



HAL
open science

Rôles des variations épigénétiques transgénérationnelles dans la résistance quantitative à la hernie chez *Arabidopsis thaliana*

Benjamin Liegard

► **To cite this version:**

Benjamin Liegard. Rôles des variations épigénétiques transgénérationnelles dans la résistance quantitative à la hernie chez *Arabidopsis thaliana*. Génétique des plantes. Agrocampus Ouest, 2018. Français. NNT : 2018NSARC137 . tel-02303309

HAL Id: tel-02303309

<https://theses.hal.science/tel-02303309v1>

Submitted on 2 Oct 2019

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

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

THESE DE DOCTORAT DE

AGROCAMPUS OUEST
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 600
Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation
Spécialité : *Génétique, génomique et bio-informatique*

Par

Benjamin LIEGARD

**Rôle des variations épigénétiques transgénérationnelles dans la
résistance quantitative à la hernie chez *Arabidopsis thaliana***

Thèse présentée et soutenue à Rennes, le 9 novembre 2018
Unité de recherche : Institut de Génétique, Environnement et Protection des Plantes
UMR 1349 INRA - AGROCAMPUS OUEST - Université de Rennes 1 - 35653 LE RHEU

Rapporteurs avant soutenance :

Laurent DESLANDES Directeur de recherche, CNRS, INRA de Toulouse
Stéphane MAURY Professeur, Université d'Orléans

Composition du Jury :

Président :	Sandrine LAGARRIGUE	Professeur, AGROCAMPUS OUEST
Examineurs :	Clémentine VITTE	Chargé de recherche, CNRS, INRA du Moulon
Dir. de thèse :	Maria MANZANARES-DAULEUX	Professeur, AGROCAMPUS OUEST, INRA du Rheu
Co-dir. de thèse :	Mélanie JUBAULT	Maitre de conférences, AGROCAMPUS OUEST, INRA du Rheu

LISTE DES TABLEAUX

CHAPITRE I

Tableau I.1.	Exemples de gènes causaux identifiés sous des QTL de résistance.	9
Tableau I.2.	Liste des modifications présentes au niveau des histones et listes des gènes encodant les protéines intervenant dans deux modifications d’histones.	14
Tableau I.3.	Gammes d’hôtes différentiels utilisées pour l’identification et la classification des différents pathotypes de <i>Plasmodiophora brassicae</i>	36

CHAPITRE II

Tableau II.1.	Liste des mutants d’ <i>Arabidopsis thaliana</i> utilisés au cours de la thèse.	49
Tableau II.2.	Liste des accessions naturelles d’ <i>Arabidopsis thaliana</i> utilisées.	51
Tableau II.3.	Caractérisation de huit pathotypes de <i>P. Brassicae</i> en utilisant la gamme d’hôtes différentiels décrite par Somé <i>et al.</i> (1996).....	52
Tableau II.4.	Différentiation des isolats de pathotype 1 par addition de deux génotypes de <i>B. oleracea</i> ssp. <i>Acephala</i> . 52	
Tableau II.5.	Liste des amorces de CHOP PCR, CHOP qPCR et RT-qPCR utilisées pour l’étude de la méthylation de l’ADN par McrBC et de l’expression des gènes.	59

CHAPITRE III

Tableau III.1.	Heritability was calculated for each trait in the epiRIL population.	72
Tableau III.2.	Summary of QTL ^{epi} detected in the epiRIL population by MQM.....	73
Tableau III.3.	Intervalles des régions étudiées et marqueurs utilisés.	83
Tableau III.4.	Sélection des epiRIL recombinantes.	84
Tableau III.5.	Liste des 30 marqueurs utilisés dans la carte épigénétique des recombinants.	85
Tableau III.6.	Tableau récapitulatif des intervalles de confiance des QTL ^{epi} détectés dans les régions sélectionnées.	86

CHAPITRE IV

Tableau IV.1.	Liste des gènes candidats potentiels sous la région en interaction épistatique avec le QTL <i>Pb-At5.2</i>	112
---------------	---	-----

LISTE DES FIGURES

CHAPITRE I

Figure I.1.	Modèle en zigzag illustrant les axes de résistance chez les plantes.	3
Figure I.2.	Reconnaissance des PAMP et DAMP par les protéines PRR menant à l'activation de voies de défense.	4
Figure I.3.	Configuration intronique / exonique et motifs protéiques de gènes NBS-LRR chez <i>Arabidopsis</i>	5
Figure I.4.	Diversité du nombre de gènes codant des NLR chez les plantes.	5
Figure I.5.....	Récapitulatif des divers mécanismes de détection des agents pathogènes par les NLR.	6
Figure I.6.	Comparaison de la résistance quantitative et de la résistance qualitative.	7
Figure I.7.	Schématisation des mécanismes de résistances qualitatives et quantitatives.	8
Figure I.8.	Représentation de Waddington (1957) de son paysage épigénétique.	10
Figure I.9.	Représentation de la structure de la chromatine.....	10
Figure I.10.	Présentation des 4 états chromatiniens identifiés au niveau de l'épigénome d' <i>Arabidopsis</i>	11
Figure I.11.	Définition et localisation au niveau des éléments du génome des 9 états chromatiniens identifiés par Sequeira-Mendes <i>et al.</i> (2014).	12
Figure I.12.	Présentation des variants H2A et localisation préférentielle des variants d'histones au niveau du génome.	13
Figure I.13.	Schéma représentant les lysines des histones H3 et H4 ciblées par les HKMT chez <i>Arabidopsis</i>	15
Figure I.14.	Représentation graphique de la cytosine méthylée et niveau de méthylation des cytosines chez différents eucaryotes.	15
Figure I.15.	Mécanisme de méthylation de l'ADN médiée par la voie RdDM (RNA-directed DNA methylation) chez <i>Arabidopsis thaliana</i>	17
Figure I.16.	Mécanismes de maintien de la méthylation de l'ADN.....	18
Figure I.17.	Méthylation de l'ADN du génome d' <i>Arabidopsis</i>	19
Figure I.18.	Schématisation des motifs de méthylation de l'ADN génique chez les plantes.....	20
Figure I.19.	Les différents types d'éléments transposables (TE).....	21
Figure I.20.	Mécanisme de méthylation des TE.	22
Figure I.21.	Phénomène de spreading au niveau de l'insertion d'un élément transposable.....	23
Figure I.22.	Les différentes classes d'épiallèles possibles.	23
Figure I.23.	Conséquences phénotypiques de différents épiallèles naturels chez les plantes.	24

Figure I.24.	Variations morphologiques observées chez des <i>Arabidopsis</i> sauvages et chez des <i>Arabidopsis</i> altérées au niveau de <i>DDM1</i> et <i>MET1</i>	25
Figure I.25.	Schéma de la construction de la population épiRIL <i>ddm1-2</i> x Col-0.....	26
Figure I.26.	Schéma de la construction de la population épiRIL <i>met1-3</i> x Col-0.....	27
Figure I.27.	Variations phénotypiques observées dans les populations épiRIL <i>ddm1-2</i> x Col-0 et <i>met1-3</i> x Col-0.	28
Figure I.28.	Carte de recombinaison des épiRIL <i>ddm1-2</i> x Col-0.	29
Figure I.29.	Distribution mondiale et française de la hernie des crucifères.	33
Figure I.30.	Cycle de vie de <i>Plasmodiophora brassicae</i>	34
Figure I.31.	Photos des symptômes observables au niveau de la partie racinaire et aérienne de <i>Brassica napus</i> lors de son infection par <i>P. brassicae</i>	37
Figure I.32.	Conséquences du traitement de l'accession Columbia d' <i>Arabidopsis</i> avec du méthyl jasmonate sur le développement de la hernie.....	38
Figure I.33.....	Conséquences de la mutation du gène <i>ARGAH2</i> sur la réponse d' <i>Arabidopsis</i> à la hernie.	39
Figure I.34.	QTL de résistance à la hernie détectés chez <i>Brassica oleracea</i> dans les populations F2/3 C10 x HDEM.	43
Figure I.35.	Localisation sur la carte génétique du colza (Lombard et Delourme, 2001) des gènes majeurs <i>Pb-Bn1</i> , <i>Pb-Bn2</i> et <i>Pb-Bn3</i> et des QTL additifs de résistance à la hernie identifiés dans les populations en ségrégation Darmor-bzh x Yudal et Stellar x Drakkar.	44
Figure I.36.	QTL additifs et épistatiques contrôlant la résistance partielle à <i>P. brassicae</i> dans la population Bur-0 x Col-0 RIL.	45

CHAPITRE II

Figure II.1.	Phénotype observé au niveau de la racine de plantes inoculées avec l'isolat eH de <i>P. brassicae</i>	48
Figure II.2.	Schématisation de l'obtention des lignées HIF recombinantes à partir des écotypes Col-0 et Bur-0.	50
Figure II.3.	Conditions de culture des plantes lors des tests pathologiques.	52
Figure II.4.	Illustration des symptômes racinaires présents pour chacune des classes de maladie définies permettant le calcul de l'indice pathologique.....	53
Figure II.5.	Carte épigénétique de la population épiRIL.....	53
Figure II.6.	Analyse du statut de méthylation des régions différenciellement méthylées.....	58

CHAPITRE III

Figure III.1. Clubroot symptoms shown by Col-0 wild-type and the <i>ddm1</i> mutant (Col-0 background).	70
Figure III.2. Distribution of phenotypic data measured in the two growth rooms used for pathological tests.....	71
Figure III.3. Epigenetic QTL identified in the epiRIL population in response to <i>Plasmodiophora brassicae</i> infection.....	74
Figure III.4. Localisation des régions 3 sélectionnées.....	83
Figure III.5. Sélection des épiRIL recombinantes, épihaplotypes et distribution des traits phénotypiques étudiés.....	84
Figure III.6. Epihaplotype des épiRIL sur 30 DMR et profils des QTL ^{épi} détectés.....	85
Figure III.7. Effet de l'épiallèle au marqueur au pic des QTL ^{épi}	86
Figure III.8. Schéma récapitulatif de la réduction des régions A et B.....	86

CHAPITRE IV

Figure IV.1. Données génomiques et de méthylation de la région de <i>RPB1</i>	90
Figure IV.2. Etat de méthylation de trois DMR candidates au niveau de la région <i>RPB1</i> et niveau d'expression des gènes de la région.....	91
Figure IV.3. Couverture génomique et état de méthylation de la région <i>RPB1</i> pour des accessions naturelles d' <i>Arabidopsis</i>	92
Figure IV.4. Fine mapping of <i>PbAt5.2</i>	97
Figure IV.5. Identification of the two candidates under the <i>Pb-At5.2</i> region.....	98
Figure IV.6. Natural variations at the <i>PbAt5.2</i> locus.....	99
Figure IV.7. Methylation profiles and disease evaluation on <i>Arabidopsis thaliana</i> accessions. ...	100
Figure IV.8. Genomic and epigenetic characterisation of <i>Pb-At5.2</i> epimutation.	101
Figure IV.9. Plot de la valeur de l'indice pathologique en fonction de l'état allélique au QTL <i>Pb-At5.2</i> et de la région épistatique sur le chromosome 1.....	112

LISTE DES ABREVIATIONS

6mA	N6 methyladenine
Ac	Activator
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
ADR1	Activated disease resistance 1
AGO	Argonaute
ARGAH2	Arginine amidohydrolase 2
ARN	Acide ribonucléique
AtFOLT	Folate transporter
ATX	<i>Arabidopsis</i> trithorax like proteins
Avr	Avirulence
BC	Back-cross
CBP	cAMP-responsive element-binding protein
CDS	Coding sequence
CLF	Curly leaf
Clk	clark kent
CMT	Chromomethylase
CmWIP1	ORF3 C. melo WIP1
CNL	CC-NLR
Cnr	Colorless non-ripenning
CNV	Copy number variant
CPM	Count per million
CR	Clubroot resistance
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DAMP	Damage associated molecular patterns
DCL	Dicer like
ddc	drm1 drm2 cmt3 triple mutant
DDM1	Decrease in DNA methylation 1
DME	Transcriptional activator demeter
DML	Demeter like
DMP	Differentially methylated position
DMR	Differentially methylated regions
DNAseq	DNA sequencing
DNMT	DNA methyltransferase
DNMT1	DNA methyltransferase 1
DRM2	Domains rearranged methyltransferase 2
dsRNA	Double stranded RNA
ECD	European clubroot differential
EMS	Ethyl methanesulfonate

epi-MA	Lignées accumulatrices d'épimutation
épiRIL	Epigenetic recombinant inbred line
ETI	Effector triggered immunity
ETS	Effector-triggered susceptibility
FLC	Flowering locus c
FWA	Flowering wageningen
gbM	Gene body methylation
GEO	Gene Expression Omnibus
GNAT	General control nondepressible 5 (GCN5) N-acetyl transferase
GWAS	Genome wide association studies
HAT	Histone acetyltransferase
HD2	Histone deacetylase 2
HDA1	Histone deacetylase 1
HDAC	Histone desacétylase
HIF	Heterogenous inbred families
HISN6	Histifine biosynthesis 6
HKMT	histone lysine methyltransferase
HMM	Hidden Markov model
HR	Hypersensitive response
IBM1	Increased in bonsai methylation 1
IGV	Integrative Genomics Viewer
IM	Intermediate methylation
IP	Indice pathologique
KYP	Kryptonite
Lf	Leaf length
LOX	Lipoxygenase
LRR	Leucine rich reapeats
M	Methylated
MA lines	Lignées accumulatrices de mutation
MAF	Minor allele frequency
MEA	Medea
MeDIP-chip	Methylated DNA immunoprecipitation microarray
MET1	methyltrasnferase 1
miRNA	MicroRNA
MYST	MOZ-YBF2/SAS3-SAS2/TIP60
NASC	Nottingham <i>Arabidopsis</i> Stock Centre
NATA1	N-acetylaminotransferase activity 1
NB	Nucleotide binding
NBS-LRR	Nucleotide-binding site-leucine rich reapeats
ncRNA	Non coding RNA
NGS	Next-generation sequencing
NLR-ID	Nucleotide-binding site-leucine rich reapeats-ID

NMR19-4	Naturally occurring DNA methylation variation region 19-4
NRG1	N requirement gene 1
NRP	NAP1-related protein
NRPE1	Nuclear RNA polymerase E1
P	Pathotype
PAI	Phosphoribosylanthanilate isomerase
PAMP	Pathogen-associated molecular patterns
PCR	Polymerase chain reaction
Pol II	RNA polymerase 2
Pol IV	RNA polymerase IV
PolV	RNA polymerase V
PPH	Pheophorbide hydrolase
PR	Pathogenesis Related
PRC2	Polycomb complex 2
PRR	Pattern recognition receptors
PTGS	Post transcriptional gene silencing
PTI	Patterns triggered immunity
QDR	Quantitative disease resistance
qPCR	Quantitative polymerase chain reaction
QQS	Qua-quine starch
QTL	Quantitative trait loci
QTL ^{épi}	Quantitative trait loci under epigenetic control
Rb	Root biomass
Rdd	ros1 dml2 dml3 triple mutant
RdDM	RNA directed DNA methylation
RDR2	RNA dependent RNA polymerase 2
RDR6	RNA dependent RNA polymerase 6
RIL	Recombinant Inbred Line
RIN	RNA integrity number
RLK	Receptor like kinase
RLP	Receptor like protein
RMG1	Resistance methylated gene 1
RNAseq	RNA sequencing
ROS	Reactive oxygen species
ROS1	Repressor of silencing 1
RPB1	Resistance to <i>Plasmodiophora brassicae</i> 1
RPD3	Reduced potassium dependency 3
RPW8	Resistance to powdery mildew 8
SAM	S-adenosinemethionine
SAR	Systemic acquired resistance
SDG	SET domain group
SET	Suppressor of variegation, enhancer of zeste and trithorax

Sga	Anti- Silencing Function 1
SIM	Simple interval mapping
SINE	Short interspersed nuclear elements
SIR2	Silent information regulator 2
siRNA	Small interfering RNA
snoRNA	Small nucleolar RNA
SNP	Single nucleotide polymorphisms
Spm	Supressor mutator
ssRNA	Single strand RNA
SUP	Superman
SUVH	SU(VAR)3-9 homolog
SUVR	SU(VAR)3-9 related
SWN	Swinger
TAF _μ 250	Transcription initiation factor TFIID 250 kDA subunit
TE	Transposable element
TE mRNA	Transposable element messenger RNA
THI2.1	Thionin 2.1
TIR	Transport inhibitor response
TMM	Trimmed mean of M-values
TNL	TIR-NLR
TSS	Transcription start site
TTS	Transcription termination site
U	Unmethylated
UM	Unmethylated
UTR	Untranslated region
UV	Ultra violet
VIM	Variant in methylation
VTE3(1)	Vitamin E 3
WT	Wild type

TABLE DES MATIERES

INTRODUCTION GENERALE	1
CHAPITRE I. SYNTHÈSE BIBLIOGRAPHIQUE.....	3
1. LES VOIES DE DEFENSE CHEZ LES VEGETAUX.....	3
1.1. <i>L'interaction hôte incompatible : cas de la résistance qualitative</i>	<i>3</i>
1.2. <i>L'interaction hôte compatible : cas de la résistance partielle ou résistance quantitative..</i>	<i>6</i>
2. L'IMPLICATION DE L'ÉPIGÉNÉTIQUE DANS LA RÉPONSE DES PLANTES AUX STRESS	10
2.1. <i>Définition générale de l'épigénétique.....</i>	<i>10</i>
2.2. <i>Épigénétique et structure de la chromatine</i>	<i>10</i>
2.3. <i>Les variants et les modifications post traductionnelles histones</i>	<i>12</i>
2.3.1. Les variants d'histones	12
2.3.2. Les modifications post traductionnelles des histones.....	13
2.4. <i>La méthylation de l'ADN</i>	<i>15</i>
2.4.1. Mise en place de la méthylation de l'ADN	16
2.4.2. Maintien de la méthylation de l'ADN	17
2.4.3. La déméthylation de l'ADN.....	18
2.4.4. Les profils de méthylation de l'ADN au niveau du génome.....	19
2.4.4.1. La méthylation de l'ADN au niveau des gènes.....	19
2.4.4.2. La méthylation de l'ADN au niveau des éléments transposables	20
2.4.5. Stabilité de la méthylation de l'ADN : l'épigénétique transgénérationnelle	22
2.4.5.1. Épiallèles naturels	24
2.4.5.2. Épiallèles induits et populations épigénétiques.....	25
2.4.5.2.1. Construction et caractérisation des populations épigénétiques.....	26
2.4.5.2.2. Populations épigénétiques et variations phénotypiques.....	27
2.4.6. Méthylation de l'ADN et réponses de la plante aux stress.....	29
2.4.6.1. Cas des stress abiotiques	29
2.4.6.2. Cas des stress biotiques.....	31
3. LE MODÈLE D'INTERACTION PLANTE PATHOGÈNE : <i>PLASMODIOPHORA BRASSICAE</i> - <i>BRASSICACEAE</i>	33
3.1. <i>Plasmodiophora brassicae, agent causal de la hernie des Crucifères</i>	<i>33</i>
3.1.1. Contexte et enjeu économique de la maladie	33
3.1.2. Biologie de <i>Plasmodiophora brassicae</i>	33
3.1.3. Cycle de développement de <i>Plasmodiophora brassicae</i>	34
3.1.4. Facteurs influençant le développement de la hernie.....	35
3.1.5. Les différents pathotypes de <i>Plasmodiophora brassicae</i> et leur caractérisation.....	35

3.2.	<i>Impact de Plasmodiophora brassicae sur la plante hôte</i>	36
3.2.1.	Modifications morphologiques	36
3.2.2.	Modifications hormonales et métaboliques	37
3.2.2.1.	Les modifications hormonales au sein de la plante.....	37
3.2.2.2.	Les modifications métaboliques au sein de la plante	39
3.3.	<i>La lutte contre Plasmodiophora brassicae</i>	40
3.3.1.	Les pratiques culturales	40
3.3.2.	La lutte chimique	41
3.3.3.	La lutte biologique.....	41
3.3.4.	La lutte génétique	41
3.4.	<i>Les résistances identifiées chez la plante hôte en réponse à P. Brassicae</i>	42
3.4.1.	Les résistances identifiées chez <i>B. rapa</i>	42
3.4.2.	Les résistances identifiées chez <i>B. oleracea</i>	43
3.4.3.	Les résistances identifiées chez <i>B. napus</i>	44
3.4.4.	Les résistances identifiées chez <i>A. thaliana</i>	45
4.	OBJECTIFS ET STRATEGIES DE LA THESE	46
	CHAPITRE II. MATERIELS ET METHODES	48
1.	MATERIELS VEGETAL ET PARASITAIRE	48
1.1.	<i>Arabidopsis thaliana</i>	48
1.1.1.	L'écotype sensible Col-0.....	48
1.1.2.	Les mutants d' <i>Arabidopsis thaliana</i>	48
1.1.3.	La population épiRIL	48
1.1.4.	L'écotype partiellement résistant Bur-0.....	49
1.1.5.	Les génotypes appartenant à des lignées quasi isogéniques HIF (Heterogenous Inbred Families)	49
1.1.6.	Autres écotypes d' <i>Arabidopsis</i> utilisés.....	50
1.2.	<i>Gamme d'hôtes différentiels</i>	50
1.3.	<i>Matériel parasitaire</i>	51
1.3.1.	Présentation des isolats de <i>P. Brassicae</i> utilisés.....	51
1.3.2.	Préparation de l'inoculum	51
2.	CONDITIONS DE CULTURE ET CARACTERISATION DE LA REPONSE DE LA PLANTE A <i>P. BRASSICAE</i>	51
2.1.	<i>Conditions de culture</i>	51
2.2.	<i>Suivi de la température</i>	51
2.3.	<i>Inoculation</i>	52
2.4.	<i>Dispositifs expérimentaux</i>	52
2.5.	<i>Caractérisation de l'impact de la présence P. Brassicae sur la plante hôte</i>	52

2.5.1.L'indice pathologique	52
2.5.2.Quantification de l'ADN du pathogène par qPCR.....	53
2.5.3.L'indice GA/LA	54
2.5.4.La longueur de la plus grande feuille de la rosette	54
2.5.5.Biomasse racinaire	54
3. ANALYSE DE LA METHYLATION ET DE L'EXPRESSION DE GENES	54
3.1. Analyse de la méthylation par CHOP PCR et CHOP qPCR.....	54
3.2. Analyse de l'expression de gènes par RT-qPCR.....	55
4. ANALYSE DE L'IMPACT DU METHYLOME DE LA PLANTE SUR LA REPOSE A <i>P. BRASSICAE</i>	55
4.1. Approche non ciblée par détection de QTL ^{épi}	55
4.1.1.Carte épigénétique de la population épiRIL	55
4.1.2.Analyses statistiques des données de phénotypage.....	55
4.1.3.Détection de QTL épigénétiques	56
4.1.4.Analyses statistiques de l'impact des éléments transposables et des variants de séquence présents dans l'intervalle de confiance des QTL ^{épi}	57
4.1.5.Réduction de l'intervalle de confiance des QTL ^{épi}	58
4.1.5.1. Recherche d'individus recombinants supplémentaires	58
4.1.5.2. Densification en marqueurs épigénétiques chez les épiRIL recombinantes	58
4.1.5.3. Cartographie fine des 4 QTL ^{épi} sélectionnés	58
4.2. Approche ciblée de l'influence du méthylome sur la réponse de la plante à <i>P. Brassicae</i> ...	59
4.2.1.Analyse du statut de méthylation au niveau de l'épimutation et de l'expression des 2 NLR sous-jacents.....	59
4.2.2.Séquençage des ARN messenger.....	59
4.2.2.1. Isolement des ARN, construction des banques d'ADNc et séquençage	59
4.2.2.2. Alignement des lectures et comptage.....	59
4.2.2.3. Détermination des gènes différentiellement exprimés	60
4.2.3.Séquençage des petits ARN	60
4.2.3.1. Isolement des petits ARN, construction des banques d'ADNc et séquençage	60
4.2.3.2. Alignement des lectures et comptage.....	60
4.2.3.3. Détermination des petits ARN différentiellement exprimés	60
4.2.4.Origine des données génomiques, d'expression et de méthylation des accessions naturelles d' <i>Arabidopsis</i> et des épiRIL.....	61
4.2.5.Traitement en visualisation des données.	61

CHAPITRE III. APPROCHE NON CIBLEE DE L'IMPACT DE LA METHYLATION DE L'ADN D'ARABIDOPSIS THALIANA SUR LA REPONSE A L'INFECTION PAR PLASMIDIOPHORA BRASSICAE 62

1.	LIEN ENTRE LES VARIATIONS DU METHYLOME D'ARABIDOPSIS ET LA REPONSE A LA HERNIE .	64
1.1.	<i>Introduction</i>	65
1.2.	<i>Materials and methods</i>	66
1.2.1.	Plant material.....	66
1.2.2.	Clubroot pathogen.....	66
1.2.3.	Growth conditions.....	66
1.2.4.	Phenotyping.....	67
1.2.5.	Pathogen DNA quantification.....	67
1.2.6.	Data analysis.....	67
1.2.7.	QTL ^{epi} detection.....	69
1.2.8.	DNA sequence variation.....	69
1.3.	<i>Results</i>	70
1.3.1.	<i>ddm1</i> mutants have reduced susceptibility to <i>P. Brassicae</i>	70
1.3.2.	Heritable differences in DNA methylation are associated with differential susceptibility to <i>P. Brassicae</i>	70
1.3.2.1.	EpiRIL response to <i>P. Brassicae</i> is quantitative.....	70
1.3.2.2.	QTL analysis of data obtained in growth room-1.....	71
1.3.2.3.	QTL analysis of data from growth room-2.....	71
1.3.3.	Temperature affects the plant response to <i>P. Brassicae</i> in the epiRIL population.....	72
1.3.4.	Effect of temperature variation on plant response to clubroot.....	72
1.3.5.	Detection of epigenotype x temperature QTL ^{epi}	73
1.3.6.	Impact of DNA sequence variation within QTL ^{epi} confidence intervals.....	74
1.4.	<i>Discussion</i>	74
1.4.1.	Epigenetic QTL control <i>Arabidopsis</i> clubroot resistance.....	75
1.4.2.	Stability and pleiotropy of clubroot epigenetic QTL.....	76
1.4.3.	Temperature modulates <i>Arabidopsis</i> clubroot responses.....	76
1.4.4.	Clubroot resistance, a sophisticated system of regulation involving genetics and epigenetics...	77
1.5.	<i>References</i>	77
2.	REDUCTION DE L'INTERVALLE DE CONFIANCE DE 3 REGIONS GENOMIQUES CONTENANT PLUSIEURS QTL ^{EPI}	83
2.1.	<i>Choix des QTL^{épi}</i>	83
2.2.	<i>Identification d'individus recombinants supplémentaires au niveau des QTL^{épi} sélectionnés</i>	84

2.3. <i>Densification en marqueurs des régions étudiées chez les épiRIL recombinantes</i>	85
3. DISCUSSION ET PERSPECTIVES	86
4. CONCLUSION PARTIELLE.....	87
CHAPITRE IV. APPROCHE CIBLEE DU ROLE POTENTIEL DES EPIMUTATIONS DANS LA RESISTANCE PARTIELLE D'ARABIDOPSIS THALIANA A PLASMODIOPHORA BRASSICAE	88
1. ANALYSE DE LA REGION DU GENE <i>RPBI</i>	89
1.1. <i>Analyse de la région au niveau des épiRIL</i>	90
1.2. <i>Analyse de la région au niveau des accessions naturelles d'A. thaliana</i>	91
2. ANALYSE DE LA REGION DU QTL ^{EPI} <i>PB-AT5.2</i>	94
2.1. <i>Introduction</i>	95
2.2. <i>Results</i>	97
2.3. <i>Discussion</i>	101
2.1. <i>Methods</i>	104
2.2. <i>References</i>	106
3. RECHERCHE D'UNE INTERACTION EPISTATIQUE AVEC LE QTL <i>PB-AT5.2</i>	112
4. VALIDATION DU ROLE DE L'UN OU DES DEUX GENES PRESENTS SOUS L'EPIMUTATION DANS LA RESISTANCE A LA HERNIE.....	114
5. CONCLUSION PARTIELLE.....	115
CHAPITRE V. DISCUSSION GENERALE ET PERSPECTIVES	117
1. RESISTANCE QUANTITATIVE ET EPIGENETIQUE	120
1.1. <i>Pourquoi prendre en compte ce type de variabilité ?</i>	120
1.2. <i>Pourquoi le modèle Arabidopsis thaliana - Plasmodiophora brassicae</i>	121
1.3. <i>Quel matériel végétal est disponible pour étudier l'impact de l'épigénétique dans le cadre de la résistance quantitative ?</i>	122
2. PERTINENCE, AVANTAGES ET INCONVENIENTS DES APPROCHES CIBLEES ET NON CIBLEES UTILISEES	123
3. INTERACTION DES QTL ^{EPI} PAR RAPPORT A L'ENVIRONNEMENT.....	126
4. TRANSFERT DES RESULTATS OBTENUS CHEZ <i>ARABIDOPSIS</i> A D'AUTRES <i>BRASSICACEAE</i>	127
5. UTILISATION DE LA VARIABILITE EPIGENETIQUE ET DES EPIALLELES EN SELECTION	128
REFERENCES BIBLIOGRAPHIQUES.....	130
ANNEXES	161

INTRODUCTION GENERALE

INTRODUCTION GENERALE

Du fait de leur caractère sessile, les plantes sont en permanence obligées de s'adapter aux modifications de leur environnement qui mettent à mal leur longévité et leur reproduction. Certaines de ces modifications environnementales sont source de stress pour la plante, avec deux grands types de stress ou contraintes : des stress d'origine abiotique, *i.e.* qui ne sont pas du domaine du vivant et des stress d'origine biotique, *i.e.* qui sont issus du vivant. Dans le cadre des changements climatiques touchant notre planète et dans un souci de limitation de l'utilisation des intrants présentant un danger pour l'environnement, de nombreuses études visent à étudier et à comprendre les capacités d'adaptation des plantes aux stress.

La capacité d'adaptation des plantes aux stress est en partie acquise au cours de l'évolution par des phénomènes de mutations qui engendrent des modifications nucléotidiques de la séquence d'ADN. Cependant de récentes études montrent que, bien que les mutations apparaissant dans le génome soient un élément crucial de la diversité des plantes au sein d'une même famille, d'autres sources de variabilité existent. Parmi ces autres sources, la variabilité de l'épigénome des plantes est décrite comme un acteur non négligeable dans les phénomènes de plasticité et d'adaptation des plantes. Des épimutations, *i.e.* les modifications héréditaires de l'expression de gènes sans modification de la séquence nucléotidique sont ainsi impliquées dans le contrôle de divers caractères complexes. D'ailleurs, ces épimutations peuvent survenir à des fréquences non négligeables. Par exemple, dans le cas de la marque épigénétique méthylation de l'ADN, l'apparition d'une épimutation au niveau d'une cytosine est 10^5 fois plus fréquente que celle d'une mutation nucléotidique au niveau des génomes d'une même population (Ossowski *et al.*, 2010 ; van der Graaf *et al.*, 2015).

L'épigénome et plus particulièrement le méthylome des plantes, est un élément à part entière de la modulation de la réponse des plantes aux stress biotiques. La rapidité avec laquelle des modifications de l'épigénome peuvent être mises en place sont en adéquation avec la dynamique de l'interaction plante/pathogène qu'elle soit de type hôte compatible ou incompatible. Dans le cadre de l'interaction hôte non compatible, l'infection de la plante par un agent pathogène peut déclencher l'expression de gènes majeurs de résistance, dits gènes R, afin de stopper rapidement la progression de l'agent pathogène. Ces gènes R présentent une très grande variabilité nucléotidique permettant de contrôler l'infection par divers agents pathogènes et par divers isolats d'un même agent pathogène. Cependant l'expression de ces gènes pouvant mener à un effet délétère chez la plante, notamment par le phénomène de HR (hypersensibilité de la plante), leur expression est finement contrôlée par différents types de facteurs incluant des modifications épigénétiques (Lai & Eulgem, 2018). Dans le cadre de l'interaction dite hôte compatible, la réponse de résistance de la plante à l'agent pathogène est le plus souvent partielle et sous contrôle polygénique. Ce type de résistance dite quantitative autorise la progression et

la multiplication de l'agent pathogène de manière plus ou moins importante au sein de la plante. La résistance quantitative faisant intervenir plusieurs gènes est théoriquement plus difficilement contournable par l'agent pathogène que la résistance qualitative du fait de la variabilité potentielle existant au niveau de chaque gène impliqué dans la résistance partielle (French *et al.*, 2016).

Actuellement, la résistance quantitative est considérée comme résultant principalement de la ségrégation de nombreux allèles à différents loci. Cependant, ces variations alléliques n'expliquent pas toujours la totalité de la variabilité phénotypique observée suggérant l'existence d'une autre source de variabilité contribuant à la réponse de résistance. Le rôle démontré de l'épigénétique dans la réponse des plantes à des stress biotiques et abiotiques, la transmissibilité de marques épigénétiques influençant des caractères complexes de plantes et des fréquences d'épimutations comparables voire plus importantes que celles des mutations nucléotidiques suggèrent qu'une partie de la résistance quantitative pourrait aussi résulter d'une ségrégation génomique de nombreux épiallèles.

Le travail effectué au cours de ces trois années de thèse s'intègre dans la problématique scientifique visant à mieux comprendre et caractériser la résistance quantitative. L'objectif de ma thèse a été de déterminer si des variations stables d'origine épigénétique étaient impliquées dans la résistance quantitative à la hernie chez *Arabidopsis thaliana*.

Mon travail de thèse a été encadré par Mélanie Jubault (Maître de Conférences, AGROCAMPUS OUEST) et Maria Manzaneres-Dauleux (Professeur, AGROCAMPUS OUEST) au sein de l'équipe « Résistance et Adaptation » de l'UMR IGEPP (Institut de Génétique, Environnement et Protection des Plantes, UMR 1349 INRA, AGROCAMPUS OUEST, Université de Rennes 1) et cofinancé par une bourse du Département Biologie et Amélioration des Plantes de l'INRA (Institut National de la Recherche Agronomique, France) et de la région Bretagne.

Mon manuscrit de thèse est structuré en 5 chapitres. Dans le premier chapitre j'ai réalisé une synthèse bibliographique ayant pour but de replacer les contextes scientifique et socioéconomique dans lesquels s'intègre mon travail. Je me suis intéressé dans cette partie i) aux voies de défense chez les végétaux, ii) à l'implication de l'épigénétique dans la réponse des plantes aux stress et iii) à la description du modèle d'interaction plante pathogène étudié, *Brassicaceae/Plasmodiophora brassicae*, l'agent causal de la Hernie des Crucifères. Dans le chapitre II sont décrits les matériels et méthodes utilisés. Les chapitres III et IV sont consacrés à la présentation des résultats obtenus au cours de ma thèse, notamment sous forme d'articles, un article soumis pour publication et accepté avec révisions majeures traitant de la détection de QTL de résistance à la hernie sous dépendance épigénétique et un article en cours de préparation traitant de l'identification d'une épimutation sous-jacente à un QTL de résistance à la hernie qui contrôle l'expression de deux gènes NLR. Le chapitre V exposera une discussion générale sur le travail que j'ai effectué.

CHAPITRE I

SYNTHESE BIBLIOGRAPHIQUE

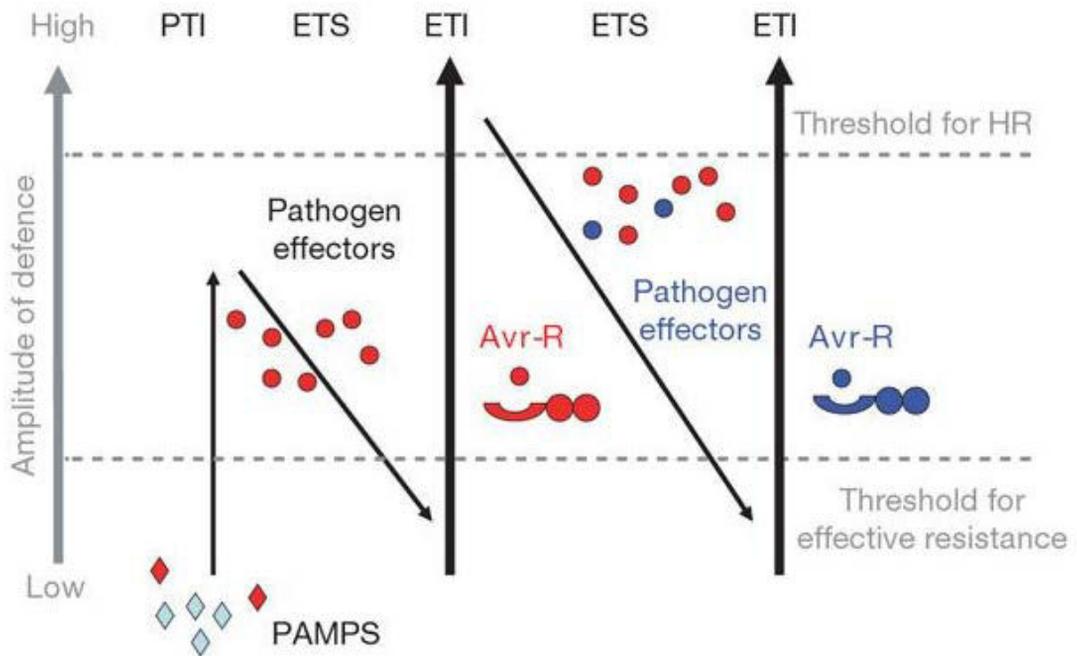


Figure I.1. Modèle en zigzag illustrant les axes de résistance chez les plantes. Pendant la phase une, les plantes détectent les motifs moléculaires associés aux microbes et aux pathogènes (MAMP/PAMP, diamants rouges) par l'intermédiaire des PRR. Cette reconnaissance déclenche la mise en place de la PAMP-triggered immunity (PTI). Durant la phase 2, les agents pathogènes vont sécréter des effecteurs qui vont interférer avec la PTI et ainsi permettre aux agents pathogènes de se développer, entraînant une susceptibilité déclenchée par les effecteurs appelée ETS. En phase 3, un effecteur (rond en rouge) est reconnu par une protéine NB-LRR, activant l'immunité déclenchée par effecteur (ETI) conduisant généralement à l'induction de la mort cellulaire hypersensible (HR). Issue de Jones & Dangl (2006)

CHAPITRE I. SYNTHÈSE BIBLIOGRAPHIQUE

1. Les voies de défense chez les végétaux

L'interaction entre une plante et un agent pathogène présent dans son environnement peut être de deux types différents, « non hôte » ou « hôte ». Les interactions plante pathogène sont qualifiées de non hôte lorsque la résistance est démontrée pour une espèce végétale entière et est présente pour tous les isolats d'un agent pathogène particulier (Mysore & Ryu, 2004). *A contrario*, les interactions plante pathogène sont qualifiées d'hôte lorsque la résistance décrite est cultivar ou accession spécifique (Mysore & Ryu, 2004). Quel que soit le type d'interaction entre l'agent pathogène et la plante, deux principaux mécanismes du système immunitaire des plantes peuvent être mis en place : la reconnaissance de motifs moléculaires associés à l'agent pathogène : les Pathogen Associated Molecular Pattern (PAMP) conduisant à un état basal de résistance chez la plante connu sous le nom de PAMP-Triggered Immunity (PTI) et la résistance induite par les effecteurs de l'agent pathogène appelée Effector-Triggered Immunity (ETI ; **Figure I.1** ; Jones & Dangl, 2006). La voie PTI représente la première ligne de réponse de défense de la plante ; l'ETI la seconde ligne de défense de la plante qui conduit à une réponse spécifique dite gène pour gène.

Dans le cadre des interactions plante pathogène de type hôte, deux catégories sont ainsi possibles i) les interactions hôte compatible où l'agent pathogène arrive à se multiplier au sein de son hôte de manière plus ou moins importante et ceci grâce à l'absence de reconnaissance de type gène pour gène, ii) les interactions de type hôte incompatible où la progression de l'agent pathogène à l'intérieur de son hôte est stoppée rapidement par des mécanismes faisant intervenir des gènes tels que les gènes R (Glazebrook, 2005). La PTI intervient dans le cadre de l'interaction hôte incompatible (Boller & Felix, 2009 ; Schulze-Lefert & Panstruga, 2011) mais peut aussi contribuer à la résistance hôte compatible dans le cas de la résistance quantitative (French *et al.*, 2016). L'ETI est présente lors des interactions hôte incompatible et mène à des réponses de défense de type qualitative (Boller & Felix, 2009).

Le travail présenté dans ce manuscrit étant appuyé sur une interaction plante pathogène de type hôte la suite des éléments décrits ne prendront en compte que le cas de ce type d'interaction.

1.1. L'interaction hôte incompatible : cas de la résistance qualitative

Les résistances qualitatives, dites aussi résistances totales, mettent en jeu les deux mécanismes du système immunitaire de la plante que sont la PTI et l'ETI. La PTI est mise en place au niveau de la plante hôte par la reconnaissance de différents motifs : i) des motifs issus de l'hôte lui-même suite aux dommages qui lui sont causés par l'agent pathogène et qui sont nommés DAMP pour Damage Associated Molecular Patterns et ii) les PAMP issus directement de l'agent pathogène (Boller & Felix,

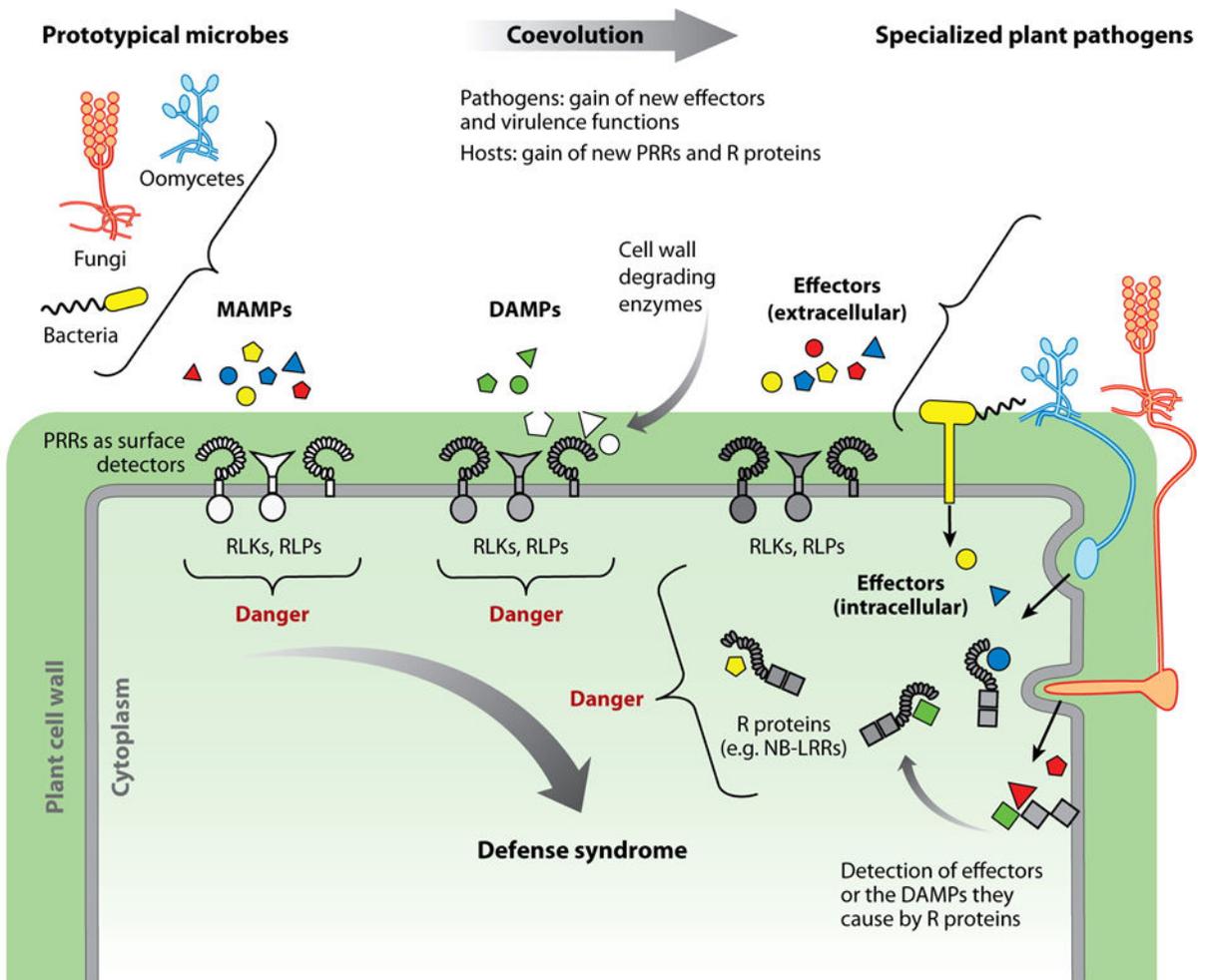


Figure I.2. Reconnaissance des PAMP et DAMP par les protéines PRR menant à l'activation de voies de défense. Les motifs moléculaires associés aux microbes (MAMP), les motifs moléculaires associés aux dommages (DAMP) ainsi que les effecteurs sont perçus comme signaux de danger. Cette reconnaissance est réalisée par des récepteurs de type PRR tels que les RLK (Receptor Like Kinase) et les RLP (Receptor Like Protein). Au cours de la co-évolution des systèmes plante- pathogène, les agents pathogènes gagnent des effecteurs appelé facteurs de virulence afin de contourner la détection par les protéines PRR. De leur côté, les plantes gagnent de nouvelles protéines PRR ainsi que des protéines de résistance R tels que les NBS-LRR (Nucleotide binding-site-leucine-rich repeat) afin de détecter les facteurs de virulence de l'agent pathogène. La détection des effecteurs de virulence par les nouvelles protéines PRR et R vont mener à la mise en place d'un nouveau signal de danger et à la mise en place d'une réponse de défense. D'après Boller et Felix (2009).

2009). L'interaction entre PAMP/DAMP et les récepteurs PRR, pour Pattern-recognition receptors, entraîne l'activation de la voie PTI qui se traduit par différentes modifications cellulaires telles que : i) un influx d'ions Ca^{2+} à travers la membrane plasmique conduisant à une alcalinisation extracellulaire et à une dépolarisation de la membrane plasmique, ii) un burst d'espèces réactives de l'oxygène (ROS), iii) l'activation des protéines 14-3-3 qui participent à la transduction du signal et iv) l'activation de protéines kinases qui vont phosphoryler des protéines et jouer un rôle dans la transduction du signal.

Les PAMP et les DAMP sont reconnus par les récepteurs PRR qui sont situés au niveau de la membrane plasmique des cellules. Les récepteurs PRR sont classés en deux types, les RLK pour receptor-like kinases et les RLP pour receptor-like proteins. Les protéines PRR sont constituées de manière globale d'un ectodomaine impliqué dans la fixation du ligand et d'un domaine transmembranaire. Les RLK possèdent en plus un domaine kinase intracellulaire à leur extrémité C terminale qui est absent chez les RLP (Macho & Zipfel, 2014). La reconnaissance des PAMP et DAMP par les protéines PRR mène à l'activation de la voie PTI, permettant ainsi d'arrêter la propagation de l'agent pathogène au sein de son hôte (**Figure I.2**). Afin de contourner cette première ligne de défense les agents pathogènes sécrètent des effecteurs dans le but de tromper ce système de défense de la plante aboutissant à un état de sensibilité de la plante vis-à-vis de l'agent pathogène (Jones & Dangl, 2006).

Afin de bloquer la propagation de l'agent pathogène suite au contournement de sa première ligne de défense la plante possède une seconde ligne de défense, l'ETI. Ce mécanisme de défense conduit de manière générale à un phénomène de mort cellulaire programmée nommé réponse hypersensible (HR). Cette HR induit un « suicide » des cellules entourant le point d'entrée de l'agent pathogène afin d'endiguer sa progression (Bashir *et al.*, 2013). La HR vient renforcer la Résistance Systémique Acquise (RSA) qui est un mécanisme de résistance qui se généralise à l'ensemble de la plante contre un large spectre d'agents pathogènes (Fu & Dong, 2013).

Le fonctionnement de cette voie a été initialement décrit suivant le modèle d'interaction gène pour gène proposé par Flor (1971). Dans ce modèle deux éléments entrent en jeu, le gène de résistance de la plante, nommé gène *R* qui code une protéine *R* et le gène d'avirulence de l'agent pathogène, qui produit une protéine nommée protéine d'avirulence (protéine *Avr*). Depuis, le terme de protéine d'avirulence a été remplacé par le terme effecteur (Boller & Felix, 2009) du fait du rôle des protéines *Avr* qui vont cibler des facteurs de la cellule hôte faisant d'eux de vrais facteurs de virulence (Jones & Dangl, 2006).

La plupart des gènes de résistance *R* codent des protéines de la famille des récepteurs de type nucleotide-binding domain leucine rich repeat (NBS-LRR ou NLR ; Dangl & Jones, 2001 ; Chisholm *et al.*, 2006 ; Jones & Dangl, 2006 ; Dodds & Rathjen, 2010 ; Maekawa *et al.*, 2011). Les protéines NLR sont constituées de manière globale i) d'une P-loop ATPase de la famille AAA + qui assure de nombreux rôles complexes tel que la régulation de l'activation des NLR, ii) d'un domaine central de liaison aux

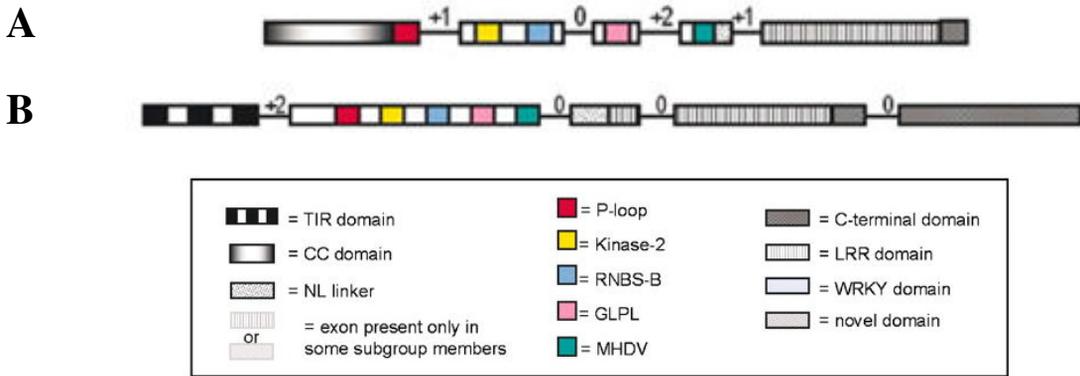


Figure I.3. Configuration intronique / exonique et motifs protéiques de gènes NBS-LRR chez *Arabidopsis*. (A) Configuration et motif des gènes CNL (CC–NBS–LRR). (B) Configuration et motif des gènes TNL (TIR–NBS–LRR). Les CNL diffèrent de TNL par la présence respective d'un domaine coiled-coil (CC domain) et d'un domaine Toll-interleukin 1 (TIR domain). Les exons sont représentés par des boîtes et les introns par les traits qui connectent les exons. Les nombres indiqués en amont des introns indiquent la phase des introns. Adaptée de Meyers *et al.* (2003).

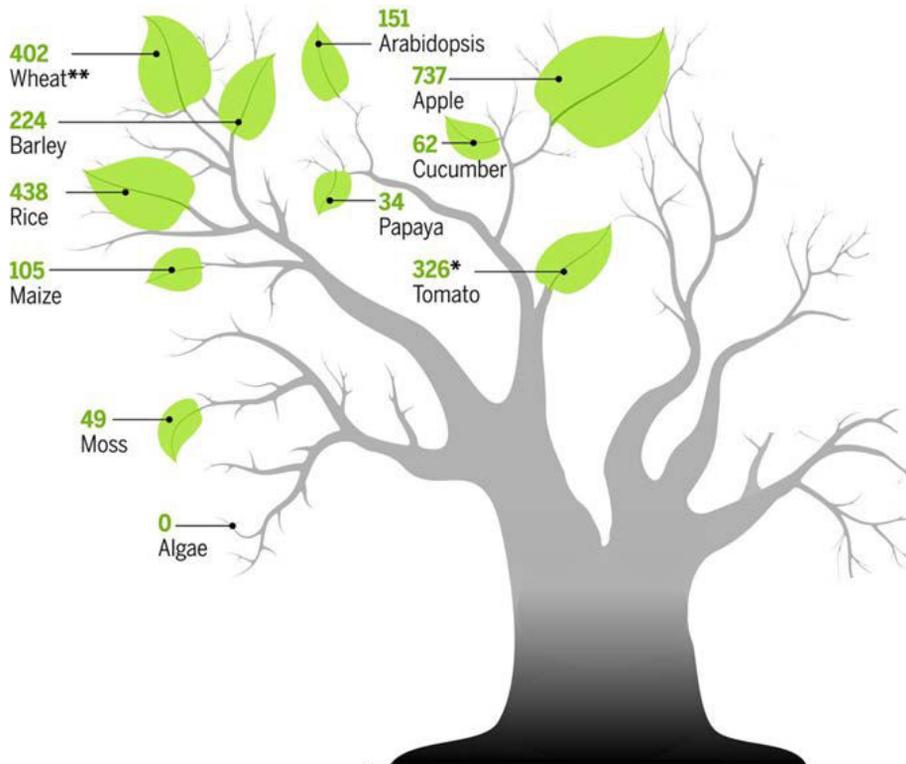


Figure I.4. Diversité du nombre de gènes codant des NLR chez les plantes. Le nombre de gènes NLR par génome identifiés par calcul varie largement, comme le montre cet arbre évolutionniste stylisé (branches non à l'échelle). Le nombre de NLR peut varier de façon marquée même à travers les génomes provenant de taxons étroitement apparentés. L'astérisque pour la tomate indique que des preuves expérimentales existent pour donner cette précision. Le double astérisque pour le blé indique le nombre de NLR par génome diploïde (le blé est hexaploïde). Adaptée de Jones *et al.* (2016).

nucléotides (NB) et iii) d'un domaine LRR à leur extrémité C-terminale. Les NLR se différencient au niveau de leurs extrémités N-terminales et sont classées en deux groupes principaux : les NLR avec un domaine Toll-interleukin 1 (TIR) qui sont appelés des TNL et les NLR avec un domaine coiled-coil (CC) qui sont appelés CNL (**Figure I.3** ; Maekawa *et al.*, 2011). Chez les génomes des plantes, le nombre de gènes codant des protéines NLR est très variable (**Figure I.4**) avec par exemple environ 150 gènes codant des protéines NLR chez *Arabidopsis thaliana* (Meyers *et al.*, 2003) alors que la pomme en possède 737 et que l'algue *Chlamydomonas* n'en possède aucun (Jones *et al.*, 2016).

Les plantes maintiennent un certain niveau de diversité des NLR leur permettant de conserver des mécanismes de défense efficaces face aux agents pathogènes. Cette diversité peut être générée par des modifications de la séquence de l'ADN reposant sur des mécanismes divers tels que i) une variation du nombre de copies (CNV) résultant de duplications en tandem menant à la création de clusters de gènes complexes, ii) la variabilité des NLR due à des mutations dans différentes espèces végétales, iii) des phénomènes de fusion de domaines par des recombinaisons inter-alléliques (Baggs *et al.*, 2017).

