex (this study) and mature leaves (Bogeat-Triboulot et al., unpublished data; GEO accession number GSE17226) are displayed

<i>P. trichocarpa</i> v2.0 gene	<i>P. trichocarpa</i> annotation	Arabidopsis homolog	Arabidopsis annotation
POPTR_0001s10040	Reticulon family protein	AT1G64090	RTNLB3 (RETICULAN LIKE PROTEIN B3)
POPTR_0003s06520	Sodium/calcium exchanger protein	AT1G54115	ATCCX4 (CATION CALCIUM EXCHANGER 4, CCX4); Involved in cation (Na and K) homeostasis
POPTR_0004s02530	Protein tyrosine kinase	AT4G23180	CRK10 (CYSTEINE-RICH RLK (RECEPTOR-LIKE PROTEIN KINASE) 10); Encodes a receptor-like protein kinase
POPTR_0005s03920	Sugar (and other) transporter	AT1G08920	ESL1 (ERD (EARLY RESPONSE TO DEHYDRATION) SIX-LIKE 1); Encodes a transporter for monosaccharides
POPTR_0006s10510	Pt-DREB2.5; similar to Dehydration responsive element binding protein 2E (DREB2E protein)	AT2G40340	DREB2C; Encodes a DREB subfamily A-2 of ERF/AP2 transcription factor family protein
POPTR_0012s05840	Pt-ARP4.2; similar to ACTIN-RELATED PROTEIN 5	AT1G18450	ARP4 (ACTIN-RELATED PROTEIN 4); Member of nuclear ARP family of genes. Component of chromatin remodeling complexes, involved in chromatin-mediated gene regulation
POPTR_0015s04880	NA	AT3G17860	JAZ3 (JASMONATE-ZIM-DOMAIN PROTEIN 3); JAZs are direct targets of the SCFCOI1 E3 ubiquitin-ligase and JA treatment induces their proteasome-mediated degradation
POPTR_0016s10780	Pt-TIP2.8; similar to Aquaporin TIP1.2	AT2G36830	TIP1.1 (TONOPLAST INTRINSIC PROTEIN 1.1); tonoplast intrinsic protein, which functions as water channel
POPTR_0017s02250	similar to oxidoreductase N-terminal domain-containing protein	AT1G66130	NAD(P)-binding Rossmann-fold superfamily protein
POPTR_0018s01990	Calcineurin-like phosphoesterase	AT5G34850	PAP26 (PURPLE ACID PHOSPHATASE 26); dual-targeted purple acid phosphatase isozyme AtPAP26. Involved in phosphate metabolism

Table S6. Differentially methylated genes (*Excel file*). Given values are fold changes between control (Ctrl), water deficit (WD) or re-irrigation (WD-RI) conditions. A gene was defined as differentially methylated if at least one of its part (5' or 3' 500 bp flanking regions, body) displayed a fold change $> |2|$ between two conditions.

Table S7. Pearson correlations between DNA methylation and expression levels. DNA methylation levels were investigated in 500 bp upstream and downstream regions as well as gene body. WD-R1: Rewatered samples; WD: water deficit sample; Ctrl: well watered sample. Significant correlations at $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ are indicated.

[illegible]

Table S8. Pearson correlations between DNA methylation and expression changes. DNA methylation changes were investigated in 500 bp upstream and downstream regions as well as gene body. WD-R1: Rewatered samples; WD: water deficit sample; Ctrl: well watered sample. Significant correlations at * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ are indicated.

[illegible]

3. Discussion, limites et perspectives

Réponses du transcriptome et du méthylome dans le méristème apical caulinaire à des variations de la disponibilité en eau du sol

Afin de préciser le rôle de la méthylation de l'ADN lors d'une contrainte hydrique chez le peuplier (Gourcilleau *et al.*, 2010), il était nécessaire de caractériser les loci affectés par des variations de méthylation de l'ADN et d'expression lors de fluctuations de la disponibilité en eau du sol. A cette fin, le génotype 'Carpaccio' (*P. deltoides* × *P. nigra*) a été soumis à un cycle de déficit hydrique/réarrosage. La réponse écophysiolgique au déficit hydrique était celle classiquement observée chez le peuplier lors de stress d'intensité modérée (Cohen *et al.*, 2010; Gourcilleau *et al.*, 2010), à savoir une diminution de la conductance stomatique et de la croissance notamment, sans altération du potentiel hydrique de base ni de l'assimilation nette de CO₂. De manière intéressante, après le réarrosage, certaines variables écophysiolgiques revenaient au niveau témoin (conductance stomatique, ajustement osmotique des feuilles) alors que d'autres comme la vitesse de croissance ou la masse foliaire spécifique restaient encore affectées comme chez les plantes sans réarrosage. Ceci peut traduire des cinétiques différentes, certaines réponses écophysiolgiques ne nécessitant pas de modifications profondes de la morphologie de la plante, étant rapides (conductance stomatique) alors que d'autres, impliquant des changements dans le programme développemental, étant plus longues à réagir aux variations de l'environnement (morphogenèse foliaire et croissance).

Les résultats de l'étude transcriptomique du méristème apical suggèrent également cette idée puisque le profil d'expression des gènes après réarrosage était plus complexe qu'un simple retour au profil de condition hydrique non limitante, avec 481 gènes différentiellement exprimés entre ces deux conditions. De plus, une grande proportion de gènes était spécifiquement sur ou sous-exprimée après le réarrosage, montrant la spécificité de cet état physiologique à l'échelle moléculaire. Cette particularité s'est retrouvée également à l'échelle du méthylome, avec un niveau de méthylation globale plus important et une distribution génomique très différente comparée aux deux autres conditions. Les variations d'expression de gènes liées à la ré-irrigation, les plus importantes en terme d'ampleur, accompagnaient ce méthylome, suggérant un fort remodelage chromatinien de l'activité transcriptionnelle. Cependant, l'observation d'un fort niveau de méthylation de l'ADN dans les chromosomes mitotiques chez le maïs, plus particulièrement en prophase et métaphase (Yang *et al.*, 2010), laisse penser à une forte activité mitotique des cellules du méristème apical lors du retour à des conditions hydriques non limitantes. Des observations

cytologiques ou des mesures de l'index mitotique seraient nécessaires pour tester cette hypothèse.

En ce qui concerne l'effet du déficit hydrique, une hypométhylation de l'ADN dans les cellules du méristème apical caulinaire a été observée chez le génotype 'Carpaccio' (*Populus deltoides* × *P. nigra*), résultat similaire à celui obtenu lors d'une étude précédente (Gourcilleau *et al.*, 2010). Cette hypométhylation semble donc une réponse à une sécheresse modérée caractéristique de 'Carpaccio', réponse présentant un caractère génotype-dépendant (Gourcilleau *et al.*, 2010). De plus, elle est retrouvée essentiellement dans les zones riches en gènes. Une étude récente avait déjà montré que les loci affectés par des variations de méthylation lors d'un stress salin étaient en majorité des corps de gènes chez le riz (Karan *et al.*, 2012). Ainsi, contrairement à ce qui a pu être suggéré chez *Arabidopsis* (Pecinka & Mittelsten Scheid, 2012), les variations de méthylation en réponse à la sécheresse chez le peuplier ne semblent pas concentrées uniquement dans les zones périceptomériques mais affectent de façon significative des gènes.

Plusieurs mécanismes moléculaires de réponse à la sécheresse sont déjà bien décrits chez les plantes (Seki *et al.*, 2007; Des Marais & Juenger, 2010), mais essentiellement dans la feuille et la racine voire à l'échelle de la plante entière. Dans cette étude, le méristème apical a été étudié. L'expression de gènes déjà connus de la réponse à la sécheresse était altérée suite aux variations de la disponibilité en eau, tels que des gènes codant pour des protéine sérine/thréonine kinases, impliqués dans la perception de signaux (Ha *et al.*, 2012), d'autres gènes impliqués dans le catabolisme des protéines, processus modifié lors d'une sécheresse (Callis & Vierstra, 2000; Cho *et al.*, 2008), ou encore des gènes codant des enzymes antioxydantes, jouant un rôle dans la protection cellulaire contre les formes actives de l'oxygène et la signalisation associée à ces molécules, à la fois lors d'une sécheresse mais aussi lors de la reprise de croissance après une réirrigation (Xu *et al.*, 2010; Filippou *et al.*, 2011). De manière plus intéressante, certains gènes étaient inhibés par le déficit hydrique et leur expression restait affectée après le réarrosage. Parmi eux, des facteurs de transcription impliqués dans la morphogenèse foliaire ou la multiplication cellulaire chez *Arabidopsis* étaient fortement représentés, comme par exemple un homologue de *PHABULOSA*, gène impliqué dans la taille du méristème apical caulinaire et la polarité des feuilles chez *Arabidopsis* (Grigg *et al.*, 2005). Aussi, le maintien dans le méristème apical du profil d'expression de ces gènes après le déficit hydrique pourrait expliquer en partie la reprise de croissance faible.

Une comparaison de la réponse transcriptomique à la sécheresse au sein du même génotype et du même dispositif expérimental (Bogeat-Triboulot *et al.*, unpublished data),

entre méristème apical et feuilles mûres, a révélé qu'une faible proportion des gènes répondait de manière commune à la sécheresse entre les deux organes (10 sur 1719 gènes identifiés par *t*-test dans le méristème apical). L'utilisation de design de sondes différents (à façon Agilent et version 2.0 des gènes de *P. trichocarpa* dans notre étude vs design Affymetrix et v1.1 pour les feuilles), de plateformes transcriptomiques et leurs normalisations propres est une autre raison possible. Par ailleurs, l'expression de nombreux gènes dans le méristème apical, liés à son fonctionnement, à la morphogenèse foliaire ou à la floraison a été modifiée par les variations de disponibilité en eau du sol. Alors que les deux premières fonctions semblent évidentes dans le cadre d'une plasticité développementale en réponse au stress, la modification de gènes impliqués dans la floraison est surprenante. Néanmoins, le déclenchement de la floraison a été observé en réponse à des stress environnementaux, considéré comme un moyen d'y échapper pour les plantes (Des Marais & Juenger, 2010).

En outre, le réarrosage après un déficit hydrique a entraîné une large gamme de variations du transcriptome et du méthylome dans le méristème apical caulinaire. Alors qu'un grand nombre de processus liés à la perception de signal ou au fonctionnement cellulaire semble commun au méristème apical caulinaire et les autres organes étudiés jusque-là, les gènes ayant une action dans le développement semblent impactés de manière spécifique dans le méristème apical caulinaire. Cela confirmerait l'hypothèse de son rôle dans la plasticité phénotypique de la tige feuillée en réponse à des variations de la disponibilité en eau du sol chez le peuplier.

Mise en évidence de relations entre transcription et méthylation des gènes dans le méristème apical caulinaire selon la disponibilité en eau du sol

Alors que plusieurs travaux ont démontré un rôle de la méthylation de l'ADN dans le contrôle de la transcription de quelques gènes en réponse à la sécheresse (González *et al.*, 2011; Bilichak *et al.*, 2012), peu d'études à l'échelle du génome sont disponibles à ce jour. Néanmoins, une étude récente a montré que les variations de méthylation de l'ADN en réponse à un stress salin affectaient préférentiellement le corps des gènes, impactant leur expression (Karan *et al.*, 2012). Dans l'étude présente, une corrélation positive chez le génotype 'Carpaccio' (*P. deltoides* × *P. nigra*) a été montrée entre niveau de méthylation du corps des gènes et leur niveau d'expression en condition hydrique non limitante, corrélation non détectée lors du déficit hydrique. De façon intrigante, cette observation rappelle que dans l'étude précédente (Gourcilleau *et al.*, 2010), une corrélation positive entre niveau de méthylation global de l'ADN dans l'apex caulinaire et production de biomasse avait été trouvée en condition hydrique non limitante mais pas en condition de déficit hydrique faible.

Une hypothèse serait alors que les variations de méthylation de gènes et de leur expression pourraient expliquer les variations de production de biomasse observées. Une étude au niveau de QTLs de production de biomasse chez le peuplier (Rae *et al.*, 2007, 2009; Monclus *et al.*, 2012) serait ainsi intéressante à réaliser pour établir un lien concret entre méthylation dans le corps des gènes et la capacité de production de biomasse.

Néanmoins, le lien quantitatif mis en évidence entre variations de méthylation de l'ADN des gènes et leur expression était faible et dépendant des conditions biologiques étudiées. Cela est cohérent avec le grand nombre d'épimutations naturelles n'ayant pas de conséquence sur le niveau de transcription de gènes chez des lignées isogéniques d'*Arabidopsis* (Schmitz *et al.*, 2011; Becker *et al.*, 2011). Cette absence de relation quantitative a été constatée à de maintes reprises (Zhang *et al.*, 2006; Vaughn *et al.*, 2007), suggérant un rôle de la méthylation dans le fonctionnement de la transcription comme l'épissage des introns ou l'élimination de sites cryptiques d'initiation dans le corps des gènes (Zilberman *et al.*, 2007; Chodavarapu *et al.*, 2010) plus que dans son activité quantitative. Plus particulièrement, la méthylation affectant préférentiellement l'ADN lié au nucléosome (Chodavarapu *et al.*, 2010), les variations de méthylation de l'ADN observées pourraient être indicatrices de la dynamique des nucléosomes (mouvement, occupation) lors de la transcription (Schones *et al.*, 2008; Gilchrist *et al.*, 2010). Ceci expliquerait pourquoi les gènes exprimés sont fortement remodelés par la méthylation de l'ADN lors des variations de disponibilité en eau du sol, possiblement sujets à une dynamique des nucléosomes importante, alors que les gènes non exprimés quelque soit la condition hydrique sont beaucoup moins affectés par les variations de méthylation. Cette dynamique de nucléosomes est également retrouvée au sein des éléments *cis* régulateurs et liée directement à la fixation de facteurs de transcription (Zhang *et al.*, 2012b). Dans ce contexte, les variations de méthylation préférentiellement observées au niveau des éléments régulateurs ABRE ou DRE sont cohérentes avec la mise en place d'une réponse moléculaire, ABA dépendante et indépendante, lors de variations de disponibilité en eau chez le peuplier (Des Marais & Juenger, 2010).

En conclusion, une corrélation positive entre niveau de méthylation du corps des gènes et niveau de leur expression a été trouvée en condition hydrique non limitante. Les remaniements de la méthylation de l'ADN en réponse à des variations de la disponibilité en eau ont affecté préférentiellement les gènes exprimés dans les cellules du méristème apical chez le peuplier. Ces données laissent ainsi supposer un rôle de cette marque épigénétique dans le contrôle de la transcription des gènes dans le méristème apical du peuplier et donc potentiellement dans la plasticité phénotypique de la tige feuillée en réponse à la sécheresse.

Un contrôle épigénétique de gènes impliqués dans le signal hormonal est mis en évidence dans le méristème apical caulinaire lors de variations de la disponibilité en eau du sol.

Un nombre croissant de travaux tend à montrer que la réponse aux stress biotiques et abiotiques fait intervenir des voies moléculaires interconnectées (Des Marais & Juenger, 2010; Chan, 2012). En effet, une majeure partie des gènes impliqués dans la réponse à la sécheresse répond également au froid ou au stress osmotique alors que certains gènes impliqués dans la réponse aux pathogènes sont induits par l'ABA (Chan, 2012). Dans l'étude présente, la surexpression, lors du déficit hydrique, de *RLK* impliqués dans la réponse aux pathogènes (comme *RLK4* ; Du & Chen, 2000) ou encore le remaniement par la méthylation de l'ADN d'éléments *cis* régulateurs liés au stress biotique (GCC box et box S ; Rushton *et al.*, 2002) en réponse aux variations de la disponibilité en eau indiquent cette interconnexion. Confirmant cette idée, trois gènes codant pour des sulfotransférases étaient surexprimés lors du déficit hydrique. Les protéines codées par ces gènes sont, d'après comparaison du site actif des mêmes enzymes chez *Arabidopsis*, impliquées dans la biosynthèse des glucosinolates, molécules impliquées notamment dans la défense des plantes contre les agents biotiques (Klein *et al.*, 2006). Ces molécules ont cependant un spectre d'action beaucoup plus large, allant de la signalisation de stress abiotiques à la modulation du signal auxine (Baskar *et al.*, 2012).

L'interconnexion des voies de réponse aux stress biotiques et abiotiques est intrinsèque au développement des plantes, contrôlé par les interactions entre plusieurs phytohormones et leur dosage (Wolters & Jürgens, 2009). En effet, la réponse à la sécheresse et au ré-arrosage a été montrée pour impliquer l'ABA, hormone clé de la réponse aux stress abiotiques, mais aussi l'éthylène, l'acide jasmonique, les brassinostéroïdes, l'auxine ou encore les gibbérellines (Oono *et al.*, 2003; Seki *et al.*, 2007; Wolters & Jürgens, 2009; Choudhary *et al.*, 2012), hormones contrôlant également la réponse aux stress biotiques. Dans la présente étude, sur les 21 gènes annotés et drastiquement affectés pour leur méthylation et expression lors des variations de disponibilité en eau du sol, 8 d'entre eux avaient un rôle connu dans la réponse à différentes phytohormones, suggérant bien que la réponse aux variations de la disponibilité en eau du sol implique l'interaction de signaux hormonaux. Notre découverte majeure est qu'un contrôle épigénétique est impliqué. Par exemple, *RNA-DEPENDENT RNA POLYMERASE 1* (*RDR1*) était hypométhylé et surexprimé lors du déficit hydrique et hyperméthylé et inhibé lors du ré-arrosage. Les mutants de tabac pour ce gène, impliqué dans le silencing *via* la production de petits ARN, ont montré une hypersensibilité à des attaques d'insectes liée à une faible production d'acide jasmonique mais à une augmentation de la concentration

d'éthylène en réponse au stress, elles-mêmes provoquées par la dérégulation de gènes ciblés par *RDR1* (Pandey *et al.*, 2008). La surexpression de ce gène lors du déficit hydrique suggère donc un contrôle positif et négatif respectivement de la production d'acide jasmonique et d'éthylène. Dans cette idée, l'inhibition de *RDR1* lors du réarrosage est cohérente avec la surexpression d'un gène codant un facteur de réponse à l'éthylène, *ERF023*, également hyperméthylé lors du réarrosage. Ce résultat est en accord avec le rôle déjà suggéré de l'éthylène dans la reprise de croissance lors d'un réarrosage après un déficit hydrique, et ce dans plusieurs études dont l'une chez le peuplier (Oono *et al.*, 2003; Berta *et al.*, 2010).

Enfin, les brassinostéroïdes ont été impliqués dans la réponse à la sécheresse, induisant la biosynthèse de l'ABA ou activant directement des gènes induits par l'ABA (Choudhary *et al.*, 2012). Les résultats de l'étude présente ont mis en évidence qu'une famille de gènes, codant des sulfotransférases dont les substrats putatifs sont des brassinostéroïdes, était fortement exprimée et méthylée lors de la ré-irrigation. La sulfonation des brassinostéroïdes par ces enzymes a été montrée pour désactiver ces phytohormones (Rouleau *et al.*, 1999; Marsolais *et al.*, 2007). Cette modification chimique, effectuée lors du réarrosage, permettrait donc l'élimination du signal hormonal lié à la sécheresse tout en maintenant l'homéostasie des brassinostéroïdes, nécessaire à une croissance normale des plantes (Choudhary *et al.*, 2012). Elle serait donc une étape essentielle à la reprise de croissance après la fin du déficit hydrique. L'inhibition du gène *EXL5* lors du déficit hydrique, induit chez *Arabidopsis* par un traitement aux brassinostéroïdes à un dosage favorisant la croissance (Schröder *et al.*, 2009), est incohérente avec l'hypothèse d'un signal brassinostéroïdien en réponse à la sécheresse complémentant l'action de l'ABA (Choudhary *et al.*, 2012). Il est cependant possible qu'*EXL5* soit inhibé par l'ABA ou qu'il réponde aux brassinostéroïdes de manière dose dépendante. Le contrôle de l'homéostasie des brassinostéroïdes par l'action des sulfotransférases (Choudhary *et al.*, 2012) lors du réarrosage pourrait ainsi expliquer l'induction dans cette condition de l'expression d'*EXL5*. L'inhibition d'*EXL5* accompagnant la perte de croissance observées en déficit hydrique et sa surexpression lors du réarrosage sont cohérentes avec la régulation de gènes impliqués dans la croissance en réponse aux brassinostéroïdes par le gène *EXO* (Coll-Garcia *et al.*, 2004), de la même famille qu'*EXL5*, dont les mutants montrent un phénotype de croissance réduite (Schröder *et al.*, 2009).

En conclusion, ce travail a montré les modifications moléculaires que subissent les cellules du méristème apical caulinaire en réponse à des variations de la disponibilité en eau du sol chez le peuplier. La succession d'un déficit hydrique et d'un retour à un régime hydrique non limitant a permis de mettre en évidence la complexité des réponses

physiologiques et moléculaires mises en place chez le peuplier et permettant une plasticité phénotypique face aux fluctuations de la disponibilité en eau du sol. Les mécanismes moléculaires permettant une reprise de croissance après un déficit hydrique ayant été peu décrits (Xu *et al.*, 2010), ces travaux contribuent ainsi à en améliorer notre compréhension. Les modifications moléculaires observées incluent des changements épigénomiques et transcriptomiques de gènes impliqués dans le développement. Ces gènes sont de potentiels déterminants de la plasticité phénotypique de la tige feuillée en réponse à des variations de la disponibilité en eau du sol chez le peuplier. Cette plasticité serait le résultat d'interactions complexes entre les différentes voies de phytohormones au niveau du méristème apical. Il serait maintenant pertinent d'étudier au cours d'une cinétique plus fine et précise ou de plusieurs cycles successifs de déficit hydrique/réarrosage ces mécanismes moléculaires ainsi que la variabilité génotypique de l'expression et la méthylation des gènes décrits dans ce travail.

Conclusions générales, limites et perspectives

Les travaux effectués dans cette thèse s'inscrivent dans une problématique de réponse des arbres aux variations environnementales et ceci dans un contexte de changements climatiques. Plus précisément, ils se basent sur les travaux entrepris au cours des 15 dernières années au sein du Laboratoire de Biologie des Ligneux et des Grandes Cultures de l'Université d'Orléans. Les recherches se sont focalisées sur le peuplier et les variations génétiques et phénotypiques de caractères complexes d'intérêt comme le potentiel de croissance, l'efficacité d'utilisation de l'eau (WUE) ou la tolérance à la sécheresse. L'axe de recherche dans lequel s'inscrit cette thèse et développé depuis 2007 par Stéphane Maury vise à caractériser les déterminants épigénétiques de ces variables et de leurs variations.