Une deuxième source de diversité des NLR repose sur les modalités de leur activation qui se fait de manière spécifique par la présence d'effecteurs qui leur sont propres. Différents modèles existent pour expliquer le mécanisme par lequel l'interaction entre les protéines NLR et les effecteurs se met en place. Historiquement le premier modèle décrit dans le contexte de l'interaction entre l'effecteur et les protéines R suggérait que la protéine R entraînait en contact direct avec l'effecteur (Keen, 1990). Cependant parmi toutes les études menées seul un nombre très restreint a réussi à démontrer une interaction directe entre les deux protéines (Jia *et al.*, 2000 ; Deslandes *et al.*, 2003 ; Dodds *et al.*, 2006 ; Krasileva *et al.*, 2010). Depuis d'autres modèles ont été proposés tels que i) le modèle de « garde » qui propose que la protéine R surveille non pas l'effecteur mais une protéine accessoire impliquée dans la signalisation de défense qui est elle la vraie cible de l'effecteur (Van der Biezen & Jones, 1998) et ii) le modèle du « leurre » qui fait intervenir une protéine leurre qui a pour rôle d'être comme la protéine de garde la cible de l'effecteur. Dans le cadre de la présence d'une protéine R fonctionnelle, le mécanisme mis en place est similaire à celui du modèle de garde. Par contre dans le cas d'une protéine R non fonctionnelle, la sélection naturelle va favoriser les mutations au niveau de la protéine leurre afin de piéger l'effecteur en l'empêchant de déclencher l'ETS (**Figure I.5**; van der Hoorn & Kamoun, 2008).

La diversité des fonctions portées par les NLR (senseur de l'effecteur, senseur de la cible de l'effecteur, participant à la voie de signalisation...) traduite par l'ensemble des modèles d'activation décrit est très importante dans la diversité des NLR. Récemment de nouveaux types de NLR ont été décrits augmentant encore la complexité et la diversité des réponses NLR : i) les NLR dénommés NLR-ID qui présentent un domaine intégré ressemblant à la cible de l'agent pathogène ; ces NLR reposent sur le modèle du « leurre intégré » où l'intégralité ou une partie de l'élément ciblé par l'effecteur est fusionnée à une protéine NLR issue de l'expression d'un gène *R* (Cesari *et al.*, 2014 ; Wu *et al.*, 2015 ; Kroj *et al.*,

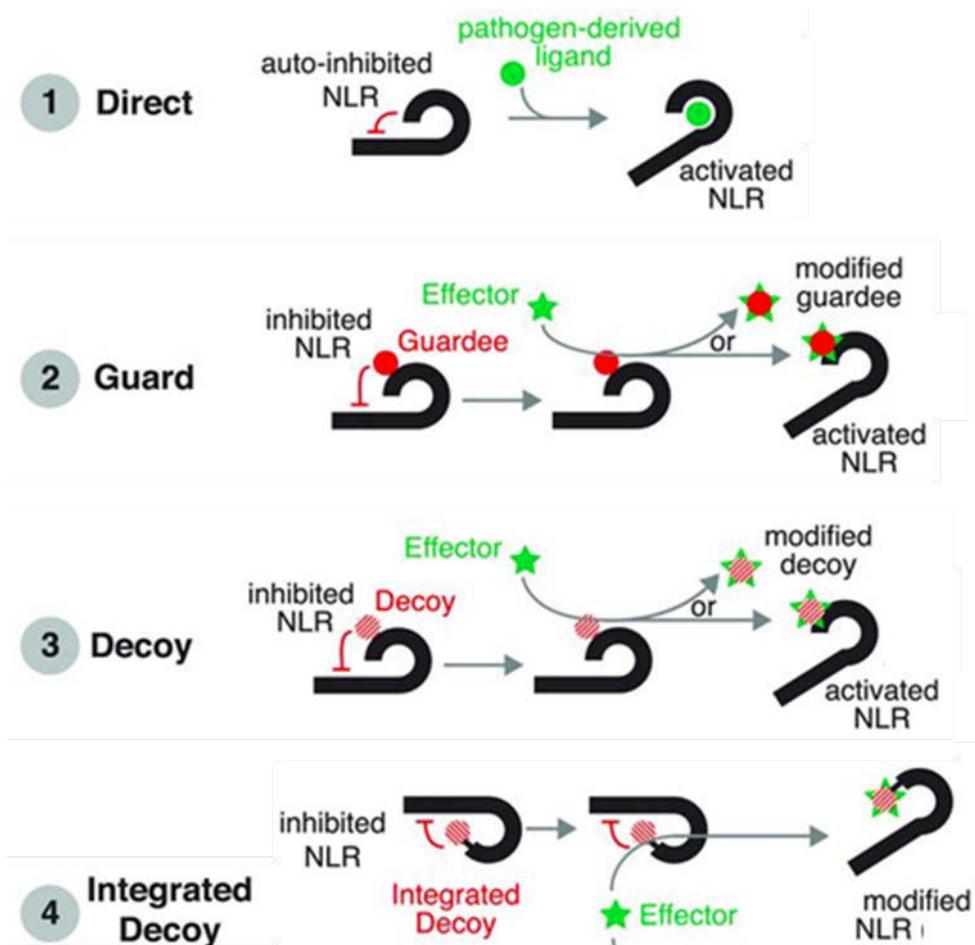


Figure I.5. Récapitulatif des divers mécanismes de détection des agents pathogènes par les NLR. Dans le premier modèle, la protéine NLR inactive interagit directement avec l'effecteur reconnu induisant alors son activation. Dans le deuxième modèle, la protéine NLR ne surveille pas la présence d'un effecteur mais d'une protéine du soi appelée « garde » (« *guard* », en rouge) qui est-elle la vraie cible de l'effecteur. Initialement la protéine garde inhibe le NLR, lors de l'interaction entre la protéine garde et l'effecteur, la structure de la protéine de garde se retrouve altérée activant alors le NLR. Dans le troisième modèle, l'interaction s'effectue comme dans le deuxième modèle entre l'effecteur et une autre protéine que le NLR qui est dans ce cas une protéine leurre (*Decoy*). Cependant à la différence du deuxième modèle, dans ce modèle si la protéine NLR n'est pas fonctionnelle, la sélection naturelle tend à favoriser l'apparition de mutations au niveau de la protéine leurre afin de bloquer sa modification par l'effecteur et la mise en place de l'ETS. Dans le quatrième modèle, la protéine leurre n'est plus indépendante de la protéine NLR mais est intégrée à cette dernière. Adaptée de Jones *et al.* (2016).

2016 ; Giolai *et al.*, 2017). ii) Les NLR helpers qui sont des NLR qui garantissent ou amplifient la réponse induite par d'autres NLR (Bonardi *et al.*, 2011). Pour le moment trois NLR helper ont été identifiés : ADR1 pour Activated Disease Resistance 1, NRG1 pour N Requirement Gene et NRC1 pour NB-LRR required for Hypersensitive Response-Associated Cell Death 1 (Gabriëls *et al.*, 2007 ; Dong *et al.*, 2016). Les NLR ADR1 et NRG1 possèdent tous les deux à leur extrémité N-terminale un domaine similaire à RPW8 (Resistance to Powdery mildew 8) qui contient un motif coiled coil. Des analyses phylogénétiques récentes suggèrent que les NLR de l'ancienne sous-clade des coiled coil (CC_R-NBS-NLR) ont évolué en NBS helper qui ont un rôle durant l'ETI (Collier *et al.*, 2011). Le rôle de NRG1 a été montré chez *Nicotiana benthamiana* où elle est nécessaire pour le fonctionnement du NLR TNL N du tabac. Chez *A. thaliana* quatre gènes codant des protéines de la famille ADR1 existent : *ADR1* (At1g33560), *ADR1-L1* (At4g33300), *ADR1-L2* (At5g04720) et *ADR1-L3* (At5g47280). Le rôle helper des trois premiers membres de la famille a été démontré par Bonardi *et al.* (2011) *via* leur importance dans la transduction du signal lors de l'ETI. En ce qui concerne le dernier membre de la famille ADR1-L3, aucun phénotype n'a été reporté dans leur étude en utilisant des mutants T-DNA de Col-0. De plus, il a été montré que ce dernier membre des ADR1 a perdu approximativement les 190 premiers nucléotides de son extrémité N-terminale modifiant ainsi son domaine RPW8 spécifique des NLR de la famille des helpers. Cependant, malgré la perte de ce domaine Bonardi *et al.* (2011) ne réfutent pas la possibilité d'une fonction spécifique de *ADR1-L3*.

La diversité des mécanismes de reconnaissance disponibles met en avant la capacité des réponses NLR d'être rapidement activées ou éteintes en fonction des conditions environnementales (stress... ; Jones *et al.*, 2016). En effet, les NLR nécessitent un système strict de contrôle de leur expression du fait d'un possible risque délétère pour la plante à des niveaux d'expression trop élevés. Leur expression est ainsi contrôlée par divers mécanismes aussi bien pré-transcriptionnels (avec par exemple la modulation de la présence de facteurs de transcription et la régulation de la structure de la chromatine par modulation de marques épigénétiques) que post transcriptionnels par l'utilisation de petits ARN conduisant à la destruction des ARN synthétisés (Lai & Eulgem, 2018).

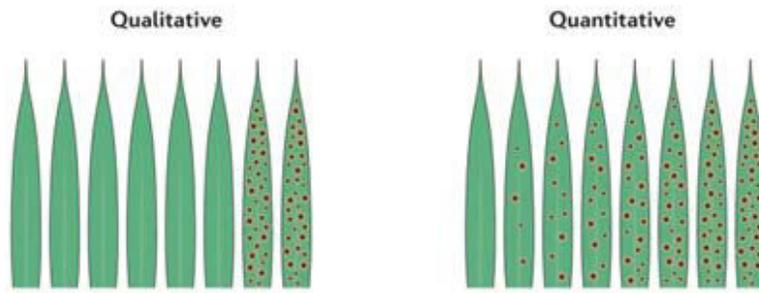
Un affaiblissement ou une absence d'ETI, due à la non fonctionnalité de gènes impliqués dans le mécanisme ou à sa suppression par d'autres éliciteurs permet à l'agent pathogène de progresser dans son hôte. Cette progression peut par la suite être stoppée par d'autres métabolites ou protéines de résistance conduisant à un état de résistance quantitatif (Kushalappa *et al.*, 2016).

1.2. L'interaction hôte compatible : cas de la résistance partielle ou résistance quantitative

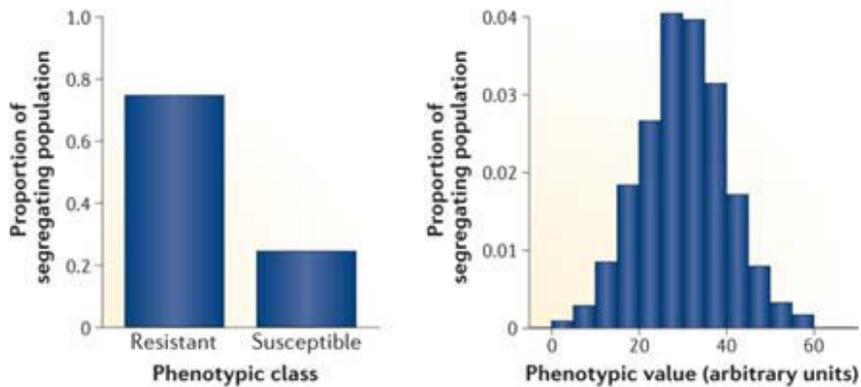
La résistance quantitative est possible dans le cadre des interactions hôte compatible et a été identifiée chez de nombreux pathosystèmes dont ci-après quelques exemples i) dans le pathosystème riz/*Magnaporthe oryzae* (Tanweer *et al.*, 2015) ii) dans les pathosystèmes *Arabidopsis/Ralstonia*

A

Phenotype

**B**

Population

**C**

Genetics

Genotype	Phenotype
AA	Resistant
Aa	
aa	Susceptible

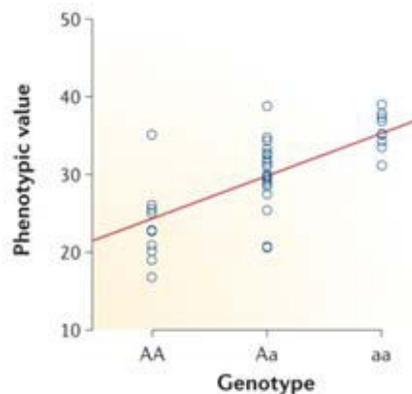


Figure I.6. Comparaison de la résistance quantitative et de la résistance qualitative. (A) Au niveau phénotypique, pour la résistance qualitative deux classes distinctes sont observées dans une population en ségrégation alors qu'une distribution continue est observée dans le cadre de la résistance quantitative. (B) Dans une population en ségrégation, la résistance qualitative portée par un gène R ségrège dans un rapport de 3 : 1, tandis que de multiples gènes de faible effet contribuent à la variation phénotypique continue de la résistance quantitative. (C) Pour la résistance qualitative, un seul gène sous-jacent engendre deux phénotypes distincts: la résistance ou la sensibilité. La ségrégation de gènes de résistance quantitative couplée à des effets environnementaux entraîne une gamme de phénotypes de résistance partielle. Figure adaptée de Nelson *et al.* (2017).

solanacearum (Aoun *et al.*, 2017), *Arabidopsis/Botrytis cinerea* (Corwin *et al.*, 2016) et *Arabidopsis/Plasmodiophora brassicae* (Jubault *et al.*, 2008c) iii) dans les pathosystèmes maïs/*Bipolaris maydis* et maïs/*Cercospora zea-maydis* (Kump *et al.*, 2011 ; Benson *et al.*, 2015) iv) dans le pathosystème soja/*Sclerotinia sclerotiorum* (Sclérotinia; Kim & Diers, 2000) v) dans les pathosystèmes colza/*Leptosphaeria maculans* (Phoma; Kumar *et al.*, 2018), colza/Sclérotinia (Zhao *et al.*, 2006) et colza/*P. Brassicae* (Manzanares-Dauleux *et al.*, 2000a) vi) dans le pathosystème tournesol/*Leptosphaeria maculans* (phoma; Darvishzadeh R. *et al.*, 2007) vii) dans le pathosystème/*Alternaria solani* (Foolad *et al.*, 2002).

La résistance quantitative ne bloque pas mais réduit la multiplication des agents pathogènes, la colonisation des plantes et/ou la gravité des symptômes. De ce fait, les phénotypes issus de la résistance quantitative se caractérisent par une distribution continue allant de phénotypes résistants à des phénotypes sensibles avec un niveau de résistance incomplet ou partiel. La distribution des phénotypes de résistance dans une descendance résulte le plus souvent de la ségrégation des allèles ayant des effets variables sur plusieurs loci (Corwin & Kliebenstein, 2017 ; **Figure I.6**). Dans la plupart des systèmes plante pathogène, les facteurs contribuant aux résistances quantitatives sont associés à des régions génomiques appelées quantitative trait loci (Lindhout, 2002 ; Stuthman *et al.*, 2007). Les effets des QTL de résistance sont souvent additifs avec certains QTL ayant des effets majeurs et d'autres des effets mineurs. Néanmoins, des QTL ayant des effets épistatiques, *i.e* des QTL ayant un effet uniquement en combinaison de deux ou plus (et non de manière individuelle), ont été identifiés (Lefebvre & Palloix, 1996 ; Manzanares-Dauleux *et al.*, 2000 ; Jubault *et al.*, 2008). De plus, les QTL de résistance peuvent parfois être spécifiquement détectés dans certaines conditions environnementales (sol, climat, population d'agents pathogènes) ou dans des contextes génétiques spécifiques (effet fonds génétique). Ainsi, lorsque plusieurs isolats d'un agent pathogène sont utilisés pour la détection de QTL dans une population, la résistance amenée par les QTL peut être spécifique (détectée pour un seul isolat) ou non spécifique (Ellis *et al.*, 2014 ; Wiesner-Hanks & Nelson, 2016). En effet, il a été montré que dans plusieurs pathosystèmes, la résistance quantitative résultait de la combinaison de QTL mineurs et de QTL à large spectre ou spécifiques à un ou plusieurs isolats (Caranta *et al.*, 1997 ; Calenge *et al.*, 2004 ; Rocherieux *et al.*, 2004).

En termes de durabilité, la résistance quantitative est décrite comme pouvant être plus stable et durable que la résistance qualitative (Poland *et al.*, 2009 ; St Clair, 2010 ; Roux *et al.*, 2014 ; Niks *et al.*, 2015). Selon French *et al.* (2016) la meilleure durabilité, *i.e* une résistance efficace sur une longue période, même en cas de culture intensive, de la résistance quantitative est due au fait qu'elle repose sur de multiples gènes conduisant ainsi à des niveaux de pressions évolutives appliqués sur le pathogène nettement moins importants. En plus de cet aspect, Pilet-Nayel *et al.* (2017) exposent dans leur revue des éléments supplémentaires permettant d'expliquer la meilleure durabilité de la résistance quantitative

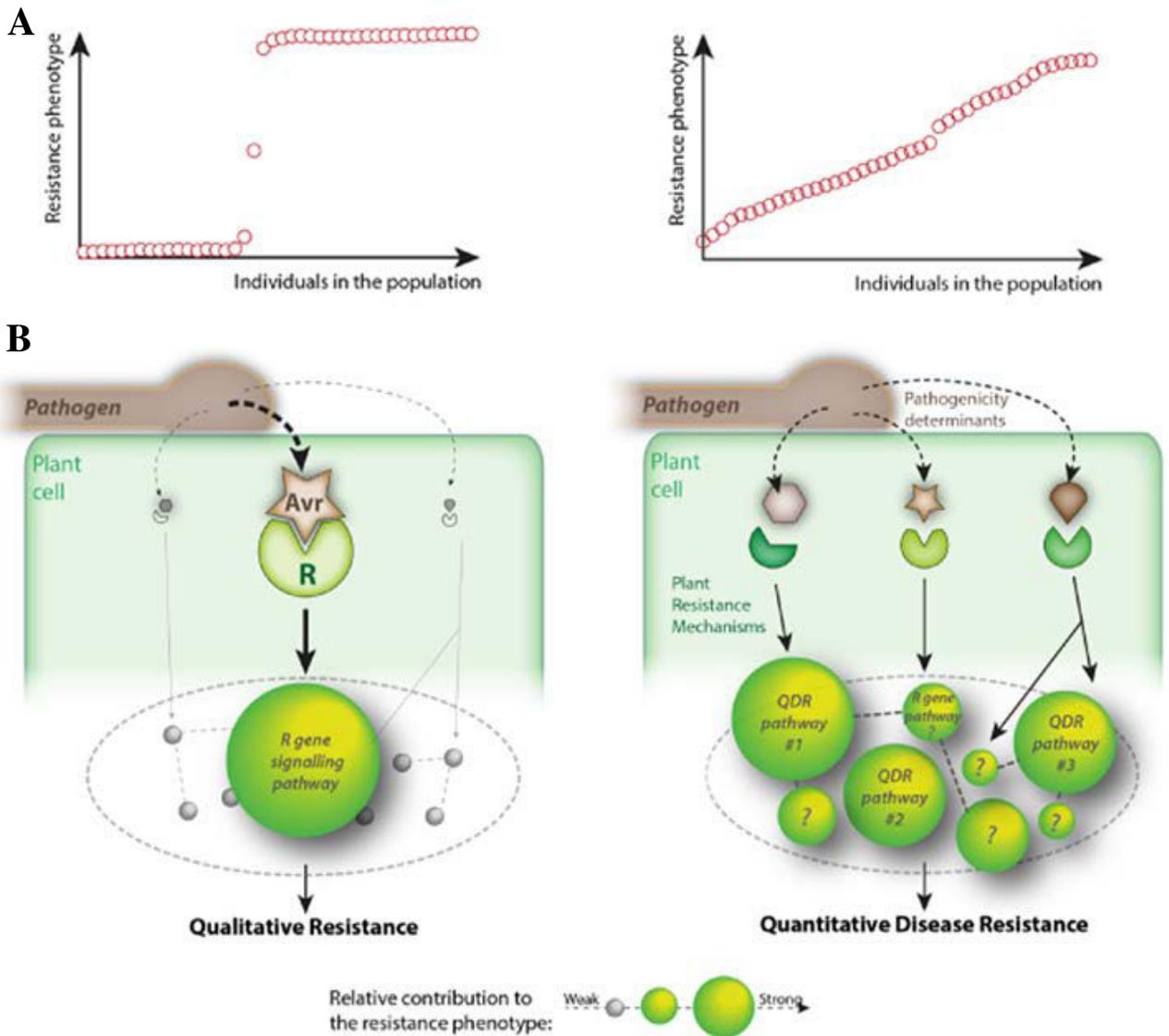


Figure I.7. Schématisation des mécanismes de résistances qualitatives et quantitatives. La figure illustre la résistance qualitative à gauche et la résistance quantitative à droite. **(A)** Au niveau des populations, le phénotype de résistance qualitative à la maladie suit une distribution binaire (résistant ou sensible) alors que la résistance quantitative suit une distribution continue allant de la sensibilité à la résistance. **(B)** Au niveau moléculaire, la résistance qualitative résulte de la perception d'un seul effecteur de l'agent pathogène (Avr) par le produit d'un gène de résistance R alors que la résistance quantitative résulte de l'intégration simultanée de plusieurs facteurs génétiques de résistance apportant des contributions différentes afin d'atteindre le phénotype de résistance globale. QDR, résistance quantitative aux maladies. (Roux *et al.*, 2014)

avec la présence i) d'une combinaison de pressions de sélection contradictoires sur l'évolution des agents pathogènes, (ii) d'une faible probabilité de mutations multiples de pathogènes nécessaires pour surmonter de multiples QTL, (iii) d'une combinaison de différents mécanismes associés à la résistance, plus difficiles à surmonter (iv) d'une combinaison de mécanismes de résistance agissant successivement à différents moments du cycle de vie de l'agent pathogène ou tout au long du développement de la plante (Palloix *et al.*, 2009 ; Mundt, 2014). Néanmoins, des études évolutives menées en appliquant successivement un agent pathogène sur des plantes portant une résistance quantitative suggèrent une possible altération ce type de résistance par les champignons et les virus (Kolmer & Leonard, 1986 ; Lehman & Shaner, 1997 ; Montarry *et al.*, 2012).

Comme l'utilisation des résistances génétiques reste l'une des meilleures méthodes pour cultiver les végétaux en utilisant un minimum d'intrants et de produits phytosanitaires, de multiples recherches ont été menées dans le but d'augmenter la durabilité de ces résistances (Delourme *et al.*, 2006 ; Stuthman *et al.*, 2007). Un des axes étudiés afin d'augmenter cette durabilité est le pyramidage de différentes résistances quantitatives entre elles ou de résistances quantitatives avec des résistances qualitatives. En effet, bien que les gènes *R* sont capables de contrôler la plupart des souches avirulentes présentes dans les populations d'agents pathogènes, les gènes impliqués dans la résistance quantitative pourraient diminuer la taille effective de la population de l'agent pathogène et ainsi diminuer le nombre des mutations au sein du génome de l'agent pathogène susceptibles de contourner les gènes *R* (Roux *et al.*, 2014). De plus, contrairement à la résistance qualitative qui apporte un désavantage en présence d'un agent pathogène nécrotrophe, la résistance quantitative est efficace contre les agents pathogènes biotrophes et nécrotrophes. Brun *et al.* (2010) proposent trois effets positifs de l'utilisation du pyramidage de gènes *R* avec des résistances quantitatives i) l'amélioration du contrôle de la maladie par la résistance quantitative et ceci grâce au contrôle des fractions avirulentes d'agents pathogènes de la population par les gènes *R* ii) la limitation de la sélection des isolats virulents *via* la résistance quantitative qui va ralentir le développement de l'épidémie et ainsi diminuer la gravité de la maladie et la taille de la population efficace iii) le maintien à un niveau satisfaisant de la protection lors du contournement de la résistance portée par le gène *R*. L'évaluation dans le pathosystème *Leptosphaeria maculans/Brassica napus* de l'effet du pyramidage a montré que 8 années de sélection récurrente ont été nécessaires pour surmonter la résistance du gène *Rlm6* introduit dans un contexte génétique partiellement résistant, contre seulement 3 ans dans un contexte génétique sensible (Brun *et al.*, 2010 ; Delourme *et al.*, 2014a).

Cependant, même si la résistance quantitative semble être plus durable que la résistance qualitative et qu'elle permet de conserver l'effet des gènes *R* avec lesquels elle est combinée, il est nécessaire de bien choisir les QTL de résistance afin de minimiser l'altération de l'effet des QTL et de créer des combinaisons optimales de résistance respectant un coût raisonnable pour la plante (Pilet-Nayel *et al.*,

Tableau I.1. Exemples de gènes causaux identifiés sous des QTL de résistance. Adapté de Nelson *et al.* (2017).

Mechanism	Gene(s) name; Disease; Host; Pathogen (type)	Evidence	Notes
Pathogen recognition	<i>Pb1</i> ; Rice blast; Rice; <i>Magnaporthe oryzae</i> (fungus)	Fine mapping, transgenic complementation	<i>Pb1</i> is an atypical NLR (CC-NBS-LRR). This class of genes is often associated with qualitative race-specific resistance. It confers adult plant resistance, likely due to its higher expression during the later stages of development. Resistance is dominant.
	<i>RCC1</i> -Anthracnose stalk rot; Maize; <i>Colletotrichum graminicola</i> (fungus) ⁴	Fine mapping, transgenic complementation	<i>RCC1</i> encodes a typical CC-NLR gene. It has an additive effect on resistance
Transcriptional response	<i>RRS1</i> and <i>RPS4</i> -Black Rot; Arabidopsis; <i>Xanthomonas campestris</i> pv. <i>campestris</i> race 2 (bacterium) ⁶	Association mapping followed by analysis of insertional mutants	The resistance gene pair <i>RRS1</i> and <i>RPS4</i> are both NLR genes and confer resistance to other bacteria such as <i>Pseudomonas syringae</i> and <i>Ralstonia solanacearum</i>
	<i>RFO3</i> ; Wilt; Arabidopsis; <i>Fusarium oxysporum</i> f.sp. <i>nathiioli</i> (fungus) ^{9,10}	Fine mapping, transgenic complementation and mutant analysis	<i>RFO3</i> is a B-lectin receptor-like kinase typical of plant pattern recognition receptors involved in recognition of MAMPs
Signal transduction	<i>C3H12</i> ; bacterial blight; Rice; <i>Xanthomonas oryzae</i> pv. <i>oryzae</i> (bacterium) ¹¹	Transgenic overexpression and silencing. Co-localization of gene with minor QTL	<i>C3H12</i> is a CCCH-type zinc finger protein and may function as an RNA-binding protein. Evidence is high suggestive but not conclusive for a causal relationship between gene and QTL.
	<i>OsWRKY13</i> ; Blast; Rice; <i>Xanthomonas oryzae</i> pv. <i>oryzicola</i> and <i>Magnaporthe grisea</i> (Bacteria and fungus) ¹²	Candidate gene strategy integrating mapping, expression profiling, and transgenic complementation	Evidence that this gene underlies any specific QTL is not conclusive. <i>OsWRKY13</i> encodes a WRKY-type transcription factor, a class that has often been implicated in plant defense response.
Signal transduction	<i>OsMPK6</i> ; Blast and blast; Rice; <i>Xanthomonas oryzae</i> pv. <i>oryzicola</i> and <i>Magnaporthe grisea</i> (Bacteria and fungus) ¹²	Candidate gene strategy integrating mapping, expression profiling, and transgenic complementation	Evidence that this gene underlies any specific QTL is not conclusive. <i>OsMPK6</i> encodes a mitogen-activated protein kinase - a type of signaling protein often implicated in the defense response
	<i>RKS1</i> ; Black Rot; Arabidopsis; <i>Xanthomonas campestris</i> pv. <i>campestris</i> (bacteria) ¹³	Fine mapping, mutant analysis	<i>RKS1</i> encodes an atypical kinase that mediates a codominant quantitative resistance mechanism in plants by restricting bacterial spread from the infection site.
Signal transduction	<i>Yr36</i> (WKS1); Stripe rust; Wheat; <i>Puccinia striiformis</i> (fungus) ¹⁴	Fine mapping, transgenic complementation and mutant analysis	<i>Yr36</i> (WKS1) is a Kinase-START gene. START is a putative lipid binding domain and the kinase is generally a signaling domain.
	<i>pan1</i> ; Northern leaf blight and Stewarts wilt; <i>Exserohilum turcicum</i> and <i>Pantoea stewartii</i> (Fungus and bacteria) ¹⁵	Fine mapping and mutant analysis.	<i>Pan1</i> is a receptor-like kinase and is known to be involved in actin organization and asymmetric cell division. Mutants are more resistant to NLB and Stewart's wilt. Resistance is dominant. Unlikely to be the only gene involved in the QTL.
Defense response	<i>Pt21</i> ; Rice blast Rice; <i>Magnaporthe oryzae</i> (fungus) ¹⁶	Fine mapping, Transgenic overexpression and silencing.	<i>Pt21</i> encodes a proline-rich protein with a putative heavy metal-binding domain and protein-protein interaction motifs. Wild-type <i>Pt21</i> appears to slow the plant's defense responses, which may support optimization of defense mechanisms. Resistance is recessive
	<i>Lr34</i> ; leaf rust, stripe rust, and powdery mildew; wheat; <i>Puccinia triticina</i> , <i>P. striiformis</i> , <i>Blumeria graminis</i> (all fungi) ¹⁷	Fine mapping, sequence comparisons of resistant and susceptible lines and mutants.	The <i>LR34</i> protein resembles adenosine triphosphate-binding cassette transporters of the pleiotropic drug resistance subfamily. It stimulates senescence-like processes in the flag leaf tips and edges of adult plants. Resistance is codominant
Antimicrobial	<i>Fhb1</i> ; Fusarium head blight; Wheat; Fusarium graminearum (fungus) ¹⁸	Fine-mapping, mutation analysis, gene silencing and transgenic overexpression.	<i>Fhb1</i> encodes pore-forming toxin-like (PFT) gene predicted to encode a chimeric lectin with two agglutinin domains and an ETX/MTX2 toxin domain. Mechanism not clear but form channels in the cell membrane of the host leading to its death.
	<i>OsPAL4</i> ; bacterial blight, sheath blight and blast; Rice; <i>Xanthomonas oryzae</i> pv. <i>oryzae</i> , <i>Rhizoctonia solani</i> , <i>Magnaporthe oryzae</i> (1st bacterium, others fungi) ¹⁹	Analysis of mutant line	<i>OsPAL4</i> encodes phenylalanine ammonia-lyase, the initial enzyme in the phenylpropanoid pathway. Several products of this pathway are associated with resistance, including lignin, salicylic acid and several phytoalexins. <i>OsPAL4</i> co-localizes with resistance QTL and a line with only one copy of this gene shows increase susceptibility to several diseases.
Nutritional restriction	<i>Magnaporthe oryzae</i> / <i>OsDR8</i> ; Blast and blast; Rice; <i>Xanthomonas oryzae</i> pv (bacterium, fungus) ¹²	Candidate gene strategy integrating mapping, expression profiling, and transgenic complementation	<i>OsDR8</i> encodes an enzyme-like protein involved in thiamine biosynthesis. Evidence that this gene underlies any specific QTL is not conclusive.
	<i>Lr67</i> ; Leaf, stripe, stem rusts, powdery mildew; Wheat; <i>Puccinia triticina</i> , <i>P. striiformis</i> f. sp. <i>tritici</i> , <i>P. graminis</i> f. sp. <i>tritici</i> , <i>Blumeria graminis</i> f. sp. <i>tritici</i> (fungi) ²¹	Fine-mapping, comparative genomics, mutagenesis and transformation	<i>Lr67</i> encodes a hexose transporter. Inhibition of apoplastix: hexose retrieval conferred by the resistance alleles may increase the hexose/sucrose ratio in the leaf apoplasm, inducing a sugar-mediated defense response.
Microbial movement	<i>Rlg4</i> ; Soybean cyst; Soybean; <i>Heterodera glycines</i> Ichinohe (nematode) ²²	Fine mapping, mutant analysis, gene silencing, transgenic complementation	<i>Rlg4</i> encodes a serine hydroxymethyltransferase, which interconverts serine and glycine and is essential for cellular one-carbon metabolism. Possibly has a role in restricting pathogen nutrition.
	<i>ZnREM6.3</i> ; Northern leaf blight; Maize; <i>Exserohilum turcicum</i> (fungus) ²⁴	Fine mapping, expression and mutant analysis.	<i>ZnREM6.3</i> encodes a remorin, thought to be part of membrane rafts that play a role in plasmodesmata function and may influence cell-to-cell movement of microbes. Evidence for a causal relationship not conclusive.
Hormone	GH3-2; Blast and blast;R; <i>Xanthomonas oryzae</i> , <i>Magnaporthe oryzae</i> (bacterium, fungus) ²⁶	Mapping, expression and transgenic analysis	GH3-2 encodes an indole-3-acetic acid (IAA)-amido synthetase. Activation of GH3-2 suppresses pathogen-induced accumulation of IAA and IAA-induced loosening of the cell wall. It confers relatively modest levels of resistance.

2017). Roux *et al.* (2014) proposent trois manières de diminuer efficacement l'effet du coût de la multi-résistance tout en maintenant une résistance durable : (i) pyramider des gènes *R* et des QTL de résistance qui sont exprimés dans différents tissus (ii) identifier des gènes associés à une amélioration compensatoire du coût de la résistance multiple (iii) utiliser une stratégie de mosaïque en alternance spatiale des gènes *R* et QDR dans les champs de culture. En plus de ces propositions, Pilet-Nayel *et al.* (2017) indiquent que le choix de QTL pourrait être basé sur leur spectre d'action sur différentes souches, leur effet sur les caractéristiques de vie des agents pathogènes et leurs mécanismes moléculaires sous-jacents (Niks *et al.*, 2015 ; Kushalappa *et al.*, 2016). Ce dernier point nécessite de connaître les bases moléculaires sous-jacentes aux QTL, or pour le moment, les mécanismes moléculaires sous-jacents à la résistance quantitative sont moins connus que ceux impliqués dans la résistance qualitative.

Roux *et al.* (2014) décrivent les mécanismes de résistance quantitative comme faisant intervenir de nombreux QTL dont les gènes sous-jacents interagissent en réseau donnant la possibilité via plusieurs voies de réponse de détecter la présence de différents effecteurs de l'agent pathogène (**Figure I.7**). Les facteurs génétiques pourraient ainsi être i) impliqués dans des fonctions de défense basale, de transduction du signal de défense, ii) des gènes *R* partiellement altérés mais aussi iii) des gènes impliqués dans le développement et la morphologie de la plante hôte (Poland *et al.*, 2009). De plus, l'effet large spectre conféré par les résistances quantitatives pourrait s'expliquer par la présence de gènes jouant un rôle au niveau des voies de perception et de réponse à des éléments spécifiques de la signature microbienne, tels que les effecteurs conservés et les PAMP (Roux *et al.*, 2014). Finalement, la diversité des gènes impliqués dans la résistance quantitative peut s'expliquer par le fait qu'elle doit pouvoir intervenir à différents niveaux de développement de l'agent pathogène. Ceci est décrit notamment pour les champignons pathogènes où l'interaction peut se faire lors de la germination des spores, lors de sa pénétration au niveau de la paroi cellulaire, lors de la colonisation de la plante hôte, lors de sa production de spores... (Niks & Rubiales, 2002).

Chez les végétaux, des analyses de liaison et d'association (Genome Wide Association Studies (GWAS)) sont utilisées pour identifier les QTL qui influencent les phénotypes de résistance (St Clair, 2010). Ces approches ont mené à l'identification de nombreux loci impliqués dans la résistance quantitative. Cependant les QTL de résistance identifiés à travers les approches de liaison incluent généralement plus d'une centaine de gènes dont plusieurs gènes candidats intéressants rendant difficile l'identification du ou des vrais gènes causaux. Cependant, des progrès considérables ont été réalisés ces dernières années dans le clonage de gènes sous-jacents aux QTL de résistance et ont permis de cloner les gènes causaux par une approche de cartographie fine (Nelson *et al.*, 2017). Au sein des gènes identifiés, beaucoup sont similaires à des gènes *NLR*, à des gènes *PRR* ou à des gènes de défense qui peuvent être contrôlés par les *NLR* et les *PRR*. Il existe également d'autres gènes qui ne sont pas apparentés à des gènes *R* mais qui sont impliqués dans des voies connues pour être associées aux

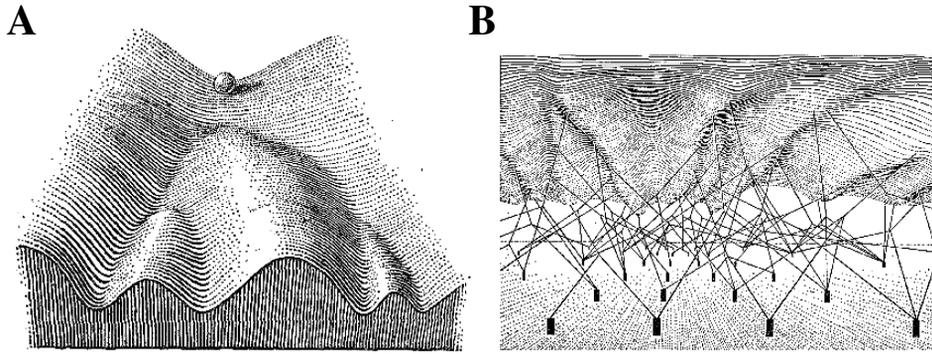


Figure I.8. Représentation de Waddington (1957) de son paysage épigénétique. (A) La bille représente un système biologique (par exemple une cellule) sur le point de suivre un ensemble d'états alternatifs plus différenciés représentés par les trois dépressions terminales à la base de la pente. (B) Une vision derrière un paysage épigénétique. La forme de la pente est déterminée par la tension de plusieurs haubans interconnectés (produits de gènes en interaction) qui sont attachés à des piquets coincés dans le sol (gènes). Issue de Fusco *et al.* (2014).

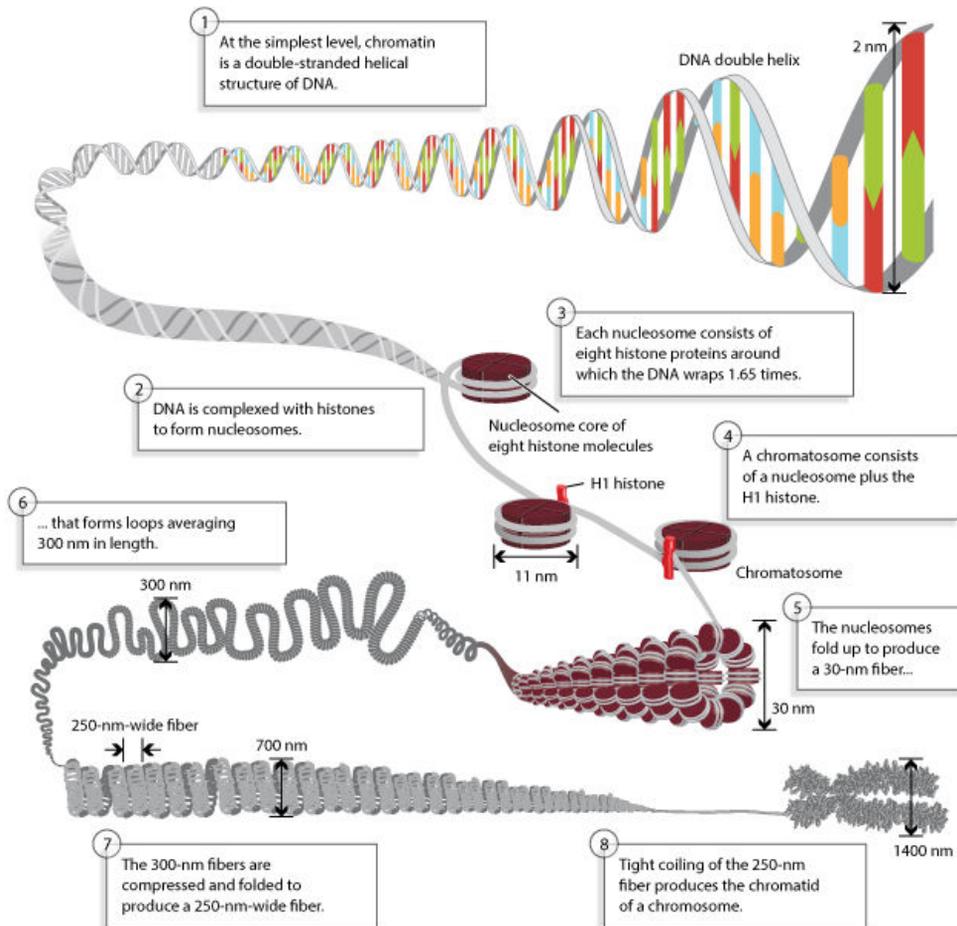


Figure I.9. Représentation de la structure de la chromatine. La molécule d'ADN libre d'un diamètre de 2 nm (1) se retrouve dans un premier temps entourée autour d'un octamère de protéines histones (2 H2A, 2 H2B, 2 H3, 2 H4) formant ainsi le nucléosome (2, 3). Les nucléosomes sont "verrouillés" par l'apposition de la protéine histone H1 formant ainsi un chromatosome (4). Par la suite, les nucléosomes se replient pour former une fibre de chromatine de 30 nanomètres, qui forme des boucles d'une longueur moyenne de 300 nanomètres (5,6). Finalement, Les fibres de 300 nm sont comprimées et pliées pour produire une fibre de 250 nm de largeur, qui est étroitement enroulée dans la chromatide d'un chromosome. D'après © 2013 Nature Education, adapted from Pierce, Benjamin. Genetics: A Conceptual Approach, 2nd ed. All rights reserved.

défenses des plantes, ainsi que des gènes qui n'avaient jusqu'à présent pas été associés à des résistances aux maladies ou à des voies de défense (**Tableau I.1**).

De manière générale, la résistance quantitative aux agents pathogènes des plantes est supposée comme issue de la combinaison de variations alléliques dues à des variations nucléotidiques au niveau de plusieurs loci. Cependant, les travaux récents menés en épigénétique, notamment au niveau de la méthylation de l'ADN, montrent l'impact de variations de l'état de méthylation de l'ADN au niveau du génome sur la régulation de l'expression des gènes, la plasticité développementale, les interactions avec l'environnement y compris des interactions avec des stress.

2. L'implication de l'épigénétique dans la réponse des plantes aux stress

L'information génétique codée par l'ADN est empaquetée dans une structure appelée chromatine. Cette dernière est sujette à de nombreuses modifications qui peuvent ou non changer la séquence nucléotidique de l'ADN. L'ensemble des modifications apposées peuvent alors avoir un impact sur l'expression de gènes. L'étude des mécanismes de régulation de l'expression des gènes a montré que ces derniers en plus d'être sous la dépendance de régions régulatrices (activatrices et inhibitrices) faisant intervenir des séquences nucléotidiques en amont (lieu de fixation de facteurs modifiant l'accès de la machinerie transcriptionnelle ; Jacob & Monod, 1961), sont sous la dépendance d'un autre niveau de régulation. Ce niveau supplémentaire, dit épigénétique (Silveira *et al.*, 2013), régule la structure de la chromatine (Roudier *et al.*, 2011 ; Ha, 2013).

2.1. Définition générale de l'épigénétique

La première définition de l'épigénétique a été donnée par Waddington en 1942 comme étant les changements de phénotypes observables sans changement du génotype (**Figure I.8** ; Waddington, 1942). Depuis cette définition historique de l'épigénétique, d'autres définitions plus complètes ont été développées. Dans ce manuscrit nous définirons l'épigénétique comme l'ensemble des processus modulant l'expression des gènes sans changer la séquence nucléotidique, par l'apposition de marques dites épigénétiques, transmissibles par mitose et/ou méiose (Johannes *et al.*, 2008). Les marques épigénétiques sont nombreuses et concernent les différents éléments constituant la chromatine tels que l'ADN et les protéines histones. Que ce soit dans le cadre des modifications touchant l'ADN ou les protéines histones, différents types de protéines sont responsables de la mise en place de ces marques épigénétiques chez les eucaryotes.

2.2. Epigénétique et structure de la chromatine

La chromatine est constituée d'un enchainement de nucléosomes qui sont issus de l'enroulement de 147 paires de bases (pb) d'ADN autour d'un octamère de protéines histones composé de 2 copies de chacune des histones H2A, H2B, H3 et H4 (Luger *et al.*, 1997). Ce nucléosome est verrouillé par l'apposition d'une autre protéine basique, l'histone H1 au niveau de l'ADN libre appelé ADN « linker ». Les

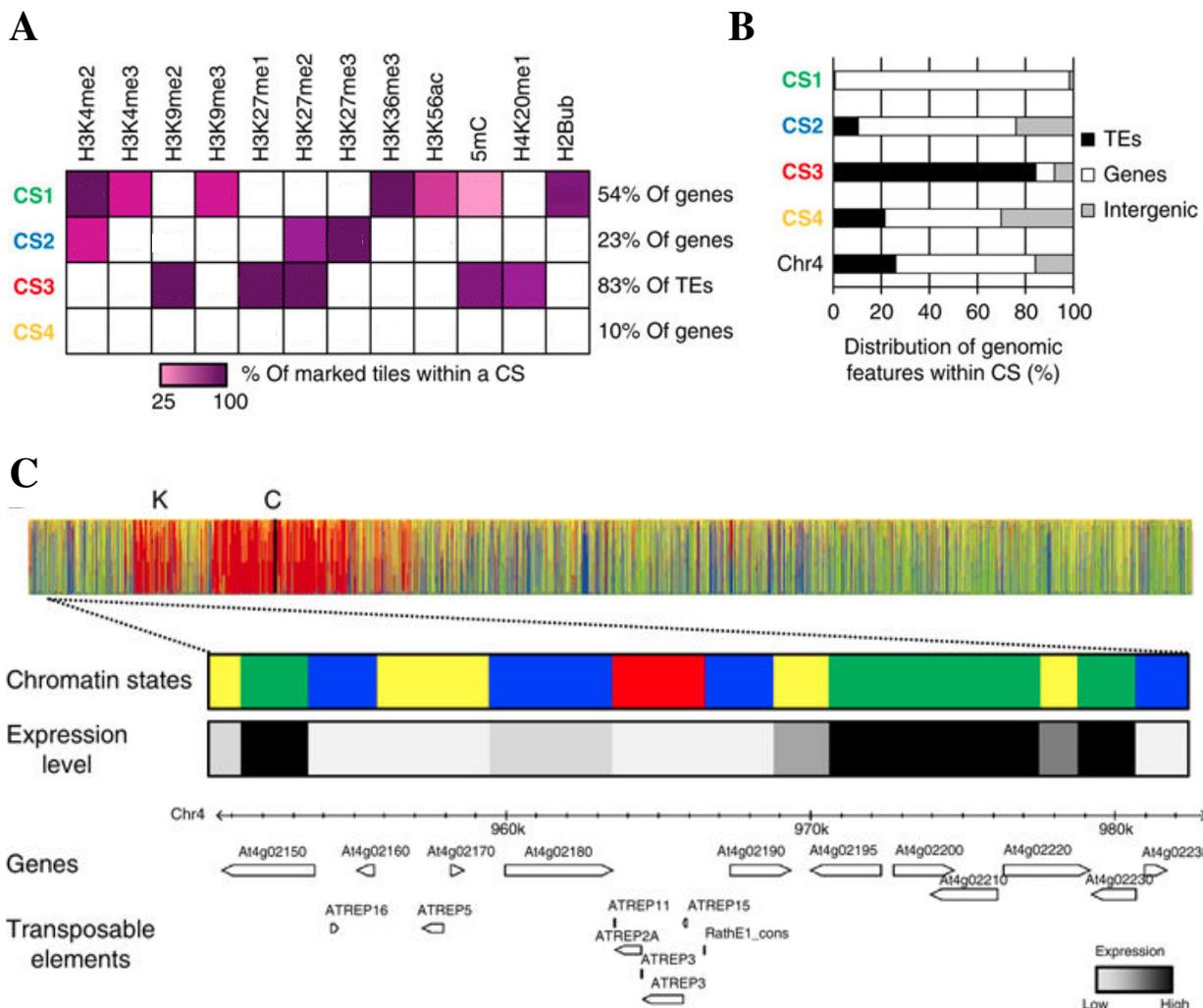


Figure I.10. Présentation des 4 états chromatinien identifiés au niveau de l'épigénome d'*Arabidopsis*. (A) Composition en marques épigénétiques des 4 états prédominants de la chromatine (CS) identifiés chez *Arabidopsis*. Chaque ligne donne la composition en marque d'un CS (CS1 à CS4). Chaque colonne correspond à l'une des 12 marques épigénétiques utilisées lors de l'étude. Une coloration des cases allant du violet clair (25 %) au violet foncé (100 %) indique la distribution des marques entre les 4 CS. Le pourcentage de gènes indexés par CS1, CS2 et CS4 et le pourcentage d'annotations TE indexés par CS3 sont également indiqués à droite. (B) Proportion relative de chaque élément génomique dans chacun des CS. (C) Représentation des différents états chromatinien le long du chromosome 4 d'*Arabidopsis*. Chaque CS est représenté avec la même couleur que celle utilisée dans les figures A et B. Le centromère est indiqué par un C et le knob hétérochromatinien par un K. La partie non séquencée du centromère est représentée par un ligne noire verticale. Le niveau d'expression des différents éléments et l'état chromatinien où ils se trouvent au niveau d'une portion du chromosome 4 sont représentés dans la partie inférieure de la figure. La coloration grise à noire du niveau d'expression indique des niveaux faibles à des niveaux fort d'expression. Adaptée de Roudier *et al.* (2011).

nucléosomes ainsi formés sont liés entre eux par de l'ADN « linker » afin de former la chromatine. La chromatine se retrouve par la suite condensée sous différents états par repliements successifs menant à la formation de chromatides puis de chromosomes (**Figure I.9**).

La chromatine est sujette à différentes modifications telles que la méthylation de l'ADN, les modifications de l'extrémité N terminale des protéines histones, l'incorporation de variants histones et à l'action de divers facteurs tels que des protéines appelées facteurs de remodelage de la chromatine et les petits ARN. L'ensemble de ces processus mènent à l'établissement de différents états de la chromatine qui régulent l'accès à l'ADN (Kouzarides, 2007 ; Berger, 2007). La transcription (Lee *et al.*, 2010) et la réplication du génome (Suzanne Dorn & Cook, 2011) sont étroitement dépendants de l'accessibilité à l'ADN. La chromatine se retrouve naturellement dans le noyau des cellules eucaryotes sous trois formes appelées euchromatine, hétérochromatine constitutive et hétérochromatine facultative. L'euchromatine est l'état de la chromatine qui est associé à des gènes actifs (exprimés) et l'hétérochromatine à des gènes inactifs (non exprimés).

Chaque état de la chromatine est étroitement lié au statut épigénétique des histones et de l'ADN la constituant. Différentes marques épigénétiques sont aujourd'hui connues au niveau de l'ADN et des protéines histones. Au niveau de protéines histones ces marques sont apposées de manière post-traductionnelle. Les combinaisons de marques histones, des variants histones et de la méthylation de l'ADN sont essentielles dans l'établissement des différents états chromatiniens qui régissent l'accessibilité ou non aux différents éléments génomiques de l'ADN.

En fonction des combinaisons de ces marques, l'épigénome conduirait à la mise en place de différents états chromatiniens. Des travaux réalisés chez *A. thaliana* confirment l'effet des marques épigénétiques sur la structure de la chromatine. Roudier *et al.* (2011) ont montré en utilisant 12 marques chromatiniennes couvrant environ 90 % du génome d'*Arabidopsis*, que l'épigénome conduirait à la mise en place de 4 états prédominants de la chromatine nommés de CS1 à CS4 (**Figure I.10A**). Dans chacun de ces états CS, on retrouverait des régions génomiques spécifiques. Par exemple, l'état CS3 est constitué majoritairement d'éléments transposables (TE) et les états CS1 et CS2 sont majoritairement associés aux gènes (**Figure I.10B**). L'état CS1 est plutôt associé à des régions transcriptionnellement actives et le CS4 à des régions transcriptionnellement réprimées (**Figure I.10C** ; Roudier *et al.*, 2011).

Une autre étude plus récente réalisée par Sequeira-Mendes *et al.* (2014) sur la totalité du génome d'*Arabidopsis* a montré l'existence de 9 états chromatiniens distincts en utilisant 11 modifications histones, la méthylation de l'ADN en CG, l'occupation des nucléosomes et trois variants histones. L'état chromatiniens 1 contient un large set de modifications d'histones associées à un état de la chromatine ouvert/actif et l'état 9 correspond aux régions hétérochromatiniennes péri-centromériques qui sont dans un état réprimé (**Figure I.11A**). Comme dans l'étude précédente, une répartition

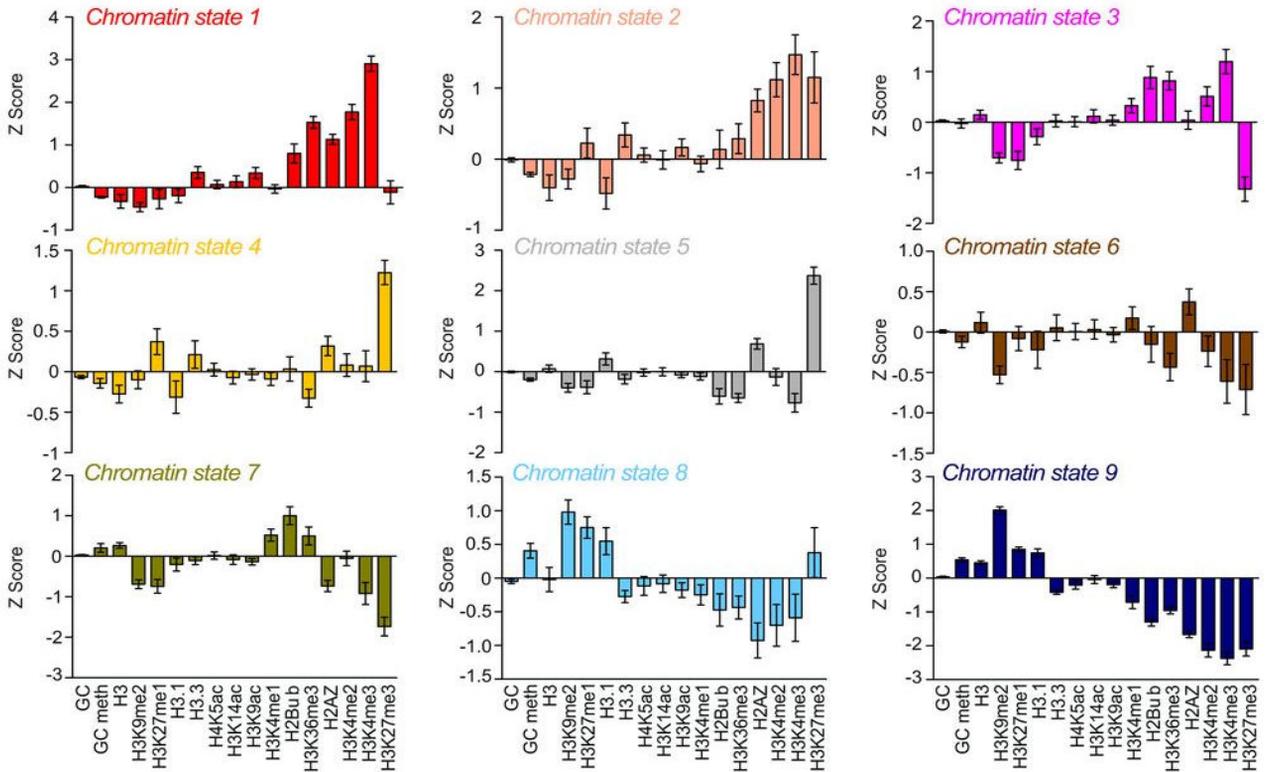
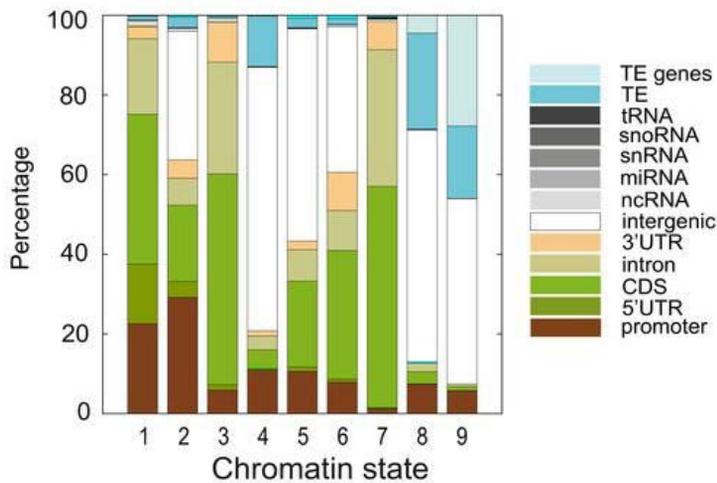
A**B**

Figure I.11. Définition et localisation au niveau des éléments du génome des 9 états chromatinien identifiés par Sequeira-Mendes *et al.* (2014). (A) Graphique des 9 états de la chromatine définis par une combinaison spécifiques de marques. Les scores z positifs et négatifs présents sur l'axe y indiquent respectivement des valeurs supérieures ou inférieures à la moyenne dans le génome de chacune des marques de la chromatine considérées. Les barres d'erreur représentent l'écart type à la moyenne. (B) Relation entre les différents éléments génomiques et les états chromatinien. Le chevauchement (en pb) entre les éléments génomiques étudiés et chaque état chromatinien a été calculé puis exprimé en pourcentage (indiqué sur l'axe y). Pour l'élément promoteur, une région de 0,65 kb a été utilisée. Dans cette figure, la classe TE fait référence aux régions génomiques qui contiennent des TE mais ne se chevauchent pas avec les gènes de type TE. Adaptée de Sequeira-Mendes *et al.* (2014).

préférentielle des différents éléments génomiques à l'intérieur de chaque état chromatinien a été décelée. Les états 1 à 3 sont très similaires au niveau des marques histones mais diffèrent au niveau des proportions dans lesquelles elles sont présentes dans les différents éléments du génome (5'UTR, CDS...). L'état chromatinien 4 contient en grande majorité les régions non codantes intergéniques. L'état chromatinien 5 correspond à des régions chromatinienne régulées par le complexe polycomb avec un haut niveau de marque H3K27me3 (cf. paragraphe 2.3.2) et un niveau modéré H2A.Z similaire à celui identifié par Roudier *et al.* (2011). Les états chromatiniens 6 et 7 sont typiques des régions intragéniques avec un enrichissement en marque H2A.Z, une forte densité en nucléosomes. En plus, l'état 6 possède des marques H3K4me1 typiques des méthylations corps gène. L'état chromatinien 8 se situe préférentiellement dans les régions intergéniques et dans les éléments transposables et enfin l'état chromatinien 9 correspond aux régions hétérochromatiniennes péri-centromériques majoritairement situées au niveau des régions intergéniques et des éléments transposables (**Figure I.11B**).

2.3. Les variants et les modifications post traductionnelles histones

Différentes modifications post-traductionnelles sont possibles au niveau de la queue des protéines histones telles que la méthylation, l'acétylation, la phosphorylation, l'ubiquitination, la sumoylation, l'ADP-ribosylation et la citrullination (Kouzarides, 2007 ; Berger, 2007). L'ensemble de ces modifications mènent à différents types de régulation qui ont été proposés comme faisant partie du code appelé le « code histone » (Jenuwein & Allis, 2001 ; Kouzarides, 2007).

2.3.1. Les variants d'histones

Les variants d'histones sont des isoformes des protéines histones qui diffèrent par des variations au niveau d'acides aminés spécifiques. La mise en place des variants d'histones au niveau du génome est régulée par des protéines chaperonnes telles que les protéines Sga1/2, Fas1, Msi1, HIRA pour les variants des histones H3-H4 et les protéines NRP1/2 pour les variants des histones H2A/H2B (Gurard-Levin *et al.*, 2014). La plupart des variants d'histones décrits chez les plantes sont ceux de la famille des histones H2A et H3. En effet, les membres de la famille H4 présentant les mêmes séquences en acides aminés et seulement quelques variants ont été décrits pour la famille H2 (Yang Hao *et al.*, 2016).

Chez *Arabidopsis* il existe 3 types majeurs de variants de l'histone H2A, les variants H2A.X et H2A.Z, présents chez les plantes et les animaux, et le variant H2A.W qui est spécifique des plantes. Les nombreuses variations en acides aminés sont en majorité localisées au niveau de trois régions de la protéine histone, la « P1 loop », le « docking domain » et la queue C terminale qui sont notamment impliqués dans l'interaction avec les autres histones influençant ainsi la stabilité et la compaction des nucléosomes (Kawashima *et al.*, 2015, **Figure I.12A**). Le variant H2A.W est très fréquent au niveau des régions hétérochromatiniennes, des régions enrichies en marques H3K9me2 et des éléments transposables (**Figure I.12B**) où il semble agir en association avec la méthylation de l'ADN afin de réprimer l'expression de ces derniers (Bernatavichute *et al.*, 2008 ; Yelagandula *et al.*, 2014). Chez

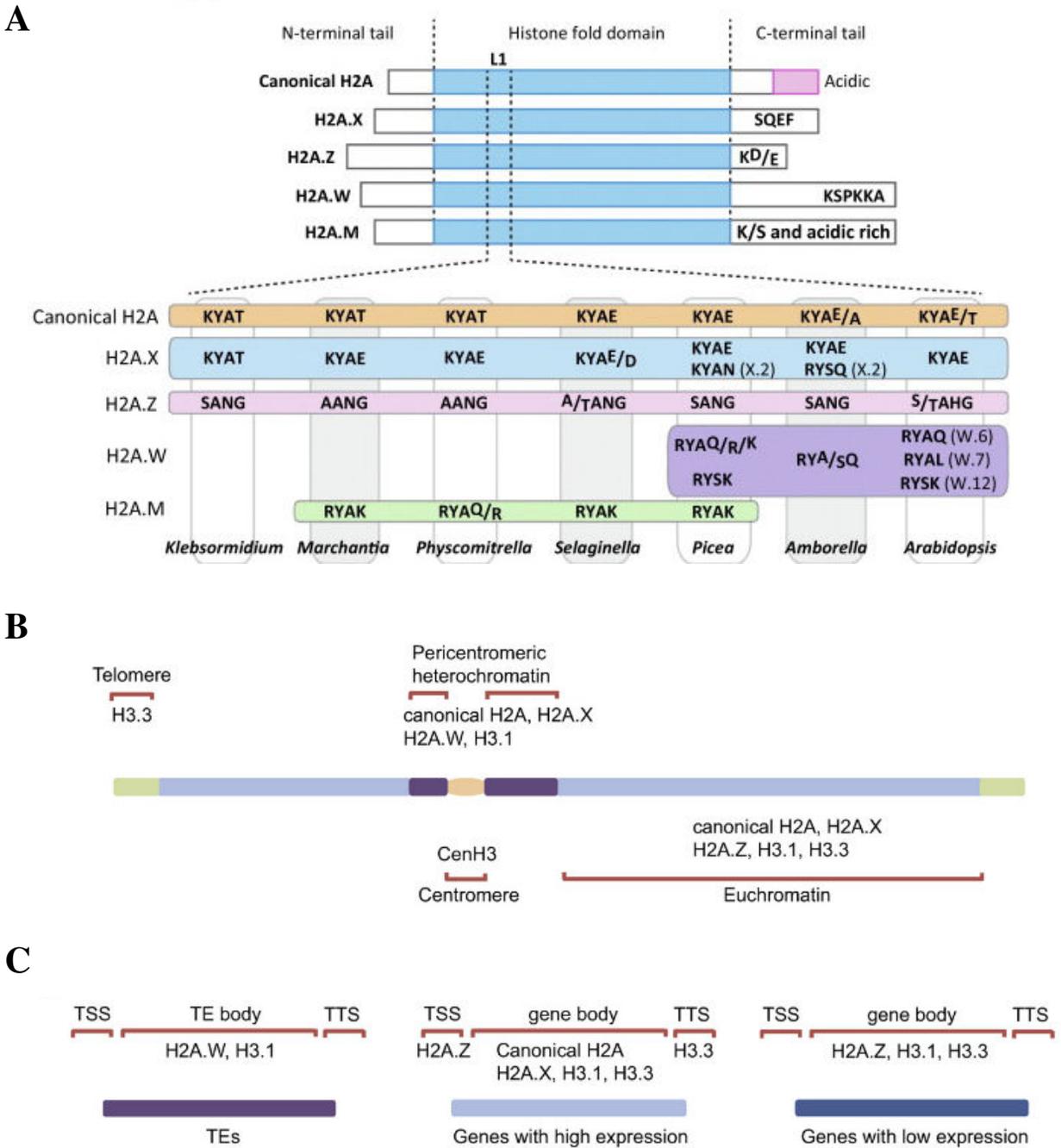


Figure I.12. Présentation des variants H2A et localisation préférentielle des variants d'histones au niveau du génome. (A) Illustration des variants histones H2A et de leurs motifs spécifiques au niveau de la boucle L1 et de l'extrémité C terminale. Les queues N et C-terminales sont indiquées par différentes tailles de rectangle blanc afin de mettre en évidence les différences de longueur. Les motifs spécifiques de chaque variants chez les plantes à fleur sont indiqués à l'intérieur de leur extrémité C terminale. Les motifs spécifiques de chaque variant H2A au niveau de la boucle L1 sont indiqués dans la partie inférieure du schéma avec le nom de l'espèce dans laquelle ils ont été identifiés. (B) Localisation préférentielle des variants d'histones. En raison des propriétés différentes des variants d'histones, ils se localisent souvent de manière préférentielle dans certaines régions de la chromatine, notamment le centromère, l'hétérochromatine péri-centromérique, l'euchromatine et le télomère. (C) Représentation de la présence et de la localisation des variants d'histones au niveau des TE et des gènes. La présence de variants d'histones dans différentes régions de gènes (par exemple, TSS, corps de gène, TTS) peut influencer de manière différente l'activité transcriptionnelle de ces derniers. Adaptée de Kawashima *et al.* (2015) et de Jiang et Berger (2017).

Arabidopsis, le variant H2A.Z est principalement associé aux régions euchromatiniennes et est très fréquent notamment au niveau du site initiateur de la transcription (Transcriptional Start : TSS) des gènes (Zilberman *et al.*, 2008 ; Coleman-Derr & Zilberman, 2012 ; Yelagandula *et al.*, 2014) promouvant ainsi l'élongation transcriptionnelle. Ce variant est également présent au niveau du corps des gènes, qui se traduit dans ce cas, par une corrélation négative entre sa présence et le niveau d'expression des gènes (Coleman-Derr & Zilberman, 2012, **Figure I.12C**). L'exemple le plus connu d'un gène d'*Arabidopsis* nécessitant la présence du variant H2A.Z aux extrémités de la région transcrite pour l'activation de sa transcription est le gène *FLC* (*FLOWERING LOCUS*) qui est responsable de la répression de la transition florale (Sheldon *et al.*, 1999 ; Deal *et al.*, 2007).

L'histone H3 présente trois types majeurs de variants, H3.1, H3.3 et CenH3. Le variant H3.1 est présent au niveau des régions enrichies en marques histones H3K9me2 et H3K27me3 suggérant leur association aux régions réprimées transcriptionnellement aussi bien au niveau de l'hétérochromatine que de l'euchromatine (Zhang *et al.*, 2007a ; Stroud *et al.*, 2012 ; **Figure I.12B**). Le variant H3.3 est enrichi dans les régions euchromatiniennes au niveau du site de terminaison de la transcription (TSS) des gènes transcrits où sa présence corrèle positivement avec l'expression des gènes (Stroud *et al.*, 2012 ; Wollmann *et al.*, 2012 ; **Figure I.12C**). Par ailleurs la présence de ce variant au niveau de la région promotrice des gènes est le plus souvent associée avec la marque H3K27me3 sur le corps du gène menant à l'inactivation du gène (Shu *et al.*, 2014). Le variant cenH3 qui diverge des autres variants au niveau de sa queue N terminale est présent au niveau des régions centromériques où il est impliqué dans l'assemblage des kinétochores et dans la ségrégation des chromosomes durant la division nucléaire (Ravi *et al.*, 2011 ; Fukagawa & Earnshaw, 2014, **Figure I.12B**).

Finalement l'histone H1 présente 3 variants chez *Arabidopsis*, les variants H1.1 et H1.2 qui sont exprimés de manière ubiquitaire et sont capables de bloquer l'accès de l'ADN aux méthyltransférases et le variant H1.3 qui est notamment induit lors de stress hydriques (Ascenzi & Gantt, 1997 ; Zemach *et al.*, 2013).

2.3.2. Les modifications post traductionnelles des histones

Différentes modifications d'histones sont impliquées dans la régulation de diverses fonctions de l'ADN telles que sa transcription, sa réparation, sa réplication et sa condensation (**Tableau I.2A**). Parmi l'ensemble des modifications possibles, la méthylation et l'acétylation des histones sont les deux marques particulièrement étudiées.

Les protéines histones peuvent être acétylées ou désacétylées au niveau de leur extrémité N terminale menant respectivement à un effet activateur ou désactivateur de l'expression des gènes (Allfrey *et al.*, 1964). L'acétylation des résidus lysine de l'extrémité N terminale des histones est réalisée par des histones acétyltransférases (HAT) qui sont regroupées au sein de quatre familles chez les plantes :

Tableau I.2. Liste des modifications présentes au niveau des histones et listes des gènes encodant les protéines intervenant dans deux modifications d'histones. (A) Différentes classes de modifications identifiées sur les histones ; Issue de Kouzarides (2007). (B) Liste des gènes codant les homologues d'histones acétyltransférases (HAC) et d'histones désacétylases (HDAC) chez *Arabidopsis*. Issue de Pandey *et al.* (2002).

A

Chromatin Modifications	Residues Modified	Functions Regulated
Acetylation	K-ac	Transcription, Repair, Replication, Condensation
Methylation (lysines)	K-me1 K-me2 K-me3	Transcription, Repair
Methylation (arginines)	R-me1 R-me2a R-me2s	Transcription
Phosphorylation	S-ph T-ph	Transcription, Repair, Condensation
Ubiquitylation	K-ub	Transcription, Repair
Sumoylation	K-su	Transcription
ADP ribosylation	E-ar	Transcription
Deimination	R > Cit	Transcription
Proline Isomerization	P-cis > P-trans	Transcription

Overview of different classes of modification identified on histones. The functions that have been associated with each modification are shown. Each modification is discussed in detail in the text under the heading of the function it regulates.

B

HDAC and HAT gene family	<i>Arabidopsis</i> gene name (synonym)	MIPS accession no.	BAC clone and genomic locus	Chromosome
RPD3/HDA1	<i>HDA2</i>	At5g26040	T1N24.9	V
	<i>HDA5</i>	At5g61060	MAF19.7	V
	<i>HDA6 (AtRPD3B)</i>	At5g63110	MDC12.7	V
	<i>HDA7</i>	At5g35600	K2K18.5	V
	<i>HDA8</i>	At1g08460	T27G7.7	I
	<i>HDA9</i>	At3g44680	T18B22.80	III
	<i>HDA10^a</i>	At3g44660	T18B22.60	III
	<i>HDA14</i>	At4g33470	F17M5.230	IV
	<i>HDA15</i>	At3g18520	MYF24.23	III
	<i>HDA17^a</i>	At3g44490	F14L2	III
	<i>HDA18</i>	At5g61070	MAF19.8	V
	<i>HDA19 (AtRPD3A)</i>	At4g38130	F20D10.250	IV
	HD2	<i>HDT1 (AtHD2A)</i>	At3g44750	T32N15.8
<i>HDT2 (AtHD2B)</i>		At5g22650	MDJ22.7	V
<i>HDT3</i>		At5g03740	F17C15.160/MED24.1	V
<i>HDT4</i>		At2g27840	F15K20.6	II
SIR2	<i>SRT1</i>	At5g55760	MDF20.20	V
	<i>SRT2</i>	At5g09230	T5E8.30	V
GNAT	<i>HAG1 (atGCN5)</i>	At3g54610	T14E10.180	III
	<i>HAG2</i>	At5g56740	MIK19.19	V
	<i>HAG3</i>	At5g50320	MXI22.3	V
MYST	<i>HAG4</i>	At5g64610	MUB3.13	V
	<i>HAG5</i>	At5g09740	F17I14.70/MTH16.20	V
CBP	<i>HAC1</i>	At1g79000	YUP8H12R.3	I
	<i>HAC2</i>	At1g67220	8 F1N21.4	I
	<i>HAC4</i>	At1g55970	F14J16.27/T6H22.23	I
	<i>HAC5</i>	At3g12980	MGH6.9	III
	<i>HAC12</i>	At1g16710	F17F16.21	I
TAF _{II} 250	<i>HAF1</i>	At1g32750	F6N18.20	I
	<i>HAF2</i>	At3g19040	K13E13.16	III

^aGenes with partial HDAC domains.

GNAT (General control nondepressible 5 (GCN5)-related N-acetyl transferase), MYST (MOZ-YBF2/SAS3-SAS2/TIP60), CBP (cAMP-responsive element-binding protein) et TAF_{II}250 (transcription initiation factor TFIID 250 kDa subunit ; Pandey *et al.*, 2002 ; **Tableau I.2B**). Le transfert du groupement acétyl sur le résidu lysine de l'extrémité N terminale des histones neutralise la charge positive du groupement amino modifiant l'interaction entre l'ADN et la protéine histone (Yang & Seto, 2008). La désacétylation des histones est assurée par des histones désacétylases (HDAC) regroupées dans trois familles : RPD3/HDA1 (Reduced Potassium Dependancy 3/Histone DeAcetylase 1) et SIR2 (pour Silent Information Regulator 2) qui sont présentes chez tous les eucaryotes (Pandey *et al.*, 2002) et HD2 qui est spécifique des plantes (Grandperret *et al.*, 2013 ; **Tableau I.2B**). L'action des HDAC conduit à l'accumulation des histones sous une forme hypo-acétylée menant à la compaction de la chromatine sous forme d'hétérochromatine et à l'inhibition de l'expression des gènes.

Parallèlement, la méthylation des protéines histones peut être effectuée sur les résidus lysine ou arginine de leur extrémité N terminale menant à différentes formes mono, di ou triméthylée pour la lysine et mono ou diméthylée pour l'arginine (Kouzarides, 2007), la majorité des modifications étant réalisées au niveau des résidus lysine. Cinq sites de méthylation ont été identifiés chez les plantes : sur les lysines 4, 9, 27 de l'histone H3 et les lysines 3 et 20 de l'histone H4 (Zhang *et al.*, 2007b ; Pfluger & Wagner, 2007). Les marques de méthylation sont apposées par des histones lysine méthyltransférases (HKMT) qui appartiennent à la super famille des protéines SDG (SET Domain Group) car possédant un domaine SET (Suppressor of variegation, Enhancer of zeste ans Trithorax) qui est responsable de leur activité catalytique (Gendler *et al.*, 2008). Les 31 protéines SDG identifiées chez *A. thaliana* présentant une activité HKMT ont été divisées en 5 classes en fonction de leur activité enzymatique et ou de l'architecture de leurs domaines (Baumbusch *et al.*, 2001 ; Springer *et al.*, 2003 ; Ng *et al.*, 2007) : i) la classe I contient des HKMT qui codent chez *Arabidopsis* trois protéines homologues de l'« Enhancer of Zeste E(Z) » de la Drosophile qui sont CURLY LEAF (CLF), MEDEA (MEA) et SWINGER (SWN). Ces trois protéines ont une activité de méthyltransférase sur la lysine 24 de l'histone H3 (H3K24) et sont des composants du complexe polycomb 2 (PRC2) qui a un rôle dans la répression de la transcription des gènes chez différents eucaryotes (Baroux *et al.*, 2007). ii) La classe II contient des HKMT qui codent des protéines impliquées majoritairement dans la méthylation de la lysine 36 de l'histone H3 (H3K36) qui est une marque enrichie au niveau des régions contenant des gènes transcrits (Lee & Shilatifard, 2007). iii) La classe III comprend des HKMT qui codent 5 gènes homologues à Trithorax qui sont les gènes *Arabidopsis Trithorax like proteins* de 1 à 5 (ATX1-5; Avramova, 2009). Parmi les 5 ATX, ATX1 conduit la triméthylation de la lysine 4 de l'histone H3 (H3K4me3) alors que ATX2 conduit à sa diméthylation (H3K4me2 ; Pien *et al.*, 2008 ; Saleh *et al.*, 2008). iv) La classe IV englobe les HKMT ATXR5 et ATXR6 qui sont impliquées dans la monométhylation de la lysine 27 de l'histone H3 (H3K27me1) qui est une marque présente au niveau de l'hétérochromatine des chromocentres (Jacob *et al.*, 2009). v) Enfin la classe V des HKMT est composée de deux groupes de protéines, les

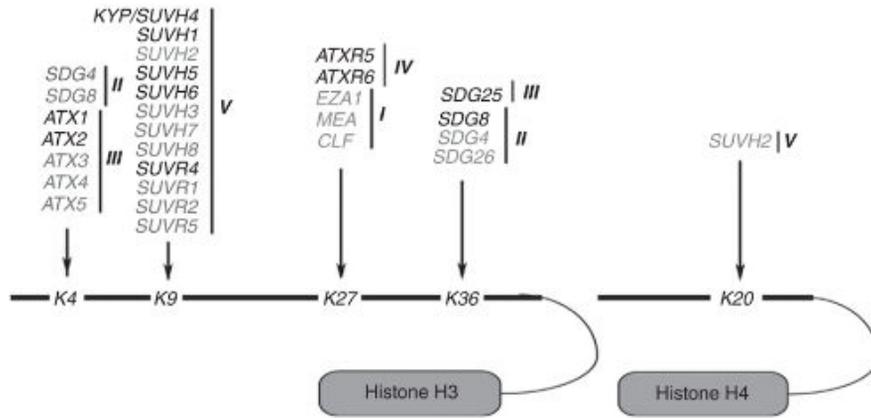
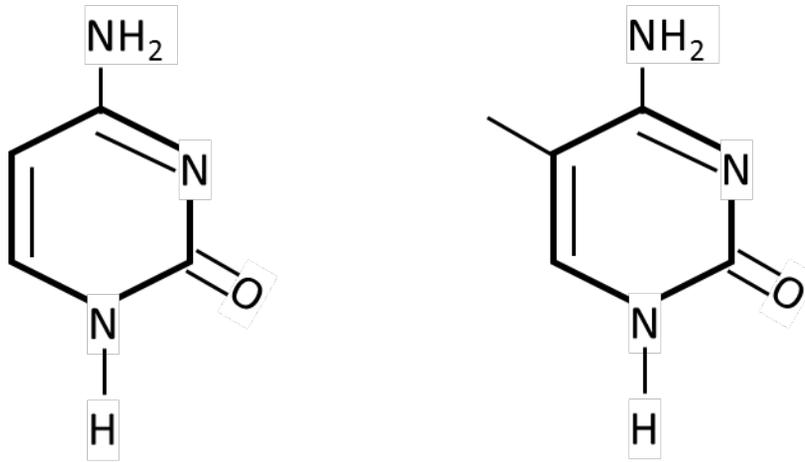


Figure I.13. Schéma représentant les lysines des histones H3 et H4 ciblées par les HKMT chez *Arabidopsis*. Les HKMT dont l'activité a été démontrée biochimiquement sont mises en évidence en utilisant le noir, tandis que les HKMT dont les spécificités n'ont pas été confirmées sont indiquées en gris. Adaptée de Pontvianne *et al.* (2010).

A



Cytosine

5 -Méthylcytosine

B

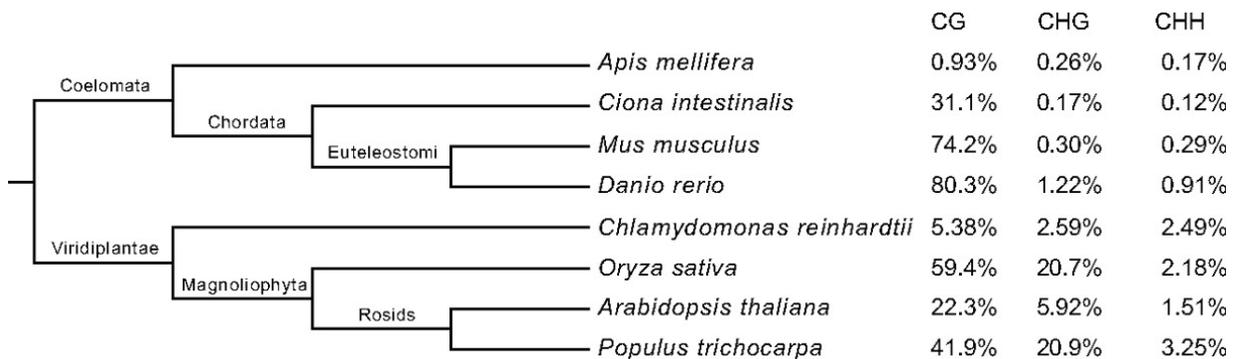


Figure I.14. Représentation graphique de la cytosine méthylée et niveau de méthylation des cytosines chez différents eucaryotes. (A) Représentation graphique de la cytosine et de sa forme modifiée la 5-méthylcytosine. (B) Exemples du niveau de méthylation de l'ADN présent chez 8 eucaryotes. Les taux de méthylation donnés en CG, CHG et CHH ont été calculés à partir des principaux chromosomes/groupes de liaison ou scaffolds et de différents tissus chez chacun des eucaryotes.

protéines SUVH (SU (VAR)3-9 homolog) et les protéines SUVR (SU (VAR)3-9 related). Parmi les protéines SUVH, SUVH4 appelée aussi KRYPTONITE a été montrée comme étant impliquée dans la répression de l'expression des gènes et le maintien de la méthylation de l'ADN en contexte CHG en méthylant la lysine 9 de l'histone H3 (H3K9 ; Jackson *et al.*, 2002 ; **Figure I.13**). Quelques auteurs suggèrent que SUVH2 méthyle la lysine 20 de l'histone H4 (H4K20) et la lysine 9 de l'histone H3 (H3K9 ; Naumann *et al.*, 2005). Cinq gènes codant des protéines SUVR (SUVR1-5) ont été identifiés chez *A. thaliana*. Parmi les protéines codées par ces gènes, SUVR4 présente une activité de mono ou diméthylation de la lysine 9 de l'histone H3 (H3K9me1 ou H3K9me2) suggérant un effet répressif des membres de cette famille sur l'expression des gènes (Thorstensen *et al.*, 2006).

Au sein des marques de méthylation apposées sur les histones, les marques H3K4me2, H3K4me3 et H3K36me3 sont associées à des régions transcrites et au site d'initiation de la transcription (TSS) des gènes alors que les marques H3K27me1 et H3K27me3 sont associées à des régions réprimées (Sequeira-Mendes *et al.*, 2014). La marque de méthylation H3K9me2 apposée par les méthyltransférases de la famille SUVH est une bonne illustration de l'interconnexion existant entre les modifications post traductionnelles des histones et la méthylation de l'ADN. Ainsi le mutant de la cytosine méthyltransférase MET1 montre une diminution de la marque H3K9me2 tandis que le mutant de la protéine SUVH4 impliquée dans l'apposition de la marque H3K9me2 présente une diminution de la méthylation des cytosines de l'ADN (Jackson *et al.*, 2002 ; Tariq *et al.*, 2003).

2.4. La méthylation de l'ADN

Actuellement la méthylation de l'ADN est la marque épigénétique la plus étudiée du fait de i) la présence de mécanismes permettant sa transmissibilité, ii) son importance pour la régulation de l'expression des gènes, iii) son rôle dans le silencing des éléments transposables, iv) son rôle dans la stabilité du génome et v) la facilité de son étude à l'échelle d'un génome entier.

Chez les eucaryotes, la méthylation s'effectue sur le 5^{ème} carbone de la cytosine donnant la 5-méthylcytosine (Vanyushin *et al.*, 1970 ; **Figure I.14A**). Cette modification site spécifique de l'ADN permet le recrutement de protéines particulières appelées « readers » qui vont influencer la structure du génome et l'expression des gènes au locus considéré. Ce motif de méthylation est incorporé ou retiré de manière dynamique contribuant ainsi de manière reprogrammable et sophistiquée à la régulation de la structure du génome et à la modulation de l'accessibilité à la séquence génomique (Raiber *et al.*, 2017).

La catalysation de la réaction de méthylation des cytosines au niveau de l'ADN est réalisée par des enzymes appelées méthyltransférases de l'ADN (DNMT) qui utilisent la S-adenosylméthionine (SAM) comme donneur de groupement méthyl. Chez les plantes, la réaction de méthylation des cytosines est catalysée au niveau de trois contextes nucléotidiques CG, CHG et CHH (où le H peut être une Adénine,

une Cytosine ou une Thymine ; Henderson & Jacobsen, 2007). Chez *A. thaliana*, le niveau de méthylation dans chacun des contextes CG, CHG et CHH est respectivement d'environ 24 %, 6,7 % et 1,7 % (Cokus *et al.*, 2008). Cependant le taux de méthylation dans chaque contexte est variable d'une espèce à l'autre avec par exemple un taux de méthylation globalement plus élevé chez *Oryza sativa* (CG: 59,4 %, CHG: 20,7 %, CHH: 2,18 %) et globalement plus faible chez *Chlamydomonas reinhardtii* (CG: 5,38 %, CHG: 2,59 %, CHH: 2,49 %) que chez *A. thaliana* (Feng *et al.*, 2010a ; **Figure I.14B**). La suppression d'une marque de méthylation apposée au niveau d'une cytosine méthylée peut se faire de manière passive, *i.e.* sans intervention d'enzyme, ou de manière active, *i.e.* par l'action d'une réaction enzymatique (Law & Jacobsen, 2010).

A contrario des eucaryotes, chez les procaryotes la méthylation de l'ADN a été rapportée comme s'effectuant de manière majoritaire sur l'azote six de l'adénine donnant des N⁶-méthyladénine (6 mA ; Dunn & Smith, 1958 ; Vanyushin *et al.*, 1968). Jusqu'à présent, cette modification n'a été montrée que chez un nombre réduit d'eucaryotes tels que l'algue verte *Chlamydomonas* (Hattman *et al.*, 1978), le protozoaire *Tetrahymena* (Gorovsky *et al.*, 1973) et plus récemment chez les métazoaires *Caenorhabditis elegans* (Greer *et al.*, 2015) et la drosophile (Zhang *et al.*, 2015).

Dans la suite du manuscrit seuls les mécanismes menant à la mise en place, le maintien et la suppression de la méthylation du 5^{ème} carbone des cytosines de l'ADN seront traités.

2.4.1. Mise en place de la méthylation de l'ADN

Chez les plantes, la méthylation *de novo* de l'ADN est médiée par la voie dépendante des ARN appelée RdDM pour RNA-directed DNA methylation. Cette voie utilise différentes protéines associées à des petits ARN interférents (siRNA) et à des ARN comme plateforme de ciblage du locus. Chez *Arabidopsis*, la méthylation *de novo* est catalysée par la méthyltransférase DRM2 pour Domains Rearranged Methyltransferase 2 (Chan *et al.*, 2005). Au niveau mécanistique, la mise en place de la méthylation a été décrite comme étant médiée par des petits ARN longs de 24 nucléotides qui sont synthétisés suite à l'action de différentes protéines. Tout d'abord un long transcrite ssRNA (ARN simple brin ; single strand RNA) est synthétisé par une ARN polymérase spécifique des plantes, l'ARN polymérase IV (Pol IV). Ce transcrite est ensuite pris en charge par la protéine RDR2 (RNA-Dependent RNA Polymerase 2) qui génère alors des ARN double brin (dsRNA). Finalement les dsRNA générés sont clivés en petits ARN de 24 nucléotides par l'action de la protéine Dicer like 3 (DCL3 ; Law & Jacobsen, 2010).

Les petits ARN de 24 nucléotides sont pris en charge par les protéines Argonaute 4 ou 6 (AGO4 ou AGO6). Ces dernières interagissent avec la sous-unité NRPE1 (Nuclear RNA Polymerase E1) d'une autre RNA polymérase spécifique des plantes, l'ARN polymérase V (Pol V). Le complexe ainsi formé entre la protéine Argonaute contenant le petit ARN de 24 nucléotides et la Pol V conduit à l'activation

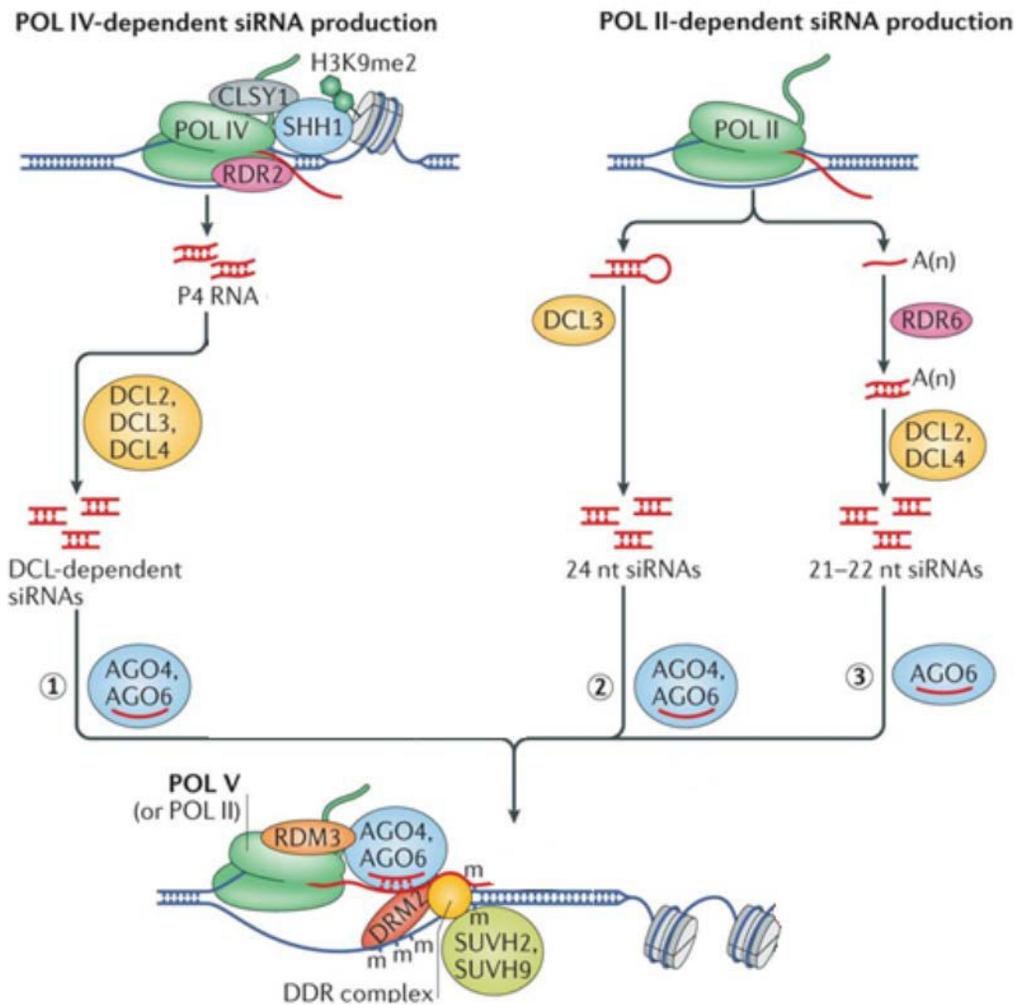


Figure I.15. Mécanisme de méthylation de l'ADN médiée par la voie RdDM (RNA-directed DNA methylation) chez *Arabidopsis thaliana*. **A gauche**, la voie canonique de la méthylation de l'ADN dirigée par l'ARN POLYMERASE IV (POL IV ; 1). Dans cette voie la POL IV génère des ARN non codants (ARN P4) servant de matrice pour la production d'ARN doubles brins (dsRNA) par la RNA-DEPENDENT RNA POLYMERASE 2 (RDR2). Les dsRNA produits sont clivés par les protéines DICER-LIKE PROTEIN 3 (DCL3), DCL2 et DCL4 afin de donner des ARN interférents (siRNAs) de 24 nucléotides. Par la suite, les siRNA sont liés par les protéines ARGONAUTE 4 (AGO4) ou AGO6 et appariés avec des plateformes d'ARN transcrits par la POL V afin de recruter la protéine DOMAINS REARRANGED METHYLASE 2 (DRM2). DRM2 est responsable de la méthylation (m) de l'ADN. La POL IV est recrutée au niveau des loci à méthyler par la voie RdDM par la protéine SAWADEE HOMEODOMAIN HOMOLOGUE 1 (SHH1), qui lie l'histone diméthylée H3 lysine 9 (H3K9me2). La protéine CLSY1, qui est une protéine de classe SNF2 qui intervient dans le remodelage de la chromatine, interagit avec la POL IV afin de permettre la production de siRNA par la POL IV. **A droite**, la voie non canonique de la méthylation de l'ADN dirigée par l'ARN POLYMERASE II (POL II). Dans cette voie, la POL II peut produire pour certains loci des siRNA de 24 nucléotides et des ARN plateforme (2) et pour d'autres loci interagir avec la protéine RDR6 afin de produire de manière collaborative des précurseurs de siRNA de 21 nucléotides ou de 22 nucléotides qui interviennent dans la méthylation de l'ADN de manière similaire aux siRNA de 24 nucléotides (3). AGO4 et/ou AGO6 s'associent directement avec la POL V, et le complexe ainsi formé est renforcé par la protéine RNA-DIRECTED DNA METHYLATION 3 (RDM3). La production des ARN plateforme par la POL V nécessite le complexe DDR, constitué du remodeleur de la chromatine DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1, DEFECTIVE IN MERISTEM SILENCING 3 et RDM1, qui s'associe à AGO4 et DRM2. Le complexe DDR interagit avec les protéines SUPPRESSOR OF VARIATION 3-9 HOMOLOGUE PROTEIN 2 (SUVH2) et SUVH9, qui se lient aux cytosines méthylées préexistantes et permettent le recrutement de la POL V. Adaptée de Zhang *et al.* (2018).

de DRM2 *via* l'interaction entre la protéine Argonaute et DRM2 (Zhong *et al.*, 2014). DRM2 méthyle ensuite le locus cible au niveau duquel la Pol V aura généré un transcrit non codant (ncRNA) servant de guide pour la méthylation (Wierzbicki *et al.*, 2009 ; **Figure I.15**).

En parallèle de cette voie dite RdDM canonique mettant en jeu des petits ARN de 24 nucléotides, la Pol IV, RDR2, DCL3, la Pol V et AGO4 ou AGO6, une autre voie de mise en place de la méthylation *de novo* dite non canonique existe. Cette voie utilise une autre ARN polymérase, l'ARN polymérase 2 (Pol II) qui est habituellement associée aux voies de silencing des gènes de manière post transcriptionnelle (PTGS). L'action de la Pol II dans cette voie non canonique a été associée à la protéine RDR6 (RNA-Dependent RNA Polymerase 6) afin de méthyler *de novo* des éléments transposables transcriptionnellement actifs (McCue *et al.*, 2015). Cette voie Pol II/RDR6 dépendante encore appelée voie RDR6-RdDM, permet la méthylation *de novo* en utilisant soit des petits ARN de 21 ou 22 nucléotides qui sont produits par les protéines Dicer like-2 (DCL2) ou Dicer like-4 (DCL4), soit des petits ARN de 24 nucléotides produits par DCL3 (Marí-Ordóñez *et al.*, 2013 ; McCue *et al.*, 2015). Les petits ARN de 21 et 22 nucléotides sont par la suite pris en charge par la protéine AGO6 alors que les petits ARN de 24 nucléotides peuvent être pris en charge par AGO4 ou AGO6 comme dans la voie canonique de la méthylation *de novo*. Finalement, la voie non canonique rejoint la voie canonique dans ses dernières étapes avec l'activation de DRM2 par interaction avec la Pol V et la méthylation du site cible guidée par un ncRNA transcrit par la Pol V (McCue *et al.*, 2015 ; **Figure I.15**).

Une fois la méthylation *de novo* mise en place cette dernière doit être maintenue au cours des multiplications et divisions cellulaires. Ce maintien est réalisé en grande partie par des mécanismes différents de ceux de la méthylation *de novo*.

2.4.2. *Maintien de la méthylation de l'ADN*

Chez les plantes, le maintien de la méthylation de l'ADN au cours des multiplications cellulaires est dépendant du contexte nucléotidique et est effectué par différentes voies (**Figure I.16**).

Le maintien de la méthylation en contexte CG est réalisé chez *Arabidopsis* par la DNA Méthyltransférase 1 (MET1) qui est un orthologue de la DNA Méthyltransférase 1 (DNMT1) présente chez les mammifères. MET1 maintient la méthylation en contexte CG *via* la reconnaissance des sites CG hémi-méthylés lors de la réplication de l'ADN. Une fois les CG hémi-méthylés reconnus, MET1 catalyse la réaction de méthylation des cytosines non méthylées du brin fille (Chan *et al.*, 2005). Ce maintien de la méthylation par MET1 nécessite également l'action de trois autres protéines de la même famille, les protéines VIM1 (VARIANT IN METHYLATION 1), VIM2 (VARIANT IN METHYLATION 2) et VIM3 (VARIANT IN METHYLATION 3 ; Woo *et al.*, 2008).

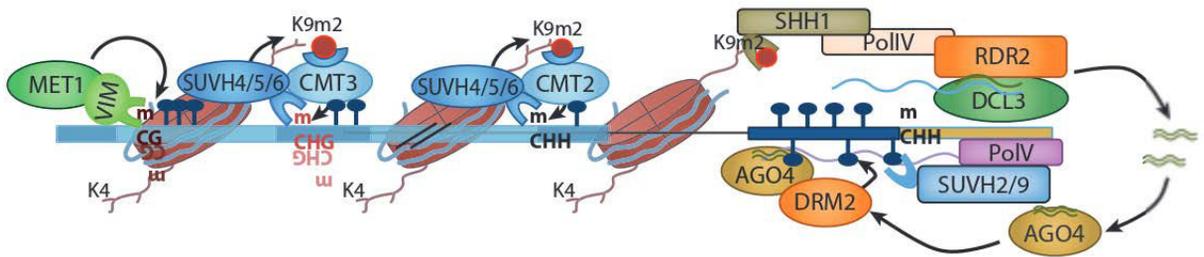


Figure I.16. Mécanismes de maintien de la méthylation de l'ADN. Le maintien de la méthylation en contexte CG est réalisé chez *Arabidopsis* par la DNA Méthyltransférase 1 (MET1) qui est recrutée par les protéines VIM qui vont reconnaître les CG hémiméthylées. Le maintien de la méthylation en contexte CHG est réalisé par deux protéines spécifiques des plantes, la chromométhylase 3 (CMT3) et dans une moindre mesure par la chromométhylase 2 (CMT2). La mise en place du maintien de la méthylation en CHG par CMT3 est contrôlée par la fixation de CMT3 au niveau de la marque H3K9me2. En tant que boucle de renforcement, les protéines SUVH (KYP) se lient au CHG méthylé pour déposer la méthylation de H3K9me2. Le maintien de la méthylation en contexte CHH fait intervenir soit CMT2 qui se fixe au niveau de la marque H3K9me2, soit la voie RdDM faisant intervenir notamment la polymérase IV (Pol IV), la Pol V, des protéines argonaute (AGO) et DRM2. Adaptée de Deleris *et al.* (2016).

Le maintien de la méthylation en contexte CHG est réalisé par deux protéines spécifiques des plantes, la chromométhylase 3 (CMT3) et dans une moindre mesure par la chromométhylase 2 (CMT2) (Stroud *et al.*, 2013 ; 2014). La mise en place du maintien de la méthylation en CHG par CMT3 est contrôlée par la fixation de cette dernière au niveau de la lysine 9 de l'histone H3 sous sa forme doublement méthylée (H3K9me₂ ; Du *et al.*, 2012). La di-méthylation de la lysine 9 de l'histone H3 est réalisée par les histones lysine méthyltransférases SUV4, SUVH5 et SUVH6 (Jackson *et al.*, 2002 ; Ebbs *et al.*, 2005 ; Ebbs & Bender, 2006).

Le maintien de la méthylation en contexte CHH est réalisé par deux mécanismes : par DRM2 de manière dépendante des siRNA (Chan *et al.*, 2005) et par CMT2 de manière indépendante des siRNA (Zemach *et al.*, 2013 ; Stroud *et al.*, 2014). La maintenance de la méthylation par DRM2 est localisée au niveau des éléments transposables récents, des éléments transposables courts, des séquences répétées des régions euchromatiniennes et des extrémités des éléments transposables longs (Law & Jacobsen, 2010). La maintenance *via* CMT2 est réalisée surtout au niveau du corps des éléments transposables et des régions péricentromériques (Zemach *et al.*, 2013) et est liée, comme pour la méthylation en CHG, à la marque H3K9me₂ et à l'action de KYP, SUVH5 et SUVH6. Chez le mutant *ddm1* la méthylation médiée par CMT2 est altérée, indiquant que le facteur de remodelage de la chromatine DDM1 (Decrease In DNA Methylation1) est nécessaire pour l'action de CMT2 (Zemach *et al.*, 2013).

Bien que les états de méthylation soient maintenus au cours de la réplication de l'ADN, la dynamique de méthylation peut aussi mener à la perte de méthylation de l'ADN par des mécanismes passifs ou actifs.

2.4.3. La déméthylation de l'ADN

L'absence de maintien de la méthylation de l'ADN par les systèmes qui viennent d'être décrits au cours des réplifications de l'ADN conduit à un phénomène de déméthylation de l'ADN de manière passive. En plus de ce système passif, une voie active de déméthylation de l'ADN par suppression de la méthylation de l'ADN de manière enzymatique existe (Meyer, 2015). Cette voie utilise plusieurs enzymes dont des ADN déméthylases qui possèdent des activités glycosylase permettant de rompre la liaison N-glycosidique liant le groupement méthyl et le 5^{ème} carbone de la cytosine (Baute & Depicker, 2008 ; Gehring *et al.*, 2009). Chez *Arabidopsis* quatre déméthylases sont décrites, ROS1 (Repressor Of Silencing 1), DME (Transcriptionnal Activator Demeter), DML2 (DEMETER-like 2) et DML3 (DEMETER-like 3). Ces dernières sont capables de catalyser la suppression du groupement méthyl dans tous les contextes nucléotidiques (Gong *et al.*, 2002 ; Morales-Ruiz *et al.*, 2006 ; Ortega-Galisteo *et al.*, 2008). Ces 4 enzymes possèdent des substrats similaires mais elles ne sont pas exprimées dans les mêmes tissus. En effet ROS1, DML2 et DML3 sont exprimées dans les tissus végétatifs de la plante, alors que DME est exprimée dans les cellules compagnes des gamètes mâle et femelle (Penterman *et al.*, 2007 ; Huh *et al.*, 2008).

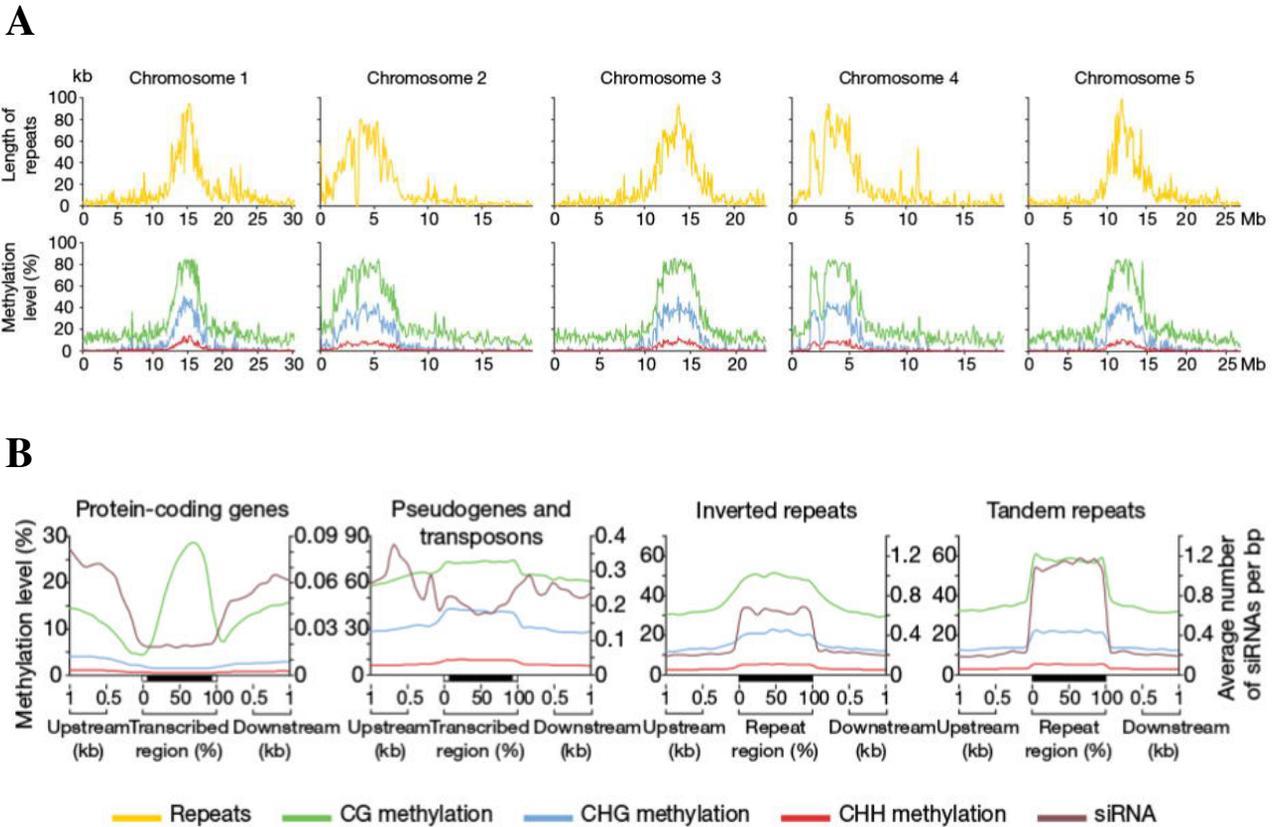


Figure I.17. Méthylation de l'ADN du génome d'*Arabidopsis*. (A) Distribution de la méthylation de l'ADN et des régions répétées le long des 5 chromosomes d'*Arabidopsis*. Le calcul du niveau de méthylation a été effectué en utilisant une fenêtre glissante de 100 kb. (B) Représentation du niveau de méthylation et de l'abondance en siRNA en fonction du type de gènes et de régions répétées. Adaptée de Cokus *et al.* (2008)

L'ensemble des processus responsables de la mise en place de la méthylation, de sa maintenance et de sa suppression mènent à l'établissement des motifs de méthylation le long du génome des plantes. Suivant les régions génomiques les motifs de méthylation sont différents et suivant leur positionnement et le type de méthylation engagée leurs conséquences seront différentes.

2.4.4. Les profils de méthylation de l'ADN au niveau du génome

Différentes régions du génome des plantes ont été montrées comme présentant des profils de méthylation qui leur sont spécifiques. Suivant le lieu et le type de méthylation apposé au niveau de l'ADN, les conséquences sur l'expression des éléments qui sont présents dans la région ciblée sont différentes. De manière globale la méthylation de l'ADN montre des fluctuations le long du génome avec la présence de régions nettement plus méthylées que d'autres (**Figure I.17A**). Chez *Arabidopsis*, un enrichissement en méthylation dans les trois contextes (CG, CHG et CHH) a été démontré au niveau des éléments répétés localisés dans les régions péracentromériques. Des profils de méthylation différents entre les gènes codant des protéines et les pseudogènes/éléments transposables existent indiquant la présence de profils de méthylation spécifiques des différents éléments génomiques (Cokus *et al.*, 2008 ; **Figure I.17B**).

2.4.4.1. La méthylation de l'ADN au niveau des gènes

L'expression du génome dépend de différentes séquences régulatrices appelées activatrices et inhibitrices se trouvant en amont et en aval des séquences codantes. Ces séquences génomiques sont le lieu de fixation en cis d'éléments transrégulateurs nécessaires à l'expression des gènes. Parallèlement, le remodelage de la structure de la chromatine réalisée par l'apposition d'un jeu de marques épigénétiques peut limiter ou faciliter la fixation de l'ARN polymérase au niveau de l'ADN. Ainsi la localisation, le niveau et le type de méthylation de l'ADN présent au niveau des gènes sont des éléments permettant d'estimer l'influence de cette marque sur l'expression du gène concerné.

Chez les plantes sans fleur telles que les mousses et les algues vertes, il y a une grande diversité de motifs de méthylation (Zemach *et al.*, 2010 ; Feng *et al.*, 2010a). Chez les angiospermes, il existe 5 classes de méthylation de gènes en fonction de l'enrichissement en méthylation des séquences exoniques (Bewick & Schmitz, 2017, **Figure I.18**) : i) la classe des gènes non méthylés (Unmethylated : UM), où il y a absence de méthylation au niveau des trois contextes CG, CHG et CHH sur l'ensemble de la région transcrite du corps du gène (**Figure I.18A**). ii) La classe des gènes corps méthylés (Gene body methylated : gbM), où il y a un enrichissement en méthylation en contexte CG de la région transcrite et une perte de méthylation au niveau du site initiateur de la transcription (Transcription Start Site : TSS) et au niveau du site terminateur (Transcription Termination Site : TTS ; **Figure I.18B**). Les gènes possédant une méthylation de ce type présentent une expression constitutive et ubiquitaire plus élevée que les gènes non gbM (Zilberman *et al.*, 2007). iii) La classe des gènes méthylés au niveau de leur TSS (**Figure I.18C**), où l'enrichissement en méthylation de type CG se fait uniquement au niveau

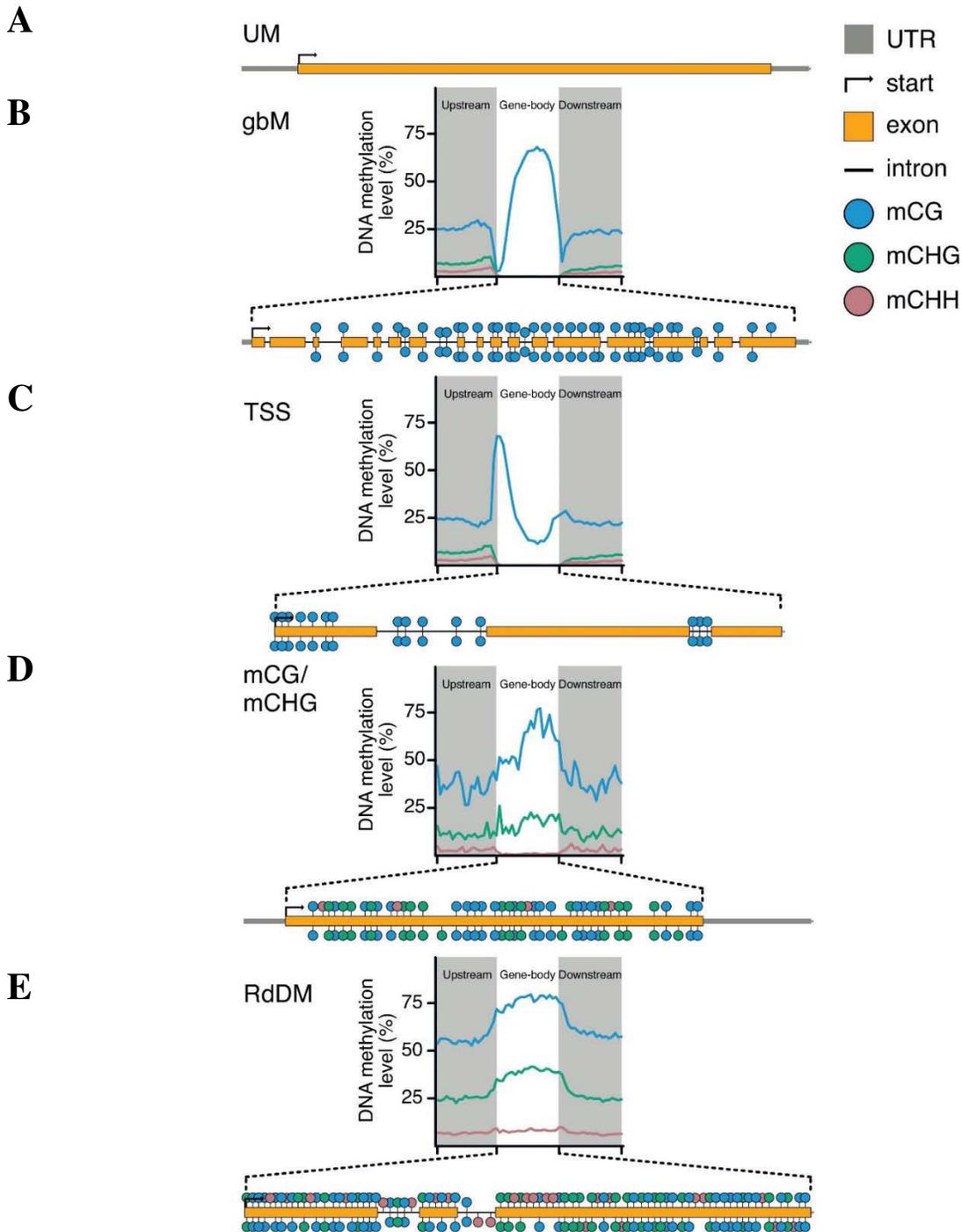


Figure I.18. Schématisation des motifs de méthylation de l'ADN génique chez les plantes. En fonction de l'enrichissement en méthylation au niveau de leurs régions exoniques les gènes peuvent être classés en différents groupes. **(A)** Les gènes non méthylés (UM : Unmethylated), où il y a absence de méthylation au niveau des trois contextes CG, CHG et CHH sur l'ensemble de la région transcrite du corps du gène. **(B)** des gènes corps méthylés (gbM : Gene body methylated), où il y a un enrichissement en méthylation en contexte CG de la région transcrite et une perte de méthylation au niveau du site initiateur de la transcription et au niveau du site terminateur. **(C)** Les gènes méthylés au niveau de leur TSS, où l'enrichissement en méthylation de type CG se fait uniquement au niveau du TSS du gène conduisant à la répression de son expression **(D)** Les gènes dits CG/CHG, où il y a un enrichissement en méthylation de CHG et une perte de méthylation de type CHH au niveau du corps du gène, avec la présence ou non de méthylation de type CG **(E)** Les gènes dits CHH/RdDM présentant un enrichissement en méthylation CHH accompagné ou non de méthylation CG et ou CHG au niveau du corps du gène. Figure issue de Bewick et Schmitz, (2017).

du TSS du gène conduisant à la répression de son expression (Niederhuth *et al.*, 2016). iv) La classe des gènes dits CG/CHG, où il y a un enrichissement en méthylation de CHG et une perte de méthylation de type CHH au niveau du corps du gène, avec la présence ou non de méthylation de type CG (**Figure I.18D**). Les gènes de cette classe sont typiquement exprimés à des niveaux très faibles par rapport aux autres gènes et aux gènes ayant une méthylation gbM chez les angiospermes (Niederhuth *et al.*, 2016). v) La classe des gènes dits CHH/RdDM présentant un enrichissement en méthylation CHH accompagné ou non de méthylation CG et ou CHG au niveau du corps du gène (**Figure I.18E**). Les gènes ayant une méthylation de ce type sont la plupart du temps réprimés dans l'ensemble de la plante chez *A. thaliana* à l'exception du pollen et des graines en développement (Schmitz *et al.*, 2013).