Dans ce contexte, l'objectif général des travaux présentés dans ce manuscrit était de caractériser chez le peuplier le rôle de la méthylation de l'ADN dans l'expression des gènes puis d'évaluer son rôle dans la plasticité développementale de la tige feuillée en réponse à la sécheresse. Le choix de la méthylation de l'ADN se justifie par le fait que cette marque épigénétique, bien que réversible et variable au cours du développement et en réponse à l'environnement, est transmise par mitose et méiose, qu'elle semble jouer un rôle déterminant entre autres sur l'expression des gènes en relation avec les autres marques épigénétiques (modifications des histones et petits ARN) et la mobilisation des transposons (variabilité génétique). Ces caractéristiques et la bibliographie disponible proposent que la méthylation de l'ADN joue un rôle majeur aussi bien dans la plasticité phénotypique que dans l'adaptation. Nous avons alors fait l'hypothèse que ces mécanismes épigénétiques pourraient moduler l'expression des gènes dans le méristème apical caulinaire contrôlant ainsi la plasticité de la morphogenèse de la tige feuillée en réponse à un déficit hydrique comme cela est bien caractérisé chez le peuplier (Marron *et al.*, 2002, 2003; Monclus *et al.*, 2006; Fichot *et al.*, 2009). Ce choix est d'autant plus pertinent que le méristème d'*Arabidopsis* est trop petit pour être étudié contrairement au peuplier.

Plus spécifiquement, les objectifs de cette thèse étaient alors de (i) caractériser le méthylome de la chromatine non condensée potentiellement riche en gènes accessibles à la machinerie de la transcription dans le méristème apical et ses relations avec l'expression des gènes, (ii) étudier l'impact d'une sécheresse sur la méthylation de l'ADN de l'apex caulinaire dans différents géotypes hybrides *Populus deltoides* × *P. nigra* en parallèle d'études écophysiologiques et (iii) identifier les loci et réseaux de gènes affectés pour leur expression et leur méthylation lors de variations de la disponibilité en eau.

Dans cette section du manuscrit, la synthèse des résultats obtenus en réponse à ces objectifs, poursuivie par la mise en lumière des limites de cette étude, permettra de proposer les perspectives majeures dégagées par ces travaux.

1. Synthèse des résultats et apports des travaux de la thèse

Hormis les travaux de Master 2 de Delphine Gourcilleau (2007) au laboratoire (Chapitre 2 ; Gourcilleau *et al.*, 2010), aucune publication sur l'épigénétique du peuplier n'était disponible au démarrage de cette thèse. En tant qu'arbre modèle et compte tenu des raisons déjà évoquées faisant du modèle arbre un objet d'étude d'intérêt majeur pour l'épigénétique, il était indispensable d'acquérir des connaissances de la méthylation de l'ADN chez le peuplier. En effet, bien que certaines caractéristiques de cette marque, déjà évoquées dans ce manuscrit, soient bien décrites et conservées au sein des Eucaryotes (Jeltsch, 2010), d'autres sont espèce-spécifiques, dépendant entre autres des voies et acteurs moléculaires établissant les marques épigénétiques ou encore de l'organisation du génome (distribution éparse des séquences répétées et des gènes ou au contraire génome compact à l'instar d'*Arabidopsis* ; Roudier *et al.*, 2011; Zhang *et al.*, 2012a). Aussi, les travaux présentés dans ce manuscrit ont permis d'acquérir des données sur la méthylation de l'ADN du peuplier, parallèlement à des études publiées récemment par d'autres équipes (Feng *et al.*, 2010a; Raj *et al.*, 2011; Vining *et al.*, 2012). De plus, ils ont permis de mieux appréhender le rôle de la méthylation de l'ADN dans la plasticité phénotypique chez le peuplier et plus généralement chez les plantes. Des travaux de cette thèse se dégagent plusieurs points.

Tout d'abord, la méthylation des sites CG dans le corps des gènes et plus précisément dans les exons, la forte méthylation des séquences répétées dans les trois contextes (CG, CHG et CHH), l'influence de la méthylation de l'ADN dans la tissu-spécificité de l'expression des gènes sont des caractéristiques communes au peuplier et aux autres plantes (Chapitre 1 ; Zhang *et al.*, 2006; Teyssier *et al.*, 2008; Feng *et al.*, 2010; Vining *et al.*, 2012). En outre, les travaux de cette thèse ont complété les connaissances du rôle de la méthylation de l'ADN dans l'expression des gènes, notamment confirmant une structure chromatinienne plus compacte dans les gènes plus longs, cohérent avec son rôle suggéré d'élimination des sites cryptiques d'initiation (Chapitre 1 ; Zilberman *et al.*, 2007). L'implication des introns dans l'expression tissu-spécifique des gènes, suggérée dans le chapitre 1, n'a jamais été montrée à ce jour. De manière générale, cette thèse a complété les connaissances existantes sur le rôle de la méthylation de l'ADN dans l'organisation et

l'expression du génome chez les plantes, particulièrement dans un état spécifique de compaction de la chromatine et ceci dans le méristème apical caulinaire, ce qui n'avait jamais été publié avant notre article (Lafon-Placette *et al.*, 2013 ; Chapitre 1). La méthylation de l'ADN des sites non-CG chez le peuplier présente certaines particularités (Chapitre 1 ; Feng *et al.*, 2010). En plus d'une méthylation en CHG supérieure à celle chez les autres plantes, les sites CHH sont affectés différemment selon la nature de la séquence (Chapitre 1). Il est possible que la forte méthylation non-CG des gènes redondants soit une conséquence à la fois de la duplication récente du génome du peuplier et de sa lente évolution (Chapitre 1 ; Tuskan *et al.*, 2006; Abrouk *et al.*, 2010). Le silencing épigénétique de ces gènes chez l'arbre modèle serait une alternative à leur élimination par mutations/délétions observée pour les organismes à temps de génération courts (*Arabidopsis*) et favoriserait leur néo-fonctionnalisation (Rodin *et al.*, 2005; Abrouk *et al.*, 2010).

Par ailleurs, un consensus concernant l'influence de l'environnement sur la méthylation de l'ADN est loin d'être acquis chez les plantes, posant ainsi de nombreuses questions sur ce phénomène chez le peuplier. Une première observation générale est que la méthylation de l'ADN chez le peuplier présente une forte variation selon les génotypes, leur provenance géographique et la disponibilité en eau du sol, ce dans plusieurs organes et à différentes étapes du développement (Chapitres 2 et 3 ; Raj *et al.*, 2011). De plus, la corrélation trouvée entre caractères liés à la production de biomasse et la méthylation de l'ADN dans le méristème apical caulinaire suggère un rôle de cette marque épigénétique dans la morphogenèse de la tige feuillée (Chapitre 2). D'après les résultats du chapitre 2, les variations de la méthylation de l'ADN du méristème apical en réponse à la sécheresse semblent se transmettre aux nouvelles feuilles et également aux générations cellulaires du méristème longtemps après le stress. Ainsi, ces données suggèrent fortement un rôle de la méthylation de l'ADN dans la réponse à la sécheresse, concordant avec des travaux existants chez des plantes annuelles (Chapitres 2 et 3 ; Boyko & Kovalchuk, 2008; Karan *et al.*, 2012).

De plus, cette thèse ajoute des précisions sur la nature des loci affectés par des remaniements de la méthylation de l'ADN lors de variations de la disponibilité eau, ainsi que l'impact des variations de la méthylation de l'ADN sur l'expression des gènes à l'échelle du génome entier. Ces variations épigénétiques, contrairement à ce que l'on pouvait attendre d'après certains auteurs (Pecinka & Mittelsten Scheid, 2012), affectent préférentiellement les gènes en leur corps dans le méristème apical caulinaire (Chapitre 3). Des travaux récents sont en accord avec cette vision (Karan *et al.*, 2012). Peu de données sur l'effet d'une sécheresse à l'échelle transcriptomique et épigénomique ont été rapportées à ce jour chez

les plantes (Boyko *et al.*, 2010). L'ampleur des variations de méthylation de l'ADN observée en réponse à la sécheresse (Chapitre 3) n'est pas en accord avec les variations plus discrètes observées chez les annuelles pour des stress abiotiques (Boyko *et al.*, 2010) mais cette différence peut provenir de l'échelle d'étude, effectuée au niveau des feuilles ou des plantes entières pour les travaux menés chez les annuelles. Les variations de méthylation de l'ADN en réponse à la sécheresse ciblées préférentiellement dans les gènes (Chapitre 3) suggèrent un rôle important dans le contrôle de la transcription en réponse aux variations de l'environnement. Bien que les variations de méthylation de l'ADN, en réponse aux variations de disponibilité en eau, n'aient pu être corrélées de manière quantitative aux variations d'expression des gènes de manière générale, la fonction des gènes affectés par ce double changement est intéressante. En effet, parmi ceux-ci, un nombre important de gènes impliqués dans la morphogenèse et la réponse aux phytohormones a été trouvé (Chapitre 3). Ce résultat suggère que le méristème apical caulinaire est, à l'échelle de la plante, le lieu d'intégration de signaux hormonaux liés à la perception des variations de l'environnement, aboutissant à terme à la modification du programme développemental et à la plasticité phénotypique de l'arbre. Ce processus serait en partie sous contrôle de la méthylation de l'ADN. L'état de chromatine non condensée dans lequel se trouve la majorité des gènes du méristème apical caulinaire du peuplier suggère une forte propension à la régulation pour ces gènes (Gilchrist *et al.*, 2010), fait cohérent avec un rôle intégrateur des signaux environnementaux aux niveaux épigénomique et transcriptomique (Chapitres 1 et 3).

2. Limites de l'étude

Certaines limites de ce travail doivent être prises en compte afin de pouvoir compléter ces travaux et conclure de manière plus précise sur le rôle de la méthylation de l'ADN dans la plasticité phénotypique chez le peuplier. Certaines d'entre elles sont évoquées ici.

La limite expérimentale principale de ce travail réside dans la disparité des environnements et des fonds génétiques étudiés. En effet, plusieurs dispositifs expérimentaux pour l'étude de l'effet de la sécheresse ont été utilisés que ce soit pour les caractéristiques du déficit hydrique, les fonds génétiques étudiés ou les conditions de culture (serre/pépinière). Même si une synthèse de ces résultats permet de dégager une réponse épigénétique spécifique du déficit hydrique et ce quelques soient les conditions étudiées, cette disparité limite les liens et les comparaisons faisables entre les données obtenues. De la même façon, alors que *Populus trichocarpa* était choisi dans le chapitre 1 (en raison de sa proximité avec le génome de référence) pour la caractérisation du méthylome, des hybrides

Populus deltoides x *P. nigra* étaient étudiés dans les chapitres 2 et 3 (en raison des travaux déjà publiés sur ces génotypes). Non seulement la méthylation de l'ADN présente une variabilité génétique importante (Gourcilleau *et al.*, 2010; Schmitz *et al.*, 2011; Becker *et al.*, 2011) mais l'hybridation interspécifique est connue pour entraîner un large spectre de modifications épigénétiques (He *et al.*, 2010; Hegarty *et al.*, 2011; Shivaprasad *et al.*, 2012). Celles-ci pourraient être impliquées dans la vigueur hybride (Goff, 2011; Smith & Weigel, 2012) et liées aux interactions entre les deux génomes parentaux, à de l'épistasie ou de l'expression monoallélique. Ceci est le cas dans le chapitre 3, où les données microarrays pour la méthylation et l'expression correspondaient à une seule séquence de référence, *P. trichocarpa*, ne correspondant à aucun des deux génomes parentaux de l'hybride étudié (*P. deltoides* x *P. nigra*) et ne permettant donc pas de les différencier.

Le choix d'étudier le méristème apical caulinaire, responsable de la morphogenèse de la tige feuillée, pour comprendre le rôle de l'épigénétique dans l'inflexion de trajectoires développementales est tout à fait cohérent. Cependant, les études du fonctionnement du méristème montrent qu'il est organisé en régions cellulaires distinctes aux profils transcriptomiques et épigénomiques propres (Traas & Vernoux, 2002; Perales & Reddy, 2012) et difficilement discernables à l'œil nu. L'isolement du méristème apical lors de cette thèse n'a pas permis l'isolement d'une zone précise. Les échantillons étudiés étaient donc probablement un ensemble de cellules hétérogènes et l'augmentation de la précision d'échantillonnage semble désormais indispensable à des études de transcriptomiques et épigénomiques plus fines.

Les variations épigénétiques chez les plantes sont à la fois très sensibles aux variations de l'environnement mais également labiles pour un certain nombre d'entre elles (Pecinka & Mittelsten Scheid, 2012). Aussi, l'observation, comme ce qui a été fait dans ces travaux, à un temps *t* de la réponse de l'arbre, est insuffisante pour savoir quelles variations de la méthylation de l'ADN dans le méristème apical impliqueront réellement des changements d'expression hérités par les organes formés permettant une plasticité phénotypique en réponse à la sécheresse.

Enfin, la rapidité de la progression des méthodes haut-débit doit être prise en compte. Au démarrage de ce travail, nous n'avons pas envisagé d'utiliser de méthodes directes de séquençage bisulfite ou de RNA-seq qui semblaient encore trop coûteuses et difficiles d'exploitation au niveau bioinformatique au vu de l'état de l'assemblage et de l'annotation du génome du peuplier en 2008-2010. Il serait maintenant pertinent de réévaluer ces possibilités dont la profondeur d'analyse est bien plus grande et qui évalueraient le polymorphisme génétique de façon concomitante.

3. Perspectives

Cette dernière partie présente de manière non exhaustive les perspectives majeures apportées par les travaux effectués lors de cette thèse.

Tout d'abord, l'étude des régions de la chromatine pauvres en nucléosomes dans le chapitre 1 pose la question de la variation de leur distribution dans le génome, que ce soit en réponse à la sécheresse ou selon les génotypes étudiés. En effet, ces régions, directement liées à la régulation de la transcription, sont très variables entre tissus et sont sensibles aux variations de l'environnement (Zhang *et al.*, 2012a,b). Elles jouent donc un rôle probable dans la plasticité phénotypique. De plus, la variation des caractéristiques de ces sites entre génotypes n'a jamais été étudiée à notre connaissance alors qu'elles pourraient être impliquées dans les variations génotypiques de traits phénotypiques. Il serait ainsi intéressant d'étudier ces variations chez le peuplier en utilisant notre approche DNase I (Lafon-Placette *et al.*, in press), aussi bien en réponse à la sécheresse qu'entre génotypes, permettant d'identifier des gènes dont l'expression est contrôlée par la dynamique des nucléosomes impliqués dans la morphogenèse de la tige feuillée chez le peuplier.

De plus, compte tenu de l'implication au sein du méristème apical caulinaire des voies de phytohormones dans la réponse du peuplier à la sécheresse (Chapitre 3), il serait pertinent d'envisager le dosage de phytohormones (brassinostéroïdes, glucosinolates...) dans l'apex et leur localisation précise comme il a été déjà décrit par exemple en analyse d'imagerie en spectrométrie de masse MALDI pour certains composés comme les glucosinolates (Shroff *et al.*, 2008; Kaspar *et al.*, 2011).

De manière plus générale, travailler sur un méristème apical caulinaire mieux caractérisé anatomiquement voire travailler sur une zone précise de ce méristème serait approprié. Par exemple, la zone centrale, comprenant des cellules indifférenciées responsables du maintien du méristème, ou la zone périphérique, à l'origine de la formation des organes latéraux (Brooks *et al.*, 2009), seraient des choix pertinents dans l'hypothèse de travail de l'axe épigénétique du LBLGC. Le développement des techniques de microdissection laser (Brooks *et al.*, 2009; Takahashi *et al.*, 2010), serait un moyen fiable et précis pour y arriver. Dans la même idée, la caractérisation systématique de l'activité mitotique, liée à la méthylation de l'ADN (Yang *et al.*, 2010), permettrait de prendre en compte ou de s'affranchir de l'effet du cycle cellulaire sur les variations épigénétiques observées. Etendre cette approche aux méristèmes des bourgeons latéraux et axillaires ou produits à partir de recépages successifs permettrait d'avoir une idée plus globale du rôle de la méthylation de l'ADN dans les phénomènes régissant l'architecture de l'arbre comme la dominance apicale.

Concevoir un dispositif expérimental unique permettrait de répondre aux questions évoquées lors de cette thèse, c'est-à-dire l'impact d'une sécheresse sur la méthylation de l'ADN dans le méristème apical caulinaire, la transmission de ces variations aux feuilles néoformées, le maintien dans le temps, même plusieurs mois ou années après le stress, des modifications de la méthylation de l'ADN dans le méristème apical caulinaire, la variation génotypique de ces phénomènes et leurs relations avec la production de biomasse. Par exemple, la mise en place d'un dispositif en pépinière avec une parcelle bien irriguée et l'autre placée en condition de déficit hydrique, avec plusieurs génotypes et le suivi de la méthylation de l'ADN dans les apex et les feuilles sur plusieurs années permettrait d'approfondir cette question (à condition d'avoir suffisamment d'individus pour pouvoir prélever chaque année le bourgeon apical d'arbres n'ayant pas subi de prélèvement les années précédentes). Les résultats du projet ANR Sylvabiom (2008-2012) sur la méthylation de l'ADN (axe 1.5 S. Maury) en fonction de la variabilité pédoclimatiques de 3 sites et sur deux modes de culture (TCR et TTCR) seront prochainement disponibles et permettront un premier bilan de ce type de dispositif. Ces études devront également se croiser avec les études de génétique quantitative (QTL biomasse par exemple ; Monclus *et al.*, 2012). Une approche complémentaire serait de créer des épimutants chez le peuplier en invalidant par une stratégie RNAi des gènes tels que *MET1* ou *DDM1* qui sont impliqués dans la méthylation de l'ADN chez les plantes. De tels travaux ont été initiés dans l'équipe de Steve Strauss (Université de l'Oregon, USA) avec la production de peuplier RNAi *DDM1* pour lequel une première publication (Zhu *et al.*, en révision) incluant des analyses de méthylation réalisées au LBLGC (S. Maury) est en cours. On pourrait alors relier le polymorphisme de méthylation à un caractère quantitatif comme cela a été fait chez *Arabidopsis* (Johannes *et al.*, 2009). Il serait alors possible d'envisager des études pour positionner des eQTL (epigenetic QTL) de la production de biomasse par exemple aussi bien dans ces lignées RNAi que dans la population en ségrégation de plus de 2100 individus [croisement *P. deltoides* (73038-62) x *P. trichocarpa* (101-74)] créée par l'INRA d'Orléans. En effet, Hauben *et al.* (2009) ont montré que même dans des lignées isogéniques une composante épigénétique pouvait encore être exploitée afin d'optimiser la sélection de traits quantitatifs. Ces études pourraient compléter les approches de génétique quantitative sur cette plante.

Prendre en compte les effets liés à l'hybridation interspécifique est un point indispensable. S'il s'agit de se concentrer sur le contrôle, par la méthylation de l'ADN, de la transcription de gènes impliqués dans la plasticité phénotypique en réponse à la sécheresse, il serait judicieux de travailler sur des génotypes ou populations non hybrides comme par exemple *P. nigra*. La variabilité épigénétique au niveau de la population est encore peu décrite et serait donc un dispositif original. Si en revanche, il s'agit d'étudier comment l'effet

hybride peut influencer la plasticité phénotypique, divers hybrides devront être analysés. Compte tenu du rôle suggéré de la méthylation de l'ADN dans les interactions entre les génomes parentaux d'un individu hybride, l'épistasie ou l'expression monoallélique de gènes, impliqués dans la vigueur hybride (Hegarty *et al.*, 2011; Goff, 2011; Shivaprasad *et al.*, 2012), il serait intéressant de savoir comment ces phénomènes sont impactés lors d'une sécheresse. Il est notamment possible que les variations épigénétiques, transcriptomiques et phénotypiques observées entre génotypes et en réponse à une sécheresse proviennent d'une « gestion » des génomes parentaux différente, comme le suggéraient les résultats obtenus dans Gourcilleau *et al.* (2010). Pour apporter des questions à cette problématique, différencier les deux génomes parentaux au sein d'un hybride donné à la fois pour leur épigénome et leur transcriptome semble indispensable, envisageable par des approches haut débit (séquençage nouvelle génération ou SNP microarrays) par exemple.

Enfin, même si les travaux de cette thèse étaient concentrés sur le rôle de l'épigénétique dans la plasticité phénotypique du peuplier en réponse à la sécheresse, la question de son implication dans l'adaptation des arbres aux changements climatiques, évoquée dans l'introduction de ce manuscrit, n'en reste pas moins sous-jacente. La transmission par méiose de variations épigénétiques pourrait être étudiée chez le peuplier qui a une descendance relativement précocement comparé aux autres espèces ligneuses. Cela pourrait être réalisé sur des arbres par classes d'âges et origine géographique (Raj *et al.*, 2011), différents types de bourgeons, en fonction de différentes conditions environnementales (déficit hydrique), associé à un génotypage et une mesure de la fitness des descendants sur un dispositif pépinière. De plus, le peuplier ayant une forte capacité à la multiplication végétative, il serait intéressant d'étudier le même dispositif expérimental pour suivre le devenir de variations épigénétiques dans des boutures (clones) issues de ces arbres.

Références

- Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003.** Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant cell* **15**: 63–78.
- Abrouk M, Murat F, Pont C, Messing J, Jackson S, Faraut T, Tannier E, Plomion C, Cooke R, Feuillet C. 2010.** Palaeogenomics of plants: synteny-based modelling of extinct ancestors. *Trends in Plant Science* **15**: 479–487.
- Agius F, Kapoor A, Zhu J-K. 2006.** Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 11796–11801.
- Akimoto K, Katakami H, Kim H-J, Ogawa E, Sano CM, Wada Y, Sano H. 2007.** Epigenetic Inheritance in Rice Plants. *Annals of Botany* **100**: 205–217.
- Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears DD, Hogg EH (Ted), et al. 2010.** A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* **259**: 660–684.
- Allis CD, Jenuwein T, Reinberg D, Caparros M-L. 2007.** *Epigenetics*. New York: Cold Spring Harbor Laboratory Press.
- Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. 2009.** Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Research* **19**: 1732–1741.
- Baek D, Jiang J, Chung J-S, Wang B, Chen J, Xin Z, Shi H. 2011.** Regulated AtHKT1 gene expression by a distal enhancer element and DNA methylation in the promoter plays an important role in salt tolerance. *Plant & cell physiology* **52**: 149–161.
- Ball MP, Li JB, Gao Y, Lee J-H, LeProust EM, Park I-H, Xie B, Daley GQ, Church GM. 2009.** Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature biotechnology* **27**: 361–368.
- Bartee L, Malagnac F, Bender J. 2001.** Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes & Development* **15**: 1753–1758.
- Baskar V, Gururani MA, Yu JW, Park SW. 2012.** Engineering Glucosinolates in Plants: Current Knowledge and Potential Uses. *Applied biochemistry and biotechnology*.