Parmi les 5 classes de méthylation possibles au niveau des gènes, la méthylation du corps des gènes (gbM) est l'une des plus étudiée afin de comprendre les origines et la fonction de ce type de méthylation au sein du génome. En effet, la méthylation gbM mène à l'expression constitutive des gènes contrairement à l'idée originale sur la méthylation de l'ADN selon laquelle celle-ci est associée à la répression de l'expression des gènes. A l'heure actuelle, les mécanismes par lesquels la méthylation gbM est mise en place restent peu connus. Bewick *et al.* (2016) ont proposé un modèle pour expliquer son accumulation. Dans ce modèle, l'histone lysine déméthylase IBM1 (Increased In Bonsai Methylation 1) ne supprime pas la marque histone H3K9me2 conduisant au recrutement de la chromométhylase CMT3 et à la méthylation des sites CHG. La méthylation en CHG provoque le recrutement des histone méthyltransférases SUVH4/5/6 menant à la diméthylation de la lysine 9 de l'histone H3 initiant la boucle de renforcement de la méthylation par CMT3 et la méthylation des sites CG par un mécanisme inconnu. Une fois les sites CG hémiméthylés, la méthyltransférase MET 1 vient maintenir la méthylation sur ces sites puis la méthylation se répand le long du gène. La fonction de la méthylation gbM n'est pas connue. Différentes hypothèses ont été proposées telles que i) un rôle dans la régulation de l'épissage des gènes et de leur expression (Maunakea *et al.*, 2010 ; Reguluski *et al.*, 2013), ii) le fait de conférer un avantage sous de nouvelles pressions de sélection avec par exemple chez *A. thaliana* l'existence pour un même gène de plusieurs épialèles possédant une classe de méthylation différente entre différentes accessions (Kawakatsu *et al.*, 2016). Bewick & Schmitz (2017) proposent également que la méthylation gbM puisse être un sous-produit de la méthylation de l'ADN responsable de la répression des éléments transposables et d'autres séquences répétées.

2.4.4.2. La méthylation de l'ADN au niveau des éléments transposables

Les éléments transposables (TE) sont des fragments mobiles de l'ADN qui ont été initialement découverts dans le génome du maïs par McClintock (1951). Ces derniers sont présents dans le génome de la plupart des eucaryotes avec une corrélation nette entre la quantité de TE présents dans le génome et la taille de ce dernier. Par exemple *A. thaliana* qui a un petit génome (environ 135 Mb), possède environ 17 % d'éléments transposables alors que chez le maïs qui a un génome de 2,3 Gb les éléments

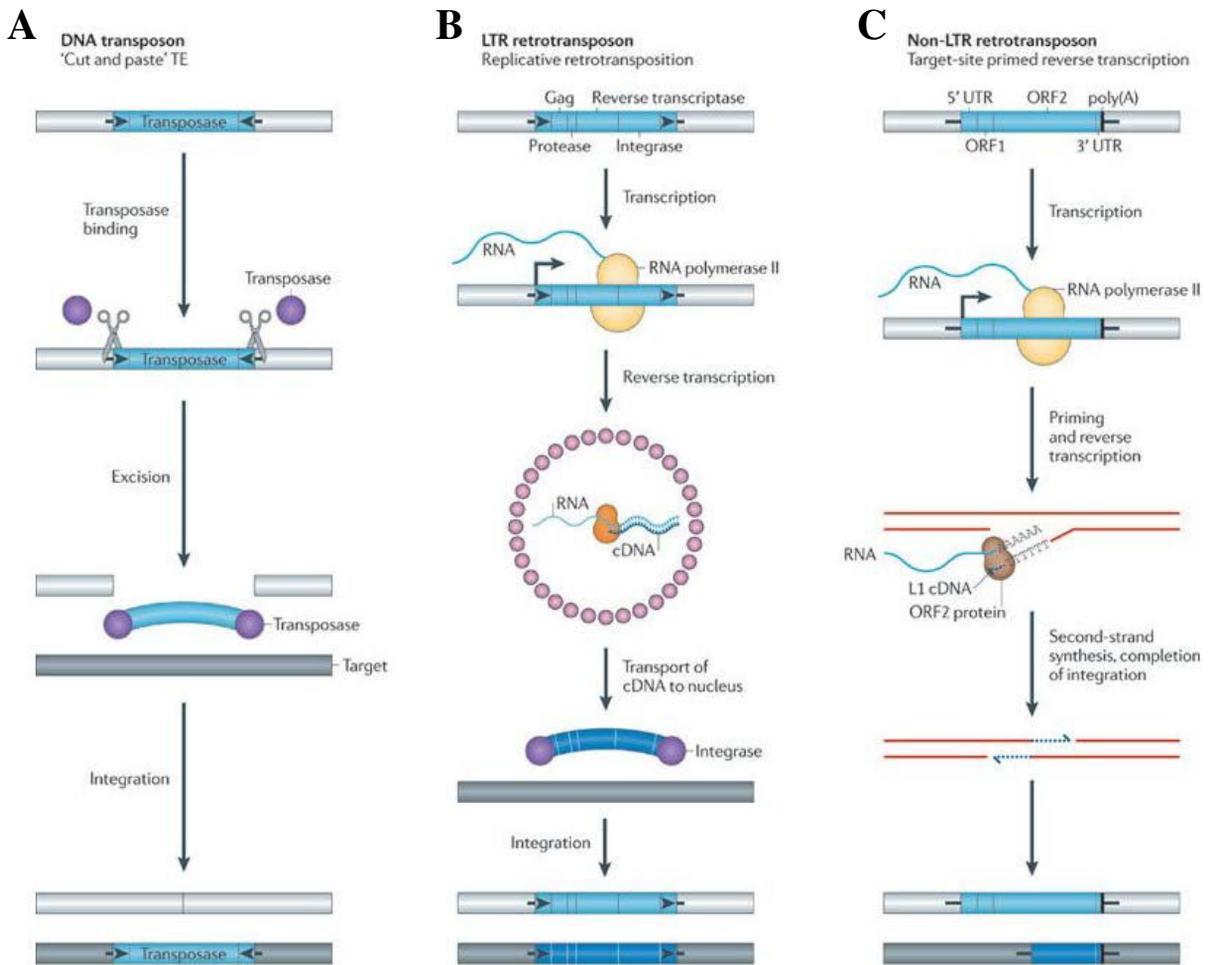


Figure I.19. Les différents types d'éléments transposables (TE). (A) De nombreux transposons à ADN sont encadrés par des régions terminales inversées (TIR ; flèches noires), encodent une transposase (TPase, cercles violets) et sont mobilisés par un mécanisme de type « couper-coller » (ciseaux). La transposase se fixe sur ou à proximité des TIR, extrait le transposon du génome (zones en gris clair) et le transfère dans une nouvelle région génomique (zones en gris foncé). Le clivage des deux brins au site ciblé entraîne un chevauchement entraînant la duplication d'une zone de 4 à 8 pb (courtes lignes noires autour du TE). (B) et (C) Les rétrotransposons sont mobilisés par un mécanisme de réplication qui nécessite un ARN intermédiaire. (B) Les LTR-rétrotransposons contiennent deux longues régions répétées (LTR ; flèches noires) et encode Gag, protéase (PR), transcriptase inverse (RT) et une intégrase (IN). La 5'LTR contient un promoteur qui est reconnu par l'ARN polymérase II de l'hôte et produit l'ARNm du TE (le début de la transcription est représenté par une ligne verticale noire). Dans la première étape, Gag (cercles roses) s'assemble en particules similaires à des virus qui contiennent l'ARNm du TE, RT et IN (cercles violets). La RT duplique l'ARNm du TE en ADN double brin (arc de cercle bleu). Lors de la seconde étape, IN introduit l'ADN dans la région cible. De la même manière que les TPases des transposons à ADN, l'activité des IN produit des TSD. (C) Les rétrotransposons non-LTR ne possèdent pas les LTR, la synthèse d'ADNc a lieu directement au site d'insertion de la nouvelle copie : un seul brin d'ADN du site cible est clivé par l'endonucléase de l'élément, l'ARNm y est rattaché, permettant la transcription inverse d'un brin. Le second brin est synthétisé après le retrait de l'ARNm (mécanisme appelé Target-Primed Reverse Transcriptase). D'après Levin et Moran (2011).

transposables représentent environ 85 % du génome (Buisine *et al.*, 2008 ; Schnable *et al.*, 2009). Les TE sont divisés en 2 classes majoritaires (Wicker *et al.*, 2007 ; **Figure I.19**). Les TE de classe I utilisent un mécanisme de « copier-coller » au travers d'une étape de réverse transcription de leur ARN messager afin de s'intégrer dans un autre endroit du génome. Les transposons de cette classe sont ainsi appelés des rétrotransposons. Les TE de classe II utilisent un système de « couper-coller » afin de s'exciser de leur région d'origine pour s'insérer dans un autre endroit du génome. Les transposons de cette classe sont appelés des transposons à ADN. Les transposons de type hélitrons utilisent un autre mécanisme appelé « cercle de roulement » ou « rolling-circle » où seulement un brin de l'élément transposable est utilisé lors de la transposition dans une autre région chromosomique. La capacité de transposition des éléments transposables leur confère un pouvoir mutateur vis-à-vis du génome. De ce fait l'expression de ces derniers est très contrôlée *via* un mécanisme de répression ou « silencing » très efficace. Alors que les mécanismes de maintien et de renforcement du silencing des éléments transposables sont bien connus, les raisons pour lesquelles un élément transposable se retrouve réprimé *via* un état hétérochromatinien de la région où il se trouve sont peu connues.

Le maintien de la répression de l'expression des TE est dépendant du facteur de remodelage de la chromatine DDM1 qui aide à la méthylation en facilitant l'accès au niveau de la région hétérochromatique des protéines servant à la méthylation de l'ADN dans les trois contextes (Zemach *et al.*, 2013). La méthyltransférase MET1 est responsable du maintien de la méthylation dans le contexte CG et la chromométhylase CMT2 et CMT3 maintiennent la méthylation en contexte CHH et CHG guidée par la marque histone H3K9me2.

Concernant l'initiation du silencing d'un élément transposable deux voies possibles ont été décrites. La première voie dépend de l'existence d'autres copies du TE déjà réprimées au sein du génome et est appelée homology-dependent initiation of silencing. La seconde voie ne dépend pas de l'existence de copies homologues du TE à réprimer est appelée homology-independent initiation silencing (Fultz *et al.*, 2015).

Dans la voie Homology-dependent initiation of silencing ce sont les siRNA de 24 nucléotides issus de la voie RdDM provenant d'une copie homologue déjà réprimée qui vont cibler le nouvel élément transposable à silencer menant ainsi à sa méthylation (**Figure I.20A**). En effet, chez toutes les plantes où cela a été vérifié, les siRNA de 24 nt produits par la répression d'un élément transposable sont omniprésents (Nobuta *et al.*, 2007 ; Cho *et al.*, 2008 ; Huang *et al.*, 2013) et servent de senseur d'homologue afin d'identifier et de permettre la répression d'éléments transposables actifs possédant une séquence similaire. Cependant une question reste en suspens, celle de comprendre comment la Pol V est recrutée au niveau de l'élément transposable actif et quel est le pourcentage d'homologie nécessaire entre les siRNA de 24 nucléotides et leur cible (Fultz *et al.*, 2015).

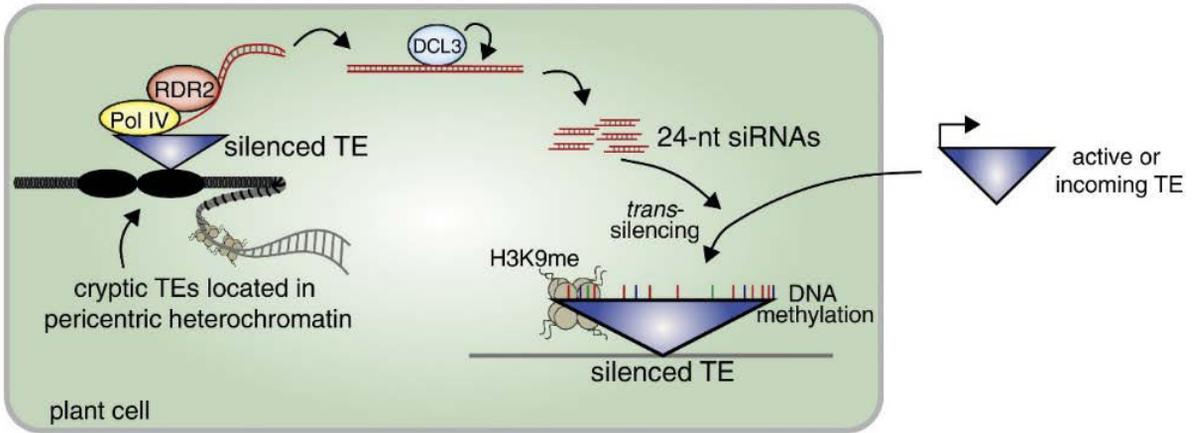
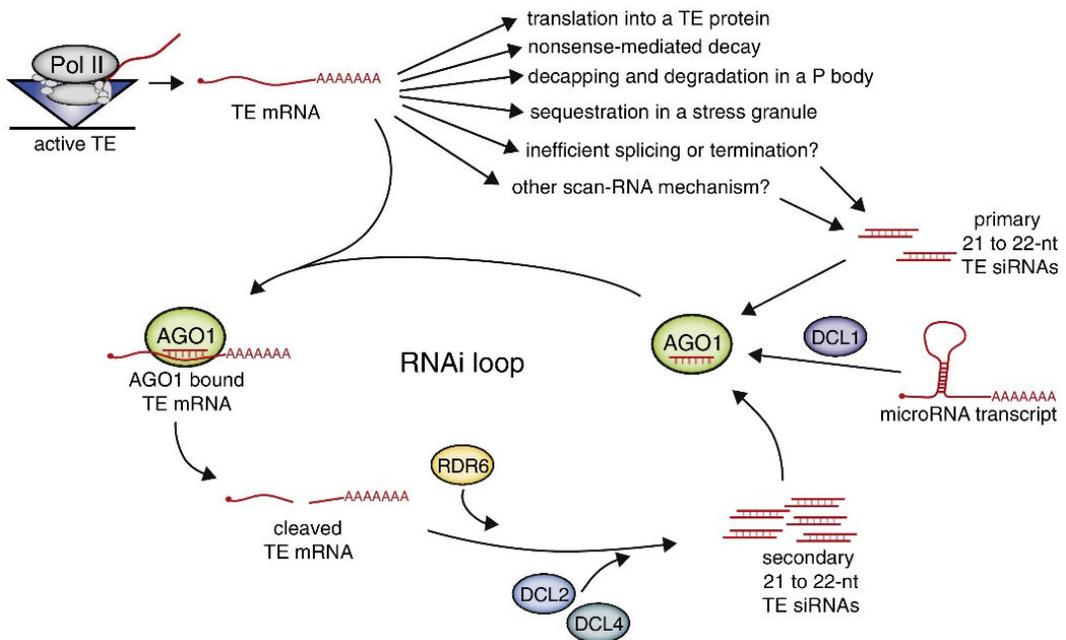
A**B**

Figure I.20. Mécanisme de méthylation des TE. (A) Mécanisme d'initiation du silencing des éléments transposables dépendant de l'homologie. Les TE précédemment silencés et dégénérés sont abondants au niveau des péricentromères du génome des plantes. La Pol IV transcrit ces TE hétérochromatiques et les ARN non codants résultant sont clivés en siARN de 24 nt par l'activité de RDR2 et DCL3. Les siARN de 24-nt peuvent fonctionner comme un capteur d'homologie pour identifier d'autres TE avec des séquences similaires et initier ou renforcer leur silencing *via* la marque H3K9me et la méthylation de l'ADN. La méthylation de l'ADN en CG est représentée en rouge, la CHG en bleu et la CHH en vert, où H = A, T ou C. Adaptée de Fultz *et al.* (2015). (B) Mécanisme d'initiation du silencing des éléments transposables indépendant de l'homologie. Dans ce schéma, de multiples mécanismes déclenchent une mise en place du silencing post transcriptionnel des TE actifs. Afin d'arriver à cela, les ARNm des TE peuvent être pris en charge par différentes voies non exclusives entraînant une réduction de l'expression du TE ciblé. Ce mécanisme passe par la dégradation des ARN messager du TE actif (TE mRNA) en un nombre abondant de siRNA de 21 et 22 nucléotides primaires *via* un mécanisme ARN interférent endogène à la plante. Ces siRNA sont ensuite pris en charge par AGO, le complexe allant cibler les TE mRNA conduisant à leur clivage. Certains TE mRNA clivés sont ensuite copiés en double brin d'ARN par RDR6 qui sont par la suite clivés par DCL2 et DCL4 produisant des siRNA secondaires qui vont participer à la boucle de dégradation des TE mRNA. Adaptée de Fultz *et al.* (2015).

Dans la voie *homology-independent initiation silencing*, l'initiation de la répression du TE actif passe tout d'abord par un mécanisme de répression post transcriptionnel (**Figure I.20B**). Ce mécanisme conduit à la dégradation des ARN messager du TE actif (TE mRNA) en un nombre abondant de siRNA de 21 et 22 nucléotides *via* un mécanisme ARN interférent endogène à la plante. Chez *Arabidopsis* ce mécanisme est composé des protéines RDR6, AGO1, DCL2 et DCL4 (Dunoyer *et al.*, 2010). Une fois ce processus de répression post transcriptionnel mis en place, deux modèles sont proposés pour expliquer la modification de la chromatine. Le premier modèle suggère que la plante mesure la quantité de transcrits aberrants synthétisés et bascule sur le clivage de ces derniers en siRNA de 24 nt par DCL3 lorsque les protéines DCL2 et DCL4 se retrouvent dépassées par le nombre de transcrits à cliver en siRNA de 21 et 22 nucléotides. Les siRNA de 24 nt synthétisés seraient utilisés de manière similaire à ceux synthétisés par la RdDM (Marí-Ordóñez *et al.*, 2013) menant à la méthylation du TE. Le second modèle utilise la voie RDR6 RdDM précédemment décrite dans le paragraphe 2.4.1.

Quelle que soit la voie menant à la méthylation d'un TE actif, un phénomène de renforcement des marques hétérochromatiniennes au niveau du TE est mis en place en utilisant la voie RdDM passant par la Pol IV afin de mettre l'élément transposable dans un état de répression permanent. Cette action de la voie Pol IV RdDM est essentielle pour marquer la jonction entre l'état hétérochromatinien au niveau du TE et l'état euchromatinien des gènes dans le cas où le TE est proche de gènes (Zhong *et al.*, 2012 ; Zemach *et al.*, 2013). De plus l'action de la voie Pol IV RdDM peut mener à l'étalement des marques de répression aux régions adjacentes du TE appelé phénomène de spreading (Daxinger *et al.*, 2009). En effet, il a été démontré que la méthylation mise en place au niveau du TE par la voie RdDM peut se propager au niveau de ses régions flanquantes sur environ 300 pb avec des conséquences possibles sur l'expression de gènes qui seraient proches du TE ciblé (Ahmed *et al.*, 2011b). L'analyse au niveau des TE nouvellement insérés chez *Arabidopsis* a permis de déterminer que le phénomène de spreading touche la méthylation dans ses 3 contextes et peut se réaliser de chaque côté du TE inséré sur des distances allant de 300 paires de bases (pb) à plus de 3500 pb (Quadrana *et al.*, 2016 ; **Figure I.21**).

L'ensemble des mécanismes précédemment évoqués mènent ainsi au profil de méthylation observé au niveau de la plupart des TE réprimés avec des taux élevés de méthylation dans les trois contextes nucléotidiques CG, CHG et CHH (Cokus *et al.*, 2008 ; Lister *et al.*, 2008).

2.4.5. Stabilité de la méthylation de l'ADN : l'épigénétique transgénérationnelle

Alors que les mécanismes de maintien des profils de méthylation au cours de la mitose sont bien connus, notamment *via* la copie de l'état de méthylation du brin mère sur le brin fille lors de la réplication de l'ADN par MET1 (Chan *et al.*, 2005), la validation de la transmissibilité des marques épigénétiques lors de la méiose est plus ambiguë. En effet l'analyse chez les mammifères a montré qu'un large processus de reprogrammation des marques épigénétiques est mis en place dans le zygote et les cellules germinales immédiatement après la fécondation (Seisenberger *et al.*, 2012), limitant ainsi l'effet

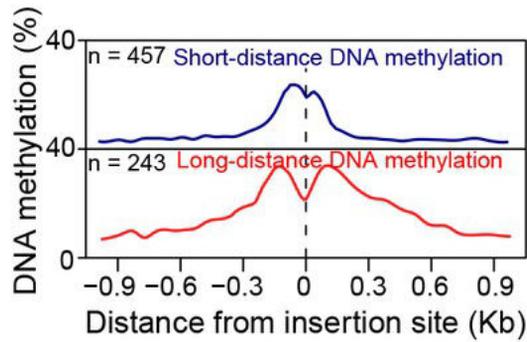


Figure I.21. Phénomène de spreading au niveau de l'insertion d'un élément transposable. Le pourcentage de méthylation présent autour du site d'insertion d'un élément transposable est représenté sur l'axe y. L'analyse sur une courte distance du niveau de méthylation entourant le site d'insertion de 457 éléments transposables (en bleu) montre que le niveau de méthylation est généralement étendu jusqu'à 300 pb du lieu de l'insertion quelque soit le contexte nucléotidique. Un phénomène de spreading plus important allant ici jusqu'à 900 pb (en rouge) est également possible. n: indique le nombre d'insertions étudiées. Adaptée de Quadrona *et al.* (2016).

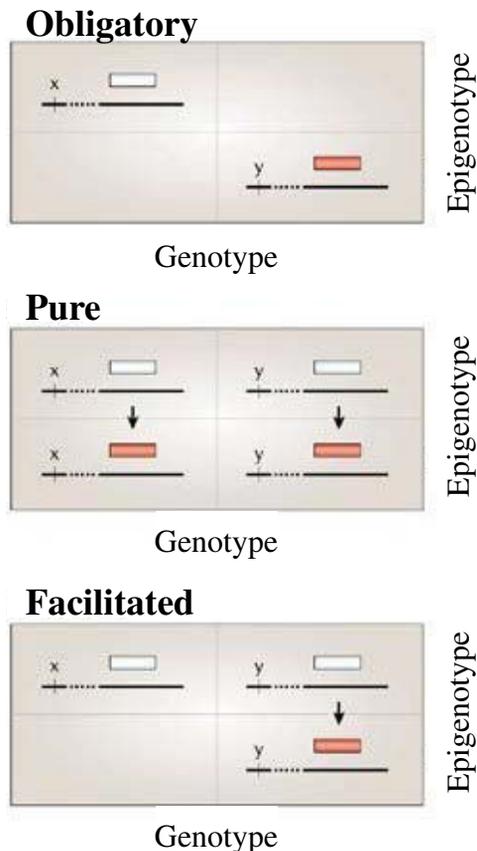


Figure I.22. Les différentes classes d'épiallèles possibles. Schématisation de trois classes d'épiallèles possibles en fonction de la relation entre l'épigénotype et le génotype. L'axe horizontal distingue deux génotypes représentés par des allèles x et y. À un emplacement génomique, soit en cis, soit en trans (ligne en pointillés), deux états épigénétiques alternatifs sont représentés sous forme de boîtes rouges ou blanches. Ainsi trois scénarii sont possibles : Obligatory, dans ce scénario, l'épigénotype du locus est strictement déterminé par le génotype ; Ainsi, l'épigénotype est un phénotype obligatoire des génotypes alternatifs ; Facilitated, où le génotype dirige ou potentialise l'épigénotype mais de manière non stricte ; Pure, dans ce scénario, l'épiallèle est généré par des évènements stochastiques largement indépendants de variations génétiques. Adaptée de Richards (2006).

transgénérationnel de ces marques. Cependant, chez les plantes le phénomène de reprogrammation pour la méthylation de l'ADN est limité au contexte asymétrique CHH dans le pollen (Calarco *et al.*, 2012). Cette différence de la reprogrammation des statuts épigénétiques rend ainsi plausible la transmissibilité des marques épigénétiques chez les plantes.

L'étude de l'épigénétique transgénérationnelle - par opposition à l'épigénétique développementale - implique l'absence de reprogrammation du statut épigénétique entre les générations (Heard & Martienssen, 2014). Le premier exemple de transmissibilité épigénétique a été mis en évidence par McClintock (1961) chez le maïs où elle a montré que l'effet des transposons Activator (Ac) et Suppressor mutator (Spm) sur un gène impliqué dans la production des anthocyanes était cyclique entre une phase active et inactive et que ces phases pouvaient être transmissibles entre les générations. La transmissibilité des marques épigénétiques lors de la méiose étant dépendante de leur contexte génétique, trois classes d'épi-allèles ont été définies par Richards (2006) i) Les épivariants obligatoires où l'épigénotype au locus est strictement déterminé par le génotype, faisant de l'épigénotype une cause directe de la variation de l'expression du gène mais pas la cause initiale. L'épigénotype peut alors être interprété comme « un phénotype obligatoire » médié par un génotype alternatif. ii) Les épivariants facilités où le génotype dirige ou potentialise l'épigénotype mais de façon non obligatoire. iii) Les épivariants purs générés par des événements stochastiques indépendants de variations génétiques tels que des erreurs aléatoires de propagation de marques de répression après la réplication de l'ADN (**Figure I.22**).

Le taux d'apparition spontané de variations de méthylation dans le génome des plantes a été estimé chez *A. thaliana* en utilisant les lignées accumulatrices de mutation (MA lines) créées par sélection monograinne de l'accession de référence *Columbia* (Shaw *et al.*, 2000). A partir de ces lignées, les analyses menées par Ossowski *et al.* (2010), Schmitz *et al.* (2011) et Van der Graaf *et al.* (2015) ont permis respectivement d'estimer le taux de mutations (de l'ordre de 10^{-9}) et d'épimutations (de l'ordre de 10^{-4}) et de montrer ainsi qu'au niveau d'une cytosine seule, le taux d'épimutation est 10^5 fois plus important que celui des mutations nucléotidiques. Par ailleurs, Becker *et al.*, 2011 ont montré que la fréquence d'apparition des DMR était du même ordre de grandeur que celui des mutations de la séquence d'ADN. En utilisant également les lignées MA, Schmitz *et al.* (2011) ont identifié différents épiallèles qui sont en majorité stables à travers les méioses et transmissibles entre les générations et qui pour certains ont un impact significatif sur la variation de la transcription des gènes au locus affecté.

A travers diverses études il a été montré que des loci présentant des états de méthylation alternatifs (épiallèles), obtenus naturellement ou induits, sont transmis au cours des générations et sont sources de variations phénotypiques.

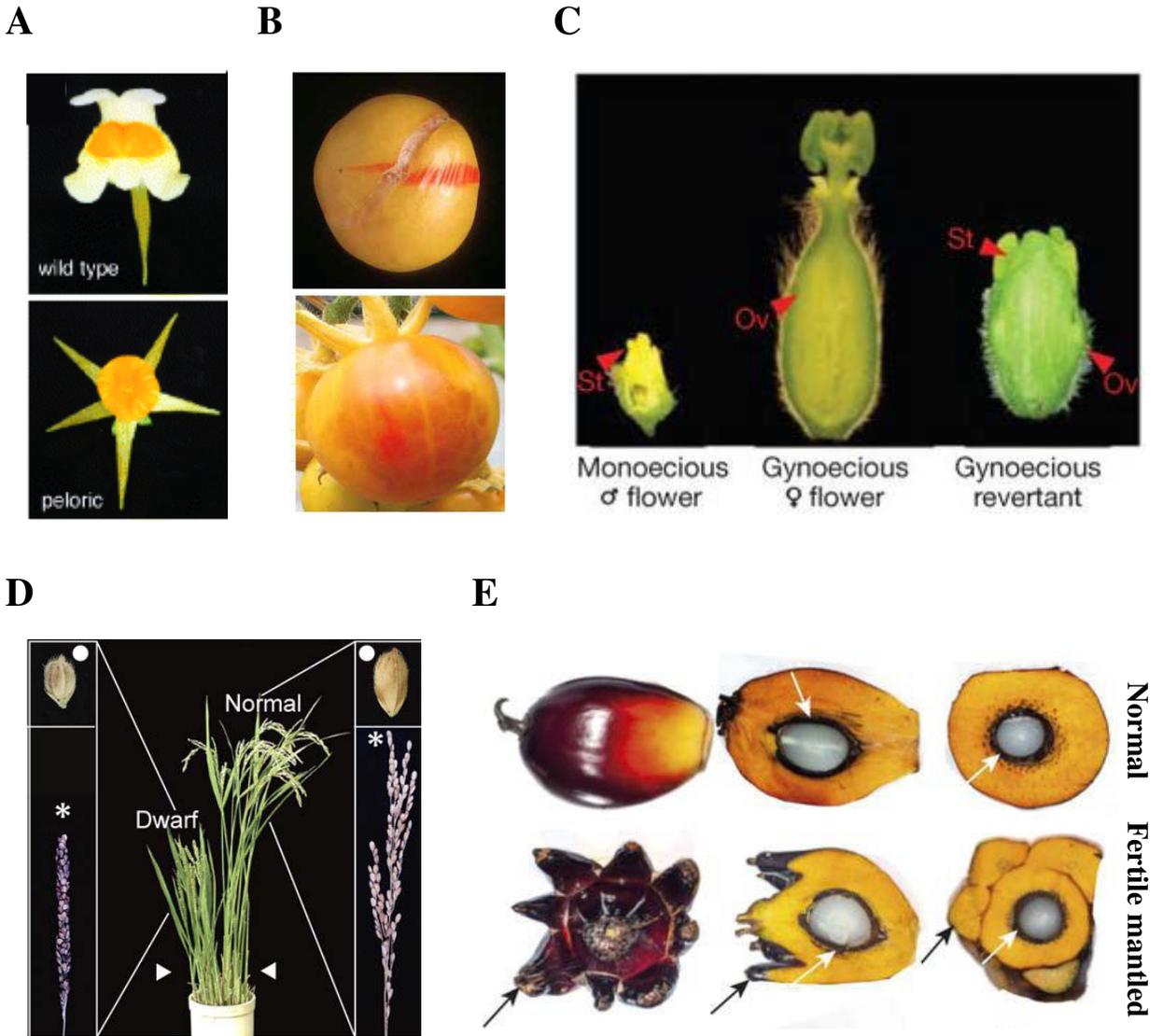


Figure I.23. Conséquences phénotypiques de différents épiallèles naturels chez les plantes. (A) Fleurs de *Linaria vulgaris* de type sauvage et du mutant pélorique vues de face. La fleur sauvage présente une symétrie dorso-ventrale alors que la fleur pélorique possède une symétrie radiale, avec toutes ses pétales ressemblant aux pétales ventrales du type sauvage. Adaptée de Cubas *et al.* (1999). (B) Modification phénotypique induite chez la tomate en réponse à un changement épigénétique au locus *Cnr*. Les deux photos illustrent le phénomène de réversion de secteurs de maturation au niveau des mutants *Cnr* avec l'apparition de phénotype de maturation de type sauvage. Adaptée de Manning *et al.* (2006). (C) Phénotypes observables au niveau des fleurs de melon. Phénotypes de fleurs mâles monoïques (Monoecious), de fleurs femelles gynoïques (gynoecious) et de fleurs gynoïques partiellement révertantes (gynoecious revertant). En accord avec l'origine épigénétique de la régulation de la détermination du sexe chez le melon, des phénomènes de réversion partielle au niveau des plantes gynoïques femelles sont observables avec la présence au niveau phénotypique de fleurs femelles avec des étamines et des ovaires réduits. Ov, ovaire ; st, étamines. Adaptée de Martin *et al.* (2009). (D) Phénotype du mutant *epi-d1* qui produit des tallages de type dwarf et normal sur la même plante. Les tallages de type dwarf (à gauche) ont des feuilles courtes vert foncé, des panicules compactes (inflorescences) et de petits grains ronds. Par ailleurs les tallages normaux ont un type sauvage (à droite). Adaptée de Miura *et al.* (2009). (E) Phénotype d'un fruit normal et d'un fruit « mantled » du palmier à huile. Le fruit est montré à gauche complet, au milieu en coupe longitudinale et à droite en coupe transversale du fruit. Les flèches noires montrent la présence de pseudo-carpelles chez les fruits mantled qui sont absents chez le fruit normal. Les flèches blanches indiquent les noyaux. Adaptée de Ong-Abdullah *et al.* (2015).

2.4.5.1. Épiallèles naturels

Un nombre restreint d'épiallèles naturels ont été identifiés chez les plantes suggérant que ce phénomène est plus rare que les variations alléliques au sein du génome. Ces épiallèles sont présents dans différentes espèces végétales et influencent différents traits phénotypiques des plantes (**Figure I.23**).

Le premier exemple d'épiallèle naturel décrit chez les plantes pour lequel les bases moléculaires n'ont pas été attribuées à un changement dans la séquence de l'ADN est le variant *pelorique* de la linaria commune (*Linaria vulgaris*). Dans ce variant, une épimutation qui persiste pendant des centaines de générations influence sa symétrie florale. En effet chez le mutant *pelorique* la symétrie florale est radiale alors que cette dernière est bilatérale dans le sauvage (**Figure I.23A**). La variation phénotypique observée est causée par des niveaux différents de méthylation au niveau de la région promotrice du gène *CYCLOIDEA* (*Lcyc* locus) qui vont alors moduler son expression. Comme aucune variation de séquençage n'a été associée à cet épiallèle celui-ci entre donc dans la classe des épiallèles purs décrit par Richards (2006).

Chez la tomate, le variant *cnr* (Colorless non-ripening) présente une altération du phénotype de maturation du fruit avec la présence de zones immatures au niveau du fruit causée par un épiallèle naturel métastable (**Figure I.23B**). En effet, il a été montré que le phénotype de *cnr* est le résultat d'une épimutation spontanée créant une hyperméthylation au niveau de la région promotrice du locus *LeSPL-CNR* conduisant à une réduction du niveau d'expression de *LeSPL-CNR* (Manning *et al.*, 2006). Chez la tomate, une seconde épimutation contrôlant l'accumulation de la vitamine E dans le fruit a été décrite par Quadrana *et al.* (2014). Cette dernière a un effet sur le taux de vitamine E *via* la régulation de l'expression du gène *VTE3(1)* (Vitamin E 3) codant une 2-méthyl-6-phytylquinol méthyltransférase qui est impliquée dans la synthèse des formes γ et α du tocophérol (Quadrana *et al.*, 2014). Quadrana *et al.* (2014) ont montré que la région promotrice du gène *VTE3(1)* pouvait exister sous différents états méthylés conduisant à la modification de l'expression de *VTE3(1)* et ceci en lien avec l'état de méthylation d'un rétrotransposon SINE localisé dans sa région promotrice.

Chez le melon, la transition des fleurs mâles vers les fleurs femelles implique le facteur de transcription *CmWIP1*. L'étude menée par Martin *et al.* (2009) a montré que ce gène présente un épiallèle stable et héritable mettant en jeu l'insertion d'un élément transposable, *Gyno-HAT*, à proximité du gène *CmWIP1*. Le TE est responsable de la mise en place et du maintien de la méthylation de la région promotrice de *CmWIP1* conduisant à la modification de son niveau d'expression et du phénotype reproducteur du melon (**Figure I.23C**).

Chez le riz, l'étude d'un mutant spontané de riz *epi-d1* montrant un phénotype de nanisme a entraîné l'identification d'une épimutation contrôlant la structure du tallage menant au phénotype *dwarf*

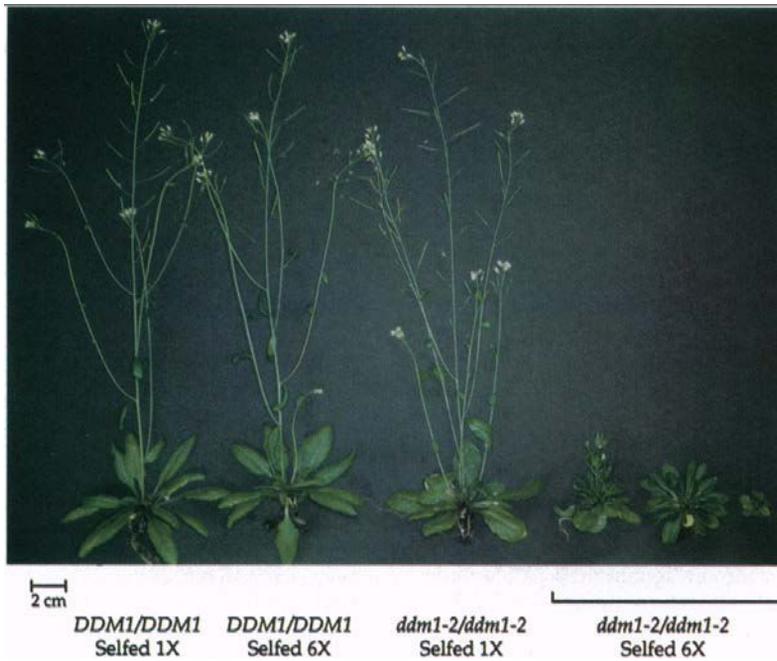
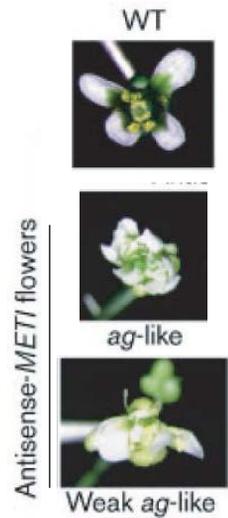
A**B****C**

Figure I.24. Variations morphologiques observées chez des *Arabidopsis* sauvages et chez des *Arabidopsis* altérées au niveau de *DDM1* et *MET1*. (A) Phénotypes morphologiques observés chez des lignées sauvages et chez le mutant *ddm1/ddm1* multipliées par autofécondation. Des altérations morphologiques importantes sont observables dans les mutants *ddm1* uniquement après plusieurs étapes de multiplication par autofécondation. En effet, l'individu mutant *ddm1* obtenu après une seule multiplication par autofécondation ne présente que de faibles altérations morphologiques (une légère réduction de la dominance apicale et des feuilles plus rondes) par rapport au sauvage. Toutes les plantes présentées sur la photo ont le même âge (40 jours) et ont été cultivées en parallèle et dans les mêmes conditions de culture. Issue de Kakutani *et al.* (1996). (B) Phénotypes foliaires observés dans des lignées sauvages et des lignées anti-sens *MET1*. Les fleurs des lignées anti-sens *MET1* présentent des niveaux réduits d'expression du gène *AGAMOUS*. Par rapport à l'individu sauvage (WT) une altération de la morphologie foliaire est clairement identifiable dans les lignées anti-sens *MET1* avec la présence de phénotypes foliaires similaires (*ag-like*) et intermédiaires (Weak *ag-like*) à celui de mutant du gène *AGAMOUS*. Adaptée de Jacobsen *et al.* (2000). (C) Phénotypes foliaires observés dans la lignée sauvage d'*Arabidopsis Landsberg erecta* et dans des mutants *sup* et *clk*. La plante sauvage *Landsberg erecta* (Ler) montre un phénotype foliaire normal avec la présence de 6 étamines et le pistil central composé de 2 carpelles fendues. Chez le mutant homozygote *sup-5* la fleur est quant à elle composée de 11 étamines et de pistils issus de la fusion incomplète de 3 carpelles. Chez le mutant homozygote *clk-3* la fleur est quant à elle composée de 9 étamines et de pistils issus de la fusion incomplète de 3 carpelles. Adaptée de Jacobsen et Meyerowitz (1997).

(**Figure I.23D**). Cet épiallèle impacte le fonctionnement du gène *DWARF1* (*DI*) qui a été décrit comme faisant varier la taille des feuilles, des panicules et des grains. L'allèle *Epi-di* a été démontré comme responsable du phénomène héritable et métastable (qui peut retourner à son état normal) de nanisme via la présence de marques histones et de méthylation répressives au niveau du promoteur de *DI* et ceci sans variation de séquence entre l'allèle *Epi-normal* et *Epi-dwarf* (Miura *et al.*, 2009).

Chez le palmier à huile, le phénotype anormal du fruit (**Figure I.23E**), dit « Mantled » est causé par la présence d'un épiallèle spontané au niveau du rétrotransposon *karma* présent au sein d'un intron du gène *DEFICIENS*. Cet épiallèle existe sous deux formes : la forme hyperméthylée du rétrotransposon appelée « the Good Karma epiallele » mène à la formation de fruit normaux et la forme hypométhylée appelée « the Bad Karma epiallele » conduit à une version altérée du fruit causant d'importantes pertes de rendement (Ong-Abdullah *et al.*, 2015).

Chez *Arabidopsis* plusieurs épiallèles héritaibles différents ont été décrits, dont des épiallèles obligatoires car issus de la duplication des gènes. Parmi ces épiallèles certains affectent i) les processus d'incompatibilité en modulant l'expression des gènes *HISN6A/HISN6B* (Blevins *et al.*, 2017) et des gènes *AtFOLT1/AtFOLT2*. Dans le cas des gènes *AtFOLT1/AtFOLT2* le réarrangement spécifique présent au niveau de *AtFLOT2* conduit à la méthylation de *AtFOLT1* via la voie RdDM (Durand *et al.*, 2012) ii) la synthèse du tryptophane en influant sur les gènes *PAI* (Phosphoribosylanthranilate isomerase ; Bender, 2004).

D'autres épiallèles sont issus de la présence d'éléments transposables à proximité d'un gène. C'est le cas du rétrotransposon *NMR19-4* (Naturally occurring DNA methylation variation region 19-4) qui contrôle la sénescence foliaire en régulant l'expression du gène *PPH* (Pheophorbide hydrolase) et l'acclimatation (He *et al.*, 2018). De plus, l'état de méthylation de *NMR19-4* a été corrélé au climat local suggérant une fonction de ce dernier dans l'adaptation locale au climat. Finalement, un épiallèle héritable décrit chez *Arabidopsis* ne montre pas de variations de séquence au niveau des régions impactées, c'est le cas du gène *QOS* (*Qua-Quine Starch*) qui est impliqué dans le métabolisme de l'amidon et qui est sujet à de nombreuses variations épigénétiques dans la nature. Le niveau d'expression de ce gène a été corrélé négativement au niveau de méthylation des séquences répétées se trouvant au niveau de l'extrémité 5' de ce gène (Silveira *et al.*, 2013).

2.4.5.2. Epiallèles induits et populations épigénétiques

Afin d'analyser les bases moléculaires et les mécanismes responsables de la mise en place et du maintien de la méthylation, différents mutants d'*Arabidopsis* altérés au niveau de gènes codant des modificateurs de la chromatine ont été utilisés. Parmi les mutants utilisés, les mutants *met1* et *ddm1*, altérés dans le maintien de la méthylation, présentent des niveaux réduits de méthylation CG ainsi

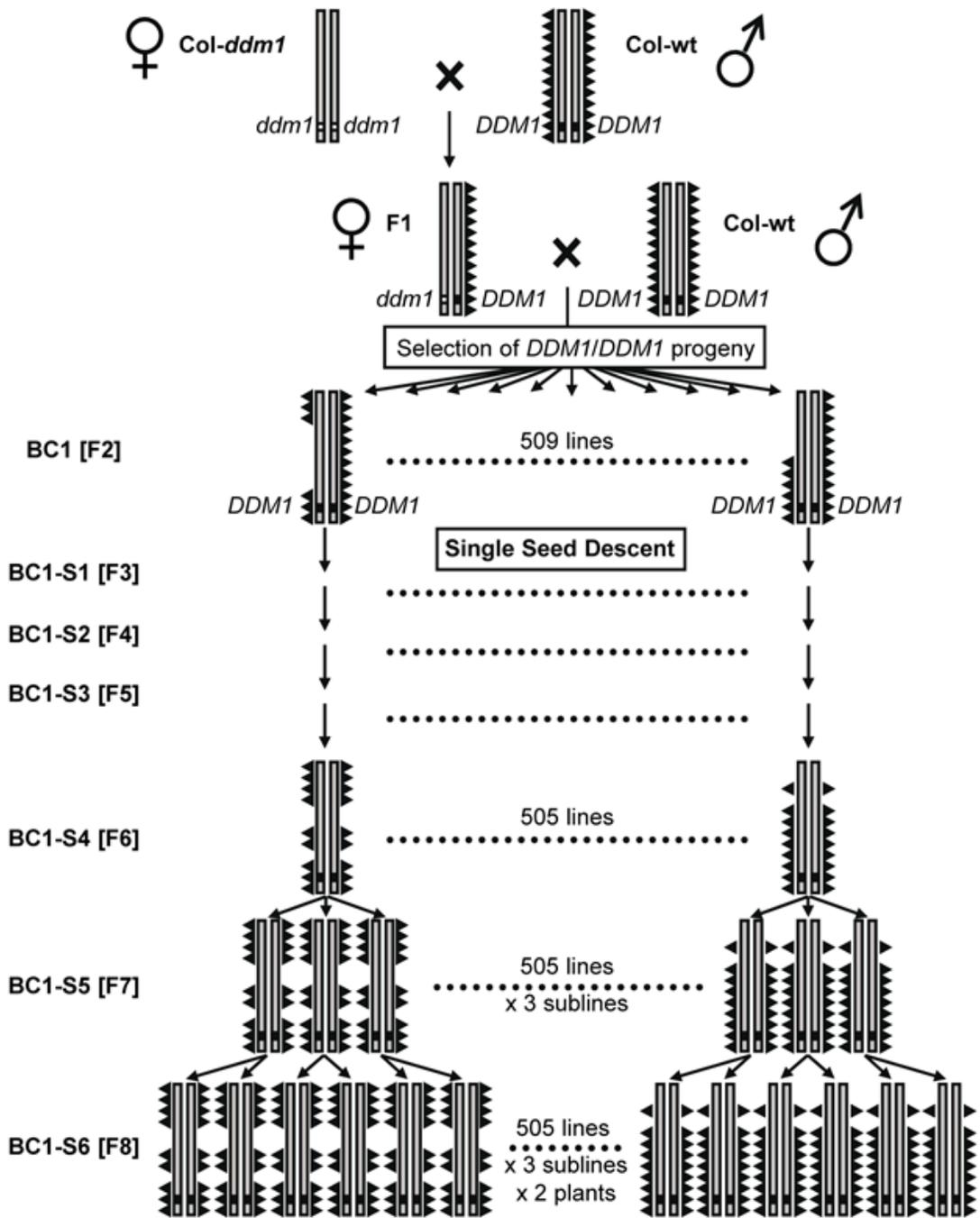


Figure I.25. Schéma de la construction de la population épIRIL *ddm1-2* x *Col-0*. Les barres grises représentent le génome d'*Arabidopsis thaliana* et les triangles représentent la méthylation de l'ADN. A l'exception du locus *DDM1* (carrés noirs et blancs) situé sur le chromosome 5, les deux parents (*Col-wt* et *Col-ddm1*) sont quasi-isogéniques; ils diffèrent dans leurs niveaux de méthylation de l'ADN. Un individu F1 a été rétrocroisé avec la lignée parentale *Col-wt* et 505 individus *DDM1 / DDM1* BC1 ont été auto-fécondés. Après trois autres autofécondations (BC1-S4), trois sous-lignées indépendantes ont été établies et auto-fécondées une fois pour obtenir la population d'épIRIL *Col-wt*. Issue de Johannes *et al.* (2009).

qu'une altération importante de leur développement (**Figure I.24A et B**). En effet la mutation du gène *DDM1* et du gène *MET1* conduit à un retardement de la floraison, à un nanisme et à une structure florale altérée par rapport au génotype sauvage (Kakutani *et al.*, 1996 ; Jacobsen *et al.*, 2000). C'est le cas par exemple des 7 épiallèles (*clark kent (clk)* 1 à 7) hyperméthylés du gène SUPERMAN (SUP) héréditaires mais instables et qui influencent le développement floral d'*Arabidopsis* en causant une augmentation du nombre d'étamines et de carpelles (Jacobsen & Meyerowitz, 1997 ; **Figure I.24C**). L'épiallèle hypométhylié de *FWA* résulte de la déméthylation de l'élément transposable SINE présent dans son promoteur et conduit à l'expression de *FWA* induisant un retard de la floraison (Soppe *et al.*, 2000). De manière intéressante, il a été montré que la perte de méthylation induite par les mutants *ddm1* et *met1* persiste dans le temps et ceci même après l'élimination des mutations causales dans la descendance (Vongs *et al.*, 1993 ; Kankel *et al.*, 2003).

L'observation d'épiallèles, qu'ils soient naturels ou induits, associés à des variations phénotypiques touchant notamment des caractères d'intérêt agronomique montre leur importance dans les différents processus développementaux des plantes. Cependant, l'étude de l'aspect épigénétique est très difficile (Johannes *et al.*, 2008) puisque les effets de(s) l'épiallèle (s) et l'allèle (s) sur le (s) phénotype (s) sont difficilement distinguables. Différentes populations recombinantes épigénétiques (épiRIL) ont été créées en utilisant soit *ddm1*, soit *met1* afin d'évaluer l'impact de modifications épigénétiques transgénérationnelles sur le phénotype végétal. L'avantage de ces populations est la présence d'une forte variabilité épigénétique avec une faible variabilité génétique. Ces populations permettent ainsi de relier plus facilement la (les) variation(s) phénotypique(s) observée(s) à l'effet(s) épigénétique(s) transgénérationnel *versus* l'effet génétique(s).

2.4.5.2.1. Construction et caractérisation des populations épigénétiques

Deux populations épiRIL dans un fonds génétique quasi isogénique ont ainsi été créées chez *Arabidopsis* i) la population épiRIL créée par Johannes *et al.* (2009) a été réalisée en croisant un individu Columbia sauvage et le mutant EMS *ddm1-2* (*Col-ddm1*, de 4^{ème} génération ; Vongs *et al.*, 1993). Une plante F1 a ensuite été rétro-croisée en tant que parent femelle avec le parent sauvage Columbia. A partir de là, 505 individus homozygotes *DDM1/DDM1* ont été sélectionnés et multipliés pendant 6 générations par autofécondation (sélection monograine ; **Figure I.25**). ii) la population épiRIL créée par Reinders *et al.* (2009) est-elle issue du croisement entre un individu Columbia sauvage et un mutant nul *met1-3* (**Figure I.26**). Les individus F1 obtenus ont été autofécondés et seuls les individus F2 présentant l'allèle sauvage de *MET1* à l'état homozygote ont été conservés. Cette sélection a été réalisée afin de restaurer immédiatement une hérédabilité transgénérationnelle stable des variations de méthylation ségrégeant au sein des individus. Finalement les 96 individus *MET1* F2 sélectionnés ont été autofécondés et multipliés par sélection monograine et autofécondations jusqu'en génération F7. Le phénotypage et l'analyse de l'ADN ont été réalisés sur les plantes F8 issues d'un bulk de 9 plantes F7.