- Beck S, Rakyan VK. 2008.** The methylome: approaches for global DNA methylation profiling. *Trends in Genetics* **24**: 231–237.
- Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. 2011.** Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature*.
- Berta M, Giovannelli A, Sebastiani F, Camussi A, Racchi ML. 2010.** Transcriptome changes in the cambial region of poplar (*Populus alba* L.) in response to water deficit. *Plant Biology* **12**: 341–354.
- Berthelot A, Bastien C, Villar M, Pinon J, Heois B, Bourlon V, Menard M. 2005.** Le GIS Peuplier, 4 ans après sa création. *Fiches Informations-Forêt AFOCEL* **2**.
- Bestor TH. 1998.** Cytosine methylation and the unequal developmental potentials of the oocyte and sperm genomes. *American journal of human genetics* **62**: 1269–1273.
- Bies-Ethève N, Gaubier-Comella P, Debures A, Lasserre E, Jobet E, Raynal M, Cooke R, Delseny M. 2008.** Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. *Plant molecular biology* **67**: 107–124.
- Bigler C, Bräker OU, Bugmann H, Dobbertin M, Rigling A.** Drought as an inciting mortality factor in scots pine stands of the Valais, Switzerland. *Ecosystems* **9**: 330–343.
- Bilichak A, Ilnytsky Y, Hollunder J, Kovalchuk I. 2012.** The progeny of *Arabidopsis thaliana* plants exposed to salt exhibit changes in DNA methylation, histone modifications and gene expression. *PloS one* **7**: e30515.
- Blödner C, Goebel C, Feussner I, Gatz C, Polle A. 2007.** Warm and cold parental reproductive environments affect seed properties, fitness, and cold responsiveness in *Arabidopsis thaliana* progenies. *Plant, cell & environment* **30**: 165–175.
- Bogeat-Triboulot M-B, Brosché M, Renaut J, Jouve L, Le Thiec D, Fayyaz P, Vinocur B, Witters E, Laukens K, Teichmann T, et al. 2007.** Gradual soil water depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in *Populus euphratica*, a poplar growing in arid regions. *Plant physiology* **143**: 876–892.
- Bonhomme L, Monclus R, Vincent D, Carpin S, Claverol S, Lomenech A-M, Labas V, Plomion C, Brignolas F, Morabito D. 2009a.** Genetic variation and drought response in two *Populus x euramericana* genotypes through 2-DE proteomic analysis of leaves from field and

glasshouse cultivated plants. *Phytochemistry* **70**: 988–1002.

Bonhomme L, Monclus R, Vincent D, Carpin S, Lomenech A-M, Plomion C, Brignolas F, Morabito D. 2009b. Leaf proteome analysis of eight *Populus xeuramericana* genotypes: genetic variation in drought response and in water-use efficiency involves photosynthesis-related proteins. *Proteomics* **9**: 4121–4142.

Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, Ilnytskyy Y, Hollander J, Meins F, Kovalchuk I. 2010. Transgenerational Adaptation of *Arabidopsis* to Stress Requires DNA Methylation and the Function of Dicer-Like Proteins (S-H Shiu, Ed.). *PLoS ONE* **5**: e9514.

Boyko A, Kovalchuk I. 2008. Epigenetic control of plant stress response. *Environmental and Molecular Mutagenesis* **49**: 61–72.

Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132**: 311–322.

Brignolas F, Paillassa E, Fichot R. 2012. Vers de nouveaux critères de sélection pour la populiculture dans un contexte de climat changeant. *Forêt-entreprise* **204**: 59–64.

Brink RA. 1956. A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. *Genetics* **41**: 872–889.

Brock RD, Davidson JL. 1994. 5-Azacytidine and gamma rays partially substitute for cold treatment in vernalizing winter wheat. *Environmental and experimental botany* **34**: 195–199.

Brooks L, Strable J, Zhang X, Ohtsu K, Zhou R, Sarkar A, Hargreaves S, Elshire RJ, Eudy D, Pawlowska T, et al. 2009. Microdissection of Shoot Meristem Functional Domains. *PLoS Genet* **5**: e1000476.

Buchanan B, Gruissem W, Jones R. 2000. *Biochemistry & Molecular Biology of Plants*. Rockville: American Society of Plant Biologists.

Bucher E, Reinders J, Mirouze M. 2012. Epigenetic control of transposon transcription and mobility in *Arabidopsis*. *Current opinion in plant biology*.

Burn JE, Bagnall DJ, Metzger JD, Dennis ES, Peacock WJ. 1993. DNA methylation, vernalization, and the initiation of flowering. *Proceedings of the National Academy of Sciences of the United States of America* **90**: 287–291.

Busov VB, Brunner AM, Meilan R, Filichkin S, Ganio L, Gandhi S, Strauss SH. 2005.

Genetic transformation: a powerful tool for dissection of adaptive traits in trees. *The New phytologist* **167**: 9–18.

Buzas DM, Tamada Y, Kurata T. 2012. FLC: a hidden polycomb response element shows up in silence. *Plant & cell physiology* **53**: 785–793.

Calarco JP, Borges F, Donoghue MTA, Van Ex F, Jullien PE, Lopes T, Gardner R, Berger F, Feijó JA, Becker JD, et al. 2012. Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA. *Cell* **151**: 194–205.

Callis J, Vierstra RD. 2000. Protein degradation in signaling. *Current opinion in plant biology* **3**: 381–386.

Causevic A, Delaunay A, Ounnar S, Righezza M, Delmotte F, Brignolas F, Hagège D, Maury S. 2005. DNA methylating and demethylating treatments modify phenotype and cell wall differentiation state in sugarbeet cell lines. *Plant Physiology and Biochemistry: PPB / Société Française De Physiologie Végétale* **43**: 681–691.

Causevic A, Gentil M-V, Delaunay A, El-Soud WA, Garcia Z, Pannetier C, Brignolas F, Hagège D, Maury S. 2006. Relationship between DNA methylation and histone acetylation levels, cell redox and cell differentiation states in sugarbeet lines. *Planta* **224**: 812–827.

Chamaillard S, Fichot R, Vincent-Barbaroux C, Bastien C, Depierreux C, Dreyer E, Villar M, Brignolas F. 2011. Variations in bulk leaf carbon isotope discrimination, growth and related leaf traits among three *Populus nigra* L. populations. *Tree physiology* **31**: 1076–1087.

Chan Z. 2012. Expression profiling of ABA pathway transcripts indicates crosstalk between abiotic and biotic stress responses in *Arabidopsis*. *Genomics* **100**: 110–115.

Chandler VL. 2007. Paramutation: from maize to mice. *Cell* **128**: 641–645.

Chandler VL. 2010. Paramutation's Properties and Puzzles. *Science* **330**: 628–629.

Chandler VL, Stam M. 2004. Chromatin conversations: mechanisms and implications of paramutation. *Nature reviews. Genetics* **5**: 532–544.

Chardenon J. 1968. Un cas très probable de mutations gemmaires chez le peuplier. *Revue Forestière Française* **1**: 33–37.

Cho SK, Ryu MY, Song C, Kwak JM, Kim WT. 2008. *Arabidopsis* PUB22 and PUB23 are homologous U-Box E3 ubiquitin ligases that play combinatory roles in response to drought stress. *The Plant cell* **20**: 1899–1914.

Chodavarapu RK, Feng S, Bernatavichute YV, Chen P-Y, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, et al. 2010. Relationship between nucleosome positioning and DNA methylation. *Nature* **466**: 388–392.

Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. 2002. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* **110**: 33–42.

Chouard P. 1960. Vernalization and its Relations to Dormancy. *Annual Review of Plant Physiology* **11**: 191–238.

Choudhary SP, Yu J-Q, Yamaguchi-Shinozaki K, Shinozaki K, Tran L-SP. 2012. Benefits of brassinosteroid crosstalk. *Trends in plant science*.

Clapier CR, Cairns BR. 2009. The biology of chromatin remodeling complexes. *Annual review of biochemistry* **78**: 273–304.

Cohen D, Bogeat-Triboulot M-B, Tisserant E, Balzergue S, Martin-Magniette M-L, Lelandais G, Ningre N, Renou J-P, Tamby J-P, Le Thiec D, et al. 2010. Comparative transcriptomics of drought responses in Populus: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. *BMC genomics* **11**: 630.

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**: 215–219.

Coll-Garcia D, Mazuch J, Altmann T, Müssig C. 2004. EXORDIUM regulates brassinosteroid-responsive genes. *FEBS letters* **563**: 82–86.

Cosgrove MS, Boeke JD, Wolberger C. 2004. Regulated nucleosome mobility and the histone code. *Nature structural & molecular biology* **11**: 1037–1043.

Crawford GE, Davis S, Scacheri PC, Renaud G, Halawi MJ, Erdos MR, Green R, Meltzer PS, Wolfsberg TG, Collins FS. 2006. DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. *Nature Methods* **3**: 503–509.

Crawford GE, Holt IE, Mullikin JC, Tai D, Blakesley R, Bouffard G, Young A, Masiello C, Green ED, Wolfsberg TG, et al. 2004. Identifying gene regulatory elements by genome-wide recovery of DNase hypersensitive sites. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 992–997.

- Cubas P, Vincent C, Coen E. 1999.** An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**: 157–161.
- Cuylen S, Haering CH. 2011.** Deciphering condensin action during chromosome segregation. *Trends in cell biology* **21**: 552–559.
- Damesin C. 2003.** Respiration and photosynthesis characteristics of current-year stems of *Fagus sylvatica*: from the seasonal pattern to an annual balance. *New Phytologist* **158**: 465–475.
- Danchin É, Charmantier A, Champagne FA, Mesoudi A, Pujol B, Blanchet S. 2011.** Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews Genetics* **12**: 475–486.
- Darwin C. 1859.** *L'Origine des Espèces*. London: J. Murray.
- Day T, Bonduriansky R. 2011.** A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. *The American Naturalist* **178**: E18–36.
- Dennis ES, Bilodeau P, Burn J, Finnegan EJ, Genger R, Helliwell C, Kang BJ, Sheldon CC, Peacock WJ. 1998.** Methylation controls the low temperature induction of flowering in *Arabidopsis*. *Symposia of the Society for Experimental Biology* **51**: 97–103.
- Dennis ES, Brettell RI. 1990.** DNA methylation of maize transposable elements is correlated with activity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **326**: 217–229.
- Dickmann DI, Kuzovkina J. 2008.** Poplars and willows of the world, with emphasis on sylviculturally important species (Chapter 2). Poplars and willows in the world: meeting the needs of society and the environment. Rome: FAO/IPC (Food and Agricultural Organization of the United Nations / International Poplar Commission), .
- Dillen SY, El Kasmioui O, Marron N, Calfapietra C, Ceulemans R. 2010.** Poplar. In: Halford NG, Karp A, eds. RSC Energy and Environment Series. Cambridge: Royal Society of Chemistry, 275–300.
- Dillen SY, Marron N, Sabatti M, Ceulemans R, Bastien C. 2009.** Relationships among productivity determinants in two hybrid poplar families grown during three years at two contrasting sites. *Tree physiology* **29**: 975–987.
- Doerfler W. 1983.** DNA methylation and gene activity. *Annual review of biochemistry* **52**: 93–124.

- Dooner HK, Robbins TP, Jorgensen RA. 1991.** Genetic and developmental control of anthocyanin biosynthesis. *Annual review of genetics* **25**: 173–199.
- Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012.** Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences of the United States of America* **109**: E2183–2191.
- Doyle J, Doyle J. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.
- Du L, Chen Z. 2000.** Identification of genes encoding receptor-like protein kinases as possible targets of pathogen- and salicylic acid-induced WRKY DNA-binding proteins in *Arabidopsis*. *The Plant journal: for cell and molecular biology* **24**: 837–847.
- Eckenwalder JE. 1996.** Systematics and evolution of *Populus*. Biology of *Populus* and its implications for management and conservation. Ottawa: NRC Research Press, .
- Ergle DR, Katterman FR. 1961.** Deoxyribonucleic acid of cotton. *Plant physiology* **36**: 811–815.
- Esteller M. 2011.** Non-coding RNAs in human disease. *Nature reviews. Genetics* **12**: 861–874.
- Exner V, Hennig L. 2008.** Chromatin rearrangements in development. *Current Opinion in Plant Biology* **11**: 64–69.
- Felsenfeld G, Groudine M. 2003.** Controlling the double helix. *Nature* **421**: 448–453.
- Feng S, Cokus SJ, Zhang X, Chen P-Y, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, et al. 2010a.** Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences* **107**: 8689–8694.
- Feng S, Jacobsen SE, Reik W. 2010b.** Epigenetic Reprogramming in Plant and Animal Development. *Science* **330**: 622–627.
- Fichot R, Barigah TS, Chamaillard S, Le Thiec D, Laurans F, Cochard H, Brignolas F. 2010.** Common trade-offs between xylem resistance to cavitation and other physiological traits do not hold among unrelated *Populus deltoides* × *Populus nigra* hybrids. *Plant, Cell & Environment*: no–no.
- Fichot R, Laurans F, Monclus R, Moreau A, Pilate G, Brignolas F. 2009.** Xylem anatomy correlates with gas exchange, water-use efficiency and growth performance under

contrasting water regimes: evidence from *Populus deltoides* x *Populus nigra* hybrids. *Tree Physiology* **29**: 1537–1549.

Filion GJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143**: 212–224.

Filippou P, Antoniou C, Fotopoulos V. 2011. Effect of drought and rewatering on the cellular status and antioxidant response of *Medicago truncatula* plants. *Plant Signaling & Behavior* **6**: 270–277.

Finnegan EJ, Genger RK, Kovac K, Peacock WJ, Dennis ES. 1998. DNA methylation and the promotion of flowering by vernalization. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 5824–5829.

Finnegan EJ, Peacock WJ, Dennis ES. 1996. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 8449–8454.

Gallé A, Haldimann P, Feller U. 2007. Photosynthetic performance and water relations in young pubescent oak (*Quercus pubescens*) trees during drought stress and recovery. *The New phytologist* **174**: 799–810.

Gehring M, Henikoff S. 2007. DNA methylation dynamics in plant genomes. *Biochimica et biophysica acta* **1769**: 276–286.

Gehring M, Huh JH, Hsieh T-F, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL. 2006. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**: 495–506.

Gehring M, Missirlian V, Henikoff S. 2011. Genomic analysis of parent-of-origin allelic expression in *Arabidopsis thaliana* seeds. *PloS one* **6**: e23687.

Gentil M-V, Maury S. 2007. Characterization of epigenetic biomarkers using new molecular approaches. In: Varshney R, Tuberosa R, eds. *Genomics-Assisted Crop Improvement: Genomics Approaches and Platforms*. Springer, 351–370.

Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, Adelman K. 2010. Pausing of RNA Polymerase II Disrupts DNA-Specified Nucleosome Organization to Enable Precise Gene Regulation. *Cell* **143**: 540–551.

Goff SA. 2011. A unifying theory for general multigenic heterosis: energy efficiency, protein

metabolism, and implications for molecular breeding. *New Phytologist* **189**: 923–937.

Goll MG, Bestor TH. 2005. Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry* **74**: 481–514.

González RM, Ricardi MM, Iusem ND. 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.

Gourcilleau D, Bogeat-Triboulot M-B, Le Thiec D, Lafon-Placette C, Delaunay A, El-Soud WA, Brignolas F, Maury S. 2010. DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity. *Annals of Forest Science* **67**: 208.

Grewal SIS, Elgin SCR. 2007. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**: 399–406.

Grigg SP, Canales C, Hay A, Tsiantis M. 2005. SERRATE coordinates shoot meristem function and leaf axial patterning in Arabidopsis. *Nature* **437**: 1022–1026.

Gross SM, Hollick JB. 2007. Multiple trans-sensing interactions affect meiotically heritable epigenetic states at the maize pl1 locus. *Genetics* **176**: 829–839.

Ha S, Vankova R, Yamaguchi-Shinozaki K, Shinozaki K, Tran L-SP. 2012. Cytokinins: metabolism and function in plant adaptation to environmental stresses. *Trends in plant science* **17**: 172–179.

Hamanishi ET, Campbell MM. 2011. Genome-wide responses to drought in forest trees. *Forestry* **84**: 273–283.

Hamanishi ET, Thomas BR, Campbell MM. 2012. Drought induces alterations in the stomatal development program in *Populus*. *Journal of experimental botany* **63**: 4959–4971.

Hauben M, Haesendonckx B, Standaert E, Van Der Kelen K, Azmi A, Akpo H, Van Breusegem F, Guisez Y, Bots M, Lambert B, et al. 2009. Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield. *Proceedings of the National Academy of Sciences*.

He G, Zhu X, Elling AA, Chen L, Wang X, Guo L, Liang M, He H, Zhang H, Chen F, et al. 2010. Global Epigenetic and Transcriptional Trends among Two Rice Subspecies and Their Reciprocal Hybrids. *THE PLANT CELL ONLINE* **22**: 17–33.

- Hegarty MJ, Batstone T, Barker GL, Edwards KJ, Abbott RJ, Hiscock SJ. 2011.** Nonadditive changes to cytosine methylation as a consequence of hybridization and genome duplication in *Senecio* (Asteraceae). *Molecular ecology* **20**: 105–113.
- Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE, Lumey LH. 2008.** Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 17046–17049.
- Henderson IR, Deleris A, Wong W, Zhong X, Chin HG, Horwitz GA, Kelly KA, Pradhan S, Jacobsen SE. 2010.** The de novo cytosine methyltransferase DRM2 requires intact UBA domains and a catalytically mutated paralog DRM3 during RNA-directed DNA methylation in *Arabidopsis thaliana*. *PLoS Genetics* **6**: e1001182.
- Ho L, Crabtree GR. 2010.** Chromatin remodelling during development. *Nature* **463**: 474–484.
- Hollick JB. 2010.** Paramutation and development. *Annual Review of Cell and Developmental Biology* **26**: 557–579.
- Holliday R. 1990.** DNA methylation and epigenetic inheritance. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **326**: 329–338.
- Holliday R. 1994.** Epigenetics: an overview. *Developmental genetics* **15**: 453–457.
- Hsieh T-F, Shin J, Uzawa R, Silva P, Cohen S, Bauer MJ, Hashimoto M, Kirkbride RC, Harada JJ, Zilberman D, et al. 2011.** Regulation of imprinted gene expression in *Arabidopsis* endosperm. *Proceedings of the National Academy of Sciences* **108**: 1755–1762.
- Ibarra CA, Feng X, Schoft VK, Hsieh T-F, Uzawa R, Rodrigues JA, Zemach A, Chumak N, Machlicova A, Nishimura T, et al. 2012.** Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science (New York, N.Y.)* **337**: 1360–1364.
- Ikeda Y, Kinoshita T. 2008.** DNA demethylation: a lesson from the garden. *Chromosoma* **118**: 37–41.
- Intergovernmental Panel on Climate Change. 2007.** *IPCC Fourth Assessment Report: Climate Change 2007 (AR4)*. Geneva, Switzerland: IPCC.
- Jackson JP, Johnson L, Jasencakova Z, Zhang X, PerezBurgos L, Singh PB, Cheng X, Schubert I, Jenuwein T, Jacobsen SE. 2004.** Dimethylation of histone H3 lysine 9 is a

critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana*. *Chromosoma* **112**: 308–315.

Jansson S, Douglas CJ. 2007. *Populus*: A Model System for Plant Biology. *Annual Review of Plant Biology* **58**: 435–458.

Jeltsch A. 2010. Phylogeny of Methylomes. *Science* **328**: 837–838.

Jiang M, Zhang Y, Fei J, Chang X, Fan W, Qian X, Zhang T, Lu D. 2010. Rapid quantification of DNA methylation by measuring relative peak heights in direct bisulfite-PCR sequencing traces. *Laboratory Investigation; a Journal of Technical Methods and Pathology* **90**: 282–290.

Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark ‘nucleosome-free regions’ of active promoters and other regulatory regions. *Nature Genetics* **41**: 941–945.

Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuissou J, Heredia F, Audigier P, et al. 2009. Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits (PM Visscher, Ed.). *PLoS Genetics* **5**: e1000530.

Johnson TB, Coghill RD. 1925. Researches on pyrimidines. C111. The discovery of 5-methylcytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus. *Journal of the American Chemical Society* **47**: 2838–2844.

Jones PA, Wolkowicz MJ, Rideout WM 3rd, Gonzales FA, Marziasz CM, Coetzee GA, Tapscott SJ. 1990. De novo methylation of the MyoD1 CpG island during the establishment of immortal cell lines. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 6117–6121.

Jullien PE, Mosquna A, Ingouff M, Sakata T, Ohad N, Berger F. 2008. Retinoblastoma and its binding partner MSI1 control imprinting in *Arabidopsis*. *PLoS biology* **6**: e194.

Jump AS, Marchant R, Peñuelas J. 2009. Environmental change and the option value of genetic diversity. *Trends in plant science* **14**: 51–58.

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ. 1996. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 12406–12411.

Kamakaka RT, Biggins S. 2005. Histone variants: deviants? *Genes & development* **19**:

295–310.

Karan R, DeLeon T, Biradar H, Subudhi PK. 2012. Salt stress induced variation in DNA methylation pattern and its influence on gene expression in contrasting rice genotypes. *PloS one* **7**: e40203.

Kaspar S, Peukert M, Svatos A, Matros A, Mock H-P. 2011. MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* **11**: 1840–1850.

Kiefer JC. 2007. Epigenetics in development. *Developmental dynamics: an official publication of the American Association of Anatomists* **236**: 1144–1156.

Kinoshita T. 2007. Reproductive barrier and genomic imprinting in the endosperm of flowering plants. *Genes & Genetic Systems* **82**: 177–186.

Klaas M, John MC, Crowell DN, Amasino RM. 1989. Rapid induction of genomic demethylation and T-DNA gene expression in plant cells by 5-azacytosine derivatives. *Plant Molecular Biology* **12**: 413–423.

Klein M, Reichelt M, Gershenzon J, Papenbrock J. 2006. The three desulfoglucosinolate sulfotransferase proteins in Arabidopsis have different substrate specificities and are differentially expressed. *The FEBS journal* **273**: 122–136.

Köhler C, Hennig L. 2010. Regulation of cell identity by plant Polycomb and trithorax group proteins. *Current opinion in genetics & development* **20**: 541–547.

Körner C, Sarris D, Christodoulakis D. 2005. Long-term increase in climatic dryness in the East-Mediterranean as evidenced for the island of Samos. *Regional Environmental Change* **5**: 27–36.

Kovarik A, Koukalova B, Bezdek M, Opatrny Z. Hypermethylation of tobacco heterochromatic loci in response to osmotic stress. *Theoretical and applied genetics* **95**: 301–306.

Labra M, Ghiani A, Citterio S, Sgorbati S, Sala F, Vannini C, Ruffini-Castiglione M, Bracale M. 2002. Analysis of Cytosine Methylation Pattern in Response to Water Deficit in Pea Root Tips. *Plant Biology* **4**: 694–699.

Lahmy S, Bies-Etheve N, Lagrange T. 2010. Plant-specific multisubunit RNA polymerase in gene silencing. *Epigenetics: official journal of the DNA Methylation Society* **5**: 4–8.

Lande R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic

plasticity and genetic assimilation. *Journal of evolutionary biology* **22**: 1435–1446.

Lang-Mladek C, Popova O, Kiok K, Berlinger M, Rakic B, Aufsatz W, Jonak C, Hauser M-T, Luschnig C. 2010. Transgenerational Inheritance and Resetting of Stress-Induced Loss of Epigenetic Gene Silencing in Arabidopsis. *Molecular Plant* **3**: 594–602.

Larsen F, Gundersen G, Prydz H. 1992. Choice of enzymes for mapping based on CpG islands in the human genome. *Genetic analysis, techniques and applications* **9**: 80–85.

Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews. Genetics* **11**: 204–220.

Leakey TI, Zielinski J, Siegfried RN, Siegel ER, Fan C-Y, Cooney CA. 2008. A simple algorithm for quantifying DNA methylation levels on multiple independent CpG sites in bisulfite genomic sequencing electropherograms. *Nucleic Acids Research* **36**: e64–e64.

Lee J-S, Smith E, Shilatifard A. 2010. The language of histone crosstalk. *Cell* **142**: 682–685.

Lejeune E, Allshire RC. 2011. Common ground: small RNA programming and chromatin modifications. *Current Opinion in Cell Biology* **23**: 258–265.

Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**: 915–926.

Li X, Wang X, He K, Ma Y, Su N, He H, Stolc V, Tongprasit W, Jin W, Jiang J, et al. 2008. High-Resolution Mapping of Epigenetic Modifications of the Rice Genome Uncovers Interplay between DNA Methylation, Histone Methylation, and Gene Expression. *THE PLANT CELL ONLINE* **20**: 259–276.

Liberloo M, Lukac M, Calfapietra C, Hoosbeek MR, Gielen B, Miglietta F, Scarascia-Mugnozza GE, Ceulemans R. 2009. Coppicing shifts CO₂ stimulation of poplar productivity to above-ground pools: a synthesis of leaf to stand level results from the POP/EUROFACE experiment. *The New phytologist* **182**: 331–346.

Lippman Z, Gendrel A-V, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, et al. 2004. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**: 471–476.

Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PCG. 2010. Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PloS one* **5**: e10326.

- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. 2008.** Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**: 523–536.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M, et al. 2009.** Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**: 315–322.
- Loustau D, Bosc A, Colin A, Ogée J, Davi H, François C, Dufrêne E, Déqué M, Cloppet E, Arrouays D, et al. 2005.** Modeling climate change effects on the potential production of French plains forests at the sub-regional level. *Tree physiology* **25**: 813–823.
- Lu Y, Rong T, Cao M. 2008.** Analysis of DNA methylation in different maize tissues. *Journal of genetics and genomics = Yi chuan xue bao* **35**: 41–48.
- Luger K. 2006.** Dynamic nucleosomes. *Chromosome Research* **14**: 5–16.
- Luo M, Taylor JM, Spriggs A, Zhang H, Wu X, Russell S, Singh M, Koltunow A. 2011.** A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. *PLoS genetics* **7**: e1002125.
- van Mantgem PJ, Stephenson NL, Byrne JC, Daniels LD, Franklin JF, Fulé PZ, Harmon ME, Larson AJ, Smith JM, Taylor AH, et al. 2009.** Widespread increase of tree mortality rates in the western United States. *Science (New York, N.Y.)* **323**: 521–524.
- Des Marais DL, Juenger TE. 2010.** Pleiotropy, plasticity, and the evolution of plant abiotic stress tolerance. *Annals of the New York Academy of Sciences* **1206**: 56–79.
- Marron N, Bastien C, Sabatti M, Taylor G, Ceulemans R. 2006.** Plasticity of growth and sylleptic branchiness in two poplar families grown at three sites across Europe. *Tree physiology* **26**: 935–946.
- Marron N, Ceulemans R. 2006.** Genetic variation of leaf traits related to productivity in a *Populus deltoides* *Populus nigra* family. *Canadian Journal of Forest Research* **36**: 390–400.
- Marron N, Delay D, Petit J-M, Dreyer E, Kahlem G, Delmotte FM, Brignolas F. 2002.** Physiological traits of two *Populus x euramericana* clones, Luisa Avanzo and Dorskamp, during a water stress and re-watering cycle. *Tree Physiology* **22**: 849–858.
- Marron N, Dreyer E, Boudouresque E, Delay D, Petit J-M, Delmotte FM, Brignolas F. 2003.** Impact of successive drought and re-watering cycles on growth and specific leaf area of two *Populus x canadensis* (Moench) clones, 'Dorskamp' and 'Luisa_Avanzo'. *Tree*

Physiology **23**: 1225–1235.

Marsolais F, Boyd J, Paredes Y, Schinas A-M, Garcia M, Elzein S, Varin L. 2007. Molecular and biochemical characterization of two brassinosteroid sulfotransferases from *Arabidopsis*, AtST4a (At2g14920) and AtST1 (At2g03760). *Planta* **225**: 1233–1244.

Martienssen R. 1996. Epigenetic phenomena: paramutation and gene silencing in plants. *Current biology: CB* **6**: 810–813.

Matzke M, Kanno T, Huettel B, Daxinger L, Matzke AJM. 2007. Targets of RNA-directed DNA methylation. *Current Opinion in Plant Biology* **10**: 512–519.

Maury S, Trap-Gentil M-V, Hébrard C, Weyens G, Delaunay A, Barnes S, Lefebvre M, Joseph C. 2012. Genic DNA methylation changes during in vitro organogenesis: organ specificity and conservation between parental lines of epialleles. *Physiologia plantarum* **146**: 321–335.

McClintock B. 1950. The Origin and Behavior of Mutable Loci in Maize. *Proceedings of the National Academy of Sciences* **36**: 344–355.

McClintock B. 1951. Chromosome organization and genic expression. *Cold Spring Harbor symposia on quantitative biology* **16**: 13–47.

McDowell N, Pockman WT, Allen CD, Breshears DD, Cobb N, Kolb T, Plaut J, Sperry J, West A, Williams DG, et al. 2008. Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? *The New phytologist* **178**: 719–739.

McKeown PC, Laouielle-Duprat S, Prins P, Wolff P, Schmid MW, Donoghue MTA, Fort A, Duszynska D, Comte A, Lao NT, et al. 2011. Identification of imprinted genes subject to parent-of-origin specific expression in *Arabidopsis thaliana* seeds. *BMC plant biology* **11**: 113.

Meehl GA, Tebaldi C. 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. *Science (New York, N.Y.)* **305**: 994–997.

Messeguer R, Ganai MW, Steffens JC, Tanksley SD. 1991. Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. *Plant molecular biology* **16**: 753–770.

Metcalfe JC, Rose KE, Rees M. 2003. Evolutionary demography of monocarpic perennials. *Trends in Ecology & Evolution* **18**: 471–480.

- Mitton JB, Grant MC. 1996.** Genetic Variation and the Natural History of Quaking Aspen. *BioScience* **46**: 25–31.
- Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T. 2001.** Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. *Nature* **411**: 212–214.
- Monclus R, Dreyer E, Villar M, Delmotte FM, Delay D, Petit J-M, Barbaroux C, Le Thiec D, Brechet C, Brignolas F. 2006.** Impact of drought on productivity and water use efficiency in 29 genotypes of *Populus deltoides* x *Populus nigra*. *New Phytologist* **169**: 765–777.
- Monclus R, Leplé J-C, Bastien C, Bert P-F, Villar M, Marron N, Brignolas F, Jorge V. 2012.** Integrating genome annotation and QTL position to identify candidate genes for productivity, architecture and water-use efficiency in *Populus* spp. *BMC plant biology* **12**: 173.
- Moore CJ, Fenech M, O’Callaghan NJ. 2011.** Telomere dynamics: the influence of folate and DNA methylation. *Annals of the New York Academy of Sciences* **1229**: 76–88.
- Morse AM, Peterson DG, Islam-Faridi MN, Smith KE, Magbanua Z, Garcia SA, Kubisiak TL, Amerson HV, Carlson JE, Nelson CD, et al. 2009.** Evolution of Genome Size and Complexity in *Pinus*. *PLoS ONE* **4**: e4332.
- Nakano Y, Steward N, Sekine M, Kusano T, Sano H. 2000.** A tobacco NtMET1 cDNA encoding a DNA methyltransferase: molecular characterization and abnormal phenotypes of transgenic tobacco plants. *Plant & cell physiology* **41**: 448–457.
- Nestler E. 2012.** Des interrupteurs cachés dans le cerveau. *Pour la Science*: 50–57.
- Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, et al. 2010.** Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**: 684–692.
- Oono Y, Seki M, Nanjo T, Narusaka M, Fujita M, Satoh R, Satou M, Sakurai T, Ishida J, Akiyama K, et al. 2003.** Monitoring expression profiles of Arabidopsis gene expression during rehydration process after dehydration using ca 7000 full-length cDNA microarray. *The Plant journal: for cell and molecular biology* **34**: 868–887.
- Ossowski S, Schneeberger K, Lucas-Lledo JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. 2009.** The Rate and Molecular Spectrum of Spontaneous Mutations in *Arabidopsis thaliana*. *Science* **327**: 92–94.

- Ou X, Zhang Y, Xu C, Lin X, Zang Q, Zhuang T, Jiang L, von Wettstein D, Liu B. 2012.** Transgenerational Inheritance of Modified DNA Methylation Patterns and Enhanced Tolerance Induced by Heavy Metal Stress in Rice (*Oryza sativa* L.). *PloS one* **7**: e41143.
- Pandey SP, Shahi P, Gase K, Baldwin IT. 2008.** Herbivory-induced changes in the small-RNA transcriptome and phytohormone signaling in *Nicotiana attenuata*. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 4559–4564.
- Papp B, Müller J. 2006.** Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes & Development* **20**: 2041–2054.
- Pavlopoulou A, Kossida S. 2007.** Plant cytosine-5 DNA methyltransferases: Structure, function, and molecular evolution. *Genomics* **90**: 530–541.
- Pecinka A, Mittelsten Scheid O. 2012.** Stress-induced chromatin changes: a critical view on their heritability. *Plant & cell physiology* **53**: 801–808.
- Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. 2007.** DNA demethylation in the Arabidopsis genome. *Proceedings of the National Academy of Sciences* **104**: 6752–6757.
- Perales M, Reddy GV. 2012.** Stem cell maintenance in shoot apical meristems. *Current opinion in plant biology* **15**: 10–16.
- Pfluger J, Wagner D. 2007.** Histone modifications and dynamic regulation of genome accessibility in plants. *Current Opinion in Plant Biology* **10**: 645–652.
- Plomion C, Lalanne C, Claverol S, Meddour H, Kohler A, Bogeat-Triboulot M-B, Barre A, Le Provost G, Dumazet H, Jacob D, et al. 2006.** Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics* **6**: 6509–6527.
- Rae AM, Street NR, Robinson KM, Harris N, Taylor G. 2009.** Five QTL hotspots for yield in short rotation coppice bioenergy poplar: the Poplar Biomass Loci. *BMC plant biology* **9**: 23.
- Rae AM, Tricker PJ, Bunn SM, Taylor G. 2007.** Adaptation of tree growth to elevated CO₂: quantitative trait loci for biomass in *Populus*. *New Phytologist* **175**: 59–69.
- Raj S, Brautigam K, Hamanishi ET, Wilkins O, Thomas BR, Schroeder W, Mansfield SD, Plant AL, Campbell MM. 2011.** Clone history shapes *Populus* drought responses. *Proceedings of the National Academy of Sciences* **108**: 12521–12526.
- Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R. 2000.** Non-CpG

methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 5237–5242.

Reinders J, Wulff BBH, Mirouze M, Mari-Ordonez A, Dapp M, Rozhon W, Bucher E, Theiler G, Paszkowski J. 2009. Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes & Development* **23**: 939–950.

Riddle NC, Richards EJ. 2002. The control of natural variation in cytosine methylation in Arabidopsis. *Genetics* **162**: 355–363.

Rodin SN, Parkhomchuk DV, Rodin AS, Holmquist GP, Riggs AD. 2005. Repositioning-dependent fate of duplicate genes. *DNA and cell biology* **24**: 529–542.

Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L, et al. 2011. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *The EMBO Journal* **30**: 1928–1938.

Roudier F, Teixeira FK, Colot V. 2009. Chromatin indexing in Arabidopsis: an epigenomic tale of tails and more. *Trends in Genetics* **25**: 511–517.

Rouleau M, Marsolais F, Richard M, Nicolle L, Voigt B, Adam G, Varin L. 1999. Inactivation of brassinosteroid biological activity by a salicylate-inducible steroid sulfotransferase from Brassica napus. *The Journal of biological chemistry* **274**: 20925–20930.

Rowland DL, Garner R, Jespersen M. 2002. A rare occurrence of seed formation on male branches of the dioecious tree, populus deltoides. *The American Midland Naturalist*.

Ruiz-García L, Cervera MT, Martínez-Zapater JM. 2005. DNA methylation increases throughout Arabidopsis development. *Planta* **222**: 301–306.

Rushton PJ, Reinstädler A, Lipka V, Lippok B, Somssich IE. 2002. Synthetic Plant Promoters Containing Defined Regulatory Elements Provide Novel Insights into Pathogen- and Wound-Induced Signaling. *The Plant Cell* **14**: 749–762.

Russo VE a, Martienssen RA, Riggs AD. 1996. *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press.

Sabbah S, Raise M, Tal M. 1995. Methylation of DNA in NaCl-adapted cells of potato. *Plant Cell Reports* **14**.

- Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, et al. 2003.** TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* **34**: 374–378.
- Sala A, Hoch G. 2009.** Height-related growth declines in ponderosa pine are not due to carbon limitation. *Plant, cell & environment* **32**: 22–30.
- Santamour FS. 1956.** Hermaphroditism in Populus. *Third Northeast Forest Tree Improvement Conference Proceedings*.
- Saze H. 2008.** Epigenetic memory transmission through mitosis and meiosis in plants. *Seminars in Cell & Developmental Biology* **19**: 527–536.
- Saze H. 2012.** Transgenerational inheritance of induced changes in the epigenetic state of chromatin in plants. *Genes & Genetic Systems* **87**: 145–152.
- Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, Schork NJ, Ecker JR. 2011.** Transgenerational Epigenetic Instability Is a Source of Novel Methylation Variants. *Science (New York, N.Y.)*.
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. 2009.** The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* **326**: 1112–1115.
- Schones DE, Cui K, Cuddapah S, Roh T-Y, Barski A, Wang Z, Wei G, Zhao K. 2008.** Dynamic Regulation of Nucleosome Positioning in the Human Genome. *Cell* **132**: 887–898.
- Schröder F, Lisso J, Lange P, Müssig C. 2009.** The extracellular EXO protein mediates cell expansion in Arabidopsis leaves. *BMC Plant Biology* **9**: 20.
- Schwartz YB, Pirrotta V. 2008.** Polycomb complexes and epigenetic states. *Current opinion in cell biology* **20**: 266–273.
- Seki M, Umezawa T, Urano K, Shinozaki K. 2007.** Regulatory metabolic networks in drought stress responses. *Current opinion in plant biology* **10**: 296–302.
- Semerçi A, Sanlı BN, Sahin O, Celik O, Balkız GB, Ceylan S, Argun N. 2008.** Examination of tree mortalities in semi-arid central Anatolian region of Turkey during last six-year period (2002–2007). Umea, Sweden, .
- Sha AH, Lin XH, Huang JB, Zhang DP. 2005.** Analysis of DNA methylation related to rice adult plant resistance to bacterial blight based on methylation-sensitive AFLP (MSAP)

analysis. *Molecular genetics and genomics: MGG* **273**: 484–490.

Shao Z. 1999. Probing Nanometer Structures with Atomic Force Microscopy. *Physiology* **14**: 142–149.

Shen JC, Rideout WM 3rd, Jones PA. 1994. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Research* **22**: 972–976.

Shivaprasad PV, Dunn RM, Santos BA, Bassett A, Baulcombe DC. 2012. Extraordinary transgressive phenotypes of hybrid tomato are influenced by epigenetics and small silencing RNAs. *The EMBO journal* **31**: 257–266.

Shroff R, Vergara F, Muck A, Svatos A, Gershenzon J. 2008. Nonuniform distribution of glucosinolates in *Arabidopsis thaliana* leaves has important consequences for plant defense. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 6196–6201.

Singer T, Yordan C, Martienssen RA. 2001. Robertson's Mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene Decrease in DNA Methylation (DDM1). *Genes & development* **15**: 591–602.

Sjödin A, Street NR, Sandberg G, Gustafsson P, Jansson S. 2009. The *Populus* Genome Integrative Explorer (PopGenIE): a new resource for exploring the *Populus* genome. *New Phytologist* **182**: 1013–1025.

Sjodin A, Wissel K, Bylesjo M, Trygg J, Jansson S. 2008. Global expression profiling in leaves of free-growing aspen. *BMC Plant Biology* **8**: 61.

Slatkin M. 2009. Epigenetic Inheritance and the Missing Heritability Problem. *Genetics* **182**: 845–850.

Smallwood SA, Kelsey G. 2012. Genome-wide analysis of DNA methylation in low cell numbers by reduced representation bisulfite sequencing. *Methods in molecular biology (Clifton, N.J.)* **925**: 187–197.

Smith LM, Weigel D. 2012. On epigenetics and epistasis: hybrids and their non-additive interactions. *The EMBO Journal* **31**: 249–250.

Soria G, Polo SE, Almouzni G. 2012. Prime, repair, restore: the active role of chromatin in the DNA damage response. *Molecular cell* **46**: 722–734.

Van Speybroeck L, De Waele D, Van de Vijver G. 2002. Theories in early embryology:

close connections between epigenesis, preformationism, and self-organization. *Annals of the New York Academy of Sciences* **981**: 7–49.

Stanton BJ. 2009. The domestication and conservation of *Populus* genetic resources (Chapter 4a). *Poplars and willows in the world: meeting the needs of society and the environment*. Rome: FAO/IPC (Food and Agricultural Organization of the United Nations / International Poplar Commission), .

Sterky F, Bhalerao RR, Unneberg P, Segerman B, Nilsson P, Brunner AM, Charbonnel-Campaa L, Lindvall JJ, Tandre K, Strauss SH, et al. 2004. A *Populus* EST resource for plant functional genomics. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 13951–13956.

Stettler RF. 1971. Variation in sex expression of black cottonwood and related hybrids. *Silvae Genetica* **20**: 42–46.

Sturtevant AH. 1925. The Effects of Unequal Crossing over at the Bar Locus in *Drosophila*. *Genetics* **10**: 117–147.

Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics* **9**: 465–476.

Takahashi H, Kamakura H, Sato Y, Shiono K, Abiko T, Tsutsumi N, Nagamura Y, Nishizawa NK, Nakazono M. 2010. A method for obtaining high quality RNA from paraffin sections of plant tissues by laser microdissection. *Journal of plant research* **123**: 807–813.

Takuno S, Gaut BS. 2012. Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. *Molecular biology and evolution* **29**: 219–227.

Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, et al. 2011. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**: 1016–1028.

Teixeira FK, Colot V. 2010. Repeat elements and the *Arabidopsis* DNA methylation landscape. *Heredity* **105**: 14–23.

Teyssier E, Bernacchia G, Maury S, How Kit A, Stammitti-Bert L, Rolin D, Gallusci P. 2008. Tissue dependent variations of DNA methylation and endoreduplication levels during tomato fruit development and ripening. *Planta* **228**: 391–399.

Thivolle-Cazat A. 2003. Disponibilité en bois de peuplier en France de 2002 à 2020. *Fiches Informations-Forêt AFOCEL* **1**.

Thomas AJ, Sherratt HS. 1956. The isolation of nucleic acid fractions from plant leaves and their purine and pyrimidine composition. *The Biochemical journal* **62**: 1–4.

Traas J, Vernoux T. 2002. The shoot apical meristem: the dynamics of a stable structure. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **357**: 737–747.

Tran RK, Henikoff JG, Zilberman D, Ditt RF, Jacobsen SE, Henikoff S. 2005. DNA Methylation Profiling Identifies CG Methylation Clusters in Arabidopsis Genes. *Current Biology* **15**: 154–159.

Trap-Gentil M-V. 2009. Contrôle épigénétique du risque de montaison chez une plante de grande culture, la betterave sucrière. Mise au point d'une stratégie de caractérisation d'épiallèles associés à la sensibilité à la montaison en vue de l'élaboration d'un test de sélection.

Trap-Gentil M-V, Hébrard C, Lafon-Placette C, Delaunay A, Hagège D, Joseph C, Brignolas F, Lefebvre M, Barnes S, Maury S. 2011. Time course and amplitude of DNA methylation in the shoot apical meristem are critical points for bolting induction in sugar beet and bolting tolerance between genotypes. *Journal of Experimental Botany* **62**: 2585–2597.

Trojer P, Reinberg D. 2007. Facultative heterochromatin: is there a distinctive molecular signature? *Molecular cell* **28**: 1–13.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al. 2006. The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Tuskan GA, DiFazio SP, Teichmann T. 2004. Poplar genomics is getting popular: the impact of the poplar genome project on tree research. *Plant biology (Stuttgart, Germany)* **6**: 2–4.

Tyree MT, Zimmermann MH. 2002. *Xylem Structure and the Ascent of Sap*. Springer.

Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sidow A. 2011. Determinants of nucleosome organization in primary human cells. *Nature* **474**: 516–520.

de Vanssay A, Bougé A-L, Boivin A, Hermant C, Teyssset L, Delmarre V, Antoniewski C, Ronsseray S. 2012. Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature*.

Vanyushin BF, Ashapkin VV. 2011. DNA methylation in higher plants: past, present and

future. *Biochimica et biophysica acta* **1809**: 360–368.

Vaucheret H, Fagard M. 2001. Transcriptional gene silencing in plants: targets, inducers and regulators. *Trends in Genetics* **17**: 29–35.

Vaughn MW, Tanurdzić M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, Dedhia N, McCombie WR, Agier N, Bulski A, et al. 2007. Epigenetic Natural Variation in *Arabidopsis thaliana*. *PLoS Biology* **5**: e174.

Vining KJ, Pomraning KR, Wilhelm LJ, Priest HD, Pellegrini M, Mockler TC, Freitag M, Strauss S. 2012. Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* **13**: 27.

Vongs A, Kakutani T, Martienssen RA, Richards EJ. 1993. *Arabidopsis thaliana* DNA methylation mutants. *Science (New York, N.Y.)* **260**: 1926–1928.

Waddington CH. 1939. *An Introduction to Modern Genetics*. London: George Allen & Unwin Ltd.

Waddington CH. 1942. The Epigenotype. *Endeavour* **1**: 18–20.

Waddington CH. 1957. *The Strategy of the Genes*. London: George Allen & Unwin Ltd.

Waddington CH. 1962. *New Patterns in Genetics and Development*. London: Columbia University Press.

Waddington CH. 1966. *Principles of Development and Differentiation*. New York: The Macmillan Company.

Wang M-B, Dennis ES. 2009. SPT5-like, a new component in plant RdDM. *EMBO reports* **10**: 573–575.

Waters AJ, Makarevitch I, Eichten SR, Swanson-Wagner RA, Yeh C-T, Xu W, Schnable PS, Vaughn MW, Gehring M, Springer NM. 2011. Parent-of-origin effects on gene expression and DNA methylation in the maize endosperm. *The Plant cell* **23**: 4221–4233.

Weber M, Schübeler D. 2007. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Current opinion in cell biology* **19**: 273–280.

Weiler KS, Wakimoto BT. 1995. Heterochromatin and gene expression in *Drosophila*. *Annual review of genetics* **29**: 577–605.