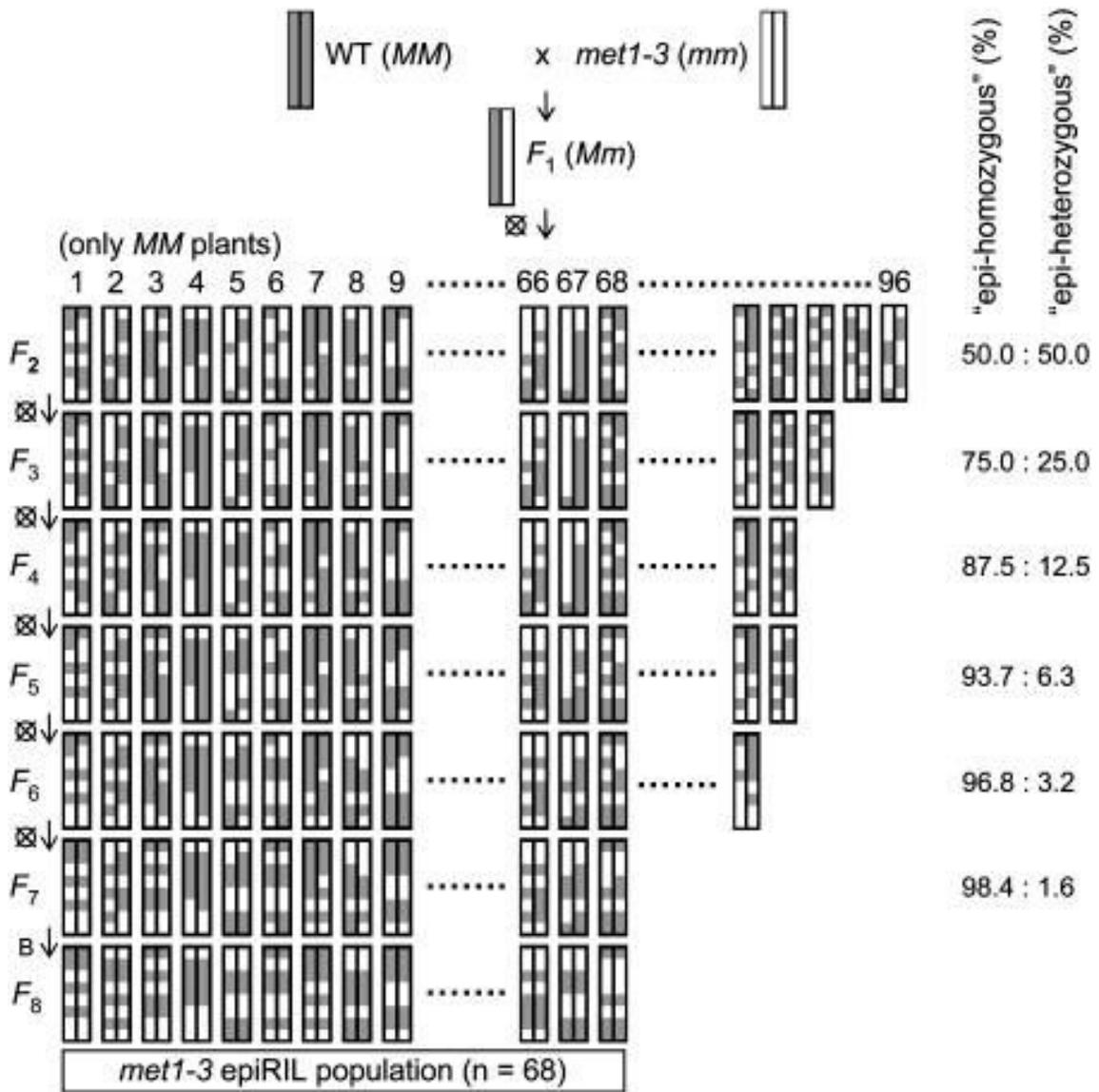


Figure I.26. Schéma de la construction de la population épiRIL *met1-3* x Col-0. La construction de la population épiRIL est illustrée à l'aide d'une paire de chromosomes. Les chromosomes des lignées parentales ainsi que leurs fragments présents au niveau des épiRIL sont annotés WT et MM ou colorés en gris pour le parent Col-0 et annotés *met1-3*, mm ou colorés en blanc pour le parent *met 1-3*. L'axe vertical montre les 8 générations de lignées épiRIL obtenues marquées de F₂ à F₈. Le cercle couplé à une croix sur la gauche des flèches indique les descendants obtenus à partir d'une seule graine. Une récolte massive des individus F₇ est illustrée par un B à la gauche de la flèche entre la F₇ et la F₈. Les niveaux prédits d'épi-homozygotes et d'épi-hétérozygotes à chaque génération sont indiqués à droite de la figure. Issue de Reinders *et al.* (2009).

Au total seulement 68 des 96 lignées épiRIL créées ont pu être amenées jusqu'en génération F8 du fait d'une perte de fitness des plantes (**Figure I.26**). Au cours de la création de leur population, Reinders *et al.* 2009 ont noté la présence de manière stochastique et instable de variations développementales au sein et entre les épiRIL.

Dans les deux populations épiRIL obtenues, des variations de l'état de méthylation entre les différentes lignées ont été observées. Chez la population de Reinders *et al.* (2009) des états de méthylation stables et hérités de chaque parent (wild-type-like et *met*-like) ont été décelés. Ils ont également détecté en proportions significatives dans la population *met1* x Col-0 des états de méthylation non parentaux ainsi que des états épihétérozygotes (**Figure I.27**). Johannes *et al.* (2009) ont analysé l'état de méthylation de 22 Col-wt épiRIL présentant des dates de floraisons extrêmes par rapport au reste de la population et ont mis en exergue la présence d'états de méthylation stables et hérités issus de chacun des deux parents avec cependant la présence de révertants de méthylation au niveau de régions initialement hypométhylés dans le mutant *ddm1*.

Les deux populations étant issues de croisement avec des mutants de modification de la chromatine, des mouvements d'éléments transposables lors de la création des populations ne pouvaient être écartés. En effet, des études précédentes avaient montré des mouvements des TE CACTA et MULES chez le mutant *ddm1* (Singer *et al.*, 2001 ; Miura *et al.*, 2001 ; Kato *et al.*, 2003) confirmés dans la population de Johannes *et al.* (2009). De manière plus inattendue des mouvements de TE CACTA ont aussi été observés dans la population de Reinders *et al.* (2009) alors qu'aucun mouvement du TE CACTA n'avait été montré dans le simple mutant *met1*.

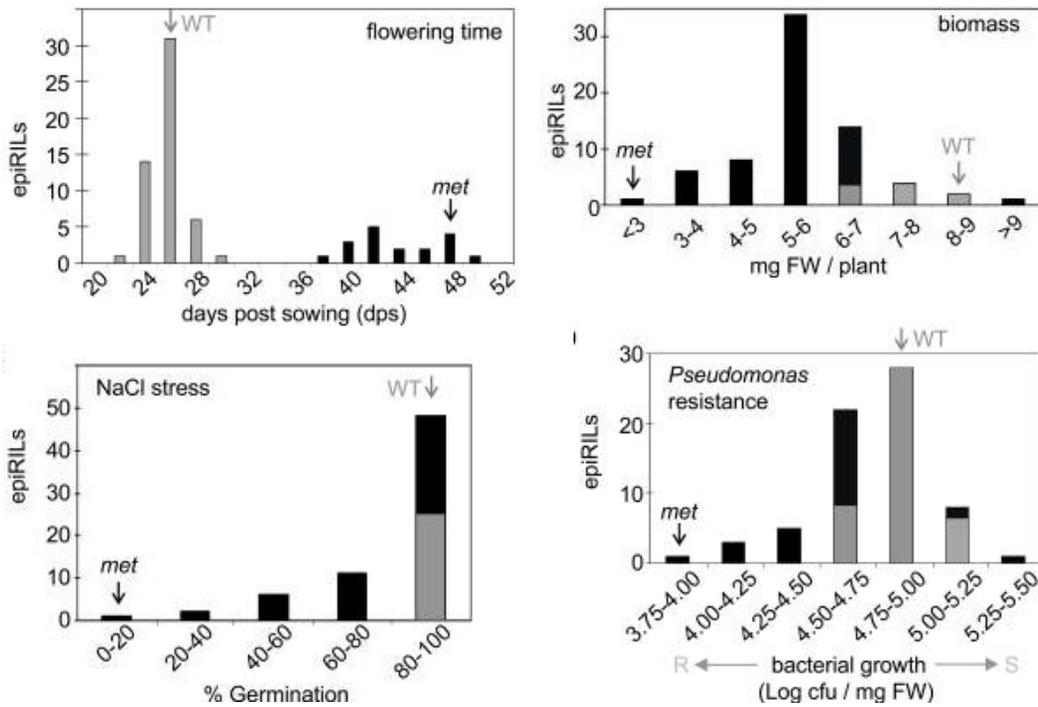
Dans les deux populations créées, des variations phénotypiques importantes ont été observées pour différents caractères.

2.4.5.2.2. Populations épigénétiques et variations phénotypiques

Différents traits phénotypiques ont été analysés dans chaque population afin de mettre en évidence le rôle des variations transgénérationnelles de méthylation dans les variations phénotypiques.

Reinders *et al.* (2009) ont ainsi analysé le temps de floraison, la croissance des plantes (biomasse), les réponses au stress salin lors de la germination et à l'infection par la bactérie pathogène *Pseudomonas syringae* pv. tomato (*Pst* ; **Figure I.27A**). Pour chacun des traits analysés, ils ont montré des variations phénotypiques en fonction des épiRIL. La distribution bimodale du caractère précocité de floraison dans la descendance épiRIL suggérait la présence d'un seul épiallèle responsable de la variation du caractère, avec 26 % des épiRIL montrant un retard de floraison. Le mutant *met1-3* étant connu pour induire un état hypométhylé de *FWA* conduisant à son expression et à un retard de floraison, ils ont analysé l'état de méthylation du TE SINE présent à l'intérieur de la région promotrice de *FWA*. Ils ont

A



B

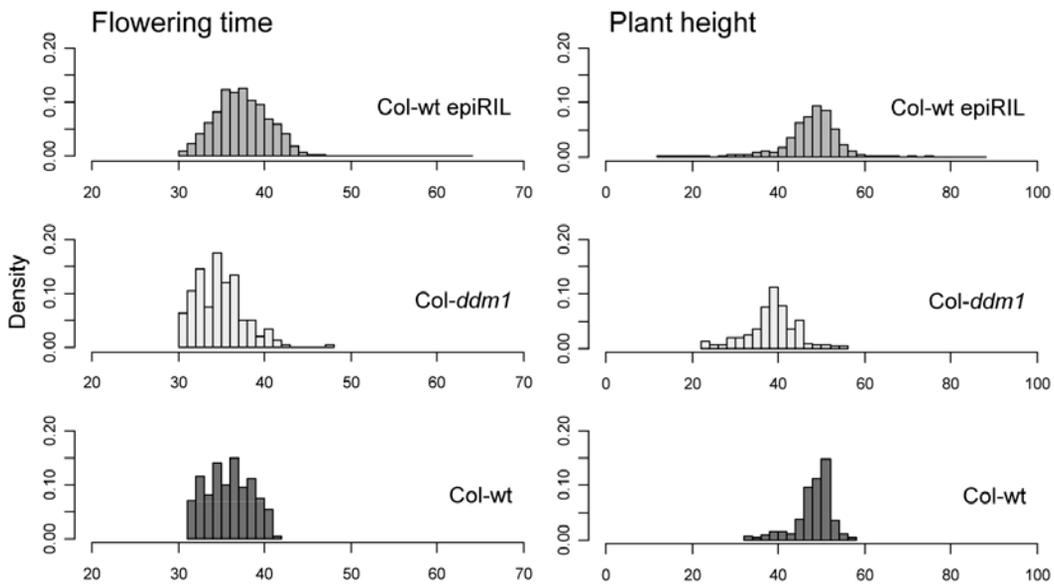


Figure I.27. Variations phénotypiques observées dans les populations épiRIL *ddm1-2* x *Col-0* et *met1-3* x *Col-0*. (A) Variations phénotypiques observées dans la population épiRIL *met1-3* x *Col-0*. Pour tous les graphiques, l'axe x montre les variations du trait et l'axe y le nombre de lignées épiRIL dans l'intervalle. Différents traits ont été analysés, le temps de floraison (flowering time), la biomasse (biomass), le taux de germination en condition de stress salin (NaCl stress) et la résistance à *Pseudomonas* (*Pseudomonas* resistance). Les épiRIL montrant une variation phénotypique de type *Col-0* ou différente de *Col-0* (niveau de confiance à 95%) sont colorés respectivement en gris et en noir. Les parents *Col-0* et *met1-3* sont indiqués respectivement par les flèches grises et noires. (dps) jours post semis ; (cfu) unités formant une colonie ; (FW) poids frais. (WT) *Col-0* sauvage ; (m) *met1-3*. Le taux de germination des graines (% Germination) a été effectué en présence de 150 mM de NaCl appliqué à 4 dps. Pour la réponse à *Pseudomonas*, les réponses de type résistante (R) et sensible (S) sont indiquées sur l'axe des x. Adaptée de Reinders *et al.* (2009). (B) Distribution des valeurs brutes des variations phénotypiques observées dans la population épiRIL *ddm1-2* x *Col-0* pour le temps de floraison et la taille des plantes. Les traits ont été mesurés sur les lignées parentales *Col-0* (*Col-wt*), *ddm1-2* (*Col-ddm1*) et les épiRIL (*Col-wt* épiRIL). Le temps de floraison (en jours) et la taille des plantes (en cm) sont représentés sur l'axe x. La densité est représentée sur l'axe y. Adaptée de Johannes *et al.* (2009).

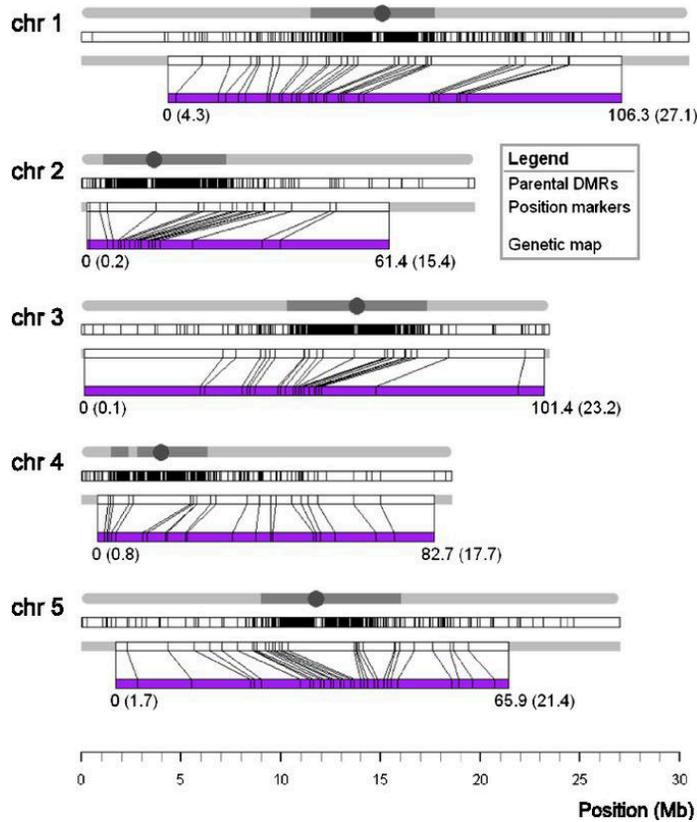
ainsi démontré que pour 11 épiRIL montrant un retard de floraison, un épiallèle de type *met1*-like était présent validant ainsi le rôle de *FWA* dans ce retard de floraison. Pour les autres caractères analysés, la distribution continue des valeurs phénotypiques dans la descendance suggérait la présence de multiples épiallèles contrôlant les traits étudiés. Un retard de germination en condition de stress salin était observé pour 60 % des épiRIL par rapport au sauvage, ainsi qu'une variation de la réponse de résistance à *Pseudomonas* avec 34 % des épiRIL plus résistants et 4 % plus sensibles que le sauvage. Cette population épiRIL de Reinders *et al.* (2009) n'a pas été utilisée par la suite pour des études d'épigénétique quantitative du fait de sa faible taille (68 individus) et des fortes instabilités des états de méthylation notamment avec la présence d'épiallèles dans des segments parentaux qui ne correspondent pas à ce parent.

Johannes *et al.* (2009) ont aussi mis en évidence un large impact du méthylome de leurs épiRIL sur les traits phénotypiques analysés (temps de floraison et taille des plantes ; **Figure I.27B**), ainsi que des valeurs modérées d'héritabilité au sens large de 0,26 et 0,32 pour le temps de floraison et la taille des plantes respectivement. Pour les deux caractères étudiés, ils ont observé une distribution continue des valeurs phénotypiques suggérant un contrôle par de multiples épiallèles. L'analyse du locus *FWA* a montré dans leur population que bien que l'épiallèle présent à ce locus joue un rôle dans le retard de floraison, ce rôle est mineur et que d'autres épiallèles expliqueraient la variabilité phénotypique observée.

Un sous ensemble de 123 lignées de la population créée par Johannes *et al.* (2009) a été utilisé par Colomé-Tatché *et al.* (2012) pour réaliser la construction d'une carte de recombinaison. Une approche de MeDIP-chip couplant immunoprécipitation ciblant les cytosines méthylées avec une hybridation des séquences capturées sur une puce à ADN a été réalisée chez les lignées parentales et les 123 épiRIL (Colomé-Tatché *et al.*, 2012). Les signaux obtenus sur la puce ont été catégorisés en trois états, non méthylé (U), méthylé de manière intermédiaire (IM) et méthylé (M) *via* un modèle HMM (Hidden Markov Model). Au total 2 611 régions différentiellement méthylées (DMR) ont été identifiées (**Figure I.28A**) entre les deux parents avec une diminution nette de l'état de méthylation chez *ddm1* avec un passage de l'état M à l'état U. Sur ces 2 611 DMR, 867 DMR parentales stables ont été identifiées dans les 123 épiRIL en F8. Les autres DMR identifiées chez les parents ont été sujettes à des phénomènes de réversion de l'état de méthylation *ddm1* vers l'état de méthylation sauvage de manière complète ou incomplète. Après une dernière étape de filtration des DMR identifiées dans les 123 épiRIL comprenant notamment la suppression des DMR donnant des marqueurs génétiquement redondants (les marqueurs situés à moins de 0,0001 cM d'écart), 126 DMR ont été retenues. Finalement la carte créée contient 126 DMR stables méiotiquement qui couvrent 81,9 % du génome (**Figure I.28A et B**).

L'évaluation morphologique du sous-ensemble de 123 lignées de la population épiRIL par Cortijo *et al.* (2014) a démontré que plusieurs DMR héritées agissent comme des loci quantitatifs sous contrôle

A



B

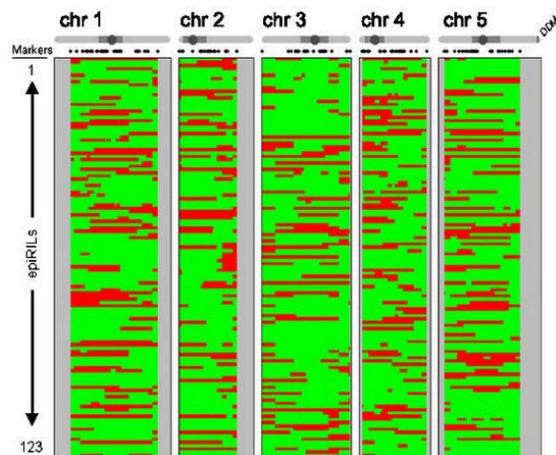


Figure I.28. Carte de recombinaison des épiRIL *ddm1-2* x Col-0. (A) Distribution le long du génome des 2 611 DMR parentales (traits noirs sur la barre du milieu) et des 126 DMR (*i.e.*, marqueurs ; traits noirs sur la barre du bas) retenues pour le construction de la carte de recombinaison (en violet) pour chacun des 5 chromosomes d'*Arabidopsis*. L'emplacement des 126 marqueurs est aussi bien montré de manière physique (échelle présente en bas de l'image) que génétique (sur les barres en violet). (B) Représentation de l'épihaplotype (WT en vert, *ddm1-2* en rouge) présent au niveau de chacun des 126 marqueurs utilisés et pour les 123 lignées épiRIL. Les épihaplotypes sont représentés sur l'axe x tandis que les 123 lignées épiRIL sont représentées sur l'axe y. Les extrémités des chromosomes non couvertes sur la carte génétique par des marqueurs sont en gris. Adaptée de Colomé-Tatché *et al.* (2012).

épigénétique (QTL^{épi}) pour deux caractères complexes. Pour les caractères temps de floraison et longueur de racine primaire, les DMR expliquent 60 % à 90 % de leur héritabilité. En outre, cette population épiRIL a été utilisée pour évaluer la contribution épigénétique dans la plasticité phénotypique des plantes, à savoir la capacité d'un génotype à exprimer des phénotypes alternatifs dans différents environnements. L'exposition de la population épiRIL à différentes contraintes abiotiques, telles que la sécheresse, la contrainte nutritionnelle et le stress salin a montré que les variations de méthylation dans le génome contribuaient à des variations phénotypiques notables et a de la plasticité phénotypique, permettant notamment l'identification de QTL^{épi} (Zhang Yuan-Ye *et al.*, 2012 ; Kooke *et al.*, 2015b) contrôlant ces différents caractères. Récemment, il a également été montré que les variations de méthylation transgénérationnelles présentes dans cette population épiRIL influencent sur les niveaux de glucosinolates au sein des plantes, avec notamment la détection de plusieurs QTL^{épi}. Cependant, l'effet des variations de méthylation était inférieur à celui induit par les variations génétiques (Aller *et al.*, 2018).

Qu'elles soient transgénérationnelles ou non, les variations de méthylation ont de nombreuses fonctions au sein de la plante et peuvent notamment influencer son phénotype, divers processus cellulaires ainsi que son interaction avec son environnement.

2.4.6. Méthylation de l'ADN et réponses de la plante aux stress

Que cela soit dans le cadre de stress abiotique ou biotique, une relation étroite entre les modifications de l'épigénome et l'exposition des plantes au stress a été identifiée. Dès lors, l'importance de la gestion du méthylome de la plante lors des interactions avec son environnement semble être un élément potentiellement important dans l'adaptation.

2.4.6.1. Cas des stress abiotiques

Un nombre important d'études ont été réalisées afin d'identifier le rôle potentiel de la méthylation de l'ADN dans la réponse des plantes aux stress d'origine abiotique. Ces études ont été réalisées chez des espèces d'intérêt agronomique telles que le riz, la tomate, le maïs, le colza mais aussi l'espèce modèle *Arabidopsis thaliana*.

Certaines études ont montré des effets de différents stress abiotiques sur la méthylation de l'ADN, tels que :

- Le stress froid. Chez *Brassica rapa*, des phénomènes de méthylation et de déméthylation du génome de la plante ont lieu pendant le processus d'acclimatation au froid (Liu *et al.*, 2017). Parmi les 29 624 DMR identifiées entre les individus en condition contrôle et ceux en acclimatation au froid, 4 DMR ont été démontrées comme ayant une influence sur l'expression de 4 gènes (Liu *et al.*, 2017). En réponse au stress froid le même phénomène de déméthylation

globale de l'ADN par rapport au contrôle est observé chez les racines de maïs (Steward *et al.*, 2002). L'analyse des conséquences du stress froid imposé aux tomates afin de les préserver lors de leur commercialisation a montré que le froid modifiait de manière transitoire la méthylation de l'ADN avec notamment l'identification de 30 918 DMR dans trois tissus différents. Contrairement aux observations faites précédemment sur l'hypométhylation induite par un stress froid chez *B. rapa* et le maïs, le froid appliqué pour le stockage des tomates conduit à une augmentation de la méthylation de gènes connus pour contribuer à la maturation des fruits qui coïncide avec leur diminution d'expression (Zhang *et al.*, 2016).

- Le stress thermique. L'étude des conséquences d'un stress thermique court sur les microspores de *Brassica napus* cv. Topas a montré qu'un stress température de 32 °C pendant 6 h mène à une hypométhylation globale de l'ADN surtout au niveau des contextes CG et CHG (Li *et al.*, 2016a). De manière intéressante, chez *Arabidopsis* la culture des plantes lors d'un stress thermique conduit à une augmentation globale du niveau de méthylation de l'ADN (Boyko *et al.*, 2010).
- Le stress salin. L'utilisation des lignées accumulatrices de mutations (MA lines) d'*Arabidopsis thaliana* en condition de stress salin (haute concentration en sel dans le sol) a mis en avant que ce stress modifie la fréquence des cytosines méthylées entre les lignées. En effet, une accumulation d'environ 45 % de cytosines différenciellement méthylées en CG (CG-DMP) pour les plantes ayant été cultivées en condition de stress salin pendant 10 générations (G10) par rapport au contrôle G1. De plus les auteurs ont observé une augmentation de la moyenne des fréquences d'apparition des DMR d'environ 200 % en condition de stress par rapport aux lignées contrôles. Finalement, l'utilisation d'une lignée G10 venant de la condition de stress salin en culture dans un sol contrôle a montré que pour la génération G11 générée, environ 75 % des CG-DMP étaient héritées (Jiang *et al.*, 2014). Deux lignées accumulatrices d'épimutations (epi-MA) de riz ont également été soumises à un stress hydrique sur 11 générations successives. Cette étude a montré que l'apposition de stress multiples conduit à une modification du méthylome des lignées de riz avec la présence d'épimutations au niveau de « hotspots ». En plus de la différence entre les plantes cultivées en stress hydrique et sans stress, la comparaison des taux de méthylation entre les lignées après une seule génération de stress (G0) et après les 11 générations de stress (G11) montre que l'accumulation des stress conduit à une diminution du niveau de méthylation entre G0 et G11 (Zheng *et al.*, 2017).

A contrario, différentes publications n'ont pas mis en avant de modifications consistantes de la méthylation de l'ADN suite à différents stress. L'exposition du maïs après 14 jours de culture en condition standard à des températures chaudes (4 traitements à 50 °C de 4 heures espacés d'une journée), températures froides (4 traitements à 4 °C de 4 heures espacés d'une journée) et aux UV (4 traitements UV à 60–64 $\mu\text{moles.m}^{-2}.\text{s}^{-1}$ de 4 heures espacés d'une journée) n'ont pas permis de mettre

en évidence des variations stables de méthylation de l'ADN entre la condition stressée et la condition contrôle (Eichten & Springer (2015). L'application d'un stress hydrique chez *Arabidopsis* pendant une seule génération et pendant 5 générations n'a entraîné que peu de variations de la méthylation de l'ADN suggérant que la méthylation de l'ADN serait peu modifiée au cours des stress hydriques (Ganguly *et al.*, 2017).

Bien que la méthylation de l'ADN semble jouer un rôle important lors des stress abiotiques, les résultats obtenus contrastés ne permettent pas une compréhension claire du mécanisme sous-jacent. Concernant les modifications de méthylation induites lors de stress, même si un maintien pendant une génération en absence de stress a pu être observé dans le cadre du stress salin, la question d'un maintien en absence de stress pendant de multiples générations se pose encore.

2.4.6.2. Cas des stress biotiques

L'utilisation de produits conduisant à la déméthylation a permis de mettre en avant un lien entre l'état de méthylation du génome de la plante et la réponse de résistance aux agents pathogènes. C'est le cas par exemple de l'augmentation de la résistance chez le riz à *Xanthomonas oryzae* pv. *oryzae* suite au traitement par le 5-azadeoxycytidine (Akimoto *et al.*, 2007). Chez des mutants altérés pour des gènes modificateurs de la chromatine tels que le mutant *met1* (modifiant la méthylation en CG), le double mutant *met1 nrpd2* et le triple mutant *ddc* (altérant à la fois la voie RdDM et le maintien de la méthylation en CG ou CHG), une augmentation de la résistance à *P. syringae* DC3000 est observée (Luna *et al.*, 2012 ; Downen *et al.*, 2012 ; Yu *et al.*, 2013). De plus, chez le double et le triple mutant, la propagation du phytopathogène au sein des tissus vasculaires était diminuée par rapport à l'individu sauvage (Yu *et al.*, 2013). L'utilisation d'un mutant du gène *ROS1*, responsable de la déméthylation active de l'ADN corrobore les résultats obtenus avec les mutants du maintien de la méthylation avec une plus grande sensibilité du mutant *ros1* à *P. syringae* DC3000 (Yu *et al.*, 2013). D'autre part, il a été montré chez *Arabidopsis* que la méthylation de l'ADN est un élément important pour la suppression du développement d'*Agrobacterium tumefaciens* ; contrairement à sa résistance accrue à Pto DC3000, le mutant *ddc* montre dans le cas d' *A. tumefaciens* un plus grande sensibilité (Bond & Baulcombe, 2015). L'inoculation du variant nain métastable *bal* dérivé du mutant *ddm1* par *P. syringae* (pv. tomato strain DC 3000) a mis en avant la présence d'une plus grande résistance due à la surexpression du gène *NLR SNC1* (Stokes *et al.*, 2002). Le mutant *ago4-2*, altéré au niveau du gène *ARGONAUTE4* impliqué dans le maintien de la méthylation en contexte CHG et CHH via la voie RdDM, montre une plus grande sensibilité à la forme virulente et avirulente de *P. syringae* pv *tomato* DC3000, ainsi qu'à *P. syringae* pv *tabaci* (Agorio & Vera, 2007). De plus, les mutants affectés pour la pol V montrent d'une part une plus grande sensibilité aux nécrotrophes *Botrytis cinerea* et *Plectosphaerella cucumerina* et d'autre part une plus grande résistance au biotrophe *P. syringae* DC300 (López *et al.*, 2011). Egalement, le mutant *nrpe1*, montrant une hypométhylation globale de l'ADN, est plus résistant à *Hyaloperonospora*

arabidopsidis alors qu'il est plus sensible au nécrotrophe *P. cucumerina*. A contrario, le mutant *ros1*, montrant une hyperméthylation globale de l'ADN, est plus sensible à *H. arabidopsidis* alors qu'il est plus résistant face à *P. cucumerina* (López Sánchez Ana *et al.*, 2016).

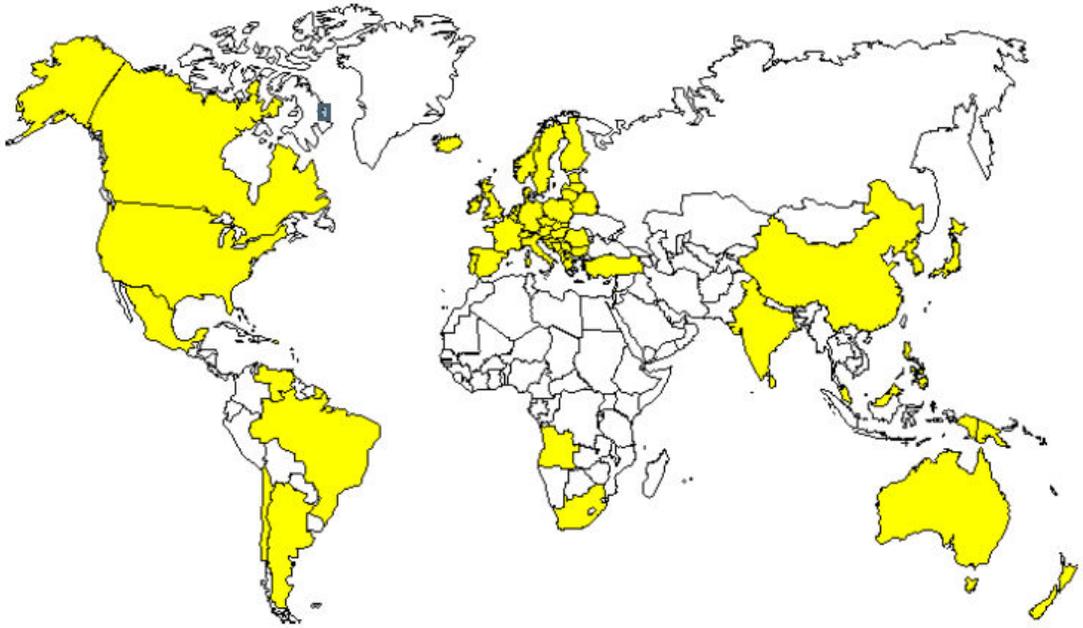
Le taux de méthylation chez la plante peut être modifié par l'infection par différents agents pathogènes. Chez *Arabidopsis*, l'infection par *P. syringae* DC300 ou lors de son interaction avec le nématode à kyste *Heterodera schachtii*, induit une hypométhylation du génome (Pavet *et al.*, 2006 ; Hewezi *et al.*, 2017). Une hypométhylation du génome lors de l'infection par le nématode à kyste a également été observée chez le soja (Rambani *et al.*, 2015).

Les éléments transposables peuvent également être un élément régulateur de la réponse des plantes à un stress biotique *via* une variation de leur niveau de méthylation. En effet, même si la transposition de TE peut être délétère pour les gènes, des cas d'effet positif d'insertions d'éléments transposables sur l'expression de gènes de défense ont été observés. Par exemple, le transposon « renovator » présent au niveau dans la région promotrice du gène *Pit* chez le riz va provoquer sa transcription et conférer une résistance au pathogène *Magnaporthe grisea* (Hayashi Keiko & Yoshida Hitoshi, 2009). L'analyse de la région promotrice du gène *RMGI* (Resistance Methylated Gene 1) présentant des éléments transposables a montré que ces derniers servaient de plateforme de contrôle de l'expression de ce gène. Cette régulation s'effectue par un jeu antagoniste de méthylation *via* la voie RdDM à l'état basal et de déméthylation lors de la réponse au stress biotique (Yu *et al.*, 2013). Le triple mutant de méthylases *rdd* (*ros1 dml2 dml3*) d'*Arabidopsis* montre également une plus grande sensibilité face à *Fusarium oxysporum*. Cette différence de sensibilité est cohérente avec la répression de l'expression de plus de 200 gènes, notamment impliqués dans la réponse au stress biotique, montrant des éléments transposables dans leurs séquences (Le *et al.*, 2014).

L'ensemble de ces observations montrent clairement l'impact de la méthylation de l'ADN dans la réponse aux stress abiotiques et biotiques par des régulations transitoires de l'état de méthylation de gènes de défense. Cependant le rôle potentiel de modifications stables de méthylation transmises au cours des générations et impliquées dans l'interaction plante pathogène reste relativement méconnu (Molinier *et al.*, 2006 ; Rasmann *et al.*, 2011 ; Slaughter *et al.*, 2011 ; Luna & Ton, 2012 ; Downen *et al.*, 2012 ; Yu *et al.*, 2013).

Lors de ma thèse je me suis intéressé à étudier cette relation entre épigénétique et réponse de résistance ou de sensibilité des plantes dans le cadre d'un pathosystème spécifique qui est le pathosystème *Brassicaceae/Plasmodiophora brassicae* qui est décrit dans la suite de cette introduction.

A



B

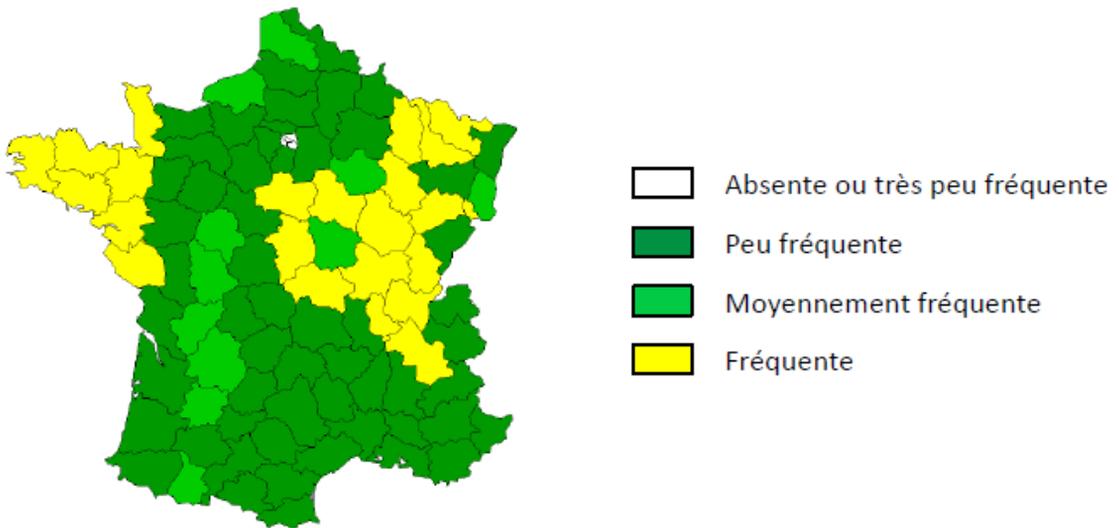


Figure I.29. Distribution mondiale et française de la hernie des crucifères. (A) Distribution mondiale de la hernie des crucifères. Les pays colorés en jaune sont ceux où la maladie a été détectée. Adaptée de Dixon (2009). (B) Représentation de la distribution de la hernie sur le sol français en 2008. Issue de Terres Inovia.

3. Le modèle d'interaction plante pathogène : *Plasmodiophora brassicae* - Brassicaceae

3.1. *Plasmodiophora brassicae*, agent causal de la hernie des Crucifères

3.1.1. Contexte et enjeu économique de la maladie

L'interaction entre les plantes de la famille des *Brassicaceae* et le protiste biotrophe obligatoire *Plasmodiophora brassicae* conduit à la maladie nommée hernie des crucifères. Cette maladie est une maladie majeure des *Brassicaceae* incluant les trois espèces *Brassica* d'importance agronomique mondiale *B. napus*, *B. rapa* et *B. oleracea* et la plante modèle *A. thaliana*. Les conséquences socio-économiques de cette maladie sont en augmentation depuis quelques années avec le classement de la hernie comme maladie émergente ou ré-émergente (Terres Inovia, 2008). Les premiers cas de hernie reportés sur des *Brassica* datent du XIII^{ème} siècle (Dixon, 2009a) mais l'identification de l'agent causal de cette maladie n'a été réalisée qu'en 1875. En effet, c'est à cette date qu'un scientifique russe, M. S. Woronin identifia un organisme plasmodiophoral comme causal de la maladie, qu'il nomma *Plasmodiophora brassicae* (Woronin, 1878). Actuellement la hernie est présente dans plus de 60 pays dans le monde (**Figure I.29A**) avec deux pays présentant des niveaux d'infection préoccupants que sont la Chine et le Canada, deux acteurs importants de la filière de production de *Brassicacées* (Chai *et al.*, 2014 ; Strelkov & Hwang, 2014). Au niveau Français, la hernie impacte les cultures légumières de *B. oleracea* ainsi que les cultures de colza avec des intensités différentes en fonction des régions françaises comme cela est illustré sur la **Figure I.29B**.

3.1.2. Biologie de *Plasmodiophora brassicae*

P. Brassicae est un eucaryote unicellulaire biotrophe obligatoire appartenant au taxon des Rhizaria dans l'ordre des Plasmodiophorida et le phylum des Cercozoaires du règne des protistes (Cavalier-Smith & Chao, 2003 ; Burki *et al.*, 2010 ; Neuhauser *et al.*, 2011 ; He *et al.*, 2014 ; Schwelm *et al.*, 2015). L'appartenance de *P. Brassicae* à l'ordre des Plasmodiophorida repose sur l'identification de la forme particulière de mitose, dite cruciforme qui caractérise les individus de cet ordre (Braselton, 1995). Au niveau génomique, les études visant à identifier la taille et le nombre de chromosomes présents chez *P. Brassicae* sont difficiles du fait de son caractère biotrophe. Le nombre de chromosomes a été estimé par électrophorèse en champ pulsé de 6 à 16 bandes chromosomiques avec des tailles allant de 1,7 Mb à 750 Kbselon les études (Ito *et al.*, 1994 ; J. Bryan *et al.*, 1996 ; Graf *et al.*, 2001 ; Fähling *et al.*, 2004 ; Siemens *et al.*, 2009), le nombre et la taille des chromosomes pouvant varier en fonction de l'isolat étudié (Graf *et al.*, 2004). Le génome de *P. Brassicae* a été séquencé récemment par Schwelm *et al.* (2015) à partir d'ADN génomique extrait de spores de repos de l'isolat monospore e3. Au total 24 Mb ont été assemblés *de novo* en 165 scaffolds représentant une taille de génome inférieure à celle mesurée pour deux autres espèces appartenant aux Rhizaria que sont *Bigelowiella natans* (~100 Mb) et *Reticulomyxa filosa* (~320 Mb). Malgré sa petite taille, le génome de *P. Brassicae* présente une

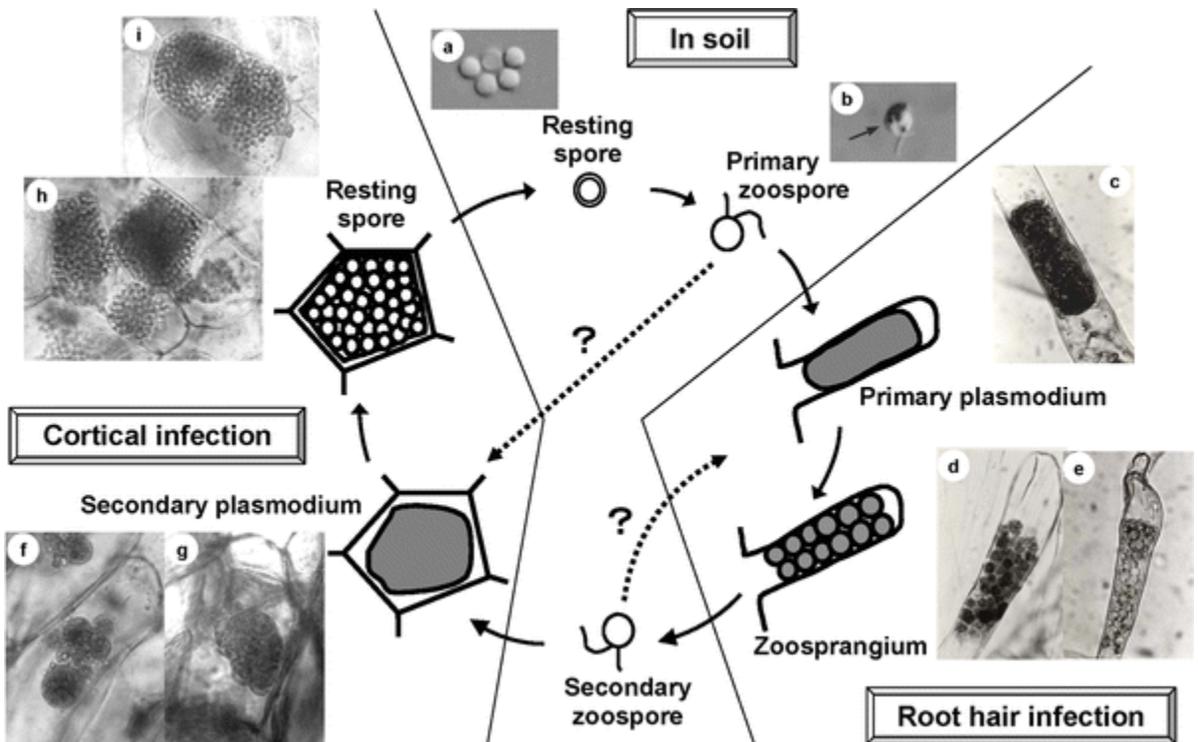


Figure I.30. Cycle de vie de *Plasmodiophora brassicae*. Le cycle de développement de *P. brassicae* se découpe en trois étapes successives, correspondant à deux phases : la survie de *P. brassicae* dans le sol et l'infection de la plante hôte. La phase d'infection de la plante est découpée en 2 étapes : l'infection des poils absorbants (infection primaire) et l'infection des cellules corticales de la racine (infection secondaire). (a) Spores de repos. (b) Zoospore primaire. (c) Plasmode primaire dans un poil absorbant. (d) Cluster de zoosporanges dans les poils absorbants. (e) Zoosporange vide. (f), (g) Plasmode secondaire dans les cellules corticales. (h) et (i) Spore de repos dans les cellules corticales. Adaptée de Kageyama et Asano (2009).

prédiction d'un total de 9 730 gènes qui est cohérente avec sa forte densité en gènes et son faible niveau de séquences répétées (Schwelm *et al.*, 2015). Cinq autres isolats de *P. Brassicae* d'origine canadienne ont été (re)-séquencés permettant de valider la structure générale du génome avec cependant, comme le suggéraient les premières études, la présence de variations entre isolats notamment concernant le nombre de gènes prédits (Rolfe *et al.*, 2016 ; Daval *et al.*, submitted). L'étude de la fonction des gènes prédits a mis en avant l'absence de gènes codant la synthèse ou l'assimilation d'éléments indispensables au développement de l'agent pathogène tels que des gènes de l'assimilation de l'azote, du soufre et de la synthèse de différents acides aminés expliquant le caractère biotrophe obligatoire de ce dernier (Schwelm *et al.*, 2015).

3.1.3. Cycle de développement de *Plasmodiophora brassicae*

Le cycle de développement de *P. Brassicae* se découpe en trois étapes successives, correspondant à deux phases : la survie de *P. Brassicae* dans le sol et l'infection de la plante hôte (**Figure I.30**). La phase d'infection de la plante est découpée en 2 étapes : l'infection des poils absorbants (infection primaire) et l'infection des cellules corticales de la racine (infection secondaire) (Ayers, 1944 ; Ingram & Tommerup, 1972 ; Naiki, 1987).

La phase de survie dans le sol commence par la dissémination de l'agent pathogène à l'état de spores de repos à partir des tissus de plantes infectées au cours d'un cycle précédent. Les spores de repos ont une forme sphérique couverte d'épines d'environ 3 µm (Williams & McNabola, 1967 ; Buczacki & Cadd, 1976 ; Ikegami *et al.*, 1978) et ont une capacité de survie très importante, la période de demi-vie d'un inoculum est estimée entre 3 et 6 ans (Wallenhammar, 1996). Dans certaines conditions, il est nécessaire d'attendre plus de 17 ans pour que le niveau d'infestation des champs les plus infectés passe sous le seuil de détection (Wallenhammar, 1996).

La seconde phase commence par la germination des spores de repos qui est induite par la présence d'exsudats racinaires issus de la plante constitués de molécules telles que les facteurs dits de germination (GSF, Germination-Stimulating Factor ; Suzuki *et al.*, 1992).

Les spores de repos germées produisent des zoospores primaires haploïdes et biflagellées (pyriformes ou fusiformes) d'une taille allant de 2,8 à 5,9 µm (Ayers, 1944). Les zoospores primaires pénètrent dans la racine de la plante au niveau des poils absorbants menant à l'infection primaire. Une fois le contenu de la zoospore primaire injecté dans les poils absorbants, un plasmode primaire plurinucléé est créé suite à des endomitoses cruciformes successives (Braselton, 1995). Par la suite, ce plasmode primaire subit des clivages entraînant la formation d'un grand nombre de zoosporanges à l'origine de zoospores secondaires (4 à 16) qui seront relarguées dans le sol.

La pénétration des zoospores secondaires au niveau des cellules corticales de la racine marque le passage à l'infection secondaire avec le développement de plasmodes secondaires qui vont permettre la prolifération de l'agent pathogène. C'est lors de cette prolifération que la formation de la galle caractéristique de la maladie au niveau racinaire a lieu. La formation de la galle est due à des processus d'hypertrophie (augmentation de la taille des cellules) et d'hyperplasie cellulaire (augmentation du nombre de cellules). Finalement, les plasmodes secondaires vont être clivés et donner des millions de spores de repos qui seront libérées dans le sol lors de la désagrégation des tissus racinaires de la plante hôte (Kageyama & Asano, 2009).

L'existence d'une reproduction sexuée chez *P. Brassicae* a été suggérée (Braselton, 1982), le grand nombre d'isolats coexistant dans le sol et au sein d'une même galle (Jones *et al.*, 1982a,b ; Fähling *et al.*, 2003) favorisant en effet la probabilité d'un brassage génétique. Toutefois aucun gène impliqué dans la reproduction sexuée n'a pu être identifié à partir de la séquence de *P. Brassicae* (Schwelm *et al.*, 2016).

3.1.4. Facteurs influençant le développement de la hernie.

Un certain nombre d'études réalisées sur *P. Brassicae* ont permis de mettre en évidence la présence de facteurs environnementaux qui vont influencer son développement. Des températures allant de 20 à 25 °C ont été montrées comme optimum pour le développement de l'agent pathogène (Wellman, 1930 ; Ikegami, 1981 ; Feng *et al.*, 2010b ; Sharma *et al.*, 2011) alors que des températures entre 10 et 15 °C mènent à un retard de l'infection et du développement des symptômes chez l'hôte (Sharma *et al.*, 2011). Des sols à pH acide (inférieur à 6,5) sont favorables à la germination des spores avec l'observation de la réduction de la germination des spores dans des sols à des pH supérieurs à 6,5, ainsi que la diminution des symptômes provoqués par *P. Brassicae* pour des pH supérieurs à 7,2. (Webster & Dixon, 1991 ; Murakami *et al.*, 2002 ; Dixon, 2009a). La combinaison de ces deux premiers facteurs environnementaux à des niveaux optima pour *P. Brassicae* (température d'environ 24 °C et un pH du sol de 6) peuvent expliquer la sévérité des symptômes observables au niveau de la plante (Xiang-Hua, 2002). Par ailleurs, une humidité importante des sols (70 %) est également un facteur favorisant le développement de *P. Brassicae* avec une meilleure dispersion des zoospores et une infection primaire favorisée (Monteith, 1924 ; Naumova, 1933 ; Larson & Walker, 1934 ; Colhoun, 1952, 1953).

3.1.5. Les différents pathotypes de *Plasmodiophora brassicae* et leur caractérisation

L'analyse du niveau de résistance de différentes espèces de crucifères vis-à-vis de différents isolats de *P. Brassicae* d'origines variées par Honig (1931) a permis de mettre en évidence la présence de variations génétiques entre les différents isolats de *P. Brassicae*. Des gammes d'hôtes différentiels constituées de génotypes présentant des niveaux de résistance variés en fonction de l'isolat ont ainsi été développées. Aujourd'hui, trois principales gammes d'hôtes sont utilisées et divergent par la nature et le nombre de génotypes de plantes qui les composent (Buczacki *et al.*, 1975 ; Some *et al.*, 199 ;

Tableau I.3. Gammes d'hôtes différentiels utilisées pour l'identification et la classification des différents pathotypes de *Plasmodiophora brassicae*. (Williams 1966 ; Buczacki *et al.*, 1975 ; Somé *et al.*, 1996)

Gamme d'hôtes	Nom	Espèce	Sous-espèce	Cultivar
Williams (1966)		<i>Brassica oleracea</i>	<i>capitata</i>	Jersey Queen Badger Shipper
		<i>Brassica napus</i>	<i>napobrassica</i>	Laurentian Wilhelmsburger
European Clubroot Differential (ECD ; Buczacki <i>et al.</i> , 1975)	ECD1	<i>Brassica rapa</i>	<i>rapifera</i>	
	ECD2		<i>rapifera</i>	
	ECD3		<i>rapifera</i>	
	ECD4		<i>rapifera</i>	
	ECD5		<i>pekinensis</i>	cv. Granaat
	ECD6	<i>Brassica napus</i>	<i>oleifera</i>	cv. Nevin
	ECD7		<i>napus</i>	
	ECD8		<i>napus</i>	
	ECD9		<i>napus</i>	
	ECD10		<i>napobrassica</i>	cv. Wilhelmsburger
	ECD11	<i>Brassica oleracea</i>	<i>capitata</i>	cv. Badger Shipper
	ECD12		<i>capitata</i>	cv. Bindsachsener
	ECD13		<i>capitata</i>	cv. Jersey Queen
	ECD14		<i>capitata</i>	cv. Septa
	ECD15		<i>acephala</i>	cv. Verheul
Somé <i>et al.</i> (1996)	ECD6	<i>Brassica napus</i>	<i>oleifera</i>	cv. Nevin
	ECD10		<i>napobrassica</i>	cv. Wilhelmsburger
	Brutor		<i>oleifera</i>	

Williams, 1966). La gamme d'hôtes de Williams (1966) est composée de deux génotypes de *B. napus* et de deux génotypes de *B. rapa*, la gamme European Clubroot Differential (ECD) de Buczacki *et al.* (1975) utilise 15 génotypes des espèces *B. napus*, *B. rapa* et *B. oleracea*. La gamme de Somé *et al.* (1996) implique trois génotypes de *B. napus*. Ces trois gammes permettent de distinguer respectivement 48, 16 et 8 pathotypes différents de *P. Brassicae* (**Tableau I.3**). En utilisant la gamme de Somé *et al.* (1996) qui permet de discriminer 8 pathotypes différents de *P. Brassicae* (P1 à P8), Manzanares-Dauleux *et al.* (2001) ont montré qu'une grande diversité de génotypes de *P. Brassicae* existe au niveau de la parcelle et au sein même de la galle. Par ailleurs cette grande diversité a été démontrée à l'échelle européenne dans les pays producteurs de *Brassica* que sont la Pologne, l'Allemagne et la République Tchèque (Zamani-Noor 2016 ; Řičařová *et al.*, 2016). En fonction des zones géographiques, la répartition des pathotypes est différente : en France les pathotypes les plus fréquents sont les P1 et P2 (gamme de Somé *et al.* (1996)), au Canada le pathotype P3 prédomine par rapport aux pathotypes P2, P5, P6 et P8 (gamme de Williams (1966) ; Howard *et al.*, 2010 ; Hwang *et al.*, 2012) et en Australie les pathotypes P3, P6 et P7 sont les plus fréquents (gamme de Williams (1966); Donald *et al.* (2006)).

3.2. Impact de *Plasmodiophora brassicae* sur la plante hôte

L'infection par *P. Brassicae* provoque chez les plantes hôtes de nombreuses modifications aussi bien au niveau morphologique, qu'au niveau hormonal et métabolique. Ces modifications induites chez l'hôte par le développement de l'agent pathogène sont responsables de lourdes pertes de rendement. A l'échelle mondiale, cette maladie entraînerait une perte de l'ordre de 10 à 15 % de rendement (Dixon, 2009b), dans certains cas, elle peut entraîner la destruction complète de la culture pouvant aller de 30 à 100 % pour des parcelles canadiennes fortement infectées (Howard *et al.*, 2010).

3.2.1. Modifications morphologiques

Aucun symptôme morphologique de l'infection par *P. Brassicae* n'est visible durant l'infection primaire. Les symptômes commencent à être visibles à partir de l'infection secondaire lors du développement de l'agent pathogène au niveau de la partie corticale de la racine de l'hôte. Au début de la phase secondaire les racines présentent de petits renflements qui vont grossir. Par la suite la prolifération de l'agent pathogène mène à la formation des galles caractéristiques de la hernie dues à des processus d'hypertrophie (augmentation de la taille des cellules) et d'hyperplasie cellulaire (augmentation du nombre de cellules ; Kageyama & Asano, 2009). Aux stades avancés de la maladie, la racine de la plante hôte sensible ne présente plus ou presque plus de système racinaire sain entraînant une perturbation importante du transport hydrique et minéral au niveau de la racine. A partir du moment où le système racinaire de la plante présente d'importantes altérations, des symptômes apparaissent au niveau de la partie aérienne de la plante avec un flétrissement des feuilles, un état de sénescence et la présence de chlorose. Aux stades avancés de la maladie les galles finissent par pourrir et entraînent la mort de la plante (**Figure I.31**).

A



B



Figure I.31. Photos des symptômes observables au niveau de la partie racinaire et aérienne de *Brassica napus* lors de son infection par *P. brassicae*. (A) Cas sévère de hernie au niveau de la racine dans laquelle la région touchée s'étend au dessus de la ligne du sol. (B) Conséquences d'un cas sévère de hernie au niveau de la partie aérienne de la plante avec la présence d'un flétrissement et un retard de croissance. Adaptée de Hwang *et al.* (2012).

3.2.2. Modifications hormonales et métaboliques

Des modifications métaboliques et hormonales sont présentes dès le stade primaire de l'infection de l'hôte par *P. Brassicae*, avec par exemple des modifications importantes au niveau protéique chez *A. thaliana* (Devos *et al.*, 2006) et ceci malgré l'absence de symptômes visibles observables. Lors de la phase secondaire d'infection de nombreuses modifications d'ordre métabolique (sucres, acides aminés et lipides ; Ludwig-Müller *et al.*, 2009) et transcriptomique (Siemens *et al.*, 2006) sont observées. Du fait de l'importance supposée de l'ensemble de ces modifications chez l'hôte pour le développement biotrophe de *P. Brassicae* différentes études visant à déterminer, comprendre et détailler ces modifications ont été réalisées dans le but de limiter l'impact de *P. Brassicae* sur le rendement des cultures.

3.2.2.1. Les modifications hormonales au sein de la plante

Les modifications structurales de la racine menant à la formation de la galle mettent en jeu des phénomènes d'hyperplasie et d'hypertrophie cellulaires suggérant un rôle crucial des hormones dans le développement des symptômes. Ainsi par exemple, les voies de signalisation auxinique et cytokinique sont suractivées dans les racines présentant des galles (Siemens *et al.*, 2006 ; Devos *et al.*, 2006 ; Jubault *et al.*, 2013). L'augmentation des niveaux d'auxine pourrait être impliquée dans le phénomène d'hypertrophie cellulaire en permettant l'augmentation de l'extensibilité des parois cellulaires de la plante hôte (Ludwig-Müller *et al.*, 2009). Cependant les résultats de la mesure de teneur en auxine chez des *Brassicaceae* infectés par *P. Brassicae* sont contradictoires, avec des augmentations (Raa, 1971 ; Butcher *et al.*, 1974 ; Ludwig-Müller *et al.*, 1993 ; Devos *et al.*, 2006) ou des diminutions des teneurs en auxine à différents points cinétiques (Butcher *et al.*, 1974 ; Kavanagh & Williams, 1981 ; Mousdale, 1981 ; Ludwig-Müller *et al.*, 1993 ; Devos *et al.*, 2006) en fonction des publications. Ces résultats mettent en avant le caractère complexe de l'interaction entre *P. Brassicae* et son hôte et peuvent être expliqués par les problèmes de quantification de l'auxine dans les tissus infectés (Ludwig-Müller *et al.*, 2009) et par les designs expérimentaux variés utilisés pour réaliser les analyses (différents temps cinétiques, isolats de *P. Brassicae* et plantes hôtes utilisés dans les analyses).