Weng Y-I, Huang TH-M, Yan PS. 2009. Methylated DNA immunoprecipitation and

microarray-based analysis: detection of DNA methylation in breast cancer cell lines. *Methods in Molecular Biology (Clifton, N.J.)* **590**: 165–176.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007. An ‘Electronic Fluorescent Pictograph’ Browser for Exploring and Analyzing Large-Scale Biological Data Sets (I Baxter, Ed.). *PLoS ONE* **2**: e718.

Wolff P, Weinhofer I, Seguin J, Roszak P, Beisel C, Donoghue MTA, Spillane C, Nordborg M, Rehmsmeier M, Köhler C. 2011. High-resolution analysis of parent-of-origin allelic expression in the Arabidopsis Endosperm. *PLoS genetics* **7**: e1002126.

Wolters H, Jürgens G. 2009. Survival of the flexible: hormonal growth control and adaptation in plant development. *Nature reviews. Genetics* **10**: 305–317.

Wright DE, Wang C-Y, Kao C-F. 2012. Histone ubiquitylation and chromatin dynamics. *Frontiers in bioscience: a journal and virtual library* **17**: 1051–1078.

Wyatt GR. 1950. Occurrence of 5-methylcytosine in nucleic acids. *Nature* **166**: 237–238.

Wyatt GR. 1951. Recognition and estimation of 5-methylcytosine in nucleic acids. *The Biochemical journal* **48**: 581–584.

Xiang H, Zhu J, Chen Q, Dai F, Li X, Li M, Zhang H, Zhang G, Li D, Dong Y, et al. 2010. Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nature biotechnology* **28**: 516–520.

Xiong LZ, Xu CG, Saghai Maroof MA, Zhang Q. 1999. Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. *Molecular & General Genetics: MGG* **261**: 439–446.

Xu Z, Zhou G, Shimizu H. 2010. Plant responses to drought and rewatering. *Plant Signaling & Behavior* **5**: 649–654.

Yang F, Zhang L, Li J, Huang J, Wen R, Ma L, Zhou D, Li L. 2010. Trichostatin A and 5-azacytidine both cause an increase in global histone H4 acetylation and a decrease in global DNA and H3K9 methylation during mitosis in maize. *BMC Plant Biology* **10**: 178.

Yin TM, DiFazio SP, Gunter LE, Riemenschneider D, Tuskan GA. 2004. Large-scale heterospecific segregation distortion in Populus revealed by a dense genetic map. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **109**: 451–463.

Yin T, Difazio SP, Gunter LE, Zhang X, Sewell MM, Woolbright SA, Allan GJ, Kelleher

CT, Douglas CJ, Wang M, et al. 2008. Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. *Genome research* **18**: 422–430.

Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. *Science* **328**: 916–919.

Zemach A, Zilberman D. 2010. Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Current biology: CB* **20**: R780–785.

Zhang M, Kimatu JN, Xu K, Liu B. 2010. DNA cytosine methylation in plant development. *Journal of genetics and genomics = Yi chuan xue bao* **37**: 1–12.

Zhang X, Shiu S, Cal A, Borevitz JO. 2008. Global Analysis of Genetic, Epigenetic and Transcriptional Polymorphisms in *Arabidopsis thaliana* Using Whole Genome Tiling Arrays (GP Copenhaver, Ed.). *PLoS Genetics* **4**: e1000032.

Zhang K, Sridhar VV, Zhu J, Kapoor A, Zhu J-K. 2007. Distinctive Core Histone Post-Translational Modification Patterns in *Arabidopsis thaliana* (F Berger, Ed.). *PLoS ONE* **2**: e1210.

Zhang W, Wu Y, Schnable JC, Zeng Z, Freeling M, Crawford GE, Jiang J. 2012a. High-resolution mapping of open chromatin in the rice genome. *Genome research* **22**: 151–162.

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW-L, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, et al. 2006. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in *Arabidopsis*. *Cell* **126**: 1189–1201.

Zhang W, Zhang T, Wu Y, Jiang J. 2012b. Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *The Plant cell* **24**: 2719–2731.

Zhang M, Zhao H, Xie S, Chen J, Xu Y, Wang K, Zhao H, Guan H, Hu X, Jiao Y, et al. 2011. Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 20042–20047.

Zhu J-K. 2009. Active DNA demethylation mediated by DNA glycosylases. *Annual review of genetics* **43**: 143–166.

Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. 2007. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics* **39**: 61–69.

Annexes et valorisation des travaux scientifiques

1. Bräutigam *et al.* in *Ecology and Evolution* (révision mineure) : article 1

Ecology and Evolution

Open Access

Epigenetic regulation of adaptive responses of forest tree species to the environment

Katharina Bräutigam¹, Kelly J. Vining², Clément Lafon-Placette³, Carl G. Fossdal⁴, Marie Mirouze⁵, José Gutiérrez Marcos⁶, Silvia Fluch⁷, Mario Fernández Fraga⁸, M. Ángeles Guevara^{9,10}, Dolores Abarca¹¹, Øystein Johnsen¹², Stéphane Maury³, Steven H. Strauss², Malcolm M. Campbell^{1,13}, Antje Rohde¹⁴, Carmen Díaz-Sala¹¹ & María-Teresa Cervera^{9,10}

¹Centre for the Analysis of Genome Evolution and Function, Department of Cell & Systems Biology, University of Toronto, Toronto, ON M5S 3B2, Canada

²Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331-5752, USA

³UFR-Faculté des Sciences, UPRES EA 1207 'Laboratoire de Biologie des Ligneux et des Grandes Cultures' (LBLGC), INRA, USC1328 'Arbres et Réponses aux Contraintes Hydrique et Environnementales' (ARCHE), University of Orléans, Rue de Chartres, BP 6759, F-45067 Orléans, France

⁴Department of Biology and Environment, Norwegian Forest and Landscape Institute, PO Box 115, N-1431 Aas, Norway

⁵Epigenetic Regulations and Seed Development, Institut de Recherche pour le Développement, UMR232 ERL5300 CNRS-IRD, 911 Av. Agropolis, 34394, Montpellier, France

⁶School of Life Sciences, University of Warwick, Wellesbourne, Warkwick, CV35 9EF, United Kingdom

⁷Platform for Integrated Clone Management (PICME), Health & Environment Department, AIT Austrian Institute of Technology GmbH, Konrad-Lorenz-Straße 24, 3430 Tulln, Austria

⁸Cancer Epigenetics Laboratory, Institute of Oncology of Asturias (IUOPA-HUCA), University of Oviedo, Spain

⁹Dpt. of Forest Ecology and Genetics, Forest Genomics and Ecophysiology group, Forest Research Centre (CIFOR). INIA, Crta. La Coruña km 7,5, 28040 Madrid, Spain

¹⁰Mixed Unit of Forest Genomics and Ecophysiology, INIA/UPM, Madrid, Spain

¹¹Department of Life Sciences, University of Alcalá, Ctra. Madrid-Barcelona Km. 33,600, 28871 Alcalá de Henares, Madrid, Spain

¹²Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, PO Box 5003, N-1432, Ås, Norway

¹³Department of Biological Sciences, University of Toronto Scarborough, University of Toronto, 1265 Military Trail, Toronto, ON M1C 1A4, Canada

¹⁴Department Plant Growth & Development, Institute of Agriculture and Fisheries Research, Caritasstraat 21, 9090 Melle, Belgium

Keywords

Adaptive response, environmental stress, epigenetic memory of stressful conditions, epigenetics, forest trees, phenotypic plasticity.

Correspondence

María-Teresa Cervera, INIA. Forest Research Centre. Department of Forest Ecology and Genetics, Forest Genomics and Ecophysiology group. Crta. La Coruña km 7,5. 28040 Madrid, Spain. Tel: +(34) 91 347 6798; Fax: +(34) 91 347 6767; E-mail: cervera@inia.es

Funding Information

The preparation of this review was supported by funding from the FP7 UE projects FoResTTraC (2440096-FP7) and ProCoGen (289841-FP7).

Received: 7 August 2012; Revised: 19 November 2012; Accepted: 27 November 2012

doi: 10.1002/ece3.461

Abstract

Epigenetic variation is likely to contribute to the phenotypic plasticity and adaptive capacity of plant species, and may be especially important for long-lived organisms with complex life cycles, including forest trees. Diverse environmental stresses and hybridization/polyploidization events can create reversible heritable epigenetic marks that can be transmitted to subsequent generations as a form of molecular “memory”. Epigenetic changes might also contribute to the ability of plants to colonize or persist in variable environments. In this review, we provide an overview of recent data on epigenetic mechanisms involved in developmental processes and responses to environmental cues in plant, with a focus on forest tree species. We consider the possible role of forest tree epigenetics as a new source of adaptive traits in plant breeding, biotechnology, and ecosystem conservation under rapid climate change.

Introduction

Epigenetics refers to the study of meiotically or mitotically heritable changes in gene function that do not result from changes in DNA sequence (Bonasio et al. 2010). At the molecular level, epigenetic phenomena are mediated by reversible marks such as DNA methylation and histone modifications, and by small RNAs that can alter regulatory states of genes or genomic regions. DNA methylation in plants occurs at cytosines in all sequence contexts (CG, CHG, and CHH where H = A, T, or C), and well-studied post-translational modifications of histone proteins at specific amino acid residues include methylation (Krauss 2008), acetylation, phosphorylation (Demidov et al. 2009) and ubiquitination (Kouzarides 2007). Genome-wide epigenetic patterns, referred to as “epigenomes”, are not static; rather, they can undergo precise changes. Epigenome modifications are involved in biological processes including genetic imprinting, transposon silencing, regulation of gene expression, and heterochromatin organization.

The influence of environmental factors on epigenetic marks, and on the resultant changes in gene expression and phenotype, has recently attracted considerable attention (Boyko and Kovalchuk 2008; Chinnusamy and Zhu 2009; Feil and Fraga 2011; Groszmann et al. 2011a; Mirouze and Paszkowski 2011). Knowledge of the regulatory mechanisms involved in adaptive epigenetic responses may help to guide management of genetic resources and plant breeding, especially in long-lived forest tree species where changes in allele frequency are expected to occur very slowly. This review provides a brief overview of recent data on epigenetic mechanisms involved in developmental processes and responses to environmental cues in forest species, as well as the implications of forest tree epigenetics to adaptation as a possible new source of beneficial traits for plant breeding and conservation in ecosystems responding to climate change.

Factors driving epigenetic regulation in plants

Epigenetic regulation in plant development

Epigenetic regulation plays important roles in multiple aspects of plant development. Two distinct roles of this regulation can be distinguished, depending on whether they concern developmentally regulated genes or transposable elements (TEs). In developmentally regulated genes, epigenetic marks allow switches in gene expression in

response to developmental transitions and/or environmental cues. After sexual reproduction, uniparental expression of parental alleles, imprinting, is associated with discrete differentially methylated regions (DMRs) that act in a genome context-independent manner (Gutierrez-Marcos et al. 2006; Jullien et al. 2006; Haun et al. 2007).

A well-characterized example of epigenetic control during post-embryonic development is vernalization, the phenomenon of cold temperature-induced competence to flower (Chouard 1960; Schmitz and Amasino 2007). In *Arabidopsis thaliana*, regulation of *FLOWERING LOCUS C (FLC)* gene expression during vernalization illustrates how environmental cues are perceived and translated into epigenetic marks that affect plant development (Bastow et al. 2004; Heo and Sung 2011; Kim and Sung 2012).

The epigenetic marks of many loci involved in plant development are normally erased or reset at each generation following meiosis, thus preventing the establishment of new “epialleles” (alleles whose expression is conditioned by their epigenetic status). On the other hand, stable, that is, heritable, epialleles can occur naturally and might confer specific phenotypes. Examples in plants include floral symmetry in *Linaria* that is influenced by DNA methylation levels at the *CYCLOIDEA* locus (Cubas et al. 1999) and absence of ripening in tomato, that is associated with hypermethylation at the *Colorless non-ripening* locus (Manning et al. 2006). Stable epialleles are potential targets for selection in evolutionary processes, or in applied plant breeding. More examples will contribute to a better understanding of their origin, their stability, and the role they might play in selection.

In contrast to the transient nature of many developmental epigenetic marks, those affecting TEs are more stable (Slotkin and Martienssen 2007; Bourc’his and Voinnet 2010; Lisch and Slotkin 2011) and the mobility of TEs is observed when these marks are alleviated in mutants affected in the epigenetic machinery (Mirouze et al. 2009; Tsukahara et al. 2009; Ito et al. 2011). However, during development, transcription of activated TEs in hypomethylated gamete companion cells is thought to produce small RNAs that migrate into the germ cell and direct the silencing machinery to TEs. Hence, at each new generation, the “immune system” against transposons is perpetuated, but also readjusted to prevent potential genome invasion of new mobile elements (Lisch and Slotkin 2011). Given the abundance of TEs in tree genomes, they should be considered as potential sources of epigenetic variation potentially affecting regulation of nearby genes.

The importance of developmentally related epigenetic modification has been underscored recently by its

potential involvement in hybrid vigor. Hybrid vigor, also known as heterosis, describes the superior performance of hybrid progeny over their parents in traits like biomass and seed production or stress resistance. Various models explaining heterotic effects at single loci have been proposed, including dominance, overdominance, and pseudo-overdominance, while interactions between genes (epistasis) have been considered as well (Birchler et al. 2010). The molecular mechanisms causing non-additive gene expression in hybrids have been the focus of studies in rice and *A. thaliana*, and epigenetic regulation has recently been associated with heterosis (Ha et al. 2009; He et al. 2010; Groszmann et al. 2011a,b). In hybrids, a number of short interfering RNAs (siRNAs) were found to accumulate to non-additive levels, which in turn can lead to changes in DNA methylation and gene expression, thus contributing to hybrid vigor (Groszmann et al. 2011a,b). Given the preponderance of hybrids in many plant taxa, including prominent tree genera like *Populus*, the putative involvement of epigenetics in heterosis is of great interest.

Epigenetic regulation in plant environmental responses

Various environmental signals and stresses can induce persistent changes in epigenetic modifications, thereby creating a flexible “memory” system for short or prolonged periods of time (Kvaalen and Johnsen 2008; Chinnusamy and Zhu 2009; Jablonka and Raz 2009; Whittle et al. 2009). Environmental conditions have an impact on a number of different epigenetic marks and mechanisms, including DNA methylation and histone modifications, or on frequencies of homologous recombination and genomic rearrangements (Bond and Finnegan 2007; Chinnusamy and Zhu 2009; Feil and Fraga 2011; Hauser et al. 2011; Mirouze and Paszkowski 2011). For example, changes in genome-wide DNA methylation patterns in response to biotic and abiotic stress treatments (pathogen, herbivore, high salt, low nutrients) occur in asexually reproduced dandelions (*Taraxacum officinale*). Notably, altered DNA methylation patterns were transmitted to the non-stressed progeny in this species and the potential role of stress-induced epigenetic inheritance in evolution has been discussed (Verhoeven et al. 2010). The involvement of a histone variant (H2A.Z) was found to mediate short-term adaptation to temperature change in *A. thaliana* (Kumar and Wigge 2010), and cold stress-induced hypomethylation and transposition of a TE (Tam-3) has been observed in *Antirrhinum* (Hashida et al. 2006).

Epigenetic recombinant inbred lines (epiRILs) have emphasized the relationship between response to environ-

mental conditions and epigenetic phenomena. In *A. thaliana*, epiRILs have nearly identical genomes, but display diverse mosaic epigenomes with variant DNA methylation patterns (Johannes et al. 2009; Reinders et al. 2009). The range of pathogen sensitivity/resistance within one isogenic epiRIL population exhibited 60% of the range of pathogen responses observed in natural, genetically varying *A. thaliana* accessions (Reinders et al. 2009). In the context of environmental challenges, such epigenetic modifications may be thought of as relatively “plastic” yet heritable marks that allow for rapid responses and adaptations and, at the same time, might avoid excessive genetic diversification (Boyko and Kovalchuk 2008; Lira-Medeiros et al. 2010).

Epigenetic control in forest tree species

Relationship between epigenetic and phenotypic plasticity

Forest trees are long-lived organisms with complex life cycles, which must contend with a variable environment over their long lifetimes (Rohde and Junttila 2008). The long generation times impose limits on natural selection under rapidly changing climate conditions (Rehfeldt et al. 1999, 2002). Consequently, trees must be highly adaptable, displaying a wide range of phenotypes as a function of their environments, known as phenotypic plasticity (Nicotra et al. 2010). Phenotypic plasticity is likely to be of great importance for both individual trees and forest populations over near- and long-term timescales. Despite this, knowledge of the extent and underlying mechanisms of phenotypic plasticity in response to a variety of stress responses and developmental traits in trees is rudimentary (Rohde 2009; Lira-Medeiros et al. 2010; Neale and Kremer 2011).

In addition to the genetic component, epigenetic variation has been suggested to contribute to the phenotypic plasticity and adaptive potential of individuals and populations (Bossdorf et al. 2008; Jablonka and Raz 2009; Herrera and Bazaga 2010; Lira-Medeiros et al. 2010; Richards et al. 2012). Insight into epigenetic variation, and its relationship to phenotypic plasticity, will contribute to the understanding of adaptive plant responses, and might help to evaluate the risk of long-lived species to both short-term and long-term fluctuations in the environment. Moreover, understanding the interplay between epigenetics and adaptation should enhance the understanding of evolutionary trajectories, as natural selection also directly targets the proportion of phenotypic variation that is shaped by epigenetic phenomena (Bossdorf et al. 2008; Herrera and Bazaga 2010).

Epigenetic and phenotypic variation in natural populations, ecotypes, and species

Despite the substantial impact that epigenetics might have in determining environmental compatibility, relatively few studies have investigated the extent of natural epigenetic variation and its relationship to phenotypic variation and adaptation potential (Cervera et al. 2002; Bossdorf et al. 2008; Jablonka and Raz 2009; Marfil et al. 2009; Herrera and Bazaga 2010; Lira-Medeiros et al. 2010; Paun et al. 2010). Among the different epigenetic mechanisms that are potentially involved in transgenerational inheritance and natural epigenetic variation, DNA methylation represents the most-studied modification (Akimoto et al. 2007; Jablonka and Raz 2009; Herrera and Bazaga 2010; Lira-Medeiros et al. 2010; Paun et al. 2010; Verhoeven et al. 2010). Two of the few published studies in higher plants that considered the interplay between genetic, epigenetic, as well as phenotypic variation and environmental factors, focused on a perennial violet species (*Viola cazorlensis*) and orchids of the *Dactylorhiza majalis* complex (Herrera 1990; Herrera and Bazaga 2010; Paun et al. 2010). The studies detected coordinated genetic–epigenetic adaptive differentiation, indicating the involvement of epigenetic processes in adaptation and evolution by influencing primary phenotypic diversity.

In tree species, natural variation in epigenetic marks and the relation to phenotypic traits is still an under-explored area. Insight into the role of epigenetics in determining tree phenotype should identify key elements in the control of growth traits and contribute to the understanding of evolutionary capacity of tree species (Grattapaglia et al. 2009; Thumma et al. 2009; Lira-Medeiros et al. 2010). In keeping with this, evidence for the correlation between tree form and epigenetics is emerging. Trees of the white mangrove (*Laguncularia racemosa*) can occur naturally in contrasting habitats and can exhibit striking differences in morphological traits (Lira-Medeiros et al. 2010). Tree-like appearance was documented in a riverside habitat with abundant fresh water and nutrient supply, whereas in a nearby salt marsh habitat, mangrove plants were characterized by abnormal growth and shrub-like morphology. Notably, despite morphological dissimilarities, analysis of DNA nucleotide sequences and methylation patterns detected greater epigenetic than genetic variation within and between populations in contrasting environments, which indicates that epigenetic variation in natural populations plays an important role in long-term adaptation to different environments (Lira-Medeiros et al. 2010).

The lasting impact of previous environmental history on a tree's capacity to respond to a current environmental stimulus has recently been explored in *Populus* (Raj et al.

2011). Poplar trees are frequently propagated vegetatively through stem cuttings of branches containing dormant buds, generating genetically identical individuals or ramets of the same genotype. Clonally propagated poplar trees can be planted in different geographic locations, thus giving rise to populations of genetically identical ramets that are characterized by their own local environment and history. To study the lasting effect of clone history on current plant performance, cuttings of the same genotype were obtained from different geographic locations and grown under common environmental conditions, after which the transcriptome response to an important environmental stress, drought, was studied. Notably, differences in transcript abundance patterns in response to drought that were based on differences in geographic origin of clonally propagated trees were detected in two of the three investigated genotypes. These transcriptome-level patterns were paralleled by differences in genome-wide DNA methylation. Genotypes with the longest time since establishment and last common propagation showed the most pronounced location-specific patterns in transcriptome response and DNA methylation indicating a possible epigenomic basis for clone history-dependent transcriptome divergence (Fig. 1). These findings underline the importance of epigenetic mechanisms related to the adaptation of long-lived species like poplar trees to the local environment (Raj et al. 2011).

The direct response of six hybrid poplar genotypes to water deficit revealed a relationship between epigenetic marks and the genotypic variability of phenotypic plasticity (Gourcilleau et al. 2010). Genotypic variation for both DNA methylation and traits related to biomass productivity was observed in hybrids (*Populus deltoids* × *P. nigra*), and a positive correlation was established among these variables in well-watered conditions (Fig. 2). While poplar genotypes showed reduced growth in water-deficit conditions, a significant genotype effect was observed for DNA methylation variations. This suggests that DNA methylation could participate in the fine-tuning of gene expression in poplar during water stress (Plomion et al. 2006; Bogeat-Triboulot et al. 2007; Bonhomme et al. 2009; Wilkins et al. 2009; Gourcilleau et al. 2010; Hamanishi and Campbell 2011).