Contrairement au rôle de l'auxine, le rôle des cytokinines semble plus clair. En effet une augmentation des cytokinines dans les galles a été identifiée dès la fin du XX^{ème} siècle (Dekhuijzen & Overeem, 1971 ; Dekhuijzen, 1981) au sein même des plasmodes de *P. Brassicae* ainsi que dans le cytoplasme des cellules de la plante hôte. Des analyses supplémentaires réalisées par Müller & Hilgenberg (1986) ont confirmé ces résultats et laissent supposer que les plasmodes seraient à l'origine de la synthèse des cytokinines. Ces cytokinines seraient par la suite libérées dans le cytoplasme des cellules de l'hôte afin de favoriser son développement en induisant la division cellulaire chez son hôte. Les données transcriptomiques obtenues par Siemens *et al.* (2006) vont dans le sens des précédentes observations avec une favorisation des voies cytokinines dans les tissus infectés qui se traduit chez la plante par une

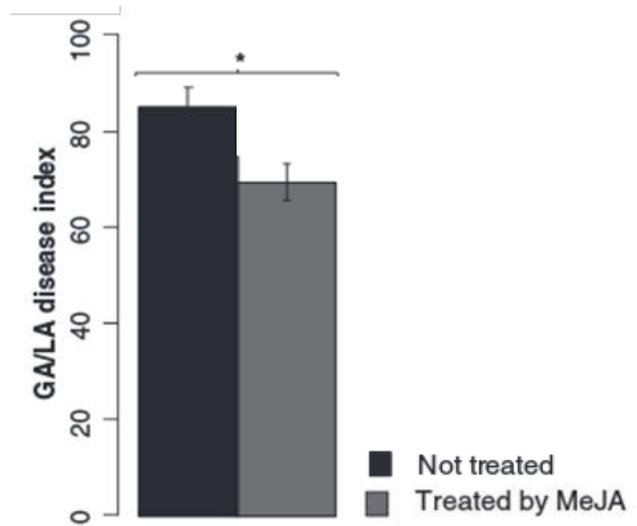
A**B**

Figure I.32. Conséquences du traitement de l'accession Columbia d'*Arabidopsis* avec du méthyl jasmonate sur le développement de la hernie. (A) Illustration des symptômes présents au niveau de l'accession Columbia en fonction de son traitement ou non avec 50 μ M de méthyljasmonate (MeJA) tous les 2 jours jusqu'à 251 jours après infection. (B) Evaluation quantitative des symptômes à 21 dpi en utilisant l'indice de la maladie Ga/La, exprimant la surface des galls par rapport à la taille des feuilles de la rosette. Les barres d'erreur représentent le SE (12 plantes par bloc, 4 blocs randomisés). Les astérisques indiquent des différences statistiquement significatives selon les tests de Wald appliqués sur un modèle linéaire mixte, * *p*-value < 0,05. Barres d'échelle = 1,6 cm. Adaptée de Lemarié *et al.* (2015).

répression des gènes de la biosynthèse et du catabolisme des cytokinines et une surexpression des gènes de la réponse aux cytokinines durant la phase secondaire d'infection. Ces résultats ont été corroborés par l'infection de mutants d'*A. thaliana* sur-exprimant des cytokinines oxydases qui présentent une moins grande sensibilité à *P. Brassicae* comparée à l'accession sauvage Columbia (Devos *et al.*, 2006). Malinowski *et al.* (2016) ont récemment apporté une nuance sur le rôle des cytokinines synthétisées par les plasmodes dans le développement de l'agent pathogène. En effet, les auteurs suggèrent que l'influence de ces dernières sur la plante hôte est minime au vu de leur faible impact sur la quantité cellulaire totale en cytokinines et sur l'expression des gènes au niveau de la plante.

Les analyses transcriptomiques réalisées chez l'accession Columbia et Burren d'*A. thaliana* (Siemens *et al.*, 2006 ; Agarwal *et al.*, 2011 ; Gravot *et al.*, 2012 ; Jubault *et al.*, 2013) ont également mis en avant la modulation de l'expression de gènes de la voie de l'acide jasmonique dans les racines des plantes infectées par *P. Brassicae*. En effet, des gènes impliqués dans la biosynthèse (*LOX2*) et dans la réponse au jasmonate (*THI2.1*, *ARGAH2* et *NATA1*) sont surexprimés dans les racines infectées. Ces résultats sont en accord avec l'augmentation du taux de jasmonate mesuré lors de la phase secondaire d'infection chez l'accession Columbia d'*A. thaliana* (Gravot *et al.*, 2012 ; Lemarié *et al.*, 2015) et ont été confirmés par une réduction de symptômes observables sur cette même accession suite à son traitement avec du méthyl-jasmonate (Lemarié *et al.*, 2015 ; **Figure I.32**). Le rôle du jasmonate dans la diminution des symptômes suite à l'infection par *P. Brassicae* a également été validé sur un autre *Brassicaceae*, *B. oleracea* var. *italica* (le brocoli) par Lovelock *et al.* (2013) avec une diminution de 25 % de la quantité de symptômes présents au niveau de la plante après traitement des plantes avec du jasmonate.

Dans la même étude Lovelock *et al.* (2013) ont testé l'effet d'un traitement par l'acide salicylique sur les symptômes observés chez le brocoli. Les résultats obtenus montrent une diminution de 25 à 65 % du développement des galls. Des résultats similaires ont été trouvés par Lemarié *et al.* (2015) avec l'observation d'un pic d'accumulation d'acide salicylique dans les racines de l'accession Burren (Bur-0) d'*A. thaliana*, partiellement résistante à *P. Brassicae*, 14 jours après infection. Ce pic d'acide salicylique conduirait à une surexpression de gènes impliqués dans la SAR (Systemic Acquired Resistance) comme *PR2* et *PR5*. De manière intéressante, Lemarié *et al.* (2015) ont montré que l'accession Columbia qui est sensible vis-à-vis de *P. Brassicae* utilise la voie du jasmonate comme voie de défense contrairement à l'accession partiellement résistante Bur-0. Ceci suggère une balance entre les deux voies en fonction des accessions et une meilleure efficacité de la voie dépendante de l'acide salicylique. Concernant ce dernier point, Ludwig-Müller *et al.* (2014) ont montré que l'isolat e3 de *P. Brassicae* était capable d'induire la méthylation de l'acide salicylique racinaire afin de le rendre inactif et de limiter son effet protecteur pour la plante.

A

Col-0

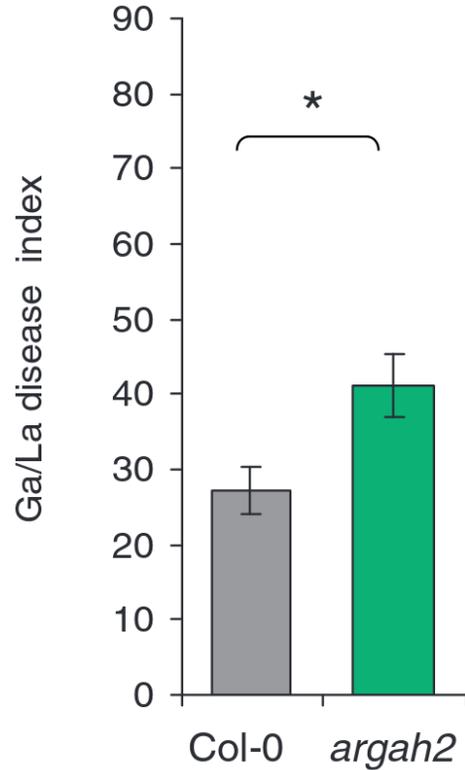
*argah2***B**

Figure I.33. Conséquences de la mutation du gène *ARGAH2* sur la réponse d'*Arabidopsis* à la hernie. (A) Illustration des symptômes de la hernie chez un individu Columbia sauvage (Col-0) et chez le mutant *argah2* à 21 jours après inoculation (21 dpi). **(B)** Evaluation quantitative des symptômes à 21 dpi en utilisant l'indice de la maladie Ga/La, exprimant la surface des galles par rapport à la taille des feuilles de la rosette. Les barres d'erreur représentent le SE (n = 8, et chaque répétition est composée d'au moins sept plantes). Les astérisques représentent des différences statistiquement significatives avec le t-tests (*p*-value < 0,05). Adaptée de Gravot *et al.* (2012).

3.2.2.2. Les modifications métaboliques au sein de la plante

Lors de l'infection par *P. Brassicae* les métabolismes primaire et secondaire de la plante sont altérés. Cette modulation du métabolisme de la plante peut s'expliquer par l'utilisation de certaines de ses ressources telles que les sucres, les acides aminés et les lipides par *P. Brassicae* afin d'assurer son développement au sein de l'hôte notamment au cours de la phase secondaire de l'infection (Ludwig-Müller *et al.*, 2009).

Au niveau du métabolisme primaire, certains sucres comme le saccharose et l'amidon sont fortement réduits au niveau des feuilles d'*A. thaliana* infectées (Evans & Scholes, 1995) alors que des accumulations de glucose, saccharose, fructose et d'amidon ont été détectées au niveau des galles (Mithen & Magrath, 1992 ; Evans & Scholes, 1995 ; Brodmann *et al.*, 2002). Ces résultats sont cohérents avec la présence de l'agent pathogène au niveau racinaire et sa dépendance à son hôte pour son développement du fait de son caractère biotrophe obligatoire. Cette modulation du métabolisme des sucres de la plante au profit de *P. Brassicae*, ainsi que son importance a été illustrée par Siemens *et al.* (2010) qui ont montré la réduction des symptômes de l'infection lors de l'inhibition d'invertases pariétales responsables de l'hydrolyse du saccharose en fructose et glucose. Parmi les sucres, le tréhalose, connu pour servir de stockage carboné chez différents organismes (Thevelein, 1984 ; Lippert *et al.*, 1993) est présent en très faibles concentrations chez les plantules d'*A. thaliana* non infectées (Blázquez *et al.*, 1998 ; Müller *et al.*, 2001 ; Vogel *et al.*, 2001) alors qu'il est fortement accumulé dans les galles (Keen & Williams, 1969 ; Brodmann *et al.*, 2002 ; Gravot Antoine *et al.*, 2011), le niveau de cette accumulation étant corrélé à la quantité de symptômes (Wagner *et al.*, 2012).

Le métabolisme des acides aminés est également altéré lors de l'infection avec une corrélation positive entre l'accumulation de 17 acides aminés et la quantité de symptômes mesurés (Wagner *et al.*, 2012). Gravot *et al.* (2012) et Jubault *et al.* (2008b) ont également montré que l'arginase (ARGAH2), qui catabolise l'arginine en L-ornithine, joue un rôle dans la limitation du développement des symptômes induits au niveau de la plante par *P. Brassicae* (**Figure I.33**).

Concernant le métabolisme secondaire les résultats obtenus dans différentes études mettent en avant une réponse complexe à interpréter. Les études réalisées chez *Arabidopsis* sur les glucosinolates, molécules considérées comme des composés de défense contre les agents pathogènes généralistes et attractifs pour les spécialistes (Halkier & Gershenzon, 2006), montrent des résultats contradictoires. En effet, le taux de glucosinolates indoliques est corrélé négativement avec la résistance dans certaines études (Butcher *et al.*, 1974 ; Buczacki & Ockendon, 1979 ; Chong *et al.*, 1985), positivement (Gu Lee & Soon Kim, 2010) ou non corrélé (Mullin *et al.*, 1980) dans d'autres. D'autre part, l'étude de Siemens *et al.* (2006) a mis en avant une induction de gènes impliqués dans la biosynthèse de flavonoïdes lors de l'infection. Cependant l'utilisation de mutants de la voie des flavonoïdes n'a pas permis de mettre

en avant une variation de la réponse à l'infection des mutants comparée à l'individu sauvage (Siemens *et al.*, 2003 ; Alix *et al.*, 2007).

3.3. La lutte contre *Plasmodiophora brassicae*

Afin de lutter contre la hernie au champ différentes stratégies ont été envisagées incluant l'utilisation i) de différentes pratiques culturales ii) d'un contrôle chimique iii) d'un contrôle biologique iv) des résistances génétiques des plantes.

3.3.1. Les pratiques culturales

Une des premières possibilités en termes de pratiques culturales au niveau du champ est la pratique de la rotation de cultures afin de lutter contre le développement de la hernie. Cependant, une telle pratique n'est pas efficace dans le cas de *P. Brassicae* car son temps de demi-vie à l'état de spore de repos est de l'ordre de 3,6 ans et plus de 17 ans sont nécessaires pour obtenir une quantité d'inoculum dans les sols suffisamment faible pour être incompatible avec une nouvelle infection (Wallenhammar, 1996). Les tentatives effectuées en utilisant des cultures intermédiaires (fraise, oignon, pomme de terre, épinard) ou d'intercultures de plantes non hôtes (le seigle, le poireau, le ray-grass) n'ont eu que des effets limités sur la réduction du nombre des spores de repos dans le sol (Ikegami, 1985) et sur la diminution des symptômes observables sur un hôte sensible comme le chou chinois (Murakami H. *et al.*, 2001 ; Friberg *et al.*, 2006 ; Ahmed *et al.*, 2011a).

Outre les pratiques visant à piéger *P. Brassicae* en le faisant se développer sur d'autres plantes ou germer sans hôte compatible, d'autres pratiques se basant sur l'altération des conditions nécessaires au développement de l'agent pathogène ont été envisagées. C'est ainsi que le chaulage, visant à alcaliniser les sols a été testé afin de modifier les pH acides du sol favorables au développement de la maladie (Gossen *et al.*, 2014). Néanmoins, l'utilisation du chaulage au champ présente deux inconvénients qui sont son coût (Hwang *et al.*, 2011 ; Gossen *et al.*, 2013) et la variabilité de son efficacité en fonction des types de sols et des conditions environnementales telles que l'hygrométrie et la température à l'application (Gossen *et al.*, 2013). Un second axe envisagé afin de limiter le développement de l'agent pathogène est l'utilisation d'un apport combiné de calcium, magnésium, potassium et bore (Dixon & Page, 1998). A lui seul, le calcium combiné à un sol présentant un pH élevé diminue la viabilité des spores (Myers & Campbell, 1985 ; Webster & Dixon, 1991 ; Lee & Hsieh, 1992) et diminue le nombre d'infections primaires (Webster, 1986) en induisant potentiellement le renforcement de la paroi cellulaire de la plante (Palm, 1963) ce qui limiterait la pénétration de *P. Brassicae* (Webster & Dixon, 1991 ; Donald & Porter, 2009). Le potassium réduirait le développement de la maladie à faibles concentrations (Walker & Hooker, 1945) car il serait un élément important du développement de *P. Brassicae* (Donald & Porter, 2009). Finalement le bore limiterait le développement de *P. Brassicae* durant ses deux phases d'infection (Webster, 1986 ; Webster & Dixon, 1991) potentiellement par son

rôle dans la régulation de la teneur en auxine endogène (Bohnsack & Albert, 1977) et le maintien de l'intégrité des parois et membranes cellulaires de la plante (Marschner, 1995).

Les pratiques culturales citées ci-dessus ayant plus comme but de limiter le développement de l'agent pathogène plutôt que de l'éradiquer d'autres moyens de lutte ont été envisagés.

3.3.2. *La lutte chimique*

Différentes matières actives ont été testées avec cependant un succès limité. En effet, à l'heure actuelle aucun produit phytosanitaire efficace contre la hernie n'est autorisé en Europe du fait de la toxicité de ces derniers pour l'Homme et l'environnement (Donald & Porter, 2009). Cependant malgré leur toxicité, certains produits montrent une efficacité contre la hernie et sont utilisés dans d'autres pays tels que i) le pentachloronitrobenzène qui a un effet fongicide, utilisé en Australie, ii) le flusulfamide qui inhibe la germination des spores (Tanaka *et al.*, 1999) utilisé au Japon et en Nouvelle Zélande (Donald & Porter, 2009), iii) le fluazinam qui stoppe la production d'énergie chez les champignons pathogènes, utilisé au Japon (Donald & Porter, 2009) et iv) le cyazofamide qui permet de réduire la germination des spores de repos de près de 80 % en inhibant le transport des électrons à travers la membrane mitochondriale (Mitani *et al.*, 2001, 2003) et qui est aussi utilisé au Japon (Donald & Porter, 2009).

Devant les problèmes de toxicité rencontrés avec les produits chimiques une autre option pour lutter contre la hernie est l'utilisation de moyens de lutte biologique.

3.3.3. *La lutte biologique*

Diverses études ont montré que les symptômes développés suite à l'infection par *P. Brassicae* peuvent être limités par l'utilisation d'un moyen de lutte biologique. Le biofongicide SerenadeTM composé de spores de *Bacillus subtilis* QST713, qui jouerait un rôle de stimulateur des défenses naturelles de la plante (Lahlali *et al.*, 2011, 2013), présenterait des effets protecteurs chez le colza vis-à-vis de la hernie (Peng *et al.*, 2011). Certains champignons endophytes racinaires tels que *Acremonium alternatum* et *Heteroconium chaetospora* montrent un impact négatif sur le développement de la maladie sur différents *Brassicaceae* tels que le chou chinois (Narisawa *et al.*, 1998, 2005 ; Doan *et al.*, 2010), le colza (Auer & Ludwig-Müller, 2014) ou *A. thaliana* (Jäschke *et al.*, 2010).

Néanmoins bien que certains moyens de lutte biologique montrent une certaine efficacité pour limiter le développement de la maladie, leur effet reste inférieur à celui de l'utilisation de variétés résistantes (Dixon, 2009b).

3.3.4. *La lutte génétique*

La lutte génétique par l'utilisation de variétés résistantes à *P. Brassicae* est aujourd'hui le moyen de lutte privilégié du fait de son faible coût et de son faible impact environnemental. A l'heure actuelle diverses variétés de choux et de colza sont disponibles sur le marché dont la variété de colza « Mendel ».

Cette variété issue d'un colza synthétique provenant du croisement entre un chou *B. oleracea* et une navette *B. rapa* présentait un fort niveau de résistance face à *P. Brassicae*. Cependant, le caractère oligogénique de cette résistance a entraîné une forte pression de sélection sur les populations pathogènes se traduisant par un rapide contournement de la résistance présente dans cette variété (Hwang *et al.*, 2012).

Afin de pouvoir développer la lutte génétique face à *P. Brassicae* il est nécessaire de diversifier et de comprendre la mise en place de la réponse de résistance chez les plantes hôtes en réponse à la hernie.

3.4. Les résistances identifiées chez la plante hôte en réponse à *P. Brassicae*

P. Brassicae est capable d'infecter la quasi-totalité des 3700 espèces appartenant à la famille des Brassicacées qu'elles soient cultivées (*B. napus*, *B. rapa*, *B. oleracea*...) ou sauvages (*Capsella bursa-pastoris*, *Thlaspi arvense*, *A. thaliana* (Buczacki & Ockendon, 1979)). L'interaction entre *P. Brassicae* et les plantes de la famille des Brassicaceae est une interaction de type hôte compatible qui peut conduire à deux types de résistance, une résistance totale ou presque totale (résistance qualitative) ou une résistance partielle (résistance quantitative) en fonction de l'hôte et des isolats de *P. Brassicae* (Manzanares *et al.*, 1996). Les analyses génétiques de résistance qui ont majoritairement été réalisées sur *B. rapa*, *B. oleracea*, *B. napus* et *A. thaliana* ont permis d'identifier des résistances qualitatives et quantitatives.

3.4.1. Les résistances identifiées chez *B. rapa*

La navette (*B. rapa*, génome AA, $2n = 20$) est une espèce diploïde possédant une taille de génome estimée de 485 Mb (Wang *et al.*, 2011) et dont 389,2 Mb ont été assemblées pour le moment (Cai *et al.*, 2017). Les différents cultivars de *B. rapa*, tels que le navet et le chou chinois, sont utilisés comme légumes, matière pour la fabrication de condiments et comme source de graines oléagineuses. L'identification et l'introduction de résistances dans des cultivars de chou chinois a été l'un des enjeux majeurs des années 90 du fait de la sensibilité des choux chinois à la hernie et l'importance économique de ces derniers en Asie de l'est (Diederichsen *et al.*, 2009). Les résistances identifiées sont majoritairement issues des navettes européennes (Crute *et al.*, 1980 ; Yoshikawa, 1993). Une grande partie des facteurs génétiques de résistance identifiés sont localisés sur le chromosome A03 de *B. rapa* dans des régions génomiques différentes. Parmi ces facteurs on retrouve les gènes majeurs *CRA*, *CRb*, *Crr3* et *CRk* (Matsumoto *et al.*, 1998 ; Piao *et al.*, 2002 ; Hirai *et al.*, 2004 ; Piao *et al.*, 2004 ; Saito *et al.*, 2006 ; Hayashida *et al.*, 2008 ; Sakamoto *et al.*, 2008 ; Kato *et al.*, 2012 ; Ueno *et al.*, 2012 ; Kato *et al.*, 2013) où *CRA* et *CRb* (Hatakeyama *et al.*, 2017) ainsi que *Crr3* et *CRk* ont été identifiés comme étant potentiellement des copies du même gène (Diederichsen *et al.*, 2009). *CRA* est présent dans près de 80 % des cultivars de *B. rapa* (Aruga *et al.*, 2013). Un deuxième locus appelé *CRAki*, lié à *CRb*, a été identifié en utilisant l'isolat N014 de *P. Brassicae* (Kato *et al.*, 2012, 2013). D'autres facteurs de

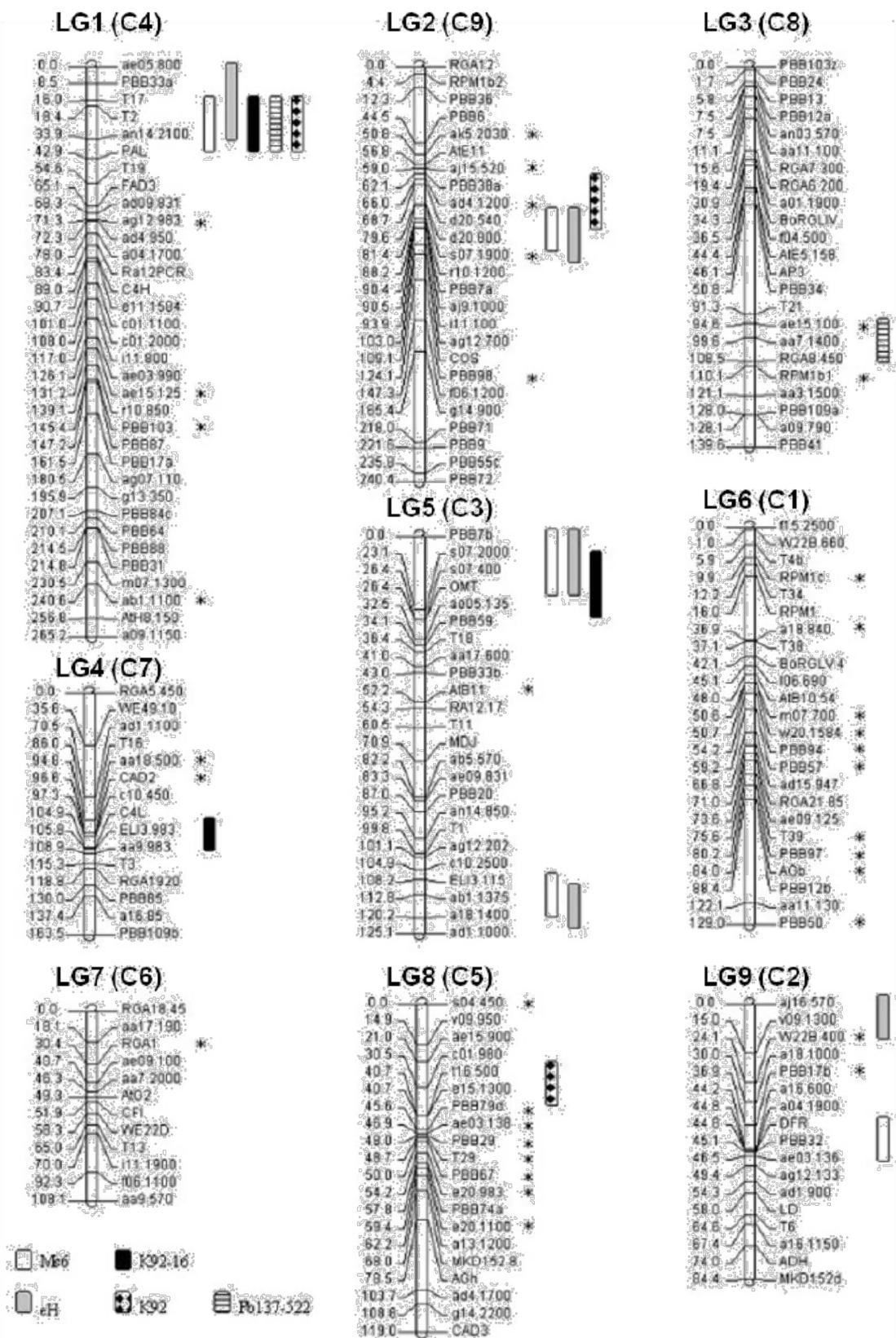


Figure I.34. QTL de résistance à la hernie détectés chez *Brassica oleracea* dans les populations F2/3 C10 x HDEM. Les rectangles représentent la position des QTL détectés pour chacun des cinq isolats avec un intervalle de confiance d'un LOD. * indique les loci présentant une distorsion de ségrégation. Adaptée de Rocherieux *et al.* (2004).

résistance ont été identifiés sur d'autres chromosomes. Ainsi *Crr1* (identifié chez la variété Siloga vis-à-vis de l'isolat Ano-01 de *P. Brassicae*) et *Crr2* (identifié chez la variété Siloga vis-à-vis de l'isolat Wakayama-01) ont été localisés respectivement sur les chromosomes A08 et A01. *Crr1* est impliqué dans la résistance aux 2 isolats, mais il doit être obligatoirement en combinaison avec le gène *Crr2* pour amener à la résistance vis-à-vis de l'isolat Wakayama-01. Le clonage du QTL *Crr1* par Hatakeyama *et al.* (2013) a permis d'identifier que ce locus contient deux gènes *Crr1a* et *Crr1b*. *Crr1a* ainsi que *CRA* qui a été également cloné, codent des NLR (TIR-NBL-LRR) (Ueno *et al.*, 2012). Récemment trois nouveaux facteurs de résistance à la hernie ont été identifiés *Rcr4* qui confère une résistance aux pathotypes 2, 3, 5, 6 et 8 canadiens (gamme de Williams (1966)) et *Rcr8* et *Rcr9* qui confèrent une résistance à un nouveau pathotype nommé 5x responsable du contournement des résistances des variétés de colza utilisées au Canada (Strelkov *et al.*, 2016 ; Yu *et al.*, 2017).

3.4.2. Les résistances identifiées chez *B. oleracea*

Le chou (*B. oleracea*, génome CC, $2n = 18$) est une espèce diploïde possédant une taille de génome d'environ 539 Mb (Liu *et al.*, 2014). L'espèce *B. oleracea* est très diversifiée avec différents cultigrupes tels que le brocoli, le chou-fleur, le chou pommé, le chou fourrager, le chou de Bruxelles, etc. Les individus de cette espèce présentent de hauts niveaux de caroténoïdes (Kopsell & Kopsell, 2006) et ils contiennent différents glucosinolates qui sont impliqués dans des voies de défense contre les champignons et les bactéries pathogènes (Halkier & Gershenzon, 2006). Les variétés de *B. oleracea* possèdent principalement de résistances partielles à la hernie et ont été peu utilisées comme source de facteurs de résistance du fait de leur association fréquente à des caractères agronomiques indésirables (Diederichsen *et al.*, 2009). La résistance quantitative chez *B. oleracea* est sous contrôle polygénique (Landry *et al.*, 1992 ; Figdore *et al.*, 1993 ; Grandclément & Thomas, 1996 ; Voorrips *et al.*, 1997 ; Moriguchi *et al.*, 1999 ; Rocherieux *et al.*, 2004 ; Nomura K. *et al.*, 2008) mais il existe aussi, comme chez *B. rapa*, des résistances de type qualitative sous contrôle mono ou oligogénique (Chiang & Crête, 1985 ; Yoshikawa, 1993). La complexité de la résistance a par exemple été illustrée par Rocherieux *et al.* (2004) qui ont montré que parmi 12 QTL identifiés contre 5 isolats de *P. Brassicae* à partir d'une population issue du croisement entre le chou C10 et un brocoli haploïde doublé HDEM, le QTL *Pb-Bo1* conférait une résistance contre les 5 isolats alors que les autres QTL étaient isolat spécifique (**Figure I.34**). Les auteurs ont par ailleurs démontré que la magnitude de l'effet du QTL *Pb-Bo1* était dépendant de l'isolat, allant d'un effet proche de celui d'un gène majeur (expliquant jusqu'à 80 % de la variabilité observée) à un effet modéré (expliquant 21 % de la variabilité observée pour l'isolat eH). Plus récemment de nouveaux QTL de résistance ont été détectés : 1 QTL chez le cultivar Green Comet et 4 QTL chez le cultivar Anju (Nagaoka *et al.*, 2010) et 3 QTL chez le chou pommé C1220 (Lee *et al.*, 2015). De manière globale la comparaison des différents QTL identifiés chez les différents cultivars de

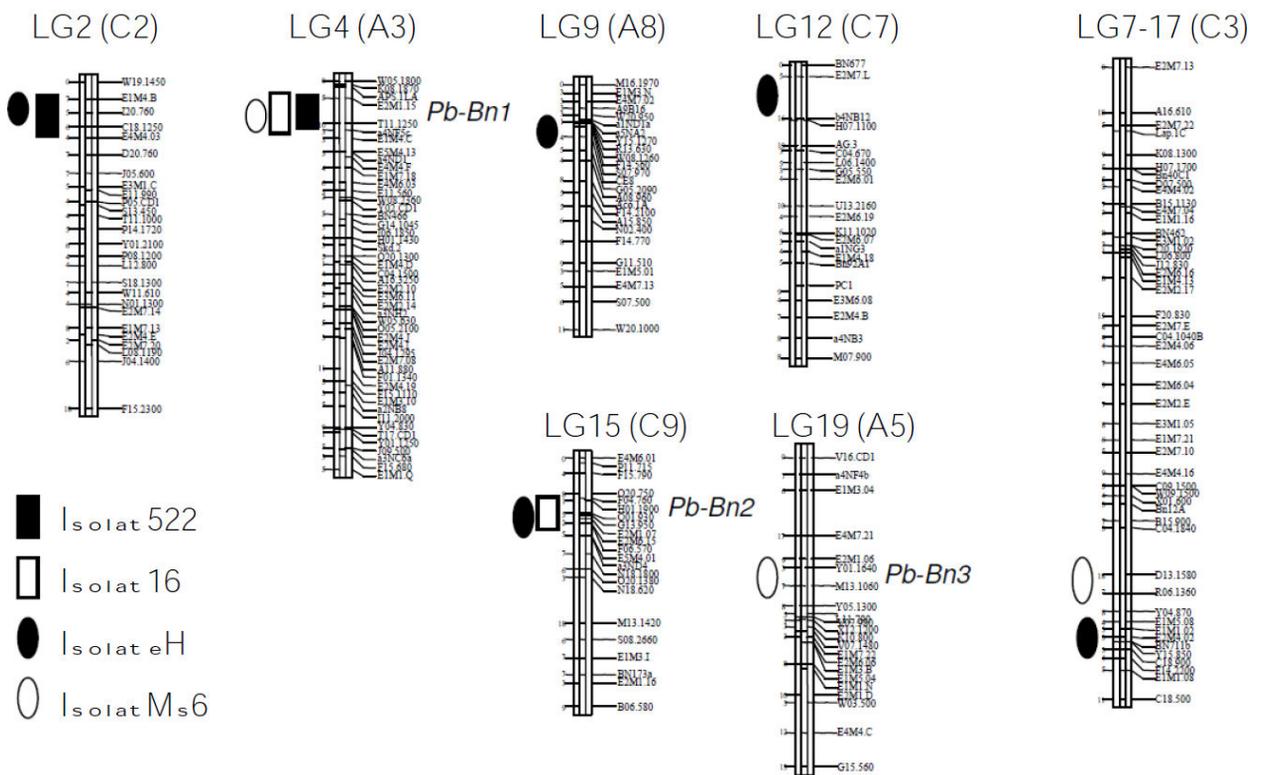


Figure I.35. Localisation sur la carte génétique du colza (Lombard et Delourme, 2001) des gènes majeurs *Pb-Bn1*, *Pb-Bn2* et *Pb-Bn3* et des QTL additifs de résistance à la hernie identifiés dans les populations en ségrégation Darmor-*bzh* x Yudal et Stellar x Drakkar. Les rectangles représentent la position des QTL détectés pour chacun des quatre isolats avec un intervalle de confiance d'un LOD (d'après Manzaneres-Dauleux *et al.* 2000, 2003). Les noms des groupes de liaison sont indiqués selon la nomenclature utilisée par les auteurs et les correspondances avec la nomenclature actuelle sont indiquées entre parenthèses.

B. oleracea reste difficile du fait de l'absence dans les cartes génétiques utilisées de marqueurs communs et l'utilisation d'isolats différents.

3.4.3. Les résistances identifiées chez *B. napus*

Le colza (*B. napus*, génome AACCC, $2n = 38$) est une espèce allotétraploïde issue de l'hybridation interspécifique naturelle il y a environ 7 500 ans (Chalhoub *et al.*, 2014) de deux parents diploïdes, la navette et le chou. Son génome possède une taille globale de 843 Mb (314,2 Mb pour le sous-génome A et 528,8 Mb pour le sous-génome C; Chalhoub *et al.*, 2014) et présente une homologie partielle à totale entre les sous-génomes A et C (Parkin *et al.*, 2014). Le colza est l'une des principales espèces oléagineuses cultivées au même titre que le soja, le tournesol et le palmier à huile (FAOSTAT, 2014). Son importance économique est liée à la production d'huile pour la consommation humaine, son utilisation dans les biocarburants et la production de tourteaux pour l'alimentation animale. Du fait de sa récente création, le colza n'a que faiblement divergé de ses progéniteurs, faisant ainsi de *B. rapa* et *B. oleracea* des sources génétiques d'intérêt pour l'introgession de résistances à la hernie. Une illustration de ce potentiel a été la création de la variété de colza Mendel obtenu à partir d'un colza synthétique provenant du croisement entre *B. rapa* ECD04 et *B. oleracea* ECD15 (Diederichsen & Sacristan, 1996).

Bien que la résistance portée par cette variété ait été très efficace pendant les premières années de sa commercialisation, de nombreux contournement de la résistance ont été décrits menant à l'abandon aujourd'hui de l'utilisation de cette variété. Des analyses menées sur la résistance de *B. napus* à la hernie ont montré que la résistance pouvait être de type qualitative et quantitative (Manzanares-Dauleux *et al.*, 2000a, 2003) et que pour un même génotype des niveaux de résistance différents existaient en fonction de l'isolat de *P. Brassicae* utilisé (Manzanares-Dauleux *et al.*, 2000a). En effet, trois niveaux de résistance ont été identifiés chez le génotype Darmor-*bzh* avec une résistance presque totale vis-à-vis des isolats eH et Pb137-522 (pathotypes 1 et 7), une résistance partielle vis-à-vis de l'isolat K92-16 (pathotype 4) et une sensibilité élevée vis-à-vis de l'isolat Ms6 (Pathotype 1 ; Manzanares-Dauleux *et al.*, 2003). Un QTL à effet majeur, *Pb-Bn1*, conférant la résistance aux isolats Pb137-522 et K92-16 et deux QTL à effet mineur spécifiques à chaque isolat ont été identifiés chez Darmor-*bzh* (Manzanares-Dauleux *et al.*, 2000a). Chez les populations issues du croisement Darmor-*bzh* x Yudal et Stellar x Drakkar, Manzanares-Dauleux *et al.* (2003) ont détecté 8 QTL dont 3 QTL avec un effet majeur (*Pb-Bn1*, *Pb-Bn2* et *Pb-Bn3*) vis-à-vis des isolats Pb137-522, eH et Ms6 et un effet résiduel pour les autres isolats (**Figure I.35**). Werner *et al.* (2007) ont identifié 19 QTL de résistance dans une population d'haploïdes doublés vis-à-vis de 7 isolats de *P. Brassicae* différents, dont certains communs à différents isolats. Cependant aucun QTL commun à tous les isolats n'a été détecté. En utilisant une approche de génétique d'association (GWAS) sur 472 accessions de *B. napus* vis-à-vis d'un isolat chinois de *P. Brassicae*, 10 QTL de résistance ont été détectés par Li *et al.* (2016b).

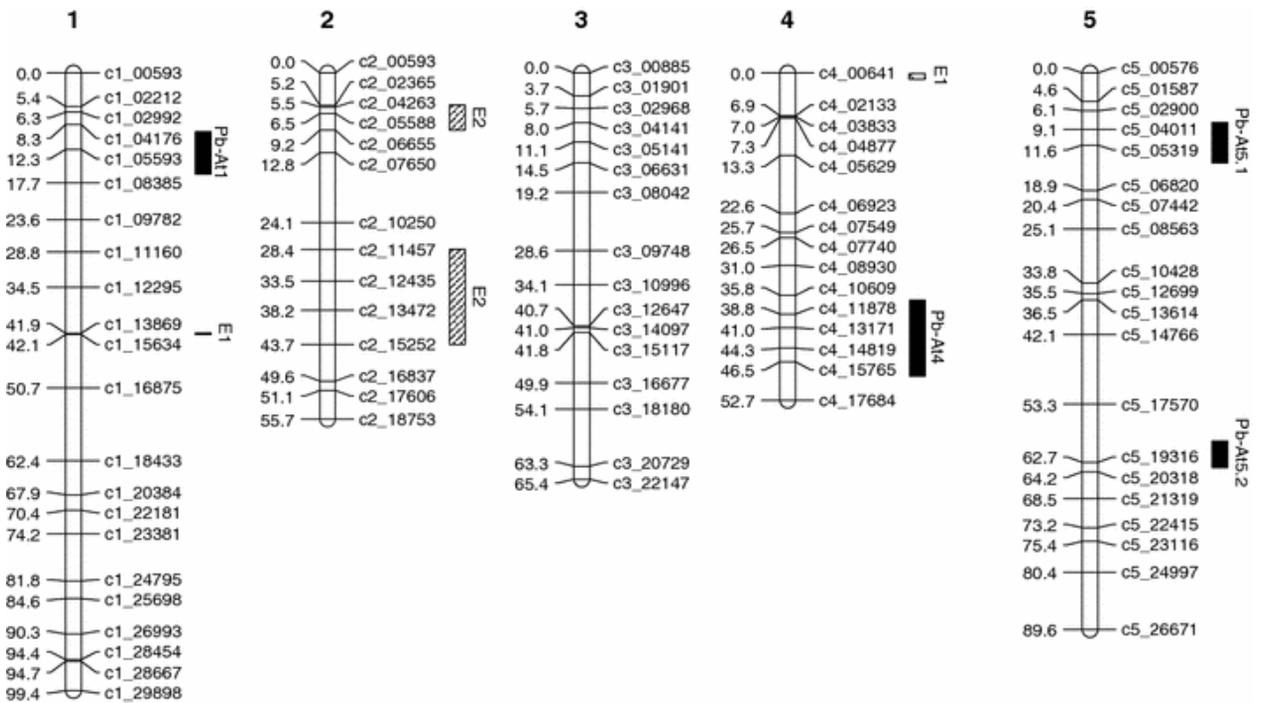


Figure I.36. QTL additifs et épistatiques contrôlant la résistance partielle à *P. brassicae* dans la population Bur-0 × Col-0 RIL. Chaque nom de marqueur est constitué du numéro du chromosome et de sa position physique en kb. Les QTL additifs détectés dans la population RIL sont indiqués par des barres de couleur noire. La longueur de la barre est égale à l'intervalle de confiance de probabilité à un LOD. Les interactions épistatiques significatives détectées par une recherche complète par paires sont indiquées par des paires de barres hachurées. Issue de Jubault *et al.* (2008).

Plus récemment, il a été montré que l'effet des résistances quantitatives à la hernie chez le colza peut être influencé par des facteurs abiotiques et biotiques. En effet, Laperche *et al.* (2017) et Aigu *et al.* (2018) ont montré que la contrainte azotée pouvait modifier les QTL de résistance à la hernie détectés pour différents traits phénotypiques. De plus, les auteurs ont montré que pour les QTL de résistance détectés dans deux conditions de culture, la variation phénotypique expliquée était différente entre les deux conditions.

3.4.4. Les résistances identifiées chez *A. thaliana*

L'Arabette des dames (*A. thaliana*, génome $2n = 10$) est une espèce diploïde de la famille des *Brassicaceae* possédant un génome d'environ 135 Mb (*Arabidopsis* Genome, 2000). De nombreux éléments tels que la petite taille de son génome, son cycle de développement court, la facilité de la réalisation de croisements et la possibilité de la muter en laboratoire ont fait de cette plante un outil important de la recherche en biologie comme plante modèle (Fink, 1998). En 50 ans cette plante a été le sujet de plus de 54033 publications couvrant 406 domaines différents de la biologie (Provar *et al.*, 2016) et est le point central de projets d'importance tel que le projet 1001 génomes ayant permis l'obtention en 2015 des données de séquençage de 1135 accessions naturelles d'*Arabidopsis* (Cao *et al.*, 2011 ; Schmitz *et al.*, 2013 ; Long *et al.*, 2013). En 1996, Fuchs & Sacristán (1996) ont mis en évidence des réponses de défense de type HR chez deux écotypes d'*Arabidopsis* Tsu-0 et Ze-0 vis-à-vis de l'isolat e3 (pathotype 1) conférées par un gène majeur présent au niveau d'un locus nommé *RPBI* pour Resistance to *Plasmodiophora brassicae* 1, localisé sur le chromosome 1. L'implication de cette région a également été mise en évidence au laboratoire dans la résistance vis-à-vis de l'isolat eH par une approche de GWAS et a permis l'identification d'un gène candidat (Jubault *et al.* in prep). En 2007, Alix *et al.* (2007) ont testé 48 accessions d'*A. thaliana* vis-à-vis des deux autres isolats de *P. Brassicae*, les isolats eH et Ms6 et identifié trois accessions partiellement résistantes aux isolats eH et/ou Ms6 Bur-0, Kn-0 et Tsu-0 et une accession sensible à tous les isolats, Col-0. La résistance portée par Bur-0 est contrôlée par 4 QTL à effet additif (*Pb-At1*, *Pb-At4*, *Pb-At5.1* et *Pb-At5.2*) localisés sur les chromosomes 1, 4 et 5 ainsi que 4 QTL à effet épistatique (**Figure I.36**). Parmi les QTL détectés, le QTL *Pb-At5.2* est un QTL à effet majeur expliquant 22,6 % de la variabilité phénotypique observée (Jubault *et al.* 2008).

Comme précédemment décrit pour des QTL de résistance à la hernie chez *B. napus* (cf paragraphe 3.4.3) pour un autre facteur environnemental, il a été montré chez *Arabidopsis* que l'effet de QTL de résistance peut être modulé en fonction du niveau d'irrigation utilisé lors la culture des plantes. En effet, Gravot *et al.* (2016) ont montré que suivant le niveau d'irrigation appliqué durant la phase secondaire d'infection par *P. Brassicae* l'effet des QTL de résistance était différent avec un effet du QTL *Pb-At5.2* sous faible irrigation alors que le QTL *Pb-At5.1* conduit à une résistance partielle sous forte irrigation.

4. Objectifs et stratégies de la thèse

Dans un contexte de développement d'une agriculture plus durable visant à réduire l'utilisation des intrants et des produits phytosanitaires, l'utilisation des résistances génétiques des plantes est l'une des composantes majeures des systèmes intégrés de protection des plantes. Le développement et l'utilisation de variétés résistantes sont une priorité notamment dans le cas des maladies telles que la hernie pour lesquelles les autres méthodes de lutte se sont avérées inefficaces ou nocives pour l'environnement. La hernie des Crucifères est aujourd'hui la principale maladie des *Brassicaceae* dans le monde, avec des émergences dans différentes parties du monde (Chine, Canada...) et une recrudescence en Europe (France, Allemagne Pologne et Royaume-Uni principalement). Le développement de variétés portant des résistances qualitatives à la hernie a montré ces limites. En effet, il a notamment été observé un rapide contournement par *P. Brassicae* de la résistance de la variété de colza Mendel, première variété commercialisée comme résistante à la hernie, basée sur un gène majeur de résistance R. Le contournement des résistances qualitatives portées par les variétés commercialisées dû à une trop forte pression de sélection sur les populations pathogènes a incité les recherches sur la résistance quantitative. L'effet de la résistance quantitative sur l'infection par l'agent pathogène n'est pas total mais partiel, ce qui a l'avantage de n'imposer qu'une faible pression sur les populations pathogènes limitant ainsi l'apparition d'isolats virulents. De plus, cette résistance est sous le contrôle de multiples gènes réduisant le risque d'effondrement brutal de la résistance car cela nécessiterait l'apparition de multiples mutations chez l'agent pathogène.

L'étude de la résistance quantitative par des approches de génétique quantitative basée sur des polymorphismes de séquence a montré que la variabilité phénotypique observée ne peut pas être totalement expliquée par les facteurs génétiques décelés (QTL) suggérant ainsi la présence d'autres sources de variabilité. Par ailleurs, diverses études ont démontré l'implication de marques épigénétiques dans le contrôle de caractères complexes et ceci de manière héritable sur différents matériels biologiques. En effet, ces études ont été menées i) sur des mutants altérés au niveau de gènes impliqués dans la structure de la chromatine ii) sur des populations expérimentales quasi isogénique possédant une mosaïque de l'épigénome parental (populations épiRIL) iii) directement sur différentes espèces végétales. Le rôle démontré de l'épigénétique dans la réponse des végétaux à des stress d'origine abiotique et biotique et la transmissibilité de marques telles que la méthylation de l'ADN au cours des générations et leur influence sur la mise en place de caractères complexes chez les plantes suggèrent qu'une source de variabilité d'ordre épigénétique pourrait moduler la résistance quantitative.

Dans ce cadre - épigénétique et réponse aux stress-, l'objectif de mon sujet de thèse concernait l'étude de l'implication des mécanismes épigénétiques dans la résistance quantitative à la hernie chez *Arabidopsis thaliana*. Afin de réaliser ce travail, différents axes d'étude ont été suivis pour répondre aux questions suivantes :

- Existe-t-il des variations d'origine épigénétique impliquées dans le contrôle de la réponse à la hernie chez *Arabidopsis thaliana* ?
- Quelle est l'architecture épigénomique (localisation de régions différentiellement méthylées) du caractère résistance à la hernie chez *A. thaliana* ?
- La variation épigénétique est-elle liée et/ou indépendante de la variation génétique contrôlant la résistance à la hernie ? Quelle est la relation entre les architectures génétiques et épigénétiques dans le contrôle du caractère résistance à la hernie chez *A. thaliana* ?

Dans un premier temps (cf chapitre III), nous avons déterminé par une approche non ciblée si des variations épigénétiques étaient impliquées dans la résistance quantitative à la hernie i) par le criblage de la réponse de différents mutants d'*A. thaliana* altérés au niveau de gènes impliqués dans la structure de la chromatine vis-à-vis de l'isolat eH de *P. Brassicae*, et ii) par l'analyse d'(épi)génétiq ue quantitative en utilisant la population épiRIL développée chez *Arabidopsis* issue du croisement d'un mutant épigénétique *ddm1-2* dans le fonds génétique Columbia et d'un individu sauvage Columbia. Afin de décomposer la réponse de résistance, différents caractères phénotypiques associés ou non à la maladie ont été mesurés en conditions contrôle (pas d'infection) et d'infection par l'isolat eH de *P. Brassicae* permettant ainsi de déterminer les facteurs génétiques associés aux différentes composantes de la maladie (contrôle de la taille de la galle, contrôle de la multiplication du pathogène, conséquences sur le développement foliaire).

Dans un deuxième temps (cf chapitre IV), nous avons effectué une approche ciblée afin de caractériser la causalité de régions épigénétiques identifiées. Pour cela nous nous sommes intéressés à deux régions (l'une décelée lors de ma thèse et l'autre identifiée lors de travaux de recherche préalables dans mon laboratoire d'accueil) présentant potentiellement de la variabilité épigénétique contrôlant la résistance à la hernie dans des écotypes d'*Arabidopsis*. L'étude de ces régions a été réalisée par des approches de génomique structurale (séquençage), de transcriptomique (RNAseq) et d'épigénomique (méthylome) en utilisant de données générées au sein du laboratoire ou publiques.

CHAPITRE II

MATERIELS ET METHODES

A



B



Figure II.1. Phénotype observé au niveau de la racine de plantes inoculées avec l'isolat eH de *P. brassicae*. (A) Accession Col-0 d'*Arabidopsis thaliana*. (B) Accession Bur-0 d'*Arabidopsis thaliana*.

CHAPITRE II. MATÉRIELS ET MÉTHODES

Afin de faciliter la lecture du manuscrit, ce chapitre regroupe l'ensemble des matériels et méthodes utilisés au cours de cette thèse, qu'ils soient détaillés ou non dans les publications.

1. Matériels végétal et parasitaire

1.1. *Arabidopsis thaliana*

1.1.1. L'écotype sensible Col-0

Trois accessions de l'écotype Américain d'*A. thaliana*, Col-0 (*Columbia*), ont été utilisées au cours de cette thèse :

- La N6000 : Cet écotpe a été utilisé comme individu contrôle pour l'ensemble des tests pathologiques effectués sur les lignées HIF (cf paragraphe II.1.1.5). Cet écotpe a été identifié au sein du laboratoire comme sensible vis-à-vis de l'isolat eH de *P. Brassicae* (Alix *et al.*, 2007, **Figure II.1A**).
- La N76778 qui provient du Nottingham *Arabidopsis* Stock Center (NASC). Cet écotpe fait partie du set des 127 écotpes d'*Arabidopsis* testés contre l'isolat eH de *P. Brassicae* (cf paragraphe II.1.1.6)
- La lignée parentale Col-WT provenant de centre de ressources de Versailles (cf paragraphe II.1.1.3). Les individus Col-WT sont les descendants homozygotes (DDM1/DDM1) issus du croisement entre un individu hétérozygote (*ddm1/DDM1*) et un individu Col-0 sauvage suivi de 6 générations de rétrocroisement avec le génotype sauvage *Columbia* (Johannes *et al.*, 2009).

1.1.2. Les mutants d'*Arabidopsis thaliana*

L'ensemble des mutants utilisés et référencés dans ce manuscrit de thèse (à l'exception du parent Col-*ddm1*, cf paragraphe 1.1.3) ont été commandés au NASC (The Nottingham *Arabidopsis* Stock center). Les mutants dont les mécanismes épigénétiques sont altérés ont été obtenus dans un fonds génétique Col-0. La liste de l'ensemble des mutants, ainsi que les informations qui leurs sont liées sont présentées dans le tableau II.1.

1.1.3. La population épiRIL

La population épiRIL résulte du croisement entre un mutant du maintien de la méthylation de l'ADN nommé *ddm1-2* et l'écotype sauvage Col-0 (Johannes *et al.*, 2009). DDM1 est une protéine appartenant à la famille des facteurs de remodelage de la chromatine Snf2 et est impliquée dans la maintenance de la méthylation en contextes CG, CHG et CHH au niveau des séquences répétées de l'ADN et dans la régulation de l'expression des éléments transposables (Bourc'his & Bestor, 2002 ; Lippman *et al.*, 2004

Tableau II.1. Liste des mutants d'*Arabidopsis thaliana* utilisés au cours de la thèse.

Numéro NASC	Référence du mutant	Mutant	Gène muté	Description du gène	LP primer (5' > 3')	RP primer (5' > 3')
N650863	SALK_150863	<i>drm2</i>	At5g14620	Putative ADN méthyltransférase avec des domaines catalytiques réarrangés; similaire aux DNMT3 méthyltransférases de mammifères; contient des domaines UBA. La partie proximale à l'extrémité 3' de la région codant pour le gène est fortement méthylée à la fois sur les résidus d'adénine et de cytosine.	AGATCGCTTCCAGAGTTAGCC	TTGTCCGCAAAAAGCAAAAAGAG
N655230	SALK_004027C	<i>hda15</i>	At3g18520	Code une protéine avec une similitude avec les histones désacétylases. Les plantes exprimant l'ARNi dirigé contre ce gène montrent une résistance modérée à la transformation des racines induite par <i>Agrobacterium</i> .	CTTCTCTGTTCATCGTTTCGC	AGCAACAATTCTCTCGTCGAAC
N667337	SALK_130607C	<i>atxr5</i>	At5g09790	Code une protéine du domaine SET, une H3K27 monométhyltransférases nécessaires pour la structure de la chromatine et l'extinction de gènes. Régule la réplication de l'ADN hétérochromatique. Contient un domaine de liaison PCNA. L'ATXR5 s'accumule préférentiellement au cours de la phase G1 ou S tardive, suggérant qu'il joue un rôle dans la régulation ou la progression du cycle cellulaire. Une lignée végétale exprimant une construction d'ARNi dirigée contre ce gène a réduit la formation de tumeurs à médiation par <i>Agrobacterium</i> .	TTTCTCTGTCCGGTGAAAATG	CCTGCAACAATCA GTGTGATG
N681376	SALK_082118C	<i>hac1</i>	At1g79000	Homologue de la protéine de liaison CREB, un co-activateur de la transcription avec l'activité histone acétyl-transférase. Une lignée végétale exprimant une construction ARNi ciblée contre HAC1 présente des taux réduits de transformation racinaire induite par <i>Agrobacterium</i> .	GACTGACAGGCTTTGCTTGG	ATGAACTTGGAATGGGAACC
N681865	SALK_149295C	<i>srt2</i>	At5g09230	Code SRT2, un membre de la famille SIRT2 des histone désacétylases.	CGCAGAGAGAGAAACAAAATCG	TTCCACATTCTGTGCTAAACCC
N681885	SALK_000590C	<i>ddm1</i>	At5g66750	Protéine similaire aux protéines de remodelage de la chromatine SWI2 / SNF2. DDM1 semble agir comme une ATPase de remodelage de la chromatine impliquée dans la méthylation de la cytosine dans des contextes CG et non-CG. Impliqué dans le silençage génique et le maintien de la méthylation de l'ADN et de la méthylation des histones. L'hypométhylation de nombreuses régions génomiques se produit chez les mutants <i>ddm1</i> causant plusieurs anomalies phénotypiques. DDM1 pourrait être impliqué dans l'établissement d'une frontière hétérochromatine. Une lignée exprimant une ARNi ciblée contre DDM1 montre une certaine résistance à la transformation des racines induite par <i>Agrobacterium</i> .	GAAAGACAGGTTTCACCTGTGC	CGTGAGAAATACTCAATGCC

; Zemach *et al.*, 2013). La F1 obtenue a été rétrocroisée par le parent sauvage Col-0 afin de restaurer la fonction de DDM1 au niveau de la génération BC1, la mutation *ddm1* étant délétère pour la plante après plusieurs autofécondations. Les individus de la BC1 ont été sélectionnés pour leur état homozygote *DDM1* et ont été autofécondés sur 7 générations (cf. Chapitre I Figure I.25). Les 505 épiRIL ainsi obtenues, possèdent une mosaïque des marques épigénétiques parentales dans le même fonds génétique (Johannes *et al.*, 2009). Les 505 individus de cette population, ainsi que les lignées parentales Col-0-wt et Col-*ddm1* ont été commandés au Centre de Ressources de Versailles (<http://publiclines.versailles.inra.fr/epirils/index>). L'analyse des données de séquençage du génome des 122 épiRIL épigénotypées à l'aide de GATK HaplotypeCaller (Gilly *et al.*, 2014 ; données non publiées) par Victoire Baillet du laboratoire de Vincent Collot a mené à l'identification de variants à petite échelle (substitutions d'une base et indels < 100 pb ; Baillet *et al.*, non publié). Les appels de variants bruts ont été filtrés en suivant les suggestions de GATK Best Practice et des scripts supplémentaires. Tous les variants ont été inspectés visuellement en utilisant IGV et un sous-ensemble des SNP détectés a été validé par PCR. Concernant les éléments transposables, l'analyse des événements de transposition a été réalisée avec TE-Tracker (Gilly *et al.*, 2014) dans 102 épiRIL de couverture génomique suffisante. Les appels bruts ont ensuite été filtrés à l'aide de scripts internes. Toutes les nouvelles insertions de TE détectées ont été inspectées visuellement en utilisant l'IGV et une partie d'entre elles a été validée par PCR (Etcheverry *et al.*, Non publié). Les variants d'insertion à petite échelle et TE partagés sont définis comme étant présents dans au moins 25 % de la population.