The potential link between natural epigenetic variation and phenotypic variability observed in trees is further supported by studies in ecotypes and individual populations of specific herbaceous plant species (Cervera et al. 2002; Marfil et al. 2009). Highly conserved DNA methylation patterns were detected within an *A. thaliana* ecotype (Ler) while clear DNA methylation differences existed between ecotypes that did not correlate with nucleotide sequence variation, but with their flowering time (Cervera et al. 2002). Furthermore, variation in the

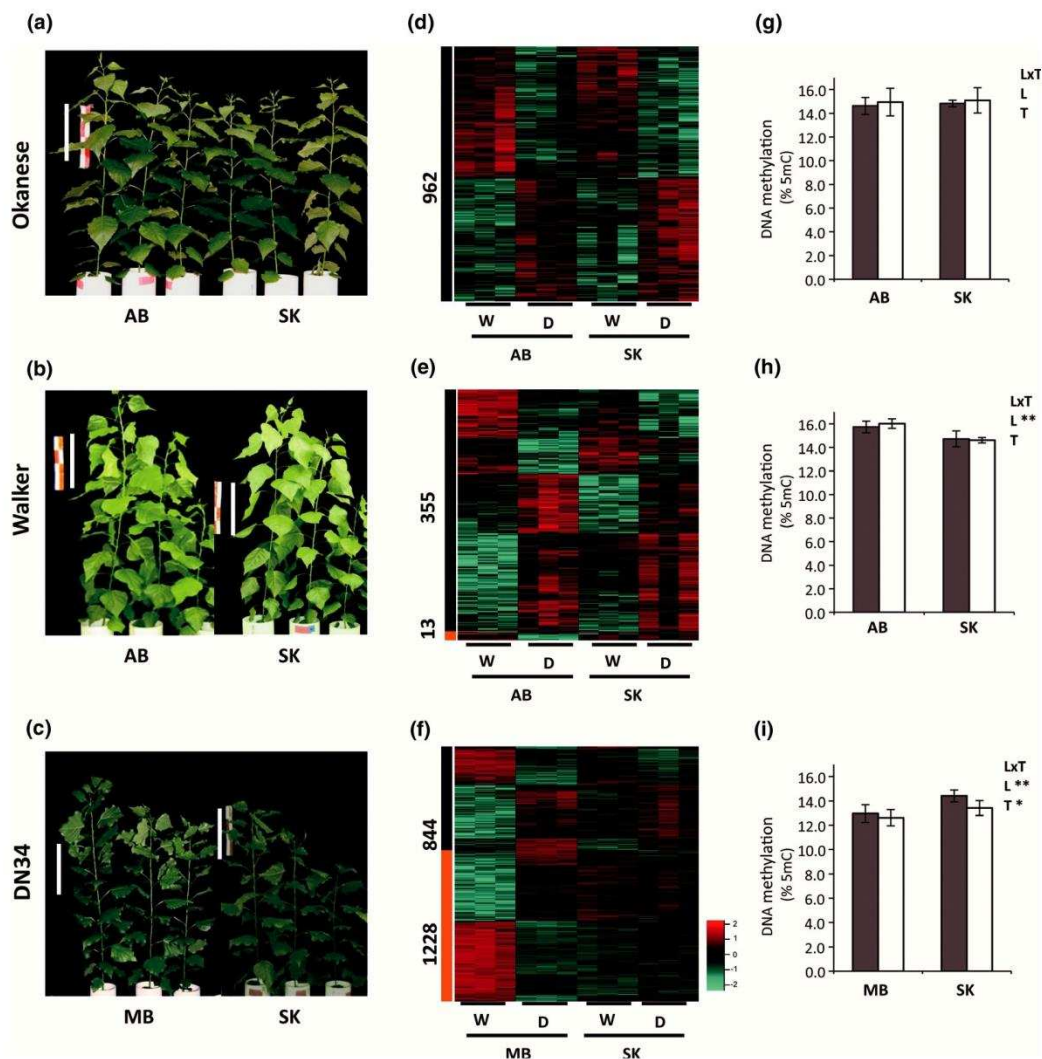


Figure 1. Clone history shapes drought responses in poplar hybrids. Transcriptome-level responses to water withholding are influenced by geographic origin for two of the three genotypes, and are paralleled by differences in total (genome-wide) DNA methylation. Ramets of hybrid poplar genotypes (a) Okanese, (b) Walker, and (c) DN34 with distinct histories were obtained from two different locations for each of the genotypes. The response to water deficit was assessed under common, controlled environmental conditions. Okanese (a, d, g); Walker (b, e, h); DN34 (c, f, i). Tree appearance (a–c). Transcriptome-level responses (d–f). Heat maps represent relative abundance of drought responsive transcripts at pre-dawn for Okanese (d), Walker (e), and Okanese (f) obtained from two locations each. The numbers indicated to the side of the heat map correspond to transcripts with significant treatment main effect only (gray) and with significant treatment: location interactions (orange bar) (BH adjusted, $P < 0.05$). W, well-watered samples; D, water-deficient samples. Global DNA methylation levels as percentage of 5 mC under well-watered (shaded bars) and water-limited conditions (white bars) for the genotypes Okanese (g), Walker (h), and DN34 (i). L, location effect; T, treatment effect; and LxT, location: treatment interaction term ($*P < 0.05$, $**P < 0.001$, $n = 6$, SD bars). Locations are abbreviated as follows: AB, Alberta; SK, Saskatchewan; MB, Manitoba. Figure is adapted from Raj et al. 2011.

floral phenotype of individuals from a single natural population of a wild hybrid potato (*Solanum ruiz-lealii*) was found to correlate with distinct DNA methylation patterns, but not with DNA sequence variation (Marfil et al. 2009).

Most studies assessing epigenetic variation in natural populations, ecotypes, or species focused on the extent of

epigenetic variability and paid less attention to the functional consequences. Indication for a functional link between a specific epigenetic mark at a specific position in the genome and variation in a quantitative trait was discovered by analyzing polymorphisms in an association population and a full-sib family of eucalypt (*Eucalyptus nitens*; Thumma et al. 2009). Making use of the low

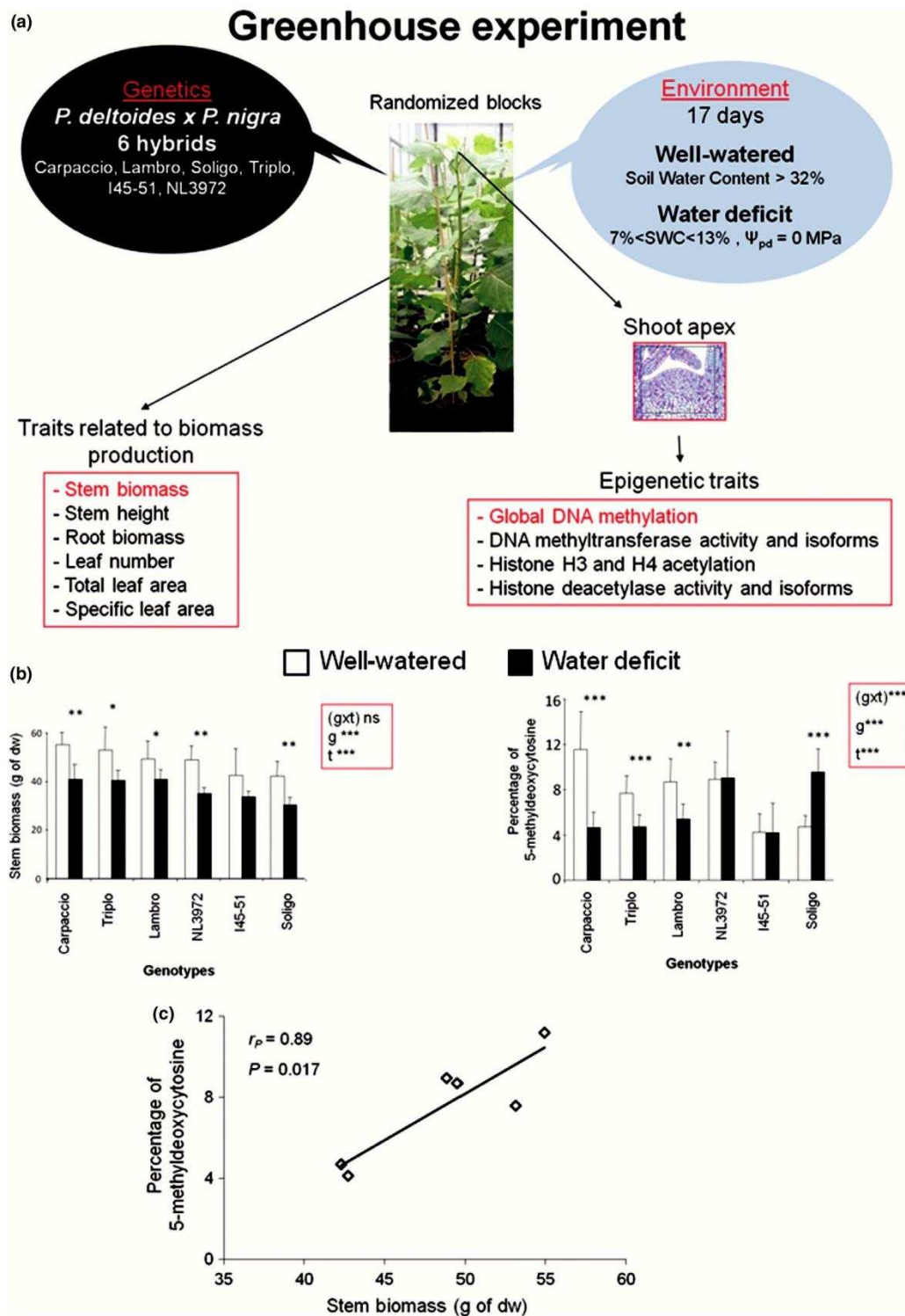


Figure 2. Relation between epigenetic marks and the genotypic variability of phenotypic plasticity under limited water availability or not in six poplar hybrids. (a). Experimental design; (b). Stem biomass and DNA methylation levels in the shoot apex (center of morphogenesis). For each graph, g indicates the genotype effect, t the treatment effect, and (gxt) genotype by treatment effect. Means are accompanied by their standard errors SE ($n = 6$). Significant differences between well-watered and water-deficit conditions are indicated by asterisk: $*P \leq 0.05$, $**P \leq 0.01$, and $***P \leq 0.001$; c. Linear correlation (Pearson, r) between stem biomass and DNA methylation levels. Adapted from Gourcilleau et al. 2010.

linkage disequilibrium in populations of forest trees, variation in cellulose content was linked to polymorphisms within a gene potentially involved in cellulose synthesis and deposition (functional polymorphisms). The COBRA-like gene *EnCOBL4A* was strongly associated with a QTL region for cellulose content and fine mapping revealed a significant association with a SNP in exon 5. Notably, allelic expression imbalance was linked to allele-specific cytosine methylation upstream of this SNP in a full-sibling family. A heritable epigenetic polymorphism is thus likely to influence phenotypic variation in cellulose content; however, further functional analyses are required (Thumma et al. 2009). The findings suggest that epigenetic variations might contribute to quantitative trait variation (Thumma et al. 2009), and it has been suggested that this phenomenon might be common (Johannes et al. 2008; Reinders et al. 2009; Thumma et al. 2009; Long et al. 2011).

To date, some prominent, shared observations have emerged from the few studies of natural epigenetic variation and phenotypic plasticity. These studies established that (1) epigenetic variation occurs in natural populations, ecotypes, and species; (2) this variation can correlate with naturally occurring phenotypic variation; and (3) there is a potential role for epigenetic variation in adaptation and potentially in evolution. Despite the commonalities that have emerged from these studies, many questions remain unresolved. For example, is epigenetic variation in natural populations a wide-spread phenomenon? Moreover, the key molecular mechanisms involved and how they are regulated remain to be determined. Finally, it remains unclear to what extent epialleles arise and how stable they are when considered in an evolutionary context (Bossdorf et al. 2008; Herrera and Bazaga 2010; Lira-Medeiros et al. 2010; Paun et al. 2010). Answers to such questions might also contribute to a better understanding of the adaptive capability of long-lived forest trees that might help to assess their susceptibility to rapidly changing environments (Grattapaglia et al. 2009).

Epigenetic plasticity in growth and development

During their relatively long lifespans, trees must make developmental adjustments while retaining flexibility to match and synchronize growth and development with prevailing environmental conditions. Epigenetic mechanisms are proposed to contribute to such flexible adjustments by generating transmittable and reversible marks that constitute temporary “memory” systems (Boyko and Kovalchuk 2008; Kvaalen and Johnsen 2008; Yakovlev et al. 2010; Jaskiewicz et al. 2011). Tissue-, organ-, and species-specific differences in DNA methylation levels are

well known (Fraga et al. 2002a,b; Valledor et al. 2007, 2010; Monteuis et al. 2009; Santamaria et al. 2009; Rodriguez Lopez et al. 2010; Vining et al. 2012; Lafon-Placette et al. 2013). Changes in epigenetic marks were found to accompany morphological and physiological changes in trees in a wide variety of processes, including aging, phase change, organ maturation, and bud set or burst (Fraga et al. 2002a,b; Santamaria et al. 2009; Valledor et al. 2010).

Bud dormancy is a vital adaptation to seasonal changes, and release and induction of bud dormancy are complex processes that largely determine length of the growth season, and thereby affect annual tree productivity. Regulation of bud burst integrates endogenous and exogenous signals such as hormone levels, day length, light quality, and temperature (Santamaria et al. 2009) and involves substantial changes in gene expression and epigenetic modifications (Ruttink et al. 2007; Rohde 2009; Santamaria et al. 2009). In apical buds of a chestnut (*Castanea sativa*), a decrease in global DNA methylation level and concomitant increase in acetylation of histone 4 were observed during bud burst when conditions were favorable for active growth. The opposite pattern (i.e., DNA hypermethylation and lower histone acetylation levels), indicative of more repressive chromatin states, was detected during bud set when environmental conditions were less favorable for growth (Santamaria et al. 2009). The observed coordinated changes in DNA methylation and histone modifications are predicted to alter the control of gene expression to shape the processes of bud burst and bud set (Santamaria et al. 2009).

Aging and maturation are characterized by altered patterns of cell differentiation and organ formation processes, and the potential role of DNA methylation in maturation has been studied in some tree species (Fraga et al. 2002a,b; Valledor et al. 2007; Monteuis et al. 2009). For example, studies in radiata pine (*Pinus radiata*) support the involvement of DNA methylation in this process. Changes in global DNA methylation levels of up to 25% during maturation have been reported in this species (Fraga et al. 2002a,b). In juvenile plants without flowering capability, young needle tissue was characterized by a markedly lower extent of DNA methylation than corresponding tissues in adult trees with reproductive ability. Regarding histone modifications, decreased levels of euchromatin-associated marks, such as histone 4 acetylation and specific histone methylation (trimethylation of histone 3 on lysine 4 or H3K4me3) have been measured in mature needles when compared with juvenile ones (Valledor et al. 2010). Moreover, the observed increase in DNA methylation levels from juvenile to mature plants in meristematic tissue could be directly linked to phase change. Conversely, an increase in the degree of tree

reinvigoration by serial grafting, measured by the recovery of morphogenetic competence, was accompanied by a decrease in global level of DNA methylation in meristematic tissue, thus pointing toward plasticity of DNA methylation marks during aging and maturation. The degree of DNA methylation, as well as additional biochemical characteristics, was proposed to serve as suitable markers for aging and reinvigoration in pine (Fraga et al. 2002a,b). However, differences between species and experimental systems might exist. In another conifer, *Larix laricina*, age-related changes in foliar traits were observed, whereas differences in DNA methylation levels between juvenile and mature scions could not be detected in DNA from whole needles (Greenwood et al. 1989).

In angiosperms, heteroblastic tree species like *Acacia mangium* with distinct leaf morphologies of juvenile and mature stages provide excellent systems to study aging. Small but significant differences between microshoots with juvenile (pinnate) and mature (phyllode) morphology were observed in this acacia species when analyzing global DNA methylation levels in physiologically active apical buds of *in vitro* grown plant material. Here, the degree of DNA methylation was higher in juvenile than in mature tissue, and might be influenced by *in vitro* culture conditions (see Epigenetic and phenotypic plasticity in artificial systems) in addition to maturation-related processes (Baurens et al. 2004; Monteuiis et al. 2009). Taken together, the aforementioned studies establish a clear relationship between DNA methylation levels and maturation for some tissue types and species in woody plants (Fraga et al. 2002a,b; Baurens et al. 2004; Valledor et al. 2007; Monteuiis et al. 2009). Observed differences might be attributable to differences in taxonomy, tissue type (meristematic vs. differentiated), or experimental system (*in vitro*, field conditions) and might also reflect underlying mechanistic differences in the relationship between aging and epigenetic marks (Fraga et al. 2002b; Monteuiis et al. 2009). Furthermore, the data indicate that DNA methylation patterns are not static and can exhibit remarkable dynamics and plasticity during development and seasonal changes (Fraga et al. 2002b; Valledor et al. 2007; Monteuiis et al. 2009).

Evidence for the remarkable dynamics and plasticity of epigenetic modification in tree species is growing. Genome-level comparative analysis of cytosine methylation among differentiated poplar tissues revealed highly heterogeneous DNA methylation profiles among chromosomes, and a number of cases of tissue-specific methylation (Fig. 3; Vining et al. 2012), many of them associated with gene bodies or promoters. Although a broadly similar chromosome methylation and gene

expression profile was observed in poplar when compared to *A. thaliana* and other plant species, significant differences were also detected. For example, only in poplar was gene body (i.e., the entire gene from the transcription start site to the end of the transcript) methylation associated with greater repression of gene expression than was promoter methylation. In addition, Vining et al. (2012) observed a distinctive pattern of transposon and gene body methylation for male catkins compared with other tissues, including female catkins. Recently, analysis of the methylome of open chromatin in poplar meristematic cells found that 74% of poplar gene models had gene body methylation, and its intensity, as well as cytosine context, varied depending on gene size, redundancy in the genome (presence of paralogs), and extent of tissue-specific gene expression (Lafon-Placette et al. 2013).

Plasticity in tree epigenetic modification has also been observed in conifer species, specifically as it relates to phenology. Phenology responses of seedlings that were produced in warm or cold years vary within the same stands (Kohmann and Johnsen 1994). In Norway spruce, a temperature-dependent epigenetic “memory” from the time of embryo development, which thereafter influences the timing of bud phenology and gene expression, has been discovered (Skrøppa and Johnsen 2000; Johnsen et al. 2005; Yakovlev et al. 2010). Colder-than-normal conditions during embryogenesis and seed development advance the timing, whereas temperatures above normal delay the onset of these adaptive processes, and the altered performance is long lasting in the progeny. This phenomenon was initially discovered when ecotypes from northern Norway were transferred to a southern seed orchard where they produced progenies with a phenology similar to that of southern ecotypes (Johnsen et al. 1996; Skråppa and Johnsen 2000). Notably, differences in day length and temperature applied during pollen formation did not affect the progeny performance. Differences in the female flowering environment did affect progeny performance. The temperature during zygotic embryogenesis and seed maturation shifted the developmental program of the seeds, resulting in significant phenotypic changes, with the effect lasting as long as over 20 years (Skrøppa and Johnsen 2000; Skråppa et al. 2010; Yakovlev et al. 2010). The traits that are affected include the timing of dehardening and bud burst in the spring; leader shoot growth cessation in the summer and bud set and cold acclimation in the autumn. All processes are thus advanced or delayed as influenced by the temperature during reproduction in progeny with identical genetic background. Similar effects have been observed in progeny from white spruce (*Picea glauca* × *Picea engelmannii*) crosses, Scots pine, *Larix* spp., and longleaf pine (Dorm-

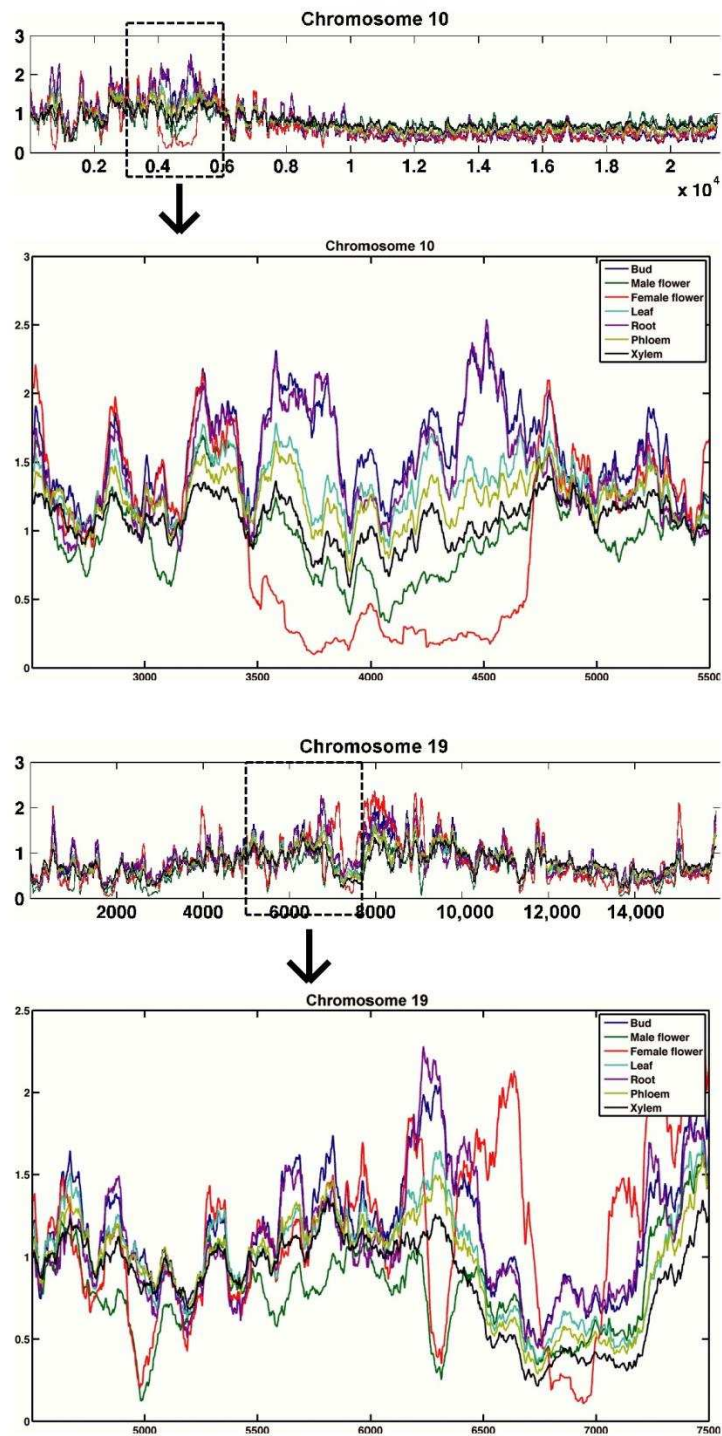


Figure 3. DNA methylation profiles vary widely among chromosomes and among tissues at selected loci in *Populus trichocarpa*. Relative DNA methylation was determined using methylated DNA immunoprecipitation followed by Illumina sequencing (MeDIP-seq). The ratio of MeDIP-seq read counts in immunoprecipitated (IP) samples versus non-IP control is plotted in 1-kb windows for chromosomes 10 and 19, and areas of tissue-differential methylation are expanded below each chromosome. Figures from Vining et al. (2012).

ling and Johnsen 1992, 1992; Greenwood and Hutchison 1996; Stoehr et al. 1998; Webber et al. 2005), but there is lack of information regarding this phenomenon in angiosperm trees (Rohde and Junttila 2008). In birch

(*Betula pendula*), a small-scale study within a population revealed a close genetic relationship between trees that had established in a year of similar temperature (Kelly et al. 2003).

The importance of plastic epigenetic modification on phenology in conifer species extends beyond the individual to encompass the ecosystem. Epigenetic effects taking place during zygote development may create phenotypic diversity at the local community level, if temperature varies considerably among successive generations. This is particularly important as phenology traits are strongly genetically differentiated.