1.1.4. L'écotype partiellement résistant Bur-0

Deux accessions de l'écotype Irlandais d'*A. thaliana*, Bur-0 ont été utilisées au cours de cette thèse :

- La N76734 qui provient du Nottingham *Arabidopsis* Stock Center (NASC). Cet écotype fait partie du set des 127 écotypes d'*Arabidopsis* testés contre l'isolat eH de *P. Brassicae* (cf paragraphe II.1.1.6). Cet écotype a été identifié au sein du laboratoire comme partiellement résistant vis-à-vis de l'isolat eH de *P. Brassicae* (Alix *et al.*, 2007, **Figure II.1B**).
- La 172AV qui provient du Centre de Ressources de Versailles et qui dérive de l'accession N1028 du NASC par filiation monograinée. Cet écotype a été utilisé comme individu contrôle pour l'ensemble des tests pathologiques effectués sur les lignées HIF (cf paragraphe II.1.1.5).

1.1.5. Les génotypes appartenant à des lignées quasi isogéniques HIF (Heterogenous Inbred Families)

Les différents individus appartenant à une famille quasi-isogénique (HIF) ont été obtenus à partir de la lignée RIL499 issue du set 20RV de Simon *et al.* (2007) elle-même dérivée du croisement Bur-0 x Col-0 (Institut Jean Pierre Bourgin, INRA Versailles, France). Elle présente un état hétérozygote résiduel au niveau de la région du QTL *Pb-At5.2* identifié par Jubault *et al.* (2008) et a été utilisée pour former

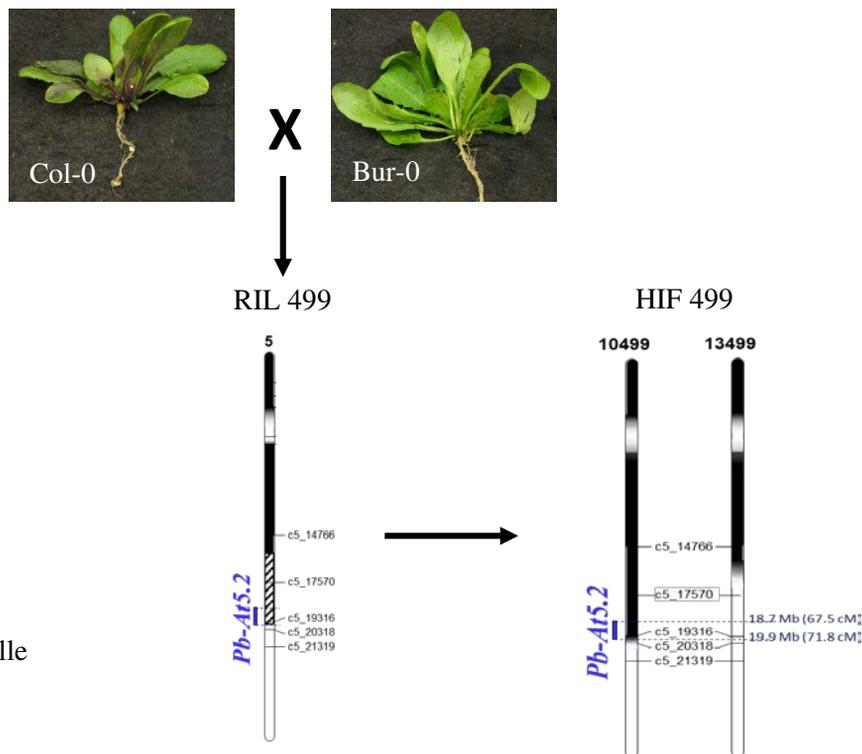
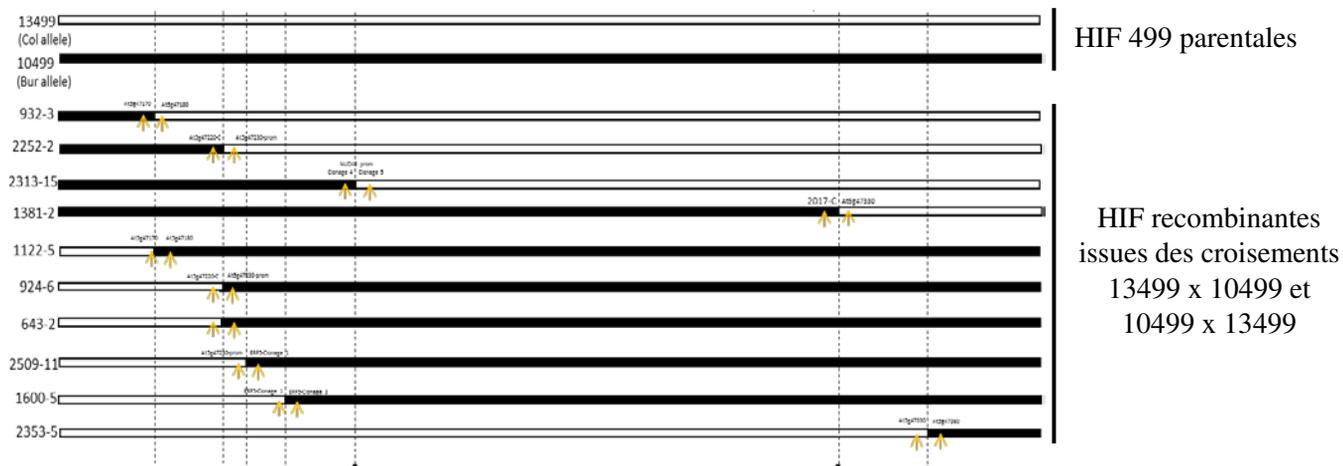
A**B**

Figure II.2 Schématisation de l'obtention des lignées HIF recombinantes à partir des écotypes Col-0 et Bur-0. (A) Schématisation de l'obtention des lignées HIF 10499 et 13499 présentant un état homozygote contrasté au niveau de la région du QTL *Pb At-5.2*. Le trait sous l'indication *PbAt-5.2* en bleu indique l'intervalle de confiance du QTL détecté par Jubault *et al.* (2008). Les indications c5_ suivies d'un numéro sur la droite des chromosomes schématisés illustrent l'emplacement des marqueurs utilisés pour effectuer le génotypage. La partie noire des chromosomes indique la présence de régions homozygotes d'origine Bur-0. La partie blanche des chromosomes indique la présence de régions homozygotes d'origine Col-0. La partie hachurée des chromosomes indique la présence de régions hétérozygotes Bur-0/Col-0 résiduelles dans la RIL 499. **(B)** Représentation du génotypage de lignées recombinantes obtenues à partir du croisement dans les 2 sens des HIF 10499 et 13499. Les numéros sur la gauche indiquent le nom de chaque lignée HIF recombinante. La partie noire des chromosomes indique la présence de régions homozygotes d'origine Bur-0. La partie blanche des chromosomes indique la présence de régions homozygotes d'origine Col-0. Les flèches jaunes indiquent les marqueurs de génotypage entourant les événements de recombinaison.

deux lignées HIF présentant un état homozygote Col-0 (HIF 10499) et Bur-0 (HIF 13499) par l'institut Jean Pierre Bourgin de l'INRA de Versailles (**Figure II.2A**). Ces deux lignées sont quasi-isogéniques, *i.e.* elles possèdent la même combinaison d'allèles Col-0 et Bur-0 à l'état homozygote le long de leur génome, à l'exception de la région du QTL *Pb-At5.2*. L'inoculation de ces deux HIF par l'isolat eH de *P. Brassicae* par le docteur Antoine Gravot au sein du laboratoire a mis en évidence une réponse à la hernie contrastée entre les deux HIF : l'une présente une résistance partielle (HIF 10499) alors que l'autre est sensible (HIF 13499). Par la suite, d'autres lignées HIF dérivées de ses deux lignées ont été obtenues par le docteur Antoine Gravot au sein du laboratoire par croisements réciproques (avec l'utilisation de l'un ou de l'autre comme parent femelle) afin d'obtenir des individus F1 dont l'état hétérozygote a été vérifié par PCR. Une plante F1 provenant de chaque croisement a été autofécondée et environ 1500 plantes F2 ont été obtenues à partir de chaque F1. Parmi ces F2, les individus recombinants dans la région du QTL *Pb-At5.2* ont été recherchés par génotypage de 10 marqueurs KASPAR. Au total, 554 F2 montrant au moins une recombinaison ont été identifiées dont 105 ont été choisies car représentatives de tous les événements de recombinaison observés dans cette région. A partir de ces F2, les lignées F3 ont été obtenues par autofécondation et ont été utilisées afin d'augmenter le nombre de marqueurs dans la région et d'améliorer la localisation des différents événements de recombinaison (**Figure II.2B**). Au total 93 marqueurs SNP ont été génotypés par la technologie KASPAR sur 88 lignées HIF F3 (correspondant à un pool de 12 plantes).

1.1.6. *Autres écotypes d'Arabidopsis utilisés*

Au cours de cette thèse une sélection de 127 accessions naturelles d'*Arabidopsis* a été testée contre l'isolat eH de *P. Brassicae*. L'ensemble de ces accessions proviennent du Nottingham *Arabidopsis* Stock Center (NASC) et ont été re-séquencées par le consortium 1001 génomes (Kawakatsu *et al.*, 2016). La liste de l'ensemble de ces accessions, ainsi que les informations qui leurs sont liées sont présentées dans le **Tableau II.2**.

1.2. Gamme d'hôtes différentiels

Afin de caractériser et de valider les différents isolats de *P. Brassicae* utilisés, la gamme de plantes hôtes définie par Somé *et al.* (1996) a été testée dans chacun des tests pathologiques réalisés. Cette gamme est composée de trois génotypes de *B. napus* : *B. napus* ssp. *oleifera* cv 'Nevin' (ECD6), *B.napus* ssp. *rapifera* cv 'Wilhelmsburger' (ECD10) et *B. napus* ssp. *oleifera* (Brutor). L'ensemble de ces génotypes permet de classer les isolats en 8 pathotypes différents (**Tableau II.3**). En complément de ces 3 génotypes, deux génotypes de *B. oleracea* ssp. *acephala* : C10 et CB151 ont été ajoutés à la gamme d'hôtes précédemment citée afin de différencier les isolats appartenant au pathotype de type 1 (P1) (**Tableau II.4**). L'ensemble de ces génotypes proviennent du Centre de Ressources Biologiques BraCySol (INRA, IGEPP).

Tableau II.2. Liste des accessions naturelles d'*Arabidopsis thaliana* utilisées.

ID	Num NASC	Génotype	Pays	ID	Num NASC	Génotype	Pays	ID	Num NASC	Génotype	Pays	ID	Num NASC	Génotype	Pays
265	N77198	PYL-6	FRA	7353	N76611	Tha-1	NED	9543	N76886	IP-Gra-0	ESP	9812	N77397	Tu-W1	GER
350	N77382	TOU-A1-88	FRA	7354	N76612	Ting-1	SWE	9545	N76920	IP-Her-12	ESP	9813	N76706	Bl-4	GER
410	N76808	Doubravnik7	CZE	7424	N76519	Jl-3	CZE	9556	N77081	IP-Men-2	ESP	9815	N76900	Ha-HBT3-11	GER
424	N76812	Draha2	CZE	7477	N78853	WAR	USA	9559	N77107	IP-Mon-5	ESP	9821	N76674	IP-Aru-0	ESP
428	N76718	Borky1	CZE	7521	N77052	Lp2-6	CZE	9584	N77283	IP-Stp-0	ESP	9823	N76681	IP-Bae-0	ESP
915	N77040	LIN_S-5	USA	8236	N76941	HSm	CZE	9586	N77340	IP-Tam-0	ESP	9825	N76714	IP-Boa-0	ESP
1872	N77100	MNF-Pot-75	USA	8239	N76976	PHW-3	GER	9587	N77344	IP-Tdc-0	ESP	9831	N76743	IP-Cas-0	ESP
2016	N77098	MNF-Pin-39	USA	8343	N76558	Na-1	FRA	9590	N77387	IP-Trs-0	ESP	9832	N76759	IP-Cat-0	ESP
5104	N78800	UKSE06-252	UK	8357	N76573	Pla-0	ESP	9599	N78846	IP-Vin-0	ESP	9841	N76836	IP-Ees-0	ESP
5720	N78781	Cal-2	UK	8365	N77201	Rak-2	CZE	9600	N78848	IP-Vis-0	ESP	9843	N76840	IP-Elp-0	ESP
5772	N78787	Set-1	UK	8386	N77268	Sr:5	SWE	9616	N76984	Krazo-1	RUS	9851	N76942	IP-Hue-3	ESP
5779	N78789	UKID74	UK	9075	N77023	Lerik1-4	AZE	9653	N76878	Giffo-1	ITA	9852	N76945	IP-Ini-0	ESP
5874	N76814	Drall-6	CZE	9078	N77024	Lerik1-7	AZE	9690	N76755	Castelfed-3-206	ITA	9855	N77008	IP-Lam-0	ESP
5984	N76822	DraIV_6-13	CZE	9089	N77119	Nar-3	AZE	9693	N76758	Castelfed-3-209	ITA	9859	N77054	IP-Lro-0	ESP
6020	N76861	FjÅr1-5	SWE	9095	N76950	Istisu-5	AZE	9722	N76890	Groch-1	BUL	9860	N77059	IP-Lum-0	ESP
6163	N77331	TAA_14	SWE	9102	N76999	Lag1-4	GEO	9726	N76853	Fanerone mi-3	GRC	9862	N77062	IP-Mad-0	ESP
6296	N78774	Udul_1-11	CZE	9104	N77001	Lag1-6	GEO	9731	N77281	Stiav-3	SVK	9877	N77165	IP-Pdl-0	ESP
6445	N78876	Zdr1_2-21	CZE	9105	N77002	Lag1-7	GEO	9733	N76697	Bela-2	SVK	9882	N77178	IP-Pil-0	ESP
6909	N76778	Col-0	USA	9113	N77006	Lag2-6	GEO	9738	N76722	Bran-1	ROU	9894	N77243	IP-Sen-0	ESP
6933	N77047	LL-0	ESP	9121	N76685	Bak-5	GEO	9762	N76487	Etna-2	ITA	9895	N77248	IP-Sfb-6	ESP
6957	N77192	Pu2-8	CZE	9125	N76876	Geg-14	ARM	9769	N76916	HE-1	GER	9897	N77257	IP-Smt-1	ESP
6960	N77211	Rennes-11	FRA	9506	N76662	IP-Alo-0	POR	9771	N77172	Pfn-N2.2-6	GER	9898	N77259	IP-Som-0	ESP
6973	N78817	UII2-3	SWE	9509	N77207	IP-Reg-0	POR	9777	N76880	Gn-1	GER	9900	N77386	IP-Tri-0	ESP
6989	N76432	Alst-1	UK	9510	N77208	IP-Rei-0	POR	9779	N76682	Bai-10	GER	9901	N78824	IP-Urd-1	ESP
7013	N76445	Bd-0	GER	9511	N78835	IP-Vav-0	POR	9780	N76856	Fel12-4	GER	9906	N77063	IP-Mah-6	ESP
7058	N76734	Bur-0	IRL	9514	N76647	IP-Adm-0	ESP	9781	N76990	Kus2-2	GER	9945	N76413	Leo-1	ESP
7094	N76791	Da-0	GER	9515	N76650	IP-AIa-0	ESP	9784	N76845	Erg2-6	GER	9947	N76415	Ped-0	ESP
7103	N76476	Dra-0	CZE	9526	N76738	IP-Cab-3	ESP	9785	N76898	Ha-HBT1-2	GER	9969	N76425	Valsi-1	ITA
7106	N78897	Dr-0	GER	9530	N76742	IP-Cat-1	ESP	9791	N76914	Haes-1	GER	9990	N76390	Lag2-2	GEO
7244	N76552	Mnz-0	GER	9533	N76763	IP-Cem-0	ESP	9798	N76902	Ha-P2-1	GER	10001	N76407	TueV-13	GER
7288	N77156	Oy-0	NOR	9534	N76774	IP-Cmo-3	ESP	9807	N77240	Schl-7	GER	10018	N76369	Dobra-1	SRB
7346	N77277	Sten-0	GER	9541	N76871	IP-Fue-2	ESP	9808	N77392	Tu-B2-3	GER				

1.3. Matériel parasitaire

1.3.1. Présentation des isolats de *P. Brassicae* utilisés

Le principal isolat utilisé tout au long de cette étude est l'isolat monospore eH de *P. Brassicae* appartenant au pathotype P1. En plus de cet isolat de *P. Brassicae*, 5 autres ont été utilisés : les isolats monospore et population de pathotype P1 Ms6 et P1*, l'isolat monospore de pathotype P4 Pb137-522 et l'isolat monospore de pathotype P7 K92-16. L'attribution du pathotype pour chacun de ces isolats a été réalisée en utilisant la gamme d'hôtes différentiels citée dans le paragraphe précédent.

1.3.2. Préparation de l'inoculum

L'inoculum a été préparé comme décrit par Manzanares-Dauleux *et al.* (2000a) à partir de galles prélevées sur *B. rapa* ssp. *pekinensis* cv. « Granaat » (chou chinois) qui est l'hôte sensible universel. Les galles congelées ont été broyées dans 100 ml d'eau distillée puis le broyat a été filtré sur un jeu de tamis de maillages décroissants (500 μm , 250 μm puis 100 μm). Le nombre de spores de repos présentes dans le filtrat obtenu a été quantifié par comptage sur cellule de Malassez pour finalement préparer un inoculum avec une concentration de 10^7 spores par ml.

2. Conditions de culture et caractérisation de la réponse de la plante à *P. Brassicae*

2.1. Conditions de culture

L'ensemble des tests effectués lors de cette étude ont été réalisés en se basant sur les conditions décrites par Alix *et al.* (2007). Dans un premier temps la germination de l'ensemble des graines utilisées a été synchronisée en plaçant ces dernières à 4 °C pendant 2 jours dans des boîtes de pétri sur un papier filtre humidifié. Les graines ont ensuite été semées individuellement à l'aide de pinceaux dans des pots de 4 cm de diamètre contenant un mélange stérilisé de terreau et de vermiculite (2 : 1, v : v). Finalement, les plantes ont été placées sur des chariots en chambre de culture avec une photopériode de 16 h de jour ($110 \pm 20 \mu\text{mol.m}^{-2}.\text{s}^{-1}$; 20 ± 1 °C) et 8 heures de nuit (18 ± 1 °C) (**Figure II.3**).

2.2. Suivi de la température

Le suivi de la température a été réalisé dans un premier temps en utilisant une sonde OCEASOFT présente au sein de la chambre. Devant les variations phénotypiques observées au sein des tests, il a été décidé d'ajouter des capteurs de température Thermo Bouton 21G référence DAL0052 au niveau des plantes sur les chariots de culture. Ainsi, pour l'ensemble des tests effectués au cours de ma thèse un minimum de 4 capteurs de température par bloc a été utilisé afin de cartographier les variations de température au sein de la chambre et de prendre en compte ces variations dans l'analyse des résultats des tests pathologiques.

Tableau II.3. Caractérisation de huit pathotypes de *P. Brassicae* en utilisant la gamme d'hôtes différentiels décrite par Somé *et al.* (1996). (-) Les plantes sont considérées comme résistantes lorsque leur indice pathologique est inférieur à 25. (+) Les plantes sont considérées comme sensibles lorsque leur indice pathologique est supérieur à 25.

Pathotypes	ECD6	ECD10	Brutor
P1	+	+	+
P2	+	-	+
P3	-	-	+
P4	-	-	-
P5	-	+	+
P6	+	-	-
P7	-	+	-
P8	+	+	-

Tableau II.4. Différentiation des isolats de pathotype 1 par addition de deux génotypes de *B. oleracea* ssp. *acephala*. (-) Les plantes sont considérées comme résistantes lorsque leur indice pathologique est inférieur à 25. (+) Les plantes sont considérées comme sensibles lorsque leur indice pathologique est supérieur à 25.

Pathotypes	ECD6	ECD10	Brutor	C10	CB151
eH	+	+	+	-	-
Ms6	+	+	+	-	+
P1*	+	+	+		



Figure II.3. Conditions de culture des plantes lors des tests pathologiques. Les plantes sont cultivées sur des chariots avec une photopériode de 16 h de jour ($110 \pm 20 \mu\text{mol.m}^{-2}.\text{s}^{-1}$) et 8 heures de nuit couplée à des températures de $20 \pm 1 \text{ }^\circ\text{C}$ le jour et $18 \pm 1 \text{ }^\circ\text{C}$ la nuit.

2.3. Inoculation

Pour chaque test pathologique réalisé, l'infection des plantes a été réalisée 10 jours après le semis au stade 1.04 décrit par Boyes *et al.* (2001) en déposant 1 ml d'inoculum au niveau du collet de la plante.

2.4. Dispositifs expérimentaux

L'ensemble des tests pathologiques réalisés au cours de cette thèse ont suivi le même dispositif expérimental en blocs complets de Fisher. En fonction des expérimentations, 2 ou 3 répétitions biologiques constituées de 2 blocs ont été réalisées en randomisant la position des génotypes. Selon l'objectif du test, la réponse à *P. Brassicae* a été évaluée en utilisant entre 6 et 12 plantes par bloc. La caractérisation de l'isolat de *P. Brassicae* a été réalisée pour chacun des tests en utilisant la gamme de Somé *et al.* (1996).

2.5. Caractérisation de l'impact de la présence *P. Brassicae* sur la plante hôte

La caractérisation de la réponse de la plante hôte a été évaluée *via* différentes variables. En fonction des expérimentations réalisées une à plusieurs variables ont été utilisées :

- L'indice pathologique (IP) défini par Manzanares-Dauleux *et al.* (2000b) pour quantifier les symptômes racinaires a été analysé pour l'ensemble des plantes
- La quantification de l'ADN de l'agent pathogène présent au sein des racines des plantes infectées par qPCR.
- L'indice GA/LA développé par Gravot *et al.* (2011).
- La longueur de la plus grande feuille de la rosette (Lf).
- La biomasse racinaire (Rb).

2.5.1. L'indice pathologique

La lecture des symptômes racinaires provoqués par *P. Brassicae* a été réalisée 21 jours après inoculation. Afin de quantifier la quantité de symptômes, les plantes ont été déterrées et lavées une par une. Le calcul de l'indice pathologique a été réalisé en suivant l'échelle décrite par Manzanares-Dauleux *et al.* (2000b). Ainsi, les plantes ont été répertoriées dans 5 classes en fonction de la localisation et de l'intensité des symptômes observés (**Figure II.4**) :

- 0 : les racines sont saines (absence de symptômes visibles)
- 1 : de petites galles sont visibles sur le système racinaire secondaire uniquement
- 2 : des galles sont présentes sur le système racinaire principal et secondaire
- 2⁺ : le système racinaire présente d'importantes galles à l'exception de quelques racines
- 3 : le système racinaire est remplacé par une galle

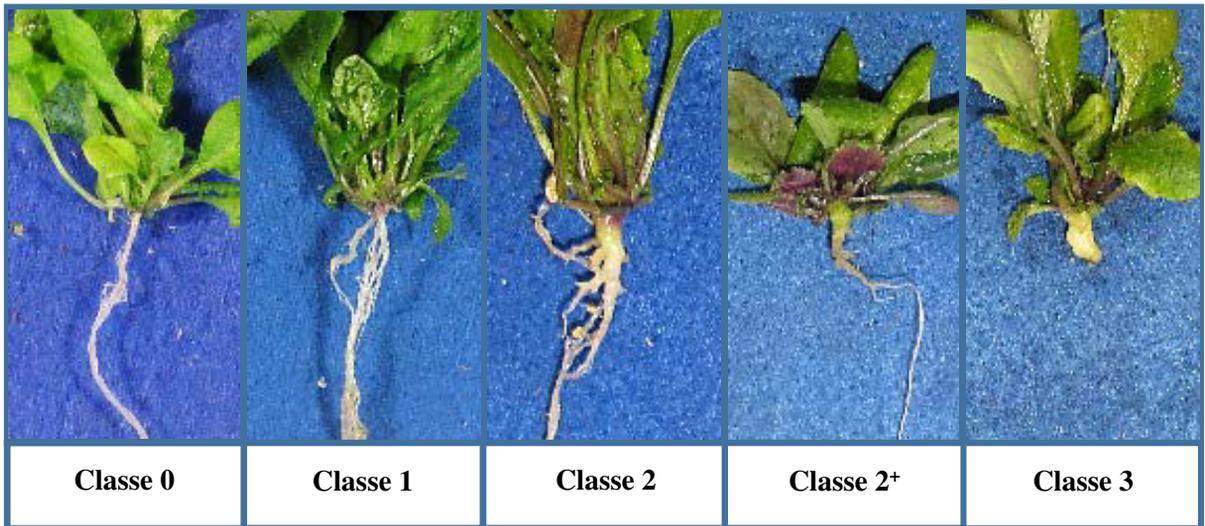


Figure II.4. Illustration des symptômes racinaires présents pour chacune des classes de maladie définies permettant le calcul de l'indice pathologique.

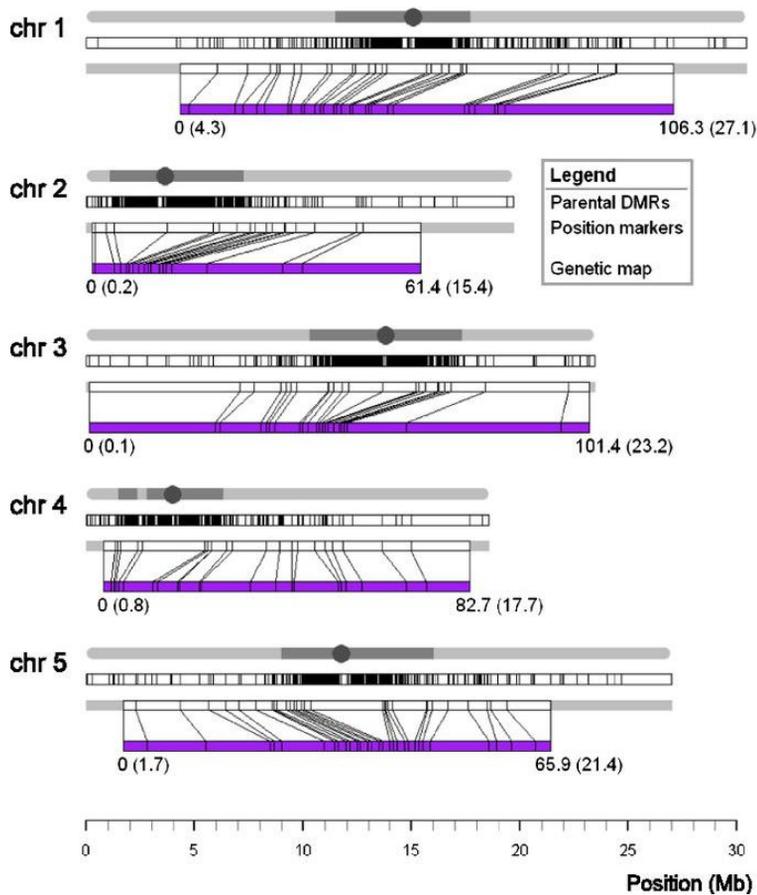


Figure II.5. Carte épigénétique de la population épiRIL. La figure représente la distribution à l'échelle du génome des 2 611 DMR parentaux sur la partie haute et des 126 DMR (c'est-à-dire les marqueurs) conservés pour la construction de la carte génétique (dans la partie violette basse) pour chacun des cinq chromosomes d'*A. thaliana*. Figure adaptée de Colomé-Tatché *et al.* (2012).

L'indice pathologique est alors calculé en affectant un coefficient aux plantes de chaque classe et en divisant par le nombre de plantes testées en utilisant la formule :

$$IP = (n_1 \times 25 + n_2 \times 50 + n_3 \times 75 + n_4 \times 100) / N$$

Avec « n_i » est le nombre de plantes de classe « i » et N le nombre total de plantes testées. Ainsi, en utilisant ce calcul, une plante présentant un IP égal à 0 sera totalement résistante et une plante présentant un IP égal à 100 sera totalement sensible.

2.5.2. Quantification de l'ADN du pathogène par qPCR

La quantification de l'ADN de l'agent pathogène présent au niveau des racines des plantes infectées a été réalisée à partir d'ADN extrait d'un pool (> 6 plantes) de racines du même génotype prélevées à 21 jours après inoculation. Les racines prélevées ont été dans un premier temps congelées dans l'azote liquide puis lyophilisées. L'extraction de l'ADN des racines lyophilisées (ADN plante + ADN agent pathogène) a été réalisée en utilisant le kit d'extraction « NucleoSpin Plant II » de Macherey-Nagel en suivant le protocole du kit. La détermination de la concentration en ADN de chacun des échantillons extraits a été réalisée en utilisant le Nanodrop ND-1000. Cinq à 10 ng d'ADN total et des amorces spécifiques de l'isolat eH de *P. Brassicae* ont été utilisés dans des réactions de qPCR afin de quantifier l'ADN du pathogène. La quantification absolue de l'ADN du pathogène présent dans les échantillons est effectuée en utilisant une gamme de dilutions d'ADN purifié de l'agent pathogène. Pour quantifier spécifiquement l'ADN de *P. Brassicae*, les amorces suivantes ont été utilisées :

- PbF-K1 : 5' TTGGGTAATTTGCGCGCCTG 3'
- PbR-K1 : 5' CAGCGGCAGGTCATTCAACA 3'

Ces amorces ont été dessinées au sein du laboratoire afin de cibler l'ARN 18S de *P. Brassicae*. Pour chaque échantillon trois répétitions techniques ont été réalisées. Les réactions de qPCR ont été réalisées dans un thermocycleur de marque Roche (LightCycler 480 II) en utilisant du Syber green (LightCycler® 480 SYBR Green I Master). Les conditions de qPCR utilisées sont les suivantes : une étape de dénaturation pendant 5 min à 95 °C suivie de 50 cycles d'amplification de l'ADN cible (dénaturation à 95 °C pendant 15 s puis hybridation à 63 °C pendant 30 s et extension à 72 °C pendant 30 s) et une étape d'extension finale à 72 °C pendant 5 min. Pour chaque échantillon et pour chaque point de gamme, une valeur de C_p a été déterminée en utilisant le logiciel Exor4 (LightCycler®480 version 1.5.0). Finalement la quantité d'ADN de pathogène dans chaque échantillon a été obtenue en faisant correspondre le C_p de l'échantillon à la quantité de pathogène correspondante en utilisant la gamme étalon de l'ADN purifié du pathogène.

2.5.3. L'indice GA/LA

L'indice GA/LA a été calculé comme décrit par Gravot *et al.* (2011) en utilisant les photographies des plantes étalées sur un support à 21 jours après inoculation. La surface des galles a été mesurée (index GA en cm²) en détournant ces dernières à l'aide du logiciel image J. La mesure de la plus grande feuille de la rosette a été réalisée soit à partir des photographies, soit à l'aide d'un pied à coulisse numérique. La longueur de la feuille obtenue a été élevée au carré donnant ainsi l'index LA en cm². Finalement l'indice GA/LA a été calculé en multipliant le ratio de l'index GA sur l'index LA par un coefficient arbitraire égal à 5000.

2.5.4. La longueur de la plus grande feuille de la rosette

Dans le but de suivre l'impact de la présence de l'agent pathogène sur le développement de la plante, la longueur de la plus grande feuille de la rosette a été mesurée avec un pied à coulisse 21 jours après inoculation. L'infection par *P. Brassicae* conduisant à un flétrissement et à un retard de croissance, cette mesure permet de quantifier l'impact de l'agent pathogène sur le développement foliaire et a été utilisée pour identifier les facteurs génétiques et épigénétiques impliqués dans la réponse de la plante à *P. Brassicae*.

2.5.5. Biomasse racinaire

Dans le but de suivre l'impact de la présence de l'agent pathogène sur le développement de la plante, la biomasse racinaire a été mesurée sur les échantillons lyophilisés de racines prélevées 21 jours après inoculation. *P. Brassicae* conduisant à une hypertrophie et à une hyperplasie cellulaire, le suivi de la biomasse racinaire dans les plantes infectées permet de quantifier l'impact de l'agent pathogène sur le système racinaire et a été utilisé pour identifier les facteurs génétiques et épigénétiques impliqués dans la réponse de la plante à *P. Brassicae*.

3. Analyse de la méthylation et de l'expression de gènes

3.1. Analyse de la méthylation par CHOP PCR et CHOP qPCR

Les profils de méthylation ont été analysés par la méthode utilisant l'enzyme McrBC (M0272L, BioLabs®). Les conditions utilisées pour traiter l'ADN avec la McrBC étaient de 0,5 µL BSA (20 mg/mL), 0,5 µL de GTP (20 mM), 5 µL NEB (10X) et 0,2 µL McrBC (10000 U/mL) pour traiter 40 ng d'ADN. Pour la CHOP PCR et qPCR 2 ng d'ADN digéré et non digéré ont été utilisés. Les conditions de températures utilisées ont été ajustées en fonction des amorces utilisées et 35 et 50 cycles d'amplification ont été effectués respectivement pour la CHOP PCR et qPCR. Pour déterminer l'état de méthylation de la région étudiée, chaque échantillon est présent sous sa forme digérée et non digérée (contrôle) par la McrBC. Le niveau de méthylation de la région ciblée a été calculé en pourcentage de molécules perdues lors de la digestion par la McrBC comme décrit par Silveira *et al.* (2013) en utilisant la formule :

$$\text{Niv. de méthylation} = (1 - 2^{-(Ct \text{ échantillon digéré} - Ct \text{ échantillon non digéré})}) * 100$$

Pour la CHOP qPCR le pourcentage de méthylation de l'ADN des gènes At5g13440 et At5g47400 a été utilisé respectivement comme contrôle hypométhylé et hyperméthylé.

3.2. Analyse de l'expression de gènes par RT-qPCR

Les ARN utilisés pour les approches de quantification d'expression par RT-qPCR ont été extraits à partir de racines préalablement lyophilisées en utilisant un protocole d'extraction au trizol. Les échantillons présentant des traces résiduelles d'ADN ont été traités à la DNase (Promega ; ref M6 10A). Avant l'étape de transcription inverse des ARN en ADNc par la SuperScript II (Invitrogen), la qualité des ARN a été vérifiée par électrophorèse sur gel d'agarose. Les RT-qPCR ont été réalisées sur les ADNc avec un thermocycler LightCycler® 480 (Roche) en utilisant le Syber green (LightCycler® 480 SYBR Green I Master) comme spécifié par le constructeur. L'expression des gènes a été normalisée en utilisant comme référence deux gènes d'*Arabidopsis* (At1g54610, At5g38470) définis comme stables durant l'infection par l'agent pathogène en utilisant des données de RNA-seq (non publiées). Pour chaque gène analysé, des jeux d'amorces spécifiques ont été dessinés.

4. Analyse de l'impact du méthylome de la plante sur la réponse à *P. Brassicae*

4.1. Approche non ciblée par détection de QTL^{épi}

4.1.1. Carte épigénétique de la population épiRIL

Afin d'identifier les régions génomiques d'*Arabidopsis* dont la variation de méthylation contrôle la réponse à la hernie (QTL^{épi}), nous avons utilisé la population épiRIL *ddm1-2* x Col-0 décrite précédemment. Pour cela, la carte épigénétique créée par Colomé-Tatché *et al.* (2012) avec un sous-set de 123 épiRIL a été utilisée (**Figure II.5**). Cette carte est basée sur la présence de régions différenciellement méthylées (DMR) entre les 2 parents et conservées dans la descendance (cf paragraphe I.2.4.5.2.2).

4.1.2. Analyses statistiques des données de phénotypage

Deux modèles linéaires généralisés ont été utilisés pour déterminer les effets du méthylome des plantes, de la température et de l'interaction méthylome x température sur la réponse de la plante à *P. Brassicae* en utilisant le package R glm2 (Marschne, 2011) de R version 3.2.2 (R Core Team 2015). Pour chaque modèle, l'option « family distribution » de la fonction glm a été adaptée en fonction de la distribution des données. Le premier modèle généralisé utilisé pour estimer l'effet du méthylome de chaque épiRIL dans chacun des environnements est décrit en (1) :

$$Y_{ijk} = \mu + G_i + R_j + B_k(R_j) + e_{ijk} \quad (1)$$

Où μ est l'effet moyen général, G_i est l'effet différentiel entre les épigénotypes, R_j est l'effet différentiel entre les répétitions biologiques, $B_k(R_j)$ est l'effet différentiel entre les blocs et les répétitions et e_{ijk} est la résiduelle.

Basé sur ce modèle, l'héritabilité au sens large a été estimée en utilisant l'équation ci-dessous :

$$H^2 = \frac{\sigma^2_G}{\sigma^2_G + \left(\frac{\sigma^2_e}{n}\right)} \quad (1.1)$$

Où σ^2_G est l'estimation de la variance épigénétique, σ^2_e est l'estimation de la variance environnementale et n est le nombre de répétitions par lignée.

Le second modèle linéaire généralisé a été utilisé pour estimer l'effet interactif entre le méthylome et la température :

$$Y_{ijkl} = \mu + G_i + T_j + R_k + B_l(R_k) + GT_{ij} + TR_{jk} + TB(R)_{jlk} + e_{ijkl} \quad (2)$$

Où μ est l'effet moyen général, G_i est l'effet différentiel entre les épigénotypes, T_j est l'effet différentiel entre les températures, R_k est l'effet différentiel entre les répétitions biologiques, $B_l(R_k)$ est l'effet différentiel entre les blocs et les répétitions, GT_{ij} est l'interaction entre l'épigénotype et la température, TR_{jk} est l'interaction entre les répétitions et la température, $TB(R)_{jlk}$ est l'interaction entre les blocs et la température et e_{ijkl} est la résiduelle.

Basé sur ce modèle, l'héritabilité au sens large a été estimée en utilisant l'équation ci-dessous :

$$H^2 = \frac{\sigma^2_G}{\sigma^2_G + \sigma^2_{GT} + \left(\frac{\sigma^2_e}{n}\right)} \quad (2.1)$$

Où σ^2_G est l'estimation de la variance épigénétique, σ^2_{GT} est l'estimation de la variance de l'interaction entre l'épigénotype et la température, σ^2_e est l'estimation de la variance environnementale et n est le nombre de répétitions par lignées.

Les données ajustées via les deux modèles ci-dessus et le package « lsmmeans » (Russell, 2016) de R ont été utilisés afin de réaliser la détection de QTL^{épi}.

4.1.3. Détection de QTL épigénétiques

Les données obtenues avec le package lsmmeans ont été traitées avec le package « qtl » (Broman *et al.*, 2003) sous l'environnement R 3.2.2 (R Core Team 2015). Le package « snow » (Tierney *et al.*, 2015)

permettant de paralléliser les opérations a été utilisé pour réduire les temps de calcul des tests de permutation.

Dans un premier temps, une approche par Simple Intervalle Mapping (SIM) a été réalisée pour déterminer les QTL^{épi} potentiellement présents en utilisant la méthode Haley-Knott (hk) (Broman & Sen, 2009) avec une taille de pas de 2 cM et une taille de fenêtre de 10 cM. Les seuils de détection ont été déterminés en réalisant 1000 permutations pour chaque condition et trait analysés. Le niveau de significativité du seuil a été fixé à $\alpha = 0,05$. Dans un second temps, la possibilité de la présence de plusieurs QTL^{épi} sur le même chromosome a été déterminée par une approche manuelle de cartographie en utilisant la méthode de QTL multiples (Broman & Sen, 2009) sur la base des résultats de l'analyse SIM. Pour cela, la fonction « stepwise » a été utilisée en sélectionnant les QTL^{épi} présents (système forward et backward, avec l'option « additive.only » = FALSE). Pour chaque trait, un minimum de deux QTL^{épi} potentiels a été utilisé dans la fonction « stepwise », même si un seul QTL^{épi} potentiel a été détecté en SIM. Les seuils LOD ont été calculés en utilisant 1000 permutations avec la fonction « scantwo » et un seuil de $\alpha = 0,05$. Une fois les QTL^{épi} sélectionnés, le modèle a été ajusté (fonction « fitqtl »), afin de calculer les scores LOD et les pourcentages de variation expliqués par chacun des QTL^{épi} (R^2). L'intervalle de confiance de chaque QTL^{épi} a été calculé avec la fonction de « Lodint » et une baisse de LOD d'une unité. L'effet de l'épiallèle a été évalué avec la fonction « effectplot ». L'interaction potentielle entre les QTL^{épi} incorporés dans le modèle a été testée avec la fonction « addint ».

4.1.4. *Analyses statistiques de l'impact des éléments transposables et des variants de séquence présents dans l'intervalle de confiance des QTL^{épi}*

Dans le but d'évaluer l'impact des éléments transposables sur la variance observée chez les épiRIL en réponse à la hernie, différents modèles de QTL ont été comparés comme décrit par Kooke *et al.* (2015). Brièvement, le résultat du modèle comprenant les marqueurs épigénétiques au pic de chaque QTL a été comparé au modèle incluant la présence ou non d'un élément transposable à la place du marqueur épigénétique. Le premier modèle a également été comparé au modèle contenant le marqueur épigénétique au pic du QTL avec l'élément transposable. Pour toutes les analyses réalisées seul l'état homozygote d'un élément transposable ou du variant de séquence a été utilisé à la place du marqueur épigénétique dans les modèles testés. Si l'élément transposable avait un effet plus important que le marqueur épigénétique au pic, alors le modèle ajusté est meilleur ou identique au modèle présentant le marqueur épigénétique au pic. La même stratégie a été appliquée sur les variants de séquence identifiés dans l'intervalle des QTL^{épi} détectés.

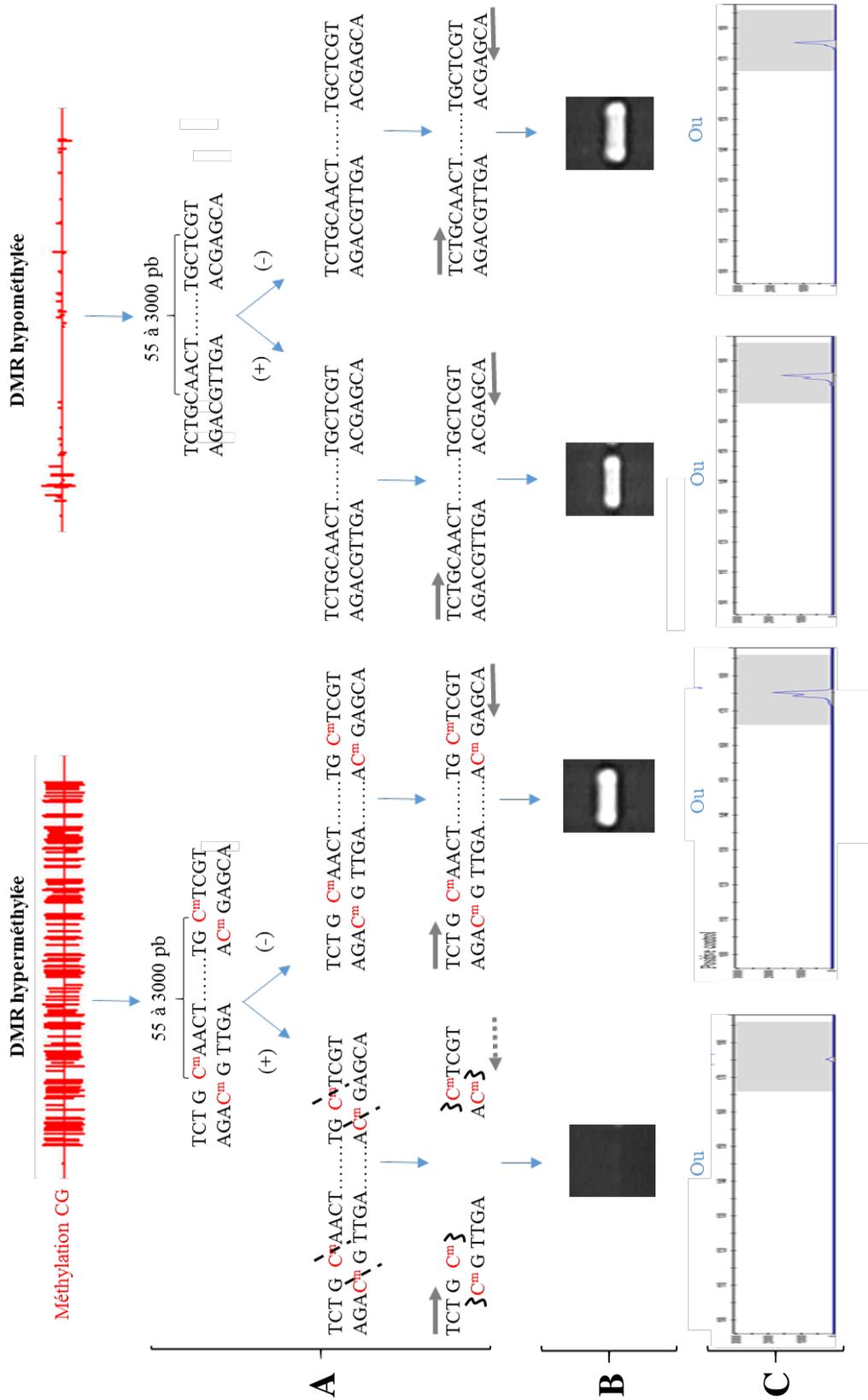


Figure II.6. Analyse du statut de méthylation des régions différentiellement méthylées. (A) Principe de fonctionnement de la McrBC. (B) Profil d'électrophorèse d'ADN digéré (+) ou non (-) par la McrBC puis amplifié par PCR. (C) Profils de fluorescence obtenus par multiplexage pour une DMR hypométhylée et une DMR hyperméthylée digérée (+) ou non (-) par la McrBC.

4.1.5. Réduction de l'intervalle de confiance des QTL^{épi}

La réduction de l'intervalle de confiance des 4 QTL^{épi} sélectionnés a été réalisée d'une part en ajoutant des marqueurs supplémentaires dans l'intervalle de confiance des QTL^{épi} et d'autre part en identifiant des individus recombinants supplémentaires dans la population entière.

4.1.5.1. Recherche d'individus recombinants supplémentaires

Quatre DMR par QTL^{épi} ont été sélectionnées sur la base de leur distribution régulière dans l'intervalle. Afin d'identifier de nouveaux recombinants dans les régions sélectionnées une enzyme de restriction coupant au niveau des cytosines méthylées, la McrBC (M0272L, BioLabs®) a été utilisée couplée à des amplifications par PCR (**Figure II.6A**, cf paragraphe 3.1). Pour chaque DMR la détection de l'état de méthylation repose sur un différentiel de signal obtenu entre un même échantillon traité et non traité par la McrBC (**Figure II.6B**). Au total 16000 amplifications ont été nécessaires afin d'obtenir le profil de méthylation des 4 DMR ciblées pour les 4 QTL^{épi} sélectionnés. Ces amplifications, ainsi que l'obtention des signaux de fluorescence issus d'une approche par multiplexage ont été réalisées en collaboration avec la plateforme GENTYANE de Clermont-Ferrand. L'analyse des résultats a été réalisée au laboratoire en utilisant le logiciel Gene Mapper® 4.0 (**Figure II.6C**) dans le cadre du stage de Master 2 de Juliette Bénéjam que j'ai co-encadré. Les profils de méthylation des 505 épiRIL ont ainsi été déterminés pour chaque épiRIL et pour les parents Col-0 et *ddm1-2*.

4.1.5.2. Densification en marqueurs épigénétiques chez les épiRIL recombinantes

L'ajout de marqueurs épigénétiques a été réalisé sur la base de la liste des régions différenciellement méthylées (DMR) déterminées par Colomé-Tatché *et al.* (2012). La détermination de l'épihaplotype a été réalisée pour chacun des recombinants comme précédemment en utilisant des étapes de digestion d'ADN à la McrBC suivi d'une amplification de la région par PCR.

4.1.5.3. Cartographie fine des 4 QTL^{épi} sélectionnés

Afin d'affiner la détection des QTL^{épi} après identification dans la population entière de recombinants supplémentaires, une nouvelle carte épigénétique a été construite combinant les 16 nouveaux marqueurs épigénétiques aux marqueurs initiaux dans les quatre régions ciblées (3.1.5.1). Au total 30 DMR (10 pour la région A, 16 pour la région B, 4 pour la région C) ont été utilisées comme marqueurs pour créer la carte épigénétique en utilisant le logiciel MAPMAKER/EXP 3.0. Les groupes de liaison ont été définis avec un LOD score minimum de 3 et une distance maximale entre les marqueurs de 40 cM. Les distances génétiques entre les marqueurs ont été calculées à l'aide de la fonction de Kosambi (Kosambi,1944).

Tableau II.5. Liste des amorces de CHOP PCR, CHOP qPCR et RT-qPCR utilisées pour l'étude de la méthylation de l'ADN par McrBC et de l'expression des gènes.

Gènes ou marqueur	Expérimentation	Amorce forward (5' > 3')	Amorce reverse (5' > 3')
MM10	Epigénotypage	AAGGGTCTGACTTGCTGGAC	AGAGCACAATTGCCTCAGGG
MM13	Epigénotypage	TTGTCCGGGTTAAGGTGCTC	ACTCCTTGTCTCGAAGCAGC
MM17	Epigénotypage	ACCTAAGAACTCGGTTCCGGC	AATGAATCAGGCTTGGGCGG
MM18	Epigénotypage	TTGCTGAAACATGCGATGCC	ATGCGGCTCTGCAAATGATG
MM20	Epigénotypage	CTGTTGCTCACTCATCTTGCC	TAGACACTTTCGGAGAGCCAC
MM22	Epigénotypage	ACTCGCAAACACGGAACAAC	AGTCGTCTCGTCGTTCTGTC
MM27	Epigénotypage	CCAGAGAGCACGGCTGATTT	AGGGTGAGCTAGTCTCCGAT
MM402	Epigénotypage	AAAGACCCATGTGCCGAGAG	TCGGTAACTTTCTGCCTCACC
MM415	Epigénotypage	AATTCCATTCAAGCCTGTCG	AACCAAGCAGCATTCATTCC
MM466	Epigénotypage	TATCCGGTTCAGCAGGTAGG	ACCTCATAACAGAGCCACCAAG
MM505	Epigénotypage	AACTGCTACTCGTCCTGCAC	TACAATCCTCCGAGCCATGC
MM510	Epigénotypage	ATGGCGAGCTATTCCTTCGG	TGGAATCGACGAGAGAAGCG
MM520	Epigénotypage	TCGGTCCCTTCTTTGAGTGC	AGCGCCTAATACACCCGATG
MM527	Epigénotypage	ACAAGGCAGCAAGTGACAATC	TCCACGAAGAAATCTGTGAGC
MM686	Epigénotypage	ATATAAGGCCCGGTTAGTTGC	TGCTCCCAACAAGACTGAAG
MM689	Epigénotypage	TCCTCCACCTCGAATCTAACC	TGTTAATGGGAAGCCCTAGC
MM691	Epigénotypage	TGTTCCGGTTGCTTCTATAACC	ATGGCACCACTTTCGTCTATG
MM694	Epigénotypage	ACGGTTGAGAGAAGGGAAGG	TTCATCGACACATCCTTTGG
<i>At1g54610</i>	qPCR	GGTCCGACAGAGGTAGAGCAG	GTATGGTTCACGGGGTTTGT
<i>At5g38470</i>	qPCR	TGTAATCGGGTATCCCTGCT	CTGGAGCTGCTGCTTGTG
<i>At5g47260</i>	qPCR	AAGGTGGTCCAATCGGGAAC	GATGGGGCAATCTGGTGTGA
<i>At5g47280</i>	qPCR	AGCATACCAAGTTCCGCAGT	TCGGGCCAAGTCATGTTGAA
<i>At5g47260</i>	CHOP qPCR	TGCGTCGACCTATCGTTACA	CCATGCCGTATCAAGCAAC
<i>At5g13440</i>	CHOP qPCR	ACAAGCCAATTTTTGCTGAGC	ACAACAGTCCGAGTGTGATGGT
<i>At5g47400</i>	CHOP qPCR	GAAGCCGAAGTCAAAGTGT	ATGGTCCGGCTCTAGGAAAA

4.2. Approche ciblée de l'influence du méthylome sur la réponse de la plante à *P. Brassicae*

4.2.1. Analyse du statut de méthylation au niveau de l'épimutation et de l'expression des 2 NLR sous-jacents

Le taux de méthylation au niveau du gène At5g47260 dans les écotypes d'*Arabidopsis* et des HIF a été quantifié en utilisant une approche de CHOP qPCR (pour les homo- et hétérozygotes). Pour cela l'ADN génomique des hétérozygotes a été digéré ou non en utilisant la McrBC comme décrit paragraphe 3.1. Pour la quantification, des amorces spécifiques des régions ciblées ont été réalisés (**Tableau II.5**). Les réactions de qPCR ont été réalisées comme décrit dans le paragraphe 3.1 (avec n = 3) avec les conditions suivantes : une étape de dénaturation pendant 5 min à 95 °C suivi de 50 cycles d'amplification l'ADN cible (dénaturation à 95 °C pendant 15 s puis hybridation à 63 °C pendant 30 s et extension à 72 °C pendant 30 s) et d'une étape d'extension finale à 72 °C pendant 5 min. La valeur de Cp a été déterminée en utilisant le logiciel Exor4 (LightCycler®480 version 1.5.0) et le taux de méthylation présent au niveau de la région ciblée a été estimé à partir des Cp obtenus comme décrit dans le paragraphe 3.1.

Pour les écotypes d'*Arabidopsis* la quantification relative de l'expression des gènes At5g47260 et At5g47280 a été réalisée en RT-qPCR comme décrit paragraphe 3.2 en utilisant des amorces spécifiques pour chacun des gènes (**Tableau II.5**). Les conditions de qPCR utilisées sont identiques que celles décrites ci-dessus. L'expression relative de chaque gène a été calculée en utilisant la méthode du $\Delta\Delta^{Ct}$.

4.2.2. Séquençage des ARN messenger

4.2.2.1. Isolement des ARN, construction des banques d'ADNc et séquençage

L'extraction des ARN totaux a été réalisée sur des racines congelées et lyophilisées (collectées 14 jours après inoculation) en utilisant le kit mirVana™ miRNA Isolation Kit (Ambion) en suivant les instructions fournies par le fabricant. Après extraction, la quantité d'ARN a été quantifiée en utilisant le NanoDrop ND-1000 et leur qualité a été vérifiée en utilisant une puce RNA 6000 assay kit total RNA (Agilent). Les échantillons possédant un RIN (RNA integrity number) supérieur ou égal à 7 ont été utilisés pour le séquençage. La construction des banques de séquençage ainsi que le séquençage ont été réalisés par la plateforme NGS de l'institut Marie Curie à Paris. Pour chaque échantillon un séquençage par Illumina HiSeq 2500 de type Paired-Ends (100x100 pb) a été réalisé.

4.2.2.2. Alignement des lectures et comptage

Pour chaque échantillon, les lectures en Paired-Ends obtenus ont été alignées sur le génome assemblé et annoté TAIR10 de l'écotype Columbia d'*A. thaliana* auquel a été concaténé le génome de l'isolat eH de *P. Brassicae*. L'alignement a été réalisé en utilisant le logiciel STAR (Version 2.5.3.a). Les paramètres d'alignement utilisés ont été sélectionnés afin de correspondre au génome d'*Arabidopsis*.