The molecular mechanism behind this striking epigenetic “memory” phenomenon is not yet clear, but transcriptional changes have been implicated (Johnsen et al. 2005; Yakovlev et al. 2010, 2011). In progeny that differ epigenetically, transcriptional analysis revealed that seedlings from full-sib families produced at different embryogenesis temperatures under long- and short-day conditions differed. Suppressive subtracted cDNA libraries revealed considerable differences in their transcriptomes. MicroRNA pathway genes *DICER-LIKE1* (*PaDCL1*), *DICER-LIKE 2* (*PaDCL2*), and *SUPPRESSOR OF GENE SILENCING 3-LIKE* (*PaSGS3*), as well as transposon-related genes, had altered transcript abundance in epigenetically different progeny with phenotypic differences in bud burst and bud set (Yakovlev et al. 2011). Norway spruce contains a set of conserved miRNAs as well as a large proportion of novel non-conserved miRNAs involved in temperature-dependent epigenetic “memory”. Most of the miRNAs were targeted to previously unknown genes, or genes with no known function. The expression of seven conserved and nine novel miRNAs showed significant differences in transcript levels in progenies showing distinct epigenetic difference in bud set, but not in the progeny from a non-responding family without differences in bud set, making them excellent candidate miRNAs. The altered transcript abundance of specific miRNAs suggests their putative participation in epigenetic regulation (Yakovlev et al. 2010). This epigenetic phenomenon is not only generated in controlled Norway spruce crosses, but such epitypes can also be produced by somatic embryogenesis (Kvaalen and Johnsen 2008). Genetically identical plants generated at different temperatures by zygotic embryogenesis expressed a difference in timing of terminal bud formation that was equivalent to a 4–6° latitudinal ecotypic difference.

The “memory” effects acting on phenological traits lasted for more than 20 years after germination and affected long-term growth under field conditions (Skrøppa et al. 2007). Notably, there was absence of any genetic segregation distortion in the progeny, strongly supporting that this “memory”, affecting the climatic adaptation in this species, is indeed an epigenetic phenomenon (Besnard et al. 2008). Thus, distinct epitypes can be produced from the same genotype in Norway spruce, a process not well documented in other tree

species so far. In view of rapid climate change, strategies to increase diversity for selection might be of prime importance for survival of species within their current geographic distribution, and therefore this epigenetic “memory” mechanism is likely of evolutionary significance and has obvious practical implications.

Epigenetic and phenotypic plasticity in artificial systems

While epigenetic phenomena are clearly important for trees in a natural context, they also could be of great consequence during specific tree production processes integrated into the wood products chain. Long generation times and the out-crossing habit of a number of forest trees can make it difficult to rapidly propagate material and maintain valuable genotypes under natural conditions. Tissue culture can provide alternative means to keep desirable genotypes by vegetative propagation and to quickly produce commercial quantities of regenerants; therefore, micropropagation is widely used in forestry. It has been observed, however, that tissue culture can introduce variation in regenerated plants. This somaclonal variation can result in subtle to drastic phenotypic variation and has been found to be attributable to genetic or epigenetic variations (e.g., reviewed in Kaeppler et al. 2000; Miguel and Marum 2011). Somaclonal variation (heritable across mitotic and meiotic cell divisions) has been considered both beneficial and disadvantageous (Jaligot et al. 2000; Kaeppler et al. 2000; Schellenbaum et al. 2008), and a number of studies have focused on elucidating underlying mechanisms (Kaeppler et al. 2000; Rival et al. 2008; Schellenbaum et al. 2008; Rodriguez Lopez et al. 2010).

A well-studied example for somaclonal variants and their relation to epigenetic marks in a tree species is the *mantled* phenotype in somatic-embryo-derived oil palm (*Elaeis guineensis*). This phenotypic variant, found in about five percent of regenerants, is characterized by abnormal inflorescence development and has been associated with global DNA hypomethylation, but not to changes in genomic structure or nucleotide sequence (Jaligot et al. 2000; Rival et al. 2008). The exact mechanisms involved in generating somaclonal variants like the *mantled* phenotype remain largely unresolved. Ongoing studies of this phenomenon might help to better understand mechanisms of epigenetic responses to tissue-culture-induced stresses (Kaeppler et al. 2000; Rival et al. 2008).

It has also been observed that the ability to generate mature somatic embryos from cultured tissue can decrease as a culture ages and that somaclonal variation can increase with culture age (Phillips et al. 1994;

Valledor et al. 2007; Krizova et al. 2009). In addition to other mechanisms, changes in DNA methylation were considered to contribute to the reduction in embryonic potential or organogenic potential in tissue culture and grafting procedures (Fraga et al. 2002b; Valledor et al. 2007). A detailed analysis of genetic and epigenetic variation in relation to callus age reports interesting plasticity in cocoa plants (*Theobroma cacao*) regenerated by somatic embryogenesis. Genetic variation was investigated using single sequence repeat (SSR) markers, and epigenetic variability was assessed by methylation-sensitive amplified polymorphism (MSAP), a method to detect genome-wide but anonymous DNA methylation patterns. Contrary to predictions, after an initial increase, a decrease in both genetic and epigenetic divergence between leaves of regenerants and the ortet plant was observed after the culture had reached an age of about 10 weeks (Rodriguez Lopez et al. 2010). One possible interpretation of the findings suggests a link between stable DNA methylation patterns and repression of *de novo* mutations during somatic embryogenesis (Rodriguez Lopez et al. 2010).

For many plant species, different physiological and developmental stages of diverse tissue explant types have been associated with distinct epigenetic characteristics, in particular DNA methylation (Fraga et al. 2002a,b; Monteuuis et al. 2009; Santamaria et al. 2009; Rodriguez Lopez et al. 2010; Valledor et al. 2010). For example, some DNA methylation patterns and levels, characteristics of the source tissue used to start an *in vitro* culture, were retained in regenerants in acacia and cocoa (Monteuuis et al. 2009; Rodriguez Lopez et al. 2010). This highlights the plasticity of DNA methylation marks under tissue culture conditions. Transitions from juvenile to adult phase are frequently accompanied by reduction or loss of morphogenetic ability in woody species (see Epigenetic plasticity in growth and development). Concomitant with maturation of pine needles, changes in epigenetic marks were measured when compared with immature needles. This finding could be in accordance with a less permissive and reprogrammable chromatin state and could account in part for the reduced organogenic capacity of explants from mature needles.

Generation of somaclonal genetic and epigenetic variants as well as plasticity in DNA methylation is widely documented outcomes of plant regeneration in tissue culture (Kaeppeler et al. 2000; Marfil et al. 2009). Studying underlying mechanisms might be of relevance for basic research and applications in plant propagation such as the understanding of differentiation and dedifferentiation processes or the selection of appropriate *in vitro* culture conditions (Kaeppeler et al. 2000; Marfil et al. 2009; Rodriguez Lopez et al. 2010).

Strategic and technical approaches to study epigenetic processes

Selection of appropriate systems

Different methods have been used in model plants to analyze epigenetic variation independently of genetic variation. These have included treatment with demethylating agents, analysis of natural epimutations, and study of DNA methylation-deficient mutants. Epigenetic recombinant inbred lines (epiRILs) have been developed in *A. thaliana* (Johannes et al. 2009; Reinders et al. 2009) using isogenic lines (wild types and mutant lines) differing only in the level and distribution of DNA methylation (see Epigenetic regulation in plant environmental responses). These lines represent a powerful tool to identify specific epigenomic regions that are associated with the observed phenotypic variation through epiQTL mapping approaches that are based on methylation-sensitive markers. The epiQTL mapping approach requires the establishment of multiple plant generations, and may be difficult to apply to tree species that require a significant amount of time to reach sexual maturation.

To discern genetic and epigenetic effects, clonally propagated plants or systems that are characterized by reduced genetic variation, such as stone pine (*Pinus pinea*), represent ideal study subjects. To separate heritable from non-heritable epigenetic variation (resulting from developmental plasticity in response to different environments), it is necessary to study, when available, clonally propagated genotypes, the progeny of different natural populations or maternal families in a common environment, and to use the resemblance of epigenetic patterns among relatives as an indication of epigenetic inheritance (Bossdorf et al. 2008).

Technical approaches

A wide variety of techniques have been developed to study epigenetic patterns and modifications. Histone modifications can be analyzed by chromatin immunoprecipitation (ChIP) using antibodies that recognize specific histone modifications, followed by either microarray hybridization (ChIP on chip) or by next generation sequencing (ChIP-Seq; Ku et al. 2011). DNA methylation at the genome level, the DNA methylome, can be investigated by methylated DNA immunoprecipitation (meDIP) or by bisulfite treatment of the DNA followed by hybridization to a microarray, or by next generation sequencing (BS-Seq; Ku et al. 2011; Krueger et al. 2012; Cokus et al. 2008). Additionally, direct detection of methylated residues using DNA synthesis technologies based on variable polymerase kinetics depending on the chemical

modification of the template nucleotide (e.g., 5-methylcytosine *vs.* cytosine) represents a novel method to directly detect DNA methylation (Flusberg et al. 2010).

Next generation sequencing technologies enable mapping of epigenetic modifications at single-base resolution. The nature and large amount of data generated by such technologies will demand new approaches in data analysis techniques. Inference of the methylation status of bisulfite-treated DNA by BS-Seq can be challenging as the data obtained do not exactly match the reference sequence. Consequently, both DNA strands must be considered separately, and methylation at a specific site can be a percentage rather than a total presence or absence. Nevertheless, a number of tools has been developed to facilitate these analyses and are now available for application to tree epigenomes (Chen et al. 2010; Lim et al. 2010; Krueger et al. 2012).

Conclusions

Many questions remain about the mechanisms and roles of epigenetic processes in enabling rapid adaptation of plants to their environment, especially in forest trees. Recently, genome-wide studies of chromatin-bound proteins and epigenetic marks in *Drosophila melanogaster* and in *A. thaliana* have substantially revised our understanding of chromatin (Roudier et al. 2011; Van Steensel 2011). The dogma of an uncompact, transcriptionally active euchromatin *versus* a compacted, silent heterochromatin is likely to be an oversimplification of the real chromatin architecture. It appears that chromatin might be composed of several types differing in their epigenetic marks as well as in their nuclear localization and chromatin-associated proteins. These types could favor or prevent association with transcription factors, thus defining gene expression patterns. Whether these chromatin types exist in perennial species is not known, and the stability of these chromatin types in long-living organisms is to be established. Similarly, the maintenance of these types during clonal and *in vitro* culture propagation will give important clues about the effect of these biotechnologies on gene expression control.

It has been observed that the induction of alternative epigenetic states not only triggers the formation of new epialleles but also promotes the movement of DNA transposons and retroelements that are very abundant in plant genomes (Mirouze and Paszkowski 2011). However, mechanisms counteracting accumulation of induced epialleles must also be in place, because otherwise we would be “constantly confronted with the inheritance of environmentally induced phenotypic variation” (Richards 2006). Additionally, in large genomes, such as those of conifer [with C estimates of DNA content ranging from

17 to 30 Gbp for pines and spruces of which more than 68% are attributed to repeated DNA (Rake et al. 1980; Ohri and Khoshoo 1986)], cytosine methylation is implicated in genomic compartmentalization, *that is*, non-coding highly repeated sequences get separated from low-copy sequence and transcriptionally active regions. The differential methylation of genic and non-genic regions observed across plant taxa may be involved in decreasing transcriptional “noise” (Rabinowicz et al. 2005). In large genomes, epigenetic mechanisms might be more prominent, as a means to control the repetitive parts of the genome. This might render their entire genomes more amenable to epigenetic regulation.

From an economic and ecological point of view, it is important to integrate information on epigenetic control of environmental and developmental processes in both forest resources management and breeding. In quantitative genetic studies, estimates of genetic variance over the total phenotypic variance are typically used to assess the heritability of a trait. Akin to other genetic characters, variance in epigenetic characters will contribute to genetic variance and/or phenotypic variance, but might go undetected in some studies, or might be confounded with normal Mendelian-based quantitative inheritance (Kalisz and Purugganan 2004). Epigenetic effects may thus inflate the true genetic variation in traits. As a consequence, the genetic clines observed for many phenology traits, even in common garden experiments, may reflect more local adaptation than DNA sequence-based genetic differences among populations.

Recent developments show that both energy efficiency and energy homeostasis, which are integral parts of yield, have an epigenetic component that can be directed and stabilized by artificial selection (i.e., selective breeding; De Block and Van Lijsebettens 2011). These findings open new possibilities for engineering plant metabolism and improving complex traits. For example, in addition to the unintended genetic and epigenetic variation imparted by *in vitro* manipulation, it may be considered and utilized as a means to amplify or release epigenetic variation of value to breeding programs. Transgenic perturbation of epigenetic mechanisms might have similar effects; however, testing such effects using a transgenic approach with forest trees at a scale relevant to application and ecological variation are, at present, constrained by government regulations (Viswanath et al. 2012).

Genome perturbation, including epigenetic components, might be important for increasing the raw material for adaptive evolution under severe stress (Kalisz and Purugganan 2004; Rapp and Wendel 2005). Rapp and Wendel (2005) suggest that a population bottleneck, while reducing genetic diversity, might simultaneously create epigenetic novelty. In contrast to genetic alleles, epialleles

might react more quickly to environmental change, be reversible, and persist for only a few generations (Kalisz and Purugganan 2004). If a new epiallele were to cause a mild phenotype through alteration of gene expression, it might experience less strong selection than a loss-of-function sequence mutation (Kalisz and Purugganan 2004) and thus enable rapid, yet fine-tuned, trait modifications. The significance of epialleles in wild populations will depend on their frequency and stability (Rohde and Junttila 2008).

The analysis of the epigenetic processes in an ecological context, known as “ecological epigenetics” is set to transform our understanding of the way in which organisms function on the landscape. Forest trees offer excellent opportunities to examine some of the most compelling questions of ecological epigenetics (Bossdorf et al. 2008), particularly those related to the interplay between epigenetic variation and phenotypic variation in natural populations, and the role of epigenetic variation in evolutionary processes. Ecological epigenetics could readily address such questions by capitalizing on the advantageous features of forest trees, including their long lifespans, their dominance of many ecosystems, their wide geographic distribution, and their life histories, especially reproductive traits like clonal propagation. Analysis of the epigenetics of forest tree species will significantly improve our understanding of the mechanisms underlying natural phenotypic variation, and the responses of organisms to environmental change, and may thereby inform efforts to manage and breed tree species to help them cope with environmental stresses.

Acknowledgments

The preparation of this review was supported by funding from the FP7 UE projects FORESTRAC (2440096-FP7) and ProCoGen (289841-FP7).

The authors are very grateful to Susana Ferrándiz for helping with the preparation of the manuscript.

Conflict of Interest

None declared.

References

- Akimoto, K., H. Katakami, H. J. Kim, E. Ogawa, C. M. Sano, Y. Wada, et al. 2007. Epigenetic inheritance in rice plants. *Ann. Bot.* 100:205–217.
- Bastow, R., J. S. Mylne, C. Lister, Z. Lippman, R. A. Martienssen, and C. Dean. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427:164–167.
- Baurens, F. C., J. Nicolleau, T. Legavre, J. L. Verdeil, and O. Monteuiis. 2004. Genomic DNA methylation of juvenile and mature *Acacia mangium* micropropagated in vitro with reference to leaf morphology as a phase change marker. *Tree Physiol.* 24:401–407.
- Besnard, G., V. Acheré, S. Jeandroz, Ø. Johnsen, P. Faivre Rampant, R. Baumann, et al. 2008. Does maternal environmental condition during reproductive development induce genotypic selection in *Picea abies*? *Ann. For. Sci.* 65:109.
- Birchler, J. A., H. Yao, S. Chudalayandi, D. Vaiman, and R. A. Vieti. 2010. Heterosis. *Plant Cell* 22:2105–2112.
- Bogeat-Triboulot, M. B., M. Brosché, J. Renaut, L. Jouve, D. Le Thiec, P. Fayyaz, et al. 2007. Gradual soil water depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in *Populus euphratica*, a poplar growing in arid regions. *Plant Physiol.* 143:876–892.
- Bonasio, R., S. Tu, and D. Reinberg. 2010. Molecular signals of epigenetic states. *Science* 330:612–616.
- Bond, D. M., and E. J. Finnegan. 2007. Passing the message on: inheritance of epigenetic traits. *Trends Plant Sci.* 12:211–216.
- Bonhomme, L., R. Monclus, D. Vincent, S. Carpin, A. M. Lomenech, C. Plomion, et al. 2009. Leaf proteome analysis of eight *Populus x euramericana* genotypes: genetic variation in drought response and in water-use efficiency involves photosynthesis-related proteins. *Proteomics* 9:4121–4142.
- Bossdorf, O., C. L. Richards, and M. Pigliucci. 2008. Epigenetics for ecologists. *Ecol. Lett.* 11:106–115.
- Bourc’his, D., and O. Voinnet. 2010. A Small-RNA perspective on gametogenesis, fertilization, and early zygotic development. *Science* 330:617–622.
- Boyko, A., and I. Kovalchuk. 2008. Epigenetic control of plant stress response. *Environ. Mol. Mutagen.* 49:61–72.
- Cervera, M. T., L. Ruiz-Garcia, and J. M. Martínez-Zapater. 2002. Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Mol. Genet. Genomics* 268:543–552.
- Chen, P. Y., S. J. Cokus, and M. B. S. Pellegrini. 2010. Seeker: precise mapping for bisulfite sequencing. *BMC Bioinformatics* 11:203.
- Chinnusamy, V., and J. K. Zhu. 2009. Epigenetic regulation of stress responses in plants. *Curr. Opin. Plant Biol.* 12:133–139.
- Chouard, P. 1960. Vernalization and its relations to dormancy. *Annu. Rev. Plant Physiol.* 11:191–238.
- Cokus, S. J., S. Feng, X. Zhang, Z. Chen, B. Merriman, C. D. Haudenschild, et al. 2008. Shotgun bisulfite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452:215–219.
- Cubas, P., C. Vincent, and E. Coen. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401:157–161.
- De Block, M., and M. Van Lijsebettens. 2011. Energy efficiency and energy homeostasis as genetic and epigenetic

- components of plant performance and crop productivity. *Curr. Opin. Plant Biol.* 14:275–282.
- Demidov, D., S. Hesse, A. Tewes, T. Rutten, J. Fuchs, R. K. Ashtiyani, et al. 2009. Aurora1 phosphorylation activity on histone H3 and its cross-talk with other post-translational histone modifications in *Arabidopsis*. *Plant J.* 59:221–230.
- Dormling, I., and Ø. Johnsen. 1992. Effects of the parental environment on full-sib families of *Pinus sylvestris*. *Can. J. For. Res.* 22:88–100.
- Feil, R., and M. F. Fraga. 2011. Epigenetics and the environment: emerging patterns and implications. *Nat. Rev. Genet.* 13:97–109.
- Flusberg, B. A., D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, A. C. Tyson, et al. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* 7:461–465.
- Fraga, M. F., M. J. Canal, and R. Rodriguez. 2002a. Phase-change related epigenetic and physiological changes in *Pinus radiata* D. Don. *Planta* 215:672–678.
- Fraga, M. F., R. Rodriguez, and M. J. Canal. 2002b. Genomic DNA methylation-demethylation during aging and reinvigoration of *Pinus radiata*. *Tree Physiol.* 22:813–816.
- Gourcilleau, D., M. B. Bogeat-Triboulot, D. Le Thiec, C. Lafon-Placette, A. Delaunay, W. A. El-Soud, et al. 2010. DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity. *Ann. For. Sci.* 67:208.
- Grattapaglia, D., C. Plomion, M. Kirst, and R. R. Sederoff. 2009. Genomics of growth traits in forest trees. *Curr. Opin. Plant Biol.* 12:148–156.
- Greenwood, M. S., and K. W. Hutchison. 1996. Genetic after effects of increased temperature in *Larix*. Pp. 56–62 in J. Hom, R. Birdsey and K. O'Brian, eds. *Proceedings of the 1995 Meeting of the Northern Global Change Program*. USDA Forest Service report, Radnor.
- Greenwood, M. S., C. A. Hopper, and K. W. Hutchison. 1989. Maturation in larch: I. Effect of age on shoot growth, foliar characteristics, and DNA methylation. *Plant Physiol.* 90:406–412.
- Groszmann, M., I. K. Greaves, N. Albert, R. Fujimoto, C. A. Helliwell, E. S. Dennis, et al. 2011a. Epigenetics in plants: vernalisation and hybrid vigour. *Biochim. Biophys. Acta* 1809:427–437.
- Groszmann, M., I. K. Greaves, Z. I. Albertyn, G. N. Scofield, W. J. Peacock, and E. S. Dennis. 2011b. Changes in 24-nt siRNA levels in *Arabidopsis* hybrids suggest an epigenetic contribution to hybrid vigor. *Proc. Natl Acad. Sci. USA* 108:2617–2622.
- Gutierrez-Marcos, J. F., L. M. Costa, M. Dal Prà, S. Scholten, E. Kranz, P. Perez, et al. 2006. Epigenetic asymmetry of imprinted genes in plant gametes. *Nat. Genet.* 38:876–878.
- Ha, M., J. Lu, L. Tian, V. Ramachandran, K. D. Kasschau, E. J. Capman, et al. 2009. Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. *Proc. Natl Acad. Sci. USA* 106:17835–17840.
- Hamanishi, E. T., and M. M. Campbell. 2011. Genome-wide responses to drought in forest trees. *Forestry* 84:273–283.
- 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.
- Haun, W. J., S. Laouelle-Duprat, M. J. O'Connell, C. Spillane, U. Grossniklaus, A. R. Phillips, et al. 2007. Genomic imprinting, methylation and molecular evolution of maize Enhancer of zeste (Mez) homologs. *Plant J* 49:325–337.
- Hauser, M., W. Aufsatz, C. Jonak, and C. Luschnig. 2011. Transgenerational epigenetic inheritance in plants. *Biochim. Biophys. Acta* 1809:459–468.
- He, G., X. Zhu, A. A. Elling, L. Chen, X. Wang, L. Guo, et al. 2010. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* 22:17–33.
- Heo, J. B., and S. Sung. 2011. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 331:76–79.
- Herrera, C. M. 1990. The adaptedness of the floral phenotype in a relict endemic, hawkmoth-pollinated violet. 2. Patterns of variation among disjunct populations. *Biol. J. Linn. Soc.* 40:275–291.
- Herrera, C. M., and P. Bazaga. 2010. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytol.* 187:867–876.
- 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* 472:115–119.
- Jablonska, E., and G. Raz. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* 84:131–176.
- Jaligot, E., A. Rival, T. Beule, S. Dussert, and J. L. Verdeil. 2000. Somaclonal variation in oil palm (*Elaeis guineensis* Jacq.): the DNA methylation hypothesis. *Plant Cell Rep.* 19:684–690.
- Jaskiewicz, M., U. Conrath, and C. Peterhansel. 2011. Chromatin modification acts as a memory for systemic acquired resistance in the plant stress response. *EMBO Rep.* 12:50–55.
- Johannes, F., V. Colot, and R. C. Jansen. 2008. Epigenome dynamics: a quantitative genetics perspective. *Nat. Rev. Genet.* 9:883–890.
- Johannes, F., E. Porcher, F. K. Teixeira, V. Saliba-Colombani, M. Simon, N. Agier, et al. 2009. Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet.* 5:e1000530.
- Johnsen, Ø., T. Skjøppa, O. Junttila, and O. G. Dæhlen. 1996. Influence of the female flowering environment on autumn

- frost-hardiness of *Picea abies* progenies. *Theor. Appl. Genet.* 92:797–802.
- Johnsen, Ø., C. G. Fossdal, N. Nagy, J. Molmann, O. G. Dælen, and T. Skrøppa. 2005. Climatic adaptation in *Picea abies* progenies is affected by the temperature during zygotic embryogenesis and seed maturation. *Plant Cell Environ.* 28:1090–1102.
- Jullien, P. E., T. Kinoshita, N. Ohad, and F. Berger. 2006. Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *Plant Cell* 18:1360–1372.
- Kaeppler, S. M., H. F. Kaeppler, and Y. Rhee. 2000. Epigenetic aspects of somaclonal variation in plants. *Plant Mol. Biol.* 43:179–188.
- Kalisz, S., and M. D. Purugganan. 2004. Epialleles via DNA methylation: consequences for plant evolution. *Trends Ecol. Evol.* 19:309–314.
- Kelly, C. K., M. W. Chase, A. De Bruijn, M. F. Fay, and F. I. Woodward. 2003. Temperature-based population segregation in birch. *Ecol. Lett.* 6:87–89.
- Kim, D. H., and S. Sung. 2012. Environmentally coordinated epigenetic silencing of FLC by protein and long noncoding RNA components. *Curr. Opin. Plant Biol.* 15:51–56.
- Kohmann, K., and Ø. Johnsen. 1994. The timing of bud-set in seedlings of *Picea abies* from seed crops of a cool versus a warm summer. *Silvae Genetica* 43:328–332.
- Kouzarides, T. 2007. Chromatin modifications and their function. *Cell* 128:693–705.
- Krauss, V. 2008. Glimpses of evolution: heterochromatic histone H3K9 methyltransferases left its marks behind. *Genetica* 133:93–106.
- Krizova, K., M. Fojtova, A. Depicker, and A. Kovarik. 2009. Cell culture-induced gradual and frequent epigenetic reprogramming of invertedly repeated tobacco transgene epialleles. *Plant Physiol.* 149:1493–1504.
- Krueger, F., B. Kreck, A. Franke, and S. R. Andrews. 2012. DNA methylome analysis using short bisulfite sequencing data. *Nat. Methods* 9:145–151.
- Ku, C. S., N. Naidoo, M. Wu, and R. Soong. 2011. Studying the epigenome using next generation sequencing. *J. Med. Genet.* 48:721–30.
- Kumar, S. V., and P. A. Wigge. 2010. H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. *Cell* 140:136–47.
- Kvaalen, H., and Ø. Johnsen. 2008. Timing of bud set in *Picea abies* is regulated by a memory of temperature during zygotic and somatic embryogenesis. *New Phytol.* 177:49–59.
- Lafon-Placette, C., P. Faivre-Rampant, A. Delaunay, N. Street, F. Brignolas, and S. Maury. 2013. Methylome of DNase I sensitive chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach revealing characteristics of gene-body DNA methylation in open chromatin state. *New Phytol.* 197:416–430.
- Lim, S. J., T. W. Tan, and J. C. Tong. 2010. Computational epigenetics: the new scientific paradigm. *Bioinformatics* 4:331–337.
- Lira-Medeiros, C. F., C. Parisod, R. A. Fernandes, C. S. Mata, M. A. Cardoso, and P. C. Ferreira. 2010. Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS ONE* 5:e10326.
- Lisch, D. R., and R. K. Slotkin. 2011. Strategies for silencing and escape: the ancient struggle between transposable elements and their hosts. *Int. Rev. Cell Mol. Biol.* 292:119–152.
- Long, Y., W. Xia, R. Li, J. Wang, M. Shao, J. Feng, et al. 2011. Epigenetic QTL mapping in *Brassica napus*. *Genetics* 189:1093–102.
- Manning, K., M. Tör, M. Poole, Y. Hong, A. J. Thompson, G. J. King, et al. 2006. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* 38:948–952.
- Marfil, C. F., E. L. Camadro, and R. W. Masuelli. 2009. Phenotypic instability and epigenetic variability in a diploid potato of hybrid origin, *Solanum ruiz-lealii*. *BMC Plant Biol.* 9:21.
- Miguel, C., and L. Marum. 2011. An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond. *J. Exp. Bot.* 62:3713–3725.
- Mirouze, M., and J. Paszkowski. 2011. Epigenetic contribution to stress adaptation in plants. *Curr. Opin. Plant Biol.* 14:267–274.
- Mirouze, M., J. Reinders, E. Bucher, T. Nishimura, K. Schneeberger, S. Ossowski, et al. 2009. Selective epigenetic control of retrotransposition in *Arabidopsis*. *Nature* 461:427–430.
- Monteuuis, O., F. C. Baurens, D. K. S. Goh, S. Doulebeau, and J. L. Verdeil. 2009. DNA methylation in *Acacia mangium* in vitro and ex-vitro buds, in relation to their within-shoot position, age and leaf morphology of the shoot. *Silvae Genetica* 58:287–292.
- Neale, D. B., and A. Kremer. 2011. Forest tree genomics: growing resources and applications. *Nat. Rev. Genet.* 12:111–122.
- Nicotra, A. B., O. K. Atkin, S. P. Bonser, A. M. Davidson, E. J. Finnegan, U. Mathesius, et al. 2010. Plant phenotypic plasticity in a changing climate. *Trends Plant Sci.* 15:684–692.
- Ohri, D., and T. N. Khoshoo. 1986. Genome size in gymnosperms. *Plant Syst. Evol.* 153:119–132.
- Paun, O., R. M. Bateman, M. F. Fay, M. Hedren, L. Civeyrel, and M. W. Chase. 2010. Stable epigenetic effects impact adaptation in allopolyploid orchids (*Dactylorhiza: Orchidaceae*). *Mol. Biol. Evol.* 27:2465–2473.
- Phillips, R. L., S. M. Kaeppler, and P. Olhoft. 1994. Genetic instability of plant tissue cultures: breakdown of

- normal controls. *Proc. Natl Acad. Sci. USA* 91:5222–5226.
- Plomion, C., C. Lalanne, S. Claverol, H. Meddour, A. Kohler, M. B. Bogeat-Triboulot, et al. 2006. Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics* 6:6509–6527.
- Rabinowicz, P. D., R. Citek, M. A. Budiman, A. Nunberg, J. A. Bedell, N. Lakey, et al. 2005. Differential methylation of genes and repeats in land plants. *Genome Res.* 15:1431–1440.
- Raj, S., K. Bräutigam, E. T. Hamanishi, O. Wilkins, W. Schroeder, S. D. Mansfield, et al. 2011. Clone history shapes *Populus* drought responses. *Proc. Natl Acad. Sci. USA* 108:12521–12526.
- Rake, A. V., J. P. Mijsche, R. B. Hall, and K. M. Hansen. 1980. DNA reassociation kinetics of four Conifers. *Can. J. Genet. Cytol.* 22:69–79.
- Rapp, R. A., and J. F. Wendel. 2005. Epigenetics and plant evolution. *New Phytol.* 168:81–91.
- Rehfeldt, G. E., C. C. Ying, D. L. Spittlehouse, and D. A. Hamilton. 1999. Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecol. Monogr.* 69:375–407.
- Rehfeldt, G. E., N. M. Tchebakova, Y. I. Parfenova, W. R. Wykoff, N. A. Kuzmina, and L. I. Milyutin. 2002. Intraspecific responses to climate in *Pinus sylvestris*. *Glob. Change Biol.* 8:912–929.
- Reinders, J., B. B. H. Wulff, M. Mirouze, A. Mari-Ordonez, M. Dapp, W. Rozhon, et al. 2009. Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes. *Genes Dev.* 23:939–950.
- Richards, E. J. 2006. Inherited epigenetic variation – revisiting soft inheritance. *Nat. Rev. Genet.* 7:395–401.
- Richards, C. L., K. J. F. Verhoeven, and O. Bossdorf. 2012. Evolutionary significance of epigenetic variation. Pp. 257–274 in J. F. Wendel, J. Greilhuber, J. Dolezel and I. J. Leitch, eds. *Plant genome diversity Volume I*. Springer, Vienna.
- Rival, A., E. Jaligot, T. Beule, and E. J. Finnegan. 2008. Isolation and expression analysis of genes encoding MET, CMT, and DRM methyltransferases in oil palm (*Elaeis guineensis* Jacq.) in relation to the ‘mantled’ somaclonal variation. *J. Exp. Bot.* 59:3271–3281.
- Rodriguez Lopez, C. M., A. C. Wetten, and M. J. Wilkinson. 2010. Progressive erosion of genetic and epigenetic variation in callus-derived cocoa (*Theobroma cacao*) plants. *New Phytol.* 186:856–868.
- Rohde, A. 2009. Bud set – a landmark of the seasonal growth cycle in poplar. Pp. 91–98 in L. V. Gusta, M. E. Wisniewski and K. K. Tanino, eds. *Plant Cold Hardiness*. CABI, Cambridge, MA.
- Rohde, A., and O. Junttila. 2008. Remembrances of an embryo: long-term effects on phenology traits in spruce. *New Phytol.* 177:2–5.
- Roudier, F., I. Ahmed, C. Bérard, A. Sarazin, T. Mary-Huard, S. Cortijo, et al. 2011. Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. *EMBO J.* 30:1928–1938.
- Ruttink, T., M. Arend, K. Morreel, V. Storme, S. Rombauts, J. Fromm, et al. 2007. A molecular timetable for apical bud formation and dormancy induction in poplar. *Plant Cell* 19:2370–2390.
- Santamaria, M. E., R. Hasbun, and M. J. Valera. 2009. Acetylated H4 histone and genomic DNA methylation patterns during bud set and bud burst in *Castanea sativa*. *J. Plant Physiol.* 166:1360–1369.
- Schellenbaum, P., V. Mohler, G. Wenzel, and B. Walter. 2008. Variation in DNA methylation patterns of grapevine somaclones (*Vitis vinifera* L.). *BMC Plant Biol.* 8:78.
- Schmitz, R. J., and R. M. Amasino. 2007. Vernalization: a model for investigating epigenetics and eukaryotic gene regulation in plants. *Biochim. Biophys. Acta* 1769: 269–275.
- Skørppa, T., and Ø. Johnsen. 2000. Patterns of adaptive genetic variation in forest tree species; the reproductive environment as an evolutionary force in *Picea abies*. Pp. 49–58 in C. Mátyás, ed. *Forest Genetics and Sustainability*. Kluwer Academic Publications, Dordrecht, The Netherlands.
- Skørppa, T., K. Kohmann, Ø. Johnsen, A. Steffenrem, and Ø. M. Edvardsen. 2007. Field performance and early test results of offspring from two Norway spruce seed orchards containing clones transferred to warmer climates. *Can. J. For. Res.* 37:515–522.
- Skørppa, T., M. Tollefsrud, C. Sperisen, and Ø. Johnsen. 2010. Rapid change in adaptive performance from one generation to the next in *Picea abies*-Central European trees in a Nordic environment. *Tree Genet. Genomes* 6:93–99.
- Slotkin, R. K., and R. A. Martienssen. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* 8:272–285.
- Stoeck, M. U., S. J. L'Hirondelle, W. D. Binder, and J. E. Webber. 1998. Parental environment after effects on germination, growth, and adaptive traits in selected spruce families. *Can. J. For. Res.* 28:418–426.
- Thumma, B. R., B. A. Matheson, D. Zhang, C. Meeske, R. Meder, G. M. Downes, et al. 2009. Identification of a Cis-acting regulatory polymorphism in a Eucalypt COBRA-like gene affecting cellulose content. *Genetics* 183:1153–1164.
- Tsukahara, S., A. Kobayashi, A. Kawabe, O. Mathieu, A. Miura, and T. Kakutani. 2009. Bursts of retrotransposition reproduced in *Arabidopsis*. *Nature* 461:423–426.

- Valledor, L., R. Hasbún, and M. Meijón. 2007. Involvement of DNA methylation in tree development and micropropagation. *Plant Cell Tiss. Org. Cult.* 91:75–86.
- Valledor, L., M. Meijon, R. Hasbun, M. Jesus Canal, and R. Rodriguez. 2010. Variations in DNA methylation, acetylated histone H4, and methylated histone H3 during *Pinus radiata* needle maturation in relation to the loss of in vitro organogenic capability. *J. Plant Physiol.* 167:351–357.
- Van Steensel, B. 2011. Chromatin: constructing the big picture. *EMBO J* 30:885–1895.
- 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.
- Vining, K. J., K. R. Pomraning, L. J. Wilhelm, H. D. Priest, M. Pellegrini, T. C. Mockler, et al. 2012. Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* 13:27.
- Viswanath, V., B. R. Albrechtsen, and S. H. Strauss. 2012. Global regulatory burden for field testing of genetically modified trees. *Tree Genet. Genomes* 8:221–226.
- Webber, J., P. Ott, J. Owens, and W. Binder. 2005. Elevated temperature during reproductive development affects cone traits and progeny performance in *Picea glauca* - *engelmannii* complex. *Tree Physiol.* 25:1219–1227.
- Whittle, C. A., S. P. Otto, M. O. Johnston, and J. E. Krochko. 2009. Adaptive epigenetic memory of ancestral temperature regime in *Arabidopsis thaliana*. *Botany* 87:650–657.
- Wilkins, O., L. Waldron, H. Nahal, N. J. Provart, and M. M. Campbell. 2009. Genotype and time of day shape the *Populus* drought response. *Plant J.* 60:703–715.
- Yakovlev, I. A., C. G. Fossdal, and Ø. Johnsen. 2010. MicroRNAs, the epigenetic memory and climatic adaptation in Norway spruce. *New Phytol.* 187:1154–1169.
- Yakovlev, I. A., D. K. A. Asante, C. G. Fossdal, O. Junttila, and Ø. Johnsen. 2011. Differential gene expression related to an epigenetic memory affecting climatic adaptation in Norway spruce. *Plant Sci.* 180:132–139.

2. Publications et communications

Publications dans des revues internationales à comité de lecture

Lafon-Placette C, Faivre-Rampant P, Street N, Delaunay A, Brignolas F and Maury S (2012). Methylome of sensitive to DNase I chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach to study gene-body DNA methylation in chromatin state. *Accepted in New Phytologist*.

Bräutigam K, Vining KJ, **Lafon-Placette C**, Fossdal C, Mirouze M, Marcos JG, Fluch S, Fraga MF, Guevara MA, Abarca DJ, Øystein, Maury S, Strauss SH, Campbell M, Rohde A, Díaz-Sala C, Cervera MT (2012). Epigenetic regulation of adaptive response of forest tree species to the environment. *Ecology and Evolution (minor revision)*.

Hébrard C, Trap-Gentil MV, **Lafon-Placette C**, Delaunay A, Joseph C, Lefebvre M, Barnes S and Maury S (2012). Identification of epialleles during vernalization unravels a control of bolting tolerance by RNA methylation. *Journal of Experimental Botany (revision)*.

Trap-Gentil MV, Hébrard C, **Lafon-Placette C**, Delaunay A, Hagège D, Joseph C, Brignolas F, Maury S (2011). Time course and amplitude of DNA methylation in shoot apex are critical points for bolting induction in sugar beet and bolting tolerance between genotypes. *J Exp Bot* 62 (8): 2585-2597.

Gourcilleau D, Bogeat-Triboulot MB, Le Thiec D, **Lafon-Placette C**, Delaunay A, El-Soud WA, Brignolas F, Maury S (2010). DNA methylation and histone acetylation: genetic variations in hybrid poplars, impact of water deficit and relationships with productivity. *Ann For Sci* 67:208.

Article en préparation

Lafon-Placette C, Hummel I, Bogeat-Triboulot MB, Delaunay A, Le Thiec D, Brignolas F and Maury S (2012). Transcriptomic and epigenomic changes in the shoot apex mediate shoot developmental plasticity of *P. x euramericana* hybrid in response to variations in water availability.

Communications orales à des congrès (le nom de l'auteur ayant présenté est souligné)

Maury S, **Lafon-Placette C**, Faivre-Rampant P, Street N, Delaunay A, Brignolas F (2012). Methylome of DNase I sensitive chromatin — A simplified approach revealing characteristics of gene-body DNA methylation in *Populus trichocarpa* open chromatin state. *Cold Spring Harbor Conference 'Plant Epigenetics, Stress & Evolution', Suzhou, Chine, 29 octobre – 2 novembre 2012*.

Lafon-Placette C, Delaunay A, Brignolas F, Maury S (2010). Méthylome et réponse à la sécheresse. *Biotechnocentre 23^{ème} colloque, Seillac, France, 21-22 octobre 2010*.

Lafon-Placette C, Gourcilleau D, Chamaillard S, Fichot R, Bogeat-Triboulot MB, Le Thiec D, Delaunay A, Villar M, Brignolas F, Maury S (2010). Epigenetics, water deficit and productivity in poplars. *Vth International Poplar Symposium, Orvieto, Italie, 20-25 septembre 2010*.

Maury S, Hébrard C, **Lafon-Placette C**, Delaunay A, Joseph C, Brignolas F (2009). Les marqueurs épigénétiques : déterminants de processus physiologiques et outils de sélection de traits d'intérêt pour la production ou multiplication végétative. *Journée scientifique Association des Sélectionneurs Français, Cappèle-en-Pévèle, France, 15 octobre 2009*.

Lafon-Placette C, Gourcilleau D, Delaunay A, Joseph C, Brignolas F, Maury S (2009). Relations entre les marqueurs épigénétiques et la production de biomasse chez le peuplier. *Journée scientifique de l'Association des Sélectionneurs Français, Cappèle-en-Pévèle, France, 15 octobre 2009*.

Lafon-Placette C, Gourcilleau D, Bogeat-Triboulot MB, Le Thiec D, Delaunay A, Brignolas F, Maury S (2008). Genetic variations and drought impact on DNA methylation and histone acetylation in hybrid poplars. *Plant Epigenetics Workshop: epigenetic control of gene expression in plants, Montpellier, France, 8-10 décembre 2008*.

Communications par affiche à des congrès

Lafon-Placette C, Faivre-Rampant P, Street N, Delaunay A, Brignolas F, Maury S (2011). Methylome of non-condensed chromatin in poplar: a new approach to study relationships between DNA methylation and chromatin organization. *ESF-EMBO Symposium 'Epigenetics in Context: From Ecology to Evolution', Sant Feliu de Guixols, Espagne, 18-23 septembre 2011*.

Maury S, **Lafon-Placette C**, Faivre-Rampant P, Street N, Delaunay A, Brignolas F (2012). Epigenomics in poplar: the methylome of non-condensed chromatin. *26th New Phytologist Symposium 'Bioenergy Trees', Nancy, France, 17-19 mai 2011*.

Lafon-Placette C, Gourcilleau D, Fichot R, Delaunay A, Brignolas F, Maury S (2010). Etude de déterminants épigénétiques de la production de biomasse chez le peuplier et sa plasticité en réponse à la sécheresse. *2^{ème} Forum ResoNat / DREAM, BRGM Orléans, France, 14 décembre 2010*.

Lafon-Placette C, Delaunay A, Brignolas F, Maury S (2010). Variations du méthylome en réponse à la sécheresse. *Biotechnocentre 23^{ème} colloque, Seillac, France, 21-22 octobre 2010*.

Lafon-Placette C, Delaunay A, Brignolas F, Maury S (2010). DNA methylation as a component of poplar phenotypic plasticity towards water availability. *Vth International Poplar Symposium, Orvieto, Italie, 20-25 septembre 2010*.

Contrôle épigénétique de la plasticité de l'appareil végétatif du peuplier en réponse à des variations de la disponibilité en eau

Au vu de l'impact croissant du changement climatique global et en particulier de la sécheresse sur les forêts, il est nécessaire de comprendre les mécanismes de réponse des arbres face à des variations de disponibilité en eau. Ces dernières années, des études ont montré un contrôle épigénétique et notamment par la méthylation de l'ADN de la plasticité phénotypique des plantes en réponse aux variations environnementales.

Dans ce contexte, cette thèse visait à évaluer le rôle de la méthylation de l'ADN des cellules du méristème apical caulinaire dans la plasticité développementale de la tige feuillée en réponse à des variations de disponibilité en eau chez le peuplier, un arbre modèle. A cette fin, le méthylome de la chromatine non condensée dans le méristème apical caulinaire de *Populus trichocarpa* a été caractérisé. Ensuite, l'impact de variations de disponibilité en eau sur la méthylation de l'ADN a été étudié dans l'apex caulinaire de différents hybrides (*P. × euramericana*). Les loci et les réseaux de gènes affectés pour leur expression et leur méthylation ont ainsi été identifiés.

Ces travaux ont montré que dans le méristème apical caulinaire, la majorité des gènes étaient dans un état non condensé de la chromatine et méthylés dans leur corps. Ils ont également mis en évidence une forte variation de la méthylation globale de l'ADN selon les génotypes et en réponse à des variations de disponibilité en eau. De plus, des corrélations ont été établies entre les niveaux de croissance des arbres et de méthylation globale de l'ADN dans l'apex caulinaire. Enfin, les variations de la méthylation de l'ADN en réponse aux variations de la disponibilité en eau s'accompagnent de variations d'expression et ont ciblé particulièrement des gènes impliqués dans la signalisation par les phytohormones ou la morphogenèse.

Ainsi, les travaux effectués lors de cette thèse suggèrent un rôle de la méthylation de l'ADN dans la plasticité phénotypique en réponse à des variations de disponibilité en eau chez le peuplier *via* le contrôle de l'expression de réseaux de gènes dans le méristème apical caulinaire.

Mots clés : chromatine, disponibilité en eau, épigénétique, expression de gènes, méristème apical caulinaire, méthylation de l'ADN, morphogenèse, peuplier, plasticité phénotypique, sécheresse, variabilité génotypique

Epigenetic control of shoot phenotypic plasticity towards variations in water availability in poplar

Predicted climate changes and particularly drought represent a major threat to forest health. Therefore, understanding mechanisms that control trees response to variations in water availability is of great interest. These last years, epigenetic marks such as DNA methylation have been involved in plant phenotypic plasticity in response to environmental stresses.

In this context, this work aimed at assessing the role of shoot apical meristem cells DNA methylation in the shoot developmental plasticity towards variations in water availability in poplar, a model tree. For this purpose, the methylome of non condensed chromatin in *Populus trichocarpa* shoot apical meristem was characterized. Then, the impact of variations in water availability on shoot apex DNA methylation in different hybrids (*P. × euramericana*) was studied. Loci and gene networks affected by DNA methylation and expression changes were thus identified.

This work showed that in shoot apical meristem, most of the genes was in non condensed chromatin state with DNA methylation in their body. A strong variation in DNA methylation depending on genotypes and water availability was highlighted. Moreover, correlations between trees growth and shoot apex DNA methylation levels were established. Lastly, DNA methylation changes in response to variations in water availability correlated to expression variations were identified for genomic loci and gene networks.

Thus, the work performed during this thesis suggests a role for DNA methylation in poplar phenotypic plasticity in response to variations in water availability through the control of gene networks transcription in the shoot apical meristem.

Keywords: chromatin, DNA methylation, drought, epigenetics, gene expression, genotypic variability, morphogenesis, poplar, phenotypic plasticity, shoot apical meristem, water availability