Un maximum de 5 alignements multiples et de 3 mésappariements ont été autorisés. Ensuite, les fichiers .bam obtenus ont été utilisés pour réaliser le comptage des reads présents au niveau de chaque gène en utilisant la fonction Counts de Feature Counts (version 1.4.6). Afin de réaliser ce comptage un fichier d'annotation concaténé des annotations TAIR10 d'*A. thaliana* et de l'annotation du génome de *P. Brassicae* a été utilisé.

4.2.2.3. Détermination des gènes différentiellement exprimés

L'analyse des gènes différentiellement exprimés a été réalisée comme décrit par Law *et al.* (2016) sous l'environnement R (3.3.0) en utilisant le package « EdgeR » (version 3.16.5 ; Robinson *et al.*, (2010) ; McCarthy *et al.*,(2012)). Les comptages bruts ont été utilisés comme données d'entrée dans EdgeR. Les comptages bruts ont été convertis en « Count Per Million » (CPM) et seul les CPM supérieurs à 1 ont été conservés. Ensuite, pour chaque gène, les données d'expression ont été normalisées en utilisant la méthode TMM (Trimmed Mean of M-values) de la fonction « CalcNormFactors » de « EdgeR ». Finalement, les gènes différentiellement exprimés ont été détectés en utilisant la fonction « decideTests » pour un fold change minimum compris entre -1,5 et 1,5.

4.2.3. Séquençage des petits ARN

4.2.3.1. Isolement des petits ARN, construction des banques d'ADNc et séquençage

L'extraction des petits ARN a été réalisée sur des racines congelées et lyophilisées (collectées 14 jours après inoculation) en utilisant le kit mirVana™ miRNA Isolation Kit (Ambion) en suivant les instructions fournies par le fabricant. La quantité et la qualité des ARN extraits a été réalisé comme dans la partie précédente. La construction des banques de séquençage ainsi que le séquençage ont été réalisés par la plateforme NGS de l'institut Marie Curie à Paris avec un Illumina HiSeq 2500 de type Single End (50 pb).

4.2.3.2. Alignement des lectures et comptage

Pour chaque échantillon les lectures en Single End obtenues ont été alignées sur le génome assemblé et annoté TAIR10 de l'écotype Columbia d'*A. thaliana* auquel a été concaténé le génome de l'isolat eH de *P. Brassicae* en utilisant le logiciel STAR (version 2.5.3.a). Le comptage et le regroupement des lectures ont été effectués en utilisant Shortstack (Axtell, 2013).

4.2.3.3. Détermination des petits ARN différentiellement exprimés

L'analyse des petits ARN différentiellement présents a été réalisée comme pour les ARNm (cf. II.4.2.2.3) à l'exception du seuil de CPM qui devait être supérieur à 1 dans au moins 3 échantillons.

4.2.4. *Origine des données génomiques, d'expression et de méthylation des accessions naturelles d'Arabidopsis et des épîRIL.*

L'analyse du pourcentage de couverture génomique, des taux de méthylation présents au niveau des régions d'intérêt et de l'expression des gènes dans les différentes accessions naturelles d'*Arabidopsis* ont été réalisées en utilisant les données générées par Alonso-Blanco *et al.* (2016) et Kawakatsu *et al.* (2016) et stockées respectivement sur le NCBI SRA number SRP056687 et sur le NCBI GEO sous les références : GSE43857 et GSE80744. Les données de bisulfite et de petits ARN de mutants d'*Arabidopsis* ont été obtenues dans l'étude de Stroud *et al.* (2014) et archivées sur le NCBI GEO sous la référence : GSE39901. Les données de bisulfite, de MedIP-CHIP et de RNAseq concernant les épîRIL nous ont été fournies par le laboratoire du docteur Vincent Colot.

4.2.5. *Traitement en visualisation des données.*

Les analyses statistiques ont été réalisées en utilisant le logiciel R (version 3.3.0, R core Team 2015). La clusterisation des écotypes d'*Arabidopsis thaliana* en fonction de leur statut de méthylation en contexte CG, CHG et CHH au niveau de l'épimutation a été réalisé en utilisant la fonction « heatmap.2 » du package « gplots » (<https://CRAN.R-project.org/package=gplots>). Les histogrammes et les boîtes à moustaches ont été générés en utilisant respectivement les fonctions « ggbarplot » et « ggboxplot » du package « ggpubr » (<https://CRAN.R-project.org/package=ggpubr>). Les représentations des taux de méthylation à l'échelle du génome chez les épîRIL et chez les mutants d'*Arabidopsis thaliana* altérés au niveau de gènes impliqués dans la structure de la chromatine ont été réalisées en utilisant le logiciel IGV (Integrative Genome Viewer). Les représentations des mapping de lecture de DNA-seq et des profils de méthylation des écotypes d'*Arabidopsis thaliana* ont été réalisées en utilisant le browser de 1001 génomes (<http://neomorph.salk.edu/1001.aj.php>). La représentation du taux de méthylation au niveau de la région *RPB1* dans l'écotype Col-0 a été réalisée en utilisant le browser du laboratoire de Jacobsen (<http://genomes.mcdb.ucla.edu/cgi-bin/hgTracks?db=araTha2&position=chr5%3A3172000-3180999>). La représentation du mapping des lectures de smRNA-seq a été réalisée à partir des fichiers BAM produits par le logiciel Shortstack (Axtell, 2013) en utilisant le logiciel IGV.

CHAPITRE III

**APPROCHE NON CIBLEE DE L'IMPACT DE LA
METHYLATION DE L'ADN D'*ARABIDOPSIS*
THALIANA SUR LA REPONSE A L'INFECTION
PAR *PLASMODIOPHORA BRASSICAE***

CHAPITRE III. APPROCHE NON CIBLEE DE L'IMPACT DE LA METHYLATION DE L'ADN D'*ARABIDOPSIS THALIANA* SUR LA REPONSE A L'INFECTION PAR *PLASMIDIOPHORA BRASSICAE*

Comme présenté dans le chapitre bibliographique, de nombreuses études ont montré l'implication de l'épigénétique dans la réponse des plantes aux stress qu'ils soient d'origine biotique ou abiotique (cf. Chapitre I, paragraphe 2.4.6.). L'interaction entre *Arabidopsis* et l'agent causal de la hernie *P. Brassicae* induit de nombreuses modifications au niveau hormonal, métabolique et morphologique au sein de la plante (cf. Chapitre I, paragraphes 3.2.1. et 3.2.2.). Le type et le niveau des modifications induites par l'agent pathogène sont différents en fonction du génotype de la plante et du génotype de l'agent pathogène. Actuellement la résistance quantitative à la hernie est considérée comme résultant principalement de la ségrégation de nombreux allèles. Cependant, les études menées jusqu'à maintenant montrent que les variations alléliques ne permettent pas d'expliquer à elles seules tout le phénotype de réponse de résistance observé, suggérant la présence d'une autre source de variabilité telle qu'une variabilité d'origine épigénétique. En effet, de récentes études ont montré que des variations héritables de méthylation de l'ADN peuvent aussi être responsables de la variation de caractères quantitatifs (cf. Chapitre I, paragraphe 2.4.5.2.2.). Ces observations nous ont conduits ainsi à émettre l'hypothèse qu'une résistance quantitative pourrait résulter de la ségrégation génomique de nombreux allèles et épiallèles. Des études préliminaires réalisées au sein du laboratoire sur des mutants d'*Arabidopsis* altérés au niveau de gènes codant des protéines impliquées dans la régulation épigénétique de la chromatine (dans un fonds génétique Col-0) en interaction avec *P. Brassicae* ont mis en évidence des niveaux de résistance/sensibilité à la hernie altérés chez ces mutants. En effet, le mutant *ddm1* a montré une sensibilité moins importante vis-à-vis de l'isolat eH de *P. Brassicae* que l'écotype Col-0. Ces résultats ont conforté notre hypothèse sur le potentiel impact de l'épigénétique, notamment du méthylome de la plante, sur la résistance quantitative de la plante à la hernie.

L'objectif de cette première partie de la thèse est d'évaluer de manière non ciblée s'il existe une variabilité épigénétique héritable à l'origine de variations de la réponse à la hernie. Pour ce faire, une approche d'épigénétique quantitative a été réalisée. Afin d'identifier la contribution de la variabilité épigénétique dans la réponse à la hernie en s'affranchissant des variations nucléotidiques, nous avons utilisé la population épiRIL *ddm1-2* x Col-0 créée par Johannes *et al.* (2009). Dans un premier temps, afin d'identifier les régions sous contrôle épigénétique impliquées dans la résistance à la hernie, différents traits phénotypiques ont été mesurés sur 123 des 505 lignées épiRIL suite à leur infection par l'isolat eH de *P. Brassicae*. Les données phénotypiques recueillies au cours de cette évaluation ont ensuite été mises en relation avec la carte épigénétique développée sur ces mêmes individus par

Colomé-Tatché *et al.* (2012) à partir de 126 régions différentiellement méthylées (DMR). Ce travail développé dans la première partie de ce chapitre - sous la forme d'un article actuellement soumis et accepté (avec révisions majeures) à *New Phytologist* - a permis l'identification de 14 QTL épigénétiques (QTL^{épi}) additifs regroupés au niveau de 6 régions génomiques réparties sur 4 des 5 chromosomes d'*Arabidopsis*. Afin d'identifier les gènes causaux sous-jacents et leurs variations épialléliques associées, j'ai ensuite réalisé la cartographie fine de 3 de ces régions génomiques en utilisant l'ensemble des 505 lignées de la population épiRIL et par l'ajout de DMR supplémentaires.

Les enjeux à terme de cette partie de ma thèse sont d'une part d'avoir une meilleure compréhension des mécanismes de résistance partielle à la hernie chez les *Brassicaceae* avec l'identification de nouvelles régions/nouveaux loci intervenant dans la résistance. D'autre part, les résultats obtenus permettront de déterminer si les facteurs de résistance contrôlés par des variations d'origine épigénétique sont différents de ceux affectés par des variations alléliques et ainsi de déterminer si les variations alléliques et épialléliques agissent de concert sur un même facteur de résistance pour le contrôle génétique/épigénétique de la réponse de la plante vis-à-vis de *P. Brassicae*.

1. Lien entre les variations du méthylome d'*Arabidopsis* et la réponse à la hernie

Cette partie de thèse est présentée sous la forme d'un article actuellement soumis et accepté (avec révisions majeures) dans la revue *New phytologist*. Des données supplémentaires sont fournies dans les paragraphes qui suivent l'article.

QUANTITATIVE RESISTANCE TO CLUBROOT INFECTION MEDIATED BY TRANSGENERATIONAL EPIGENETIC VARIATION IN *ARABIDOPSIS*

Benjamin Liégard¹, Victoire Baillet², Mathilde Etcheverry², Evens Joseph¹, Christine Lariagon¹, Jocelyne Lemoine¹, Aurélie Evrard¹, Vincent Colot², Antoine Grivot¹, Maria J. Manzanares-Dauleux¹, Mélanie Jubault¹.

¹ IGEPP, INRA, AGROCAMPUS OUEST, Université de Rennes, F-35000 Rennes, France

² Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure, Centre National de la Recherche Scientifique (CNRS), Institut National de la Santé et de la Recherche Médicale (INSERM), F-75005 Paris, France

SUMMARY

- Quantitative disease resistances, often influenced by environmental factors, are thought to be the result of DNA sequence variants segregating at multiple loci. However, heritable differences in DNA methylation, so-called transgenerational epigenetic variants, could also contribute to quantitative traits. Here, we tested this possibility using the well-characterized quantitative resistance of *Arabidopsis* to clubroot, a *Brassica* major disease caused by *Plasmodiophora brassicae*.
- For that, we used the epigenetic recombinant inbred lines (epiRIL) derived from the cross *ddm1-2* x Col-0, which show extensive epigenetic variation but limited DNA sequence variation. Quantitative loci under epigenetic control (QTL^{epi}) mapping was carried out on 123 epiRIL infected with *P. Brassicae* and using various disease-related traits.
- EpiRIL displayed a wide range of continuous phenotypic responses. Seventeen QTL^{epi} were detected across the five chromosomes, with a *bona fide* epigenetic origin for fourteen of them. The effect of five QTL^{epi} was dependent on temperature conditions. Thirteen QTL^{epi} co-localized with previously identified clubroot resistance genes and QTL in *Arabidopsis*.
- Co-localization of clubroot resistance QTL^{epi} with previously detected DNA-based QTL reveals a complex model in which a combination of allelic and epiallelic variations interacts with the environment to lead to variation in clubroot quantitative resistance.

Keywords: *Arabidopsis*, Clubroot, Epigenetics, EpiRIL, Methylome, *Plasmodiophora brassicae*, Quantitative resistance, Transgenerational

1.1. Introduction

Clubroot caused by the protist *Plasmodiophora brassicae* is a major disease of *Brassicaceae* including the three most economically important *Brassica* species: *B. napus*, *B. rapa*, *B. oleracea* (Dixon, 2009) and the model plant *Arabidopsis thaliana* (Koch *et al.*, 1991). *P. Brassicae* infection leads to tumorous club formation on roots resulting from cell hyperplasia and hypertrophy (Ingram & Tommerup, 1972). Cropping practices and crop protection products have limited success in controlling clubroot (Dixon, 2009). Currently, one of the most effective ways to limit the impact of this disease is to use resistant varieties (Diederichsen *et al.*, 2009). To date, both qualitative and quantitative (QTL) clubroot resistance loci have been identified in different *Brassicaceae* species (Manzanares-Dauleux *et al.*, 2000a; Rocherieux *et al.*, 2004; Alix *et al.*, 2007; Jubault *et al.*, 2008; Piao *et al.*, 2009; Lee *et al.*, 2016). Current approaches to generate resistant varieties rely mainly on a few loci controlling qualitative resistance, with the inevitable outcome of rapid adaptation of the pathogen populations (Diederichsen *et al.*, 2009). In this context, diversification and access to other sources of clubroot resistance variability is becoming necessary.

Numerous studies (Downen *et al.*, 2012; Luna *et al.*, 2012; Zhang *et al.*, 2013; Liu *et al.*, 2015; Zheng *et al.*, 2017; Aoun *et al.*, 2017) have reported that plant responses to abiotic (temperature, drought) and biotic stresses could be associated with epigenetic variation in addition to nucleotidic variation. For instance, Downen *et al.* (2012) and Yu *et al.* (2013) have shown that *Arabidopsis* mutants altered in the maintenance of DNA methylation in the CG, CHG and CHH contexts, showed strong resistance to *Pseudomonas syringae* pv, *tomato* strain DC3000. The role of epigenetics in the expression of adaptive plant traits thus suggests that epigenetic variability could be used for generating stress-tolerant or resistant plants. Furthermore, the occurrence of natural DNA methylation variants (epialleles) in plants (Becker *et al.*, 2011, Schmitz *et al.*, 2011) and their implication in evolution (Weigel & Colot, 2012) suggest that epialleles could be considered as a source of variability in plant breeding. In this “epigenetic breeding” approach (Gallusci *et al.*, 2017), two conditions are needed: transgenerational inheritance of epialleles and a clear connection between epigenotype and observed phenotype. Previous studies demonstrated that epialleles could be stably transmitted across at least 8 generations (Teixeira *et al.*, 2009; Johannes *et al.*, 2009), and that such heritable differences in DNA methylation could be associated with heritable phenotypic variation for several complex traits (Johannes *et al.*, 2009; Reinders *et al.*, 2009). However, linking heritable phenotypic variation to epigenetic variation remains challenging because of the difficulty in teasing apart its effects to that of DNA sequence variation in natural settings (Johannes *et al.*, 2008; Quadrana & Colot, 2016). This problem can however be greatly alleviated in *Arabidopsis* by using experimental populations of so-called epigenetic recombinant inbred lines (epiRIL), which show extensive epigenetic variation but limited DNA sequence variation (Johannes *et al.*, 2009; Reinders *et al.*, 2009). One such population (Johannes *et al.*, 2009) was indeed used to build a genetic map solely based on heritable differences in DNA methylation (differentially

methyated regions, DMR) (Colomé-Tatché *et al.*, 2012) and to identify epigenetic quantitative trait loci (QTL^{epi}) for several complex traits (Cortijo *et al.*, 2014; Kooke *et al.*, 2015; Aller *et al.*, 2018).

In the present study, we used this same epiRIL population to determine the impact of heritable differences in DNA methylation on the response of *Arabidopsis* to clubroot. We first showed that *ddm1* mutants were less susceptible to clubroot than the wild-type Col-0 and that the assessed subset of 123 epiRIL displayed a wide range of continuous phenotypic responses to clubroot. Seventeen QTL^{epi} were detected across the five chromosomes, with a *bona fide* epigenetic origin for fourteen of them. We have thus demonstrated that heritable differences in DNA methylation could also contribute to quantitative resistance to clubroot. Thirteen QTL^{epi} co-localized with previously identified clubroot resistance genes and QTL in *Arabidopsis* revealing that quantitative resistance to clubroot in natural accessions could be controlled by both nucleotidic and epigenetic variations.

1.2. Materials and methods

1.2.1. Plant material

Mutant plants altered in genes encoding chromatin modifiers were ordered from the NASC. All plants were obtained from T-DNA insertion in the Columbia (Col-0) genetic background and showed a homozygous insertion in the plant genome. The mutant lines used were *ddrm2-2* (SALK_150863), *hda15* (SALK_004027C), *atxr5* (SALK_130607C), *hac1* (SALK_082118C), *srt2* (SALK_149295C) and *ddm1* (SALK_000590C). T-DNA insertions and homozygosity were confirmed by PCR using the set of appropriate primers designed with the T-DNA Primer Design interface (<http://signal.salk.edu/tdnaprimers.2.html>, **Annexe I.A**). The epiRIL population is that derived from a cross between the *A. thaliana* ecotype Col-0 and a Col-0 line carrying the *ddm1-2* mutant allele (Johannes *et al.*, 2009). Note that the *ddm1-2* allele was obtained by EMS mutagenesis, not T-DNA transformation (Vongs *et al.*, 1993).

1.2.2. Clubroot pathogen

All clubroot tests were performed with the *P. Brassicae* eH isolate described by Fahling *et al.* (2003). Isolate eH belongs to the P1 pathotype according to the classification using the differential host set of Some *et al.* (1996). One ml of resting spore suspension (10^7 spores.ml⁻¹) prepared according to Manzaneres-Dauleux *et al.* (2000b) was used for pathogen inoculation 10 days (d) after germination (stage 1.04 (Boyes *et al.*, 2001)). The inoculum was applied to the bottom of the stem base of each seedling.

1.2.3. Growth conditions

The *Arabidopsis* accession Col-0 and the six mutant lines (Col-0 background) were evaluated in a randomised complete block design (with three blocks) against the eH *P. Brassicae* isolate. For QTL^{epi}

detection, the 123 epiRIL that have been previously epigenotyped (Colomé-Tatché *et al.*, 2012) together with the two parental lines were evaluated in four biological replicates, split in two growth rooms, in a randomized complete block design (with two blocks per replicate). In growth room-2, four temperature sensors per block were placed at the height of plants. For each pathological test, seed germination was synchronized by placing seeds on wet blotting paper in Petri dishes for two days at 4 °C. Seeds were sown individually in pots (four cm diameter) containing a sterilised mix (by autoclaving) composed of 2/3 compost and 1/3 vermiculite. Per block, six plants per genotype (T-DNA mutants, Col-0, *ddm1-2* and epiRIL) were grown in controlled conditions of 16 h light ($110 \mu\text{mol m}^{-2}\text{s}^{-1}$) at 20 °C and 8 h dark at 18 °C.

1.2.4. Phenotyping

Phenotyping was performed three weeks after inoculation (21 days post inoculation (dpi)). Plants were thoroughly rinsed with water and photographed. Infected roots were removed and frozen in liquid nitrogen. For the epiRIL and their parents, four disease-related traits were assessed: longest leaf length (Lfi), disease index (DI), Root biomass (Rbi) and pathogen DNA quantity (Pb), whereas for mutants, only DI was evaluated. DI was calculated according to Manzanares-Dauleux *et al.* (2000b) and pathogen DNA quantification was determined by quantitative PCR (qPCR). These phenotypic traits were chosen due to their relevance for the study of the quantitative response to eH isolate on *Arabidopsis* (Jubault *et al.*, 2008; unpublished).

1.2.5. Pathogen DNA quantification

Total genomic DNA (plant and pathogen) was extracted from lyophilised infected roots with the NucleoSpin Plant II kit according to manufacturer's instructions. After the evaluation of DNA concentration by Nanodrop ND-1000, 5 to 10 ng of total genomic DNA extract and specific primers were used to quantify by qPCR pathogen DNA quantity in each epiRIL and the parental lines. The pathogen specific primers designed on the 18S rRNA of *P. Brassicae* used were PbF-K1 5' TTGGGTAATTTGCGCGCCTG 3'; PbR-K1 5' CAGCGGCAGGTCATTCAACA 3'. qPCR reactions were carried out in a thermocycler (LightCycler 480 II Roche), with Syber green (LightCycler® 480 SYBR Green I Master). The qPCR conditions used were: 50 cycles of denaturation at 95 °C for 15 s, annealing/extension at 63 °C for 30 and 72 °C for 30 s. Absolute quantification of pathogen DNA was carried out using a standard curve obtained based on a series of known amounts of pathogen DNA. The average of the pathogen quantity present in each epiRIL (ratio of pathogen DNA quantity to total DNA used for qPCR) was used in further statistical analyses and QTL^{epi} detection.

1.2.6. Data analysis

Two generalized linear models were used to determine the effects of epigenome, temperature and epigenome x temperature interaction with the R package "glm2" (Marschne, 2011) in R software

version 3.2.2 (R Core Team 2015). For each model, the family distribution option of the glm function was adapted according to the data distribution. The first generalized linear model described in (1) was used to estimate the epigenotype effect of each epiRIL across biological replicates in each growth room:

$$Y_{ijk} = \mu + G_i + R_j + B_k(R_j) + e_{ijk} \quad (1)$$

Where μ is the mean general effect, G_i is the differential effect between epigenotypes, R_j is the differential effect between replicates, $B_k(R_j)$ is the interaction between blocks and replicates and e_{ij} for the residual variance. Based on this model, broad sense heritability was estimated using the following equation:

$$H^2 = \frac{\sigma^2_G}{\sigma^2_G + \left(\frac{\sigma^2_e}{n}\right)} \quad (1.1)$$

Where σ^2_G is the estimated epigenetic variance, σ^2_e is the estimated environmental variance and n is the number of replicates per line.

The second generalized linear model described in (2) was used to estimate the epigenotype x temperature effect in the growth room-2 for each epiRIL across the biological replicates:

$$Y_{ijkl} = \mu + G_i + T_j + R_k + B_l(R_k) + GT_{ij} + TR_{jk} + TB(R)_{jlk} + e_{ijkl} \quad (2)$$

Where μ is the mean general effect, G_i is the differential effect between epigenotypes, T_j is the differential temperature conditions, R_k is the differential effect between replicates, $B_l(R_k)$ is the interaction between blocks and replicates, GT_{ij} is the interaction between epigenotype and temperature condition, TR_{jk} is the interaction between replicate and temperature condition, $TB(R)_{jlk}$ is the interaction between block and temperature condition and e_{ijkl} is the residual variance. Based on this model, broad sense heritability was estimated using the following equation:

$$H^2 = \frac{\sigma^2_G}{\sigma^2_G + \sigma^2_{GT} + \left(\frac{\sigma^2_e}{n}\right)} \quad (2.1)$$

Where σ^2_G is the estimated epigenetic variance, σ^2_{GT} is the estimated epigenetic x temperature variance, σ^2_e is the estimated environmental variance and n is the number of replicates per line.

For all traits, least square means on each effect according to the two models described above were estimated with the function lsmeans of “lsmeans” package (Russell, 2016) in R software version 3.2.2 (R Core Team 2015). Lsmeans function was also used to extract the epigenotype effect (G) and the interaction epigenotype x temperature (GT) of each trait according to the generalized model used.

1.2.7. *QTL^{epi} detection*

Epigenotype effect (G) and the interaction epigenotype x temperature (GT) of each trait were treated with the package “qtl” (Broman *et al.*, 2003) of R software version 3.2.2 (R Core Team 2015). The package “snow” (Tierney *et al.*, 2015) allowing the use of processor cores as cluster was used to reduce permutation calculation time. Simple Interval Mapping (SIM) was first carried out to identify potential QTL^{epi} with the Haley-Knott (hk) method (Broman & Sen, 2009), using a step size of 2 cM and a window size of 10 cM. One thousand permutations with the hk method were carried out in order to determine SIM threshold levels for each condition and trait analyzed. The significance level of threshold was fixed at $\alpha = 0.05$. In order to integrate the possibility of the presence of multiple QTL^{epi} on the same chromosome, a manual multiple QTL mapping (MQM) approach (Broman & Sen, 2009) was used based on the results of SIM analysis. For this, the stepwise function was used in order to select the QTL^{epi} (forward and backward system, option “additive.only = FALSE”) based on the preliminary putative QTL^{epi} identified by SIM. For each trait, a minimum of two potential QTL^{epi} was used in the stepwise function even if only one potential QTL^{epi} was detected in SIM. LOD thresholds were calculated using 1000 permutations with the function scantwo with a significance level of $\alpha = 0.05$. Once the QTL^{epi} selected, the model was fitted (fitqtl function), in order to calculate the LOD scores and the percentage of variation explained by each and all QTL^{epi} (R^2). The confidence interval of each QTL^{epi} was calculated with the lodint function with a LOD drop of one parameter. The epiallele effect was evaluated with the function effectplot. Putative interactions among the QTL^{epi} incorporated in the model were tested with the function addint according to Browman and Sen (2009).

1.2.8. *DNA sequence variation*

Whole genome sequence data are available for the 123 epigenotyped epiRILs (Gilly *et al.*, 2014; unpublished data). The joint identification of small-scale variants (single-base substitutions and indels < 100 bp) was performed using GATK HaplotypeCaller on the whole-genome sequencing data available for 122 epiRIL (Baillet *et al.*, unpublished). Raw variant calls were then filtered following GATK Best Practice suggestions and additional scripts. All variants were visually inspected using IGV and a subset of the detected SNP was validated by PCR. For the analysis of transposition events, TE-Tracker (Gilly *et al.*, 2014) was used to identify non-reference TE insertions in 102 epiRIL of sufficient genomic coverage. Raw calls were then filtered using in-house scripts. All of the detected new TE insertions were visually inspected using IGV and a fraction of them were validated by PCR (Etcheverry *et al.*, unpublished). Shared small-scale and TE insertion variants are defined as present in at least 25 % of the population. In order to evaluate the impact of shared sequence variants on the heritable variation observed in the epiRIL in response to clubroot, we compared various QTL models as described by Kooke *et al.* (2015). Three different models were tested i) using all QTL peak epigenetic markers (MM), ii) using each shared DNA sequence variant (SNP, indel, TE insertion) located in the confidence interval instead of the epigenetic marker at the peak, and iii) using both peak epigenetic markers and the shared



Figure III.1: Clubroot symptoms shown by Col-0 wild-type and the *ddm1* mutant (Col-0 background). Col-0 and mutants were inoculated 10 days after germination and phenotyped three weeks after inoculation. (A) Col-0 ecotype; (B) *ddm1* mutant. The *ddm1* mutant showed a significant decrease in disease symptoms (Dunnett test, p -value < 0.02) compared to Col-0.

DNA sequence variants included in the confidence interval. QTL detected using the three models were compared. If the DNA-based markers had a more significant effect than the peak QTL epigenetic markers, then the model fitted was considered to be better or identical to the model with the peak epigenetic markers.

1.3. Results

1.3.1. *ddm1* mutants have reduced susceptibility to *P. Brassicae*

To determine if epigenome variations could affect clubroot responses, we first assessed several *Arabidopsis* epigenetic mutants in genes encoding chromatin modifiers. Infection of Col-0 and six T-DNA insertion mutants in genes encoding chromatin modifiers (*atxr5*, *ddm1*, *drm2*, *hac1*, *hdc15* and *srt2*) with the *P. Brassicae* eH isolate led to a range of phenotypic responses. A statistically significant Disease Index (DI) effect (ANOVA: $F = 4.99$, $p\text{-value} = 0.005$) (**Annexe I Table S2A**) was identified 21 days after infection suggesting that epigenetics was involved in plant response to infection by *P. Brassicae*. A Dunnett's post hoc test (**Annexe I Table S2B**) revealed a significant difference between *ddm1* and Col-0 ($t = -4.437$, $p\text{-value} = 0.003$) for the DI trait, *ddm1* showing reduced symptoms compared to Col-0 (**Figure III.1**).

1.3.2. Heritable differences in DNA methylation are associated with differential susceptibility to *P. Brassicae*

1.3.2.1. *EpiRIL* response to *P. Brassicae* is quantitative

To identify epialleles involved in clubroot response, QTL^{epi} detection was carried out on the subset of 123 lines of the *EpiRIL* population previously used in Colomé-Tatché *et al.* (2012). In total, each genotype was assessed against *P. Brassicae* isolate eH in four biological replicates, split in two growth rooms, each biological replicate being composed of two blocks. Distribution of the four disease-related traits assessed on the 123 *EpiRIL* showed continuous distribution suggesting polygenic control of these traits (**Figure III.2**). However, significant differences were observed for all traits between the biological replicates (Mann-Whitney U-test) set up in the two growth rooms (DI: $W = 3685.5$, $p\text{-value} < 2.2e-16$; Lfi: $W = 233530$, $p\text{-value} < 2.2e-16$; Pb: $W = 4826$, $p\text{-value} < 2.2e-16$), suggesting an influence of the growth conditions on the *EpiRIL* response to clubroot (**Annexe I Table S3**). Indeed, higher levels of disease symptoms were observed on the Col-0 parent line (growth room-1: $DI = 53.25 \pm 2.36$; growth room-2: $DI = 90.75 \pm 9.64$) and on the *EpiRIL* population (growth room-1: $DI = 51.17 \pm 11.42$; growth room-2: $DI = 86.27 \pm 10.74$) growing in growth room-2 compared to growth room-1. Consequently, we decided to analyze data from the two growth rooms independently. Analysis of biological replicates grown in each growth room showed a significant epigenotype effect (glm, model 1) on all phenotypic traits measured (growth room-1: Lfi: $\chi^2 = 126.32$, $p\text{-value} < 2.2e-16$; DI: $\chi^2 = 22256.3$, $p\text{-value} = 1.57e-06$; Pb: $\chi^2 = 71.73$, $p\text{-value} = 0.02$; growth room-2 : Lfi: $\chi^2 = 32.74$, $p\text{-value} < 2.2e-16$;

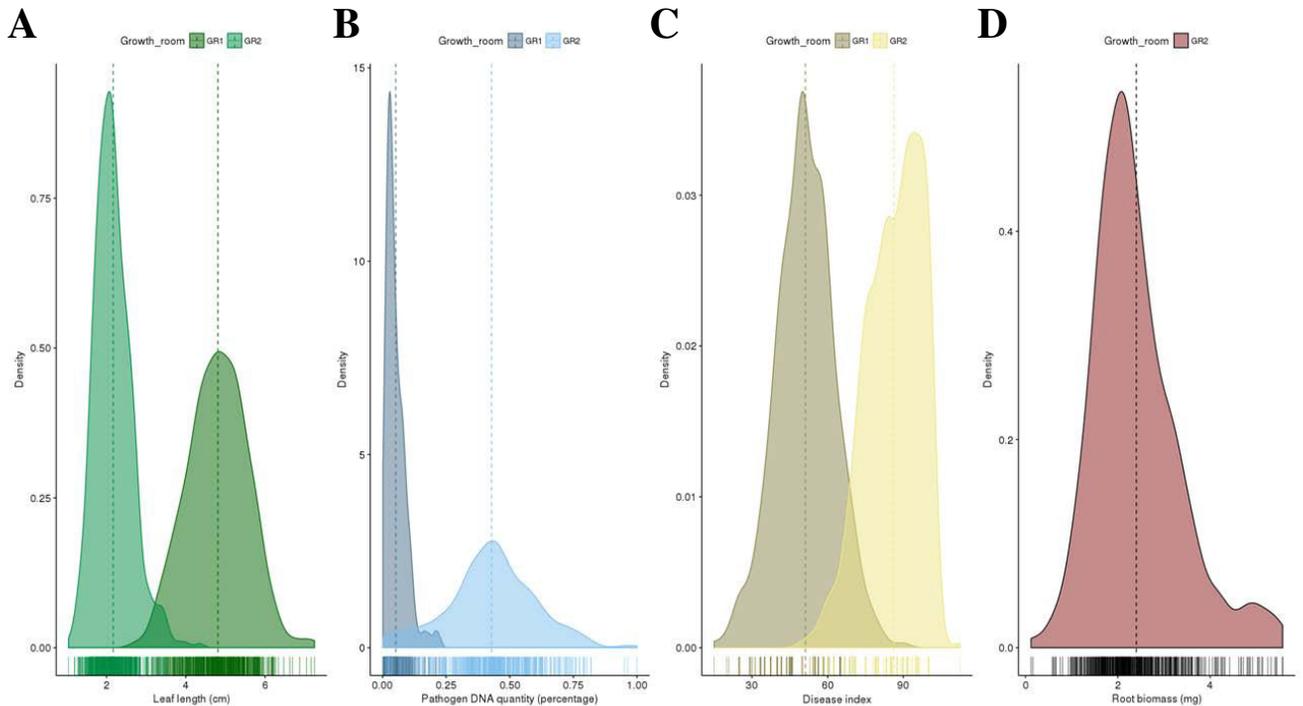


Figure III.2: Distribution of phenotypic data measured in the two growth rooms used for pathological tests. For each trait studied, data are coloured to show in which growth chamber the plants were grown. Vertical coloured dashed lines indicate means of traits in each growth room. Data distribution is shown for leaf length (A), pathogen DNA quantity (B), disease index (C) and root biomass (D). GR1 corresponds to growth room-1. GR2 corresponds to growth room-2.

DI: $\chi^2 = 2.69$, $p\text{-value} = 1.65\text{e-}4$; Pb: $\chi^2 = 3.75$, $p\text{-value} = 0.35$, Rbi: $\chi^2 = 28.05$, $p\text{-value} < 2.2\text{e-}16$). Broad-sense heritability (H^2) was estimated for each trait using the formula (1.1) and ranged from 0.46 to 0.76 in the growth room-1 and from 0.33 to 0.62 in the growth room-2 depending on the trait studied (**Table III.1**).

1.3.2.2. QTL analysis of data obtained in growth room-1

Phenotypic data measured on the two biological replicates set up in growth room-1 were used in the generalized linear model (1) to extract the epigenotype effect (G) with the lsmean function. The epigenotype effect G identified for each trait was then used for the QTL^{epi} analysis. In total, five QTL^{epi} were detected (**Figure III.3**). Two QTL^{epi} were identified for pathogen DNA quantity (Pb) on chromosomes 1 and 4 (Pb1^{epi}-At1, Pb1^{epi}-At4) explaining 14.64 and 9.65 % of the phenotypic variability, respectively. Three QTL^{epi} were detected for leaf length (Lfi) on chromosomes 1, 3 and 5 (Lfi1^{epi}-At1, Lfi1^{epi}-At3, Lfi1^{epi}-At5) explaining 15.49, 15.01 and 8.11 % of phenotypic variation, respectively (**Table III.2**). The variance explained by each fitted QTL model was of 43.82 and 19.59 % for Pb and Lfi, respectively. Surprisingly, no QTL^{epi} was identified for DI despite a significant effect of epigenotype on this trait. Confidence intervals of the QTL^{epi} detected ranged from 6.23 to 36 cM (**Table III.2**). No epistatic interaction was found between QTL^{epi} for either trait. For the three QTL^{epi} detected for Lfi (markers nearest of the peak LOD score: MM52, MM427 and MM728), WT epialleles were associated with an increase in the trait values. The *ddm1*-derived epiallele was associated with an increase in pathogen quantity at QTL^{epi} on chromosome 1 (peak marker: MM123) whereas it was associated with a decrease in the Pb value on chromosome 4 (peak marker MM550) (**Annexe I Figure S1, Table III.2**).

1.3.2.3. QTL analysis of data from growth room-2

As above, phenotypic data measured on the two biological replicates in growth room-2 were used in the generalized linear model (1) to extract the epigenotype effect with the lsmean function. Once again, the epigenotype effect G identified for each trait was then used for the QTL^{epi} analysis. In total, seven QTL^{epi} were detected (**Figure III.3**). One QTL^{epi} was detected on chromosome 4 for Pb (Pb2^{epi}-At4) and two QTL^{epi} on chromosomes 1 and 3 were detected for Lfi (Lfi2^{epi}-At1, Lfi2^{epi}-At3). QTL mapping also revealed four QTL^{epi} controlling the trait root biomass (Rbi), which was measured only in this growth room: one on chromosome 1, two on chromosome 2 and one on chromosome 4 (Rbi^{epi}-At1, Rbi^{epi}-At2a, Rbi^{epi}-At2b, Rbi^{epi}-At4). Again, no QTL^{epi} was identified for DI despite a significant effect of the epigenotype. The variance explained by each fitted QTL model was of 10.86, 25.61 and 44.59 % for Pb, Lfi and Rbi, respectively. The QTL^{epi} identified for Pb on chromosome 4 explained 10.86 % of the variability. The two QTL^{epi} detected on chromosomes 1 and 3 for Lfi explained respectively 14.97 and 12.61 % of the phenotypic variation (**Table III.2**). Confidence intervals of the detected QTL^{epi} ranged from 7.62 to 54.33 cM (**Table III.2**). No epistatic interaction was found between QTL^{epi} for Lfi

Table III.1. Heritability was calculated for each trait in the epiRIL population.

Trait	Heritability equation	H ² Growth room-1	H ² Growth room-2
DI	1	0.55	0.50
Pb	1	0.46	0.44
Lfi	1	0.76	0.33
Rbi	1	NA	0.62
DI	2	NA	0.50
Pb	2	NA	0.50
Lfi	2	NA	0.67
Rbi	2	NA	0.50

H² is the broad sense heritability calculated with the formula 1: $H^2 = \sigma^2_G / [\sigma^2_G + (\sigma^2_e/n)]$ and 2: $H^2 = \sigma^2_G / [\sigma^2_G + \sigma^2_{GT} + (\sigma^2_e/n)]$ including the variance of the temperature x epigenotype interaction. σ^2_G is the estimated epigenetic variance, σ^2_{GT} is the estimated temperature x epigenetic interaction variance, σ^2_e is the estimated environmental variance and n is the number of replicates per line. DI is the disease index. Pb is the pathogen DNA quantity. Lfi is the leaf length. Rbi is the Root biomass.

and Rbi traits. The additive allele effect of the two QTL^{epi} detected for Lfi (peak markers: MM91 and MM515) and of three of the four QTL^{epi} detected for Rbi (peak markers: MM147, MM383, MM686) was in the same direction: WT-derived epialleles were associated with an increase in trait value. For the Pb QTL (peak marker: MM693) and one of the four QTL^{epi} of Rbi (peak marker: MM385), the *ddm1*-derived epiallele was associated with a decrease in trait value (**Annexe I Figure S1, Table S2**).

For Lfi, QTL^{epi} detected on chromosomes 1 and 3 in growth room-1 (Lfi1^{epi}-At1, Lfi1^{epi}-At3) co-localized with the QTL^{epi} detected on chromosomes 1 and 3 in growth room-2 (Lfi2^{epi}-At1, Lfi2^{epi}-At3). The QTL^{epi} on chromosome 5 was only detected in growth room-1. The Pb QTL^{epi} detected in the two growth rooms were different. These results indicated that the QTL detection was dependent on the growth room conditions (**Figure III.3, Table III.2**).

1.3.3. *Temperature affects the plant response to P. Brassicae in the epiRIL population*

To explain the differences observed between the two growth rooms, we paid specific attention to the temperature conditions, as this parameter was shown to be critical for the development of clubroot disease (Sharma *et al.*, 2011; Siemens *et al.*, 2002). While similar values of global mean temperature were noted in both growth rooms (growth room-1 = 19.95 °C, growth room-2 = 20.06 °C), significant differences ($F = 2.67$, $p\text{-value} = 0.002$) in global temperature variances were registered between the two rooms (**Figure S2A**). We then analyzed mean and variance temperature values for each photoperiod (day and night). Day and night mean temperatures were similar in both chambers (growth room-1: day temperature = 20.98 °C, night temperature = 17.83 °C; growth room-2: day temperature = 20.85 °C, night temperature = 18.51 °C) and were very close to the required values (day temperature = 20 °C, night temperature = 18 °C). Conversely, the temperature variance in the two growth rooms and for both periods was significantly different (Fisher permutation test; day period: $F = 9.05$, $p\text{-value} = 0.002$, night period: $F = 3.67$, $p\text{-value} = 0.002$). The temperature range in growth room-1 was for the day period = 10.83 °C, and for the night period = 4.98 °C; the temperature range in growth room-2 was for the day period = 3.51 °C, and for the night period = 1.88 °C (**Figure S2B,C**).

1.3.4. *Effect of temperature variation on plant response to clubroot*

Based on these observations, a more detailed analysis was carried out in growth room-2 to evaluate the impact of the temperature variability on epiRIL response to clubroot. For this, we used data from temperature sensors placed at the height of the plants in growth room-2. Significant differences (Kruskal-Wallis test) in the median temperatures were observed ($\chi^2 = 4115.3$, $df = 15$, $p\text{-value} < 2.2e-16$) according to the location of the sensors. Pair-wise analysis (Dunnett's test) of the differences between temperatures measured by the sensors showed significant temperature variations at several positions in the growth room (**Annexe I Table S4**) and a temperature gradient from 21.4 to 24.1 °C (**Annexe I Table S5**).

Table III.2: Summary of QTL^{epi} detected in the epiRIL population by MQM.

Trait	Test Location	QTL	Chr	Pos	Lod	Closest peak marker	Confidence interval (cM)	R ² (%)	R ² for all QTL by trait	Favourable allele
G-Rbi	Growth room-2	Rb1 ^{epi} -At1	1	56.00	3.04	MM147	8.00	62.33	6.67	WT
	Growth room-2	Rb1 ^{epi} -At2a	2	32.00	7.48	MM383	28.00	35.62	17.91	WT
	Growth room-2	Rb1 ^{epi} -At2b	2	40.00	6.18	MM385	36.00	46.00	14.42	44.59 <i>ddm1-2</i>
	Growth room-2	Rb1 ^{epi} -At4	4	28.00	4.79	MM686	22.02	36.00	10.88	WT
G-Lfi	Growth room-1	Lfi1 ^{epi} -At1	1	33.60	6.51	MM52	20.00	44.00	15.50	WT
	Growth room-1	Lfi1 ^{epi} -At3	3	46.10	6.32	MM427	44.36	50.59	15.01	43.82 WT
	Growth room-1	Lfi1 ^{epi} -At5	5	37.80	3.60	MM728	16.00	52.00	8.11	WT
G-Pb	Growth room-2	Lfi2 ^{epi} -At1	1	37.04	4.90	MM91	30.73	60.00	14.97	25.61 WT
	Growth room-2	Lfi2 ^{epi} -At3	3	49.45	4.04	MM515	44.37	52.00	12.16	WT
	Growth room-1	Pb1 ^{epi} -At1	1	42.00	4.36	MM123	20.00	45.93	14.64	19.59 WT
	Growth room-1	Pb1 ^{epi} -At4	4	0.00	2.95	MM550	0.00	8.00	9.65	<i>ddm1-2</i>
GT-Lfi	Growth room-2	Pb2 ^{epi} -At4	4	42.99	3.05	MM693	32.00	50.00	10.86	10.86 <i>ddm1-2</i>
	Growth room-2	Lfi ^{epi} GT-At1	1	52.00	3.32	MM128	33.07	61.20	10.67	19.34 WT
	Growth room-2	Lfi ^{epi} GT-At3	3	49.45	2.47	MM547	6.00	82.00	7.81	WT
GT-Rbi	Growth room-2	Rb1 ^{epi} GT-At3	3	46.65	2.52	MM432	40.00	50.59	9.06	9.06 WT
	Growth room-2	Pb ^{epi} GT-At4a	4	22.00	2.38	MM679	1.73	30.00	8.04	15.89 <i>ddm1-2</i>
GT-Pb	Growth room-2	Pb ^{epi} GT-At4b	4	36.00	4.50	MM689	22.02	42.00	15.88	<i>ddm1-2</i>

Confidence intervals are in cM. Peak positions are indicated in cM and with the marker nearest to the LOD score peak. R²: phenotypic variation explained by the QTL^{epi}. Chr: chromosome. For Lfi and Rbi, favourable alleles translate an increase in the value. For Pb, favourable alleles translate a decrease in the value. G-Rbi, G-Lfi and G-Pb represent epigenetic QTL for each trait. GT - Rbi, GT - Lfi and GT Pb represent epigenetic x temperature QTL for each trait.

A significant temperature effect (glm, model 2) was shown for two of the phenotypic traits measured (Lfi: $\chi^2 = 10.14$, $p\text{-value} < 2.2\text{e-}16$; DI: $\chi^2 = 0.005$, $p\text{-value} = 0.54$; Pb: $\chi^2 = 0.01$, $p\text{-value} = 0.53$; Rbi: $\chi^2 = 2.98$, $p\text{-value} = 2.94\text{e-}09$). Moreover, a significant interaction between temperature and epigenotype effects (GT interaction) on the phenotypic traits measured was identified for all traits (Lfi: $\chi^2 = 23.58$, $p\text{-value} < 2.2\text{e-}16$; DI: $\chi^2 = 1.79$, $p\text{-value} = 0.44$; Pb: $\chi^2 = 4.20$, $p\text{-value} = 0.01$; Rbi: $\chi^2 = 14.51$, $p\text{-value} = 0.002$). Heritability (H^2) in the case of temperature and epigenotype interaction was estimated for each trait using the formula (2.1). Heritability ranged from 0.50 to 0.67 according to the dataset used (**Table III.1**).

1.3.5. Detection of epigenotype x temperature QTL^{epi}

In order to determine whether temperature variation could influence plant response to clubroot in the epiRIL population, we associated each epiRIL with the mean temperature data measured with the temperature sensor placed in the block where the epiRIL was grown. These data were used in the generalized linear model (2) to extract the interaction epigenotype x temperature (GT) using the lsmean function. The interaction GT values identified for each trait were then used for the QTL^{epi} analysis. Two QTL^{epiGT} (epigenotype x temperature QTL^{epi}) were detected on the chromosomes 1 and 3 for Lfi (Lfi^{epiGT}-At1, Lfi^{epiGT}-At3), and one QTL^{epiGT} on chromosome 3 was detected for Rbi (Rbi^{epiGT}-At3). No QTL^{epi} was detected for Pb at a 5 % significance level using the SIM by stepwise approach but two QTL^{epi} were detected on the chromosome 4 using a 10 % significance level (Pb^{epiGT}-At4a, Pb^{epiGT}-At4b). Again, no QTL^{epi} could be detected for DI (**Figure III.3**). The variance explained by the fitted QTL model was 19.34 % for Lfi, 9.06 % for Rbi and 15.89 % for Pb. The QTL^{epi} found for Lfi on chromosomes 1 and 3 explained 10.67 % and 7.81 % of the variability. The QTL^{epi} found for Rbi explained 9.06 % of phenotypic variation. The two QTL^{epi} found for Pb on chromosome 4 explained 8.04 and 15.88 % of the variation (**Table III.2**). Confidence intervals of the detected QTL^{epi} ranged from 10.59 to 76 cM (**Table III.2**). No epistatic interaction was found between QTL^{epi} for Lfi and Pb traits. For all QTL^{epi} detected, the wild-type epi-alleles were associated with an increase in the values (**Annexe I Figure S1, Table III.2**). Comparison of the QTL^{epi} and QTL^{epiGT} (taking into account the interaction with temperature) showed that all QTL^{epiGT} co-localized totally or partially with QTL^{epi} detected in at least one growth room (**Figure III.3**). Indeed, the comparison of the confidence intervals of QTL^{epi} detected using the leaf length trait showed that the QTL^{epi} Lfi1^{epi}-At1 and Lfi2^{epi}-At1 overlapped with the QTL^{epiGT} Lfi^{epiGT}-At1. Similarly, Lfi1^{epi}-At3 and Lfi2^{epi}-At3 confidence intervals co-localized with the confidence interval of Lfi^{epiGT}-At3 (**Figure III.3 and Table III.2**). For the QTL^{epi} detected using the quantification of the pathogen DNA (Pb) two overlaps were identified on chromosome 4, one between Pb1^{epi}-At4 and Pb^{epiGT}-At4a in the region extending from the short arm to the pericentromeric region and another between Pb2^{epi}-At4, Pb^{epiGT}-At4a and Pb^{epiGT}-At4b on the long arm of chromosome 4 (**Figure III.3 and Table III.2**).

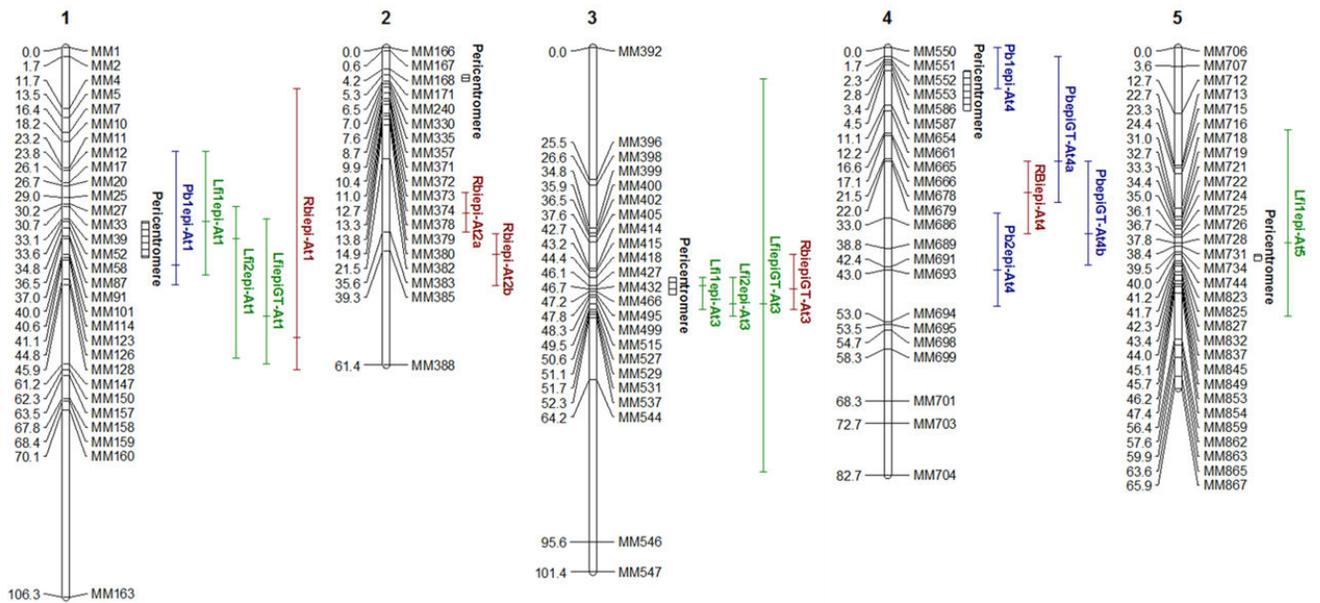


Figure III.3: Epigenetic QTL identified in the epiRIL population in response to *Plasmodiophora brassicae* infection. Numbers above chromosomes indicate *Arabidopsis* chromosomes. MM indicates markers. For each marker, position in cM is indicated. The vertical bar length is equal to the one-LOD likelihood confidence interval. The dash in the bar indicates the peak position. QTL name indicates the trait associated and chromosomal localisation.

1.3.6. Impact of DNA sequence variation within QTL^{epi} confidence intervals

Although the epiRIL population was designed to minimize as much as possible DNA sequence variation, the presence of a small number of segregating nucleotidic variants cannot be avoided, notably as a result of the known effect of loss of DNA methylation on TE activity. In order to investigate the potential contribution of parental DNA sequence variants to the differential susceptibility to *P. Brassicae* associated with the 17 QTL^{epi}, we identified using whole genome sequencing data all sequence variants shared by more 25 % of the epiRIL. In total, 63 small scale and 11 TE insertion variants were located within the 17 QTL^{epi}, respectively. Eleven shared insertions of TEs were detected in the QTL^{epi} confidence intervals, with 15 QTL^{epi} showing at least one insertion and two QTL^{epi} showing no insertion. All QTL^{epi} detected included at least one small-scale sequence polymorphism in their confidence interval (**Annexe I File S1**).

Effects of the shared TEs and sequence variants were tested as described in Kooke *et al.* (2015). For 14 of the 17 QTL^{epi}, the effect of the epigenetic marker at the QTL peak was more significant than the effect of the TE or the small-scale sequence variant (**Annexe I File S3**). However, for Lfi^{epi}GT-At1, Lfi^{epi}GT-At3 and Lfi^{epi}-At5, the significant effect observed for a SNP or a TE was greater than for the epigenetic marker at the QTL peak (**Annexe I File S3**). In this case, we considered that these three QTL^{epi} were actually caused by DNA sequence variants. However, a linkage disequilibrium between the significant genetic markers and a causal epiallele could also be possible.

1.4. Discussion

The aim of the present study was to investigate the role of epigenetic modifications in the *Arabidopsis* response to *P. Brassicae*. To determine whether epigenetic regulation does play a role in clubroot quantitative resistance, a reverse genetics approach, using six T-DNA insertional mutants (Col-0 background) in genes involved in epigenetic pathways, was first carried out. For the six mutants tested, the DI was only significantly reduced in the *ddm1* mutant compared to Col-0. This finding suggests that *ddm1* confers decreased susceptibility of *Arabidopsis* to *P. Brassicae*. This observation is in agreement with previous results obtained by Kellenberger *et al.* (2016) and Sharma *et al.* (2017) in the *Brassica* genus, which linked plant responses to biotic and abiotic stress to epigenetic regulations. Moreover, the involvement of *ddm1* in the *Arabidopsis*-*P. Brassicae* interaction further supports the hypothesis that DNA methylation plays a role in plant pathogen infections (Dowen *et al.*, 2012; López Sánchez *et al.*, 2016; Hewezi *et al.*, 2017). To decipher the epigenetic architecture of the clubroot resistance in *Arabidopsis*, we then carried out a QTL^{epi} detection experiment using the epiRIL population. Four disease-related traits (disease index, pathogen DNA quantity, leaf length and root biomass), mostly previously used in *Arabidopsis* (Jubault *et al.*, 2008; unpublished) to evaluate plant response to clubroot, were monitored. These traits were chosen in order to characterize disease development (disease index

and pathogen DNA quantity) and the consequences of *P. Brassicae* infection on plant development (root biomass and leaf length).

1.4.1. Epigenetic QTL control *Arabidopsis clubroot* resistance

From our experiments using four biological replicates, we have shown a significant epigenetic effect on the response to *P. Brassicae* infection. Thus, heritable plant responses induced by DNA methylation appear to be involved in the plant response to *P. Brassicae*. The moderate to high heritability values (from 0.33 to 0.76) observed were similar to those described in control and abiotic stress conditions for this population (Johannes *et al.*, 2009; Colomé-Tatché *et al.*, 2012; Kooke *et al.*, 2015). Moreover, these heritability values were also similar to those described by Jubault *et al.* (2008) on a RIL population infected with *P. Brassicae*. The moderate heritability levels calculated for the traits mean that this epigenetic variability could be considered for use in breeding. As the plant response to *P. Brassicae* varied depending on the growth room used, we carried out the QTL^{epi} detection experiments using data from each growth room. In total, fourteen additive QTL^{epi} grouped in six genomic regions were identified distributed throughout four *A. thaliana* chromosomes. For three of the 17 QTL^{epi} initially detected (Lfi^{epi}GT-At1, Lfi^{epi}GT-At3 and Lfi^{epi}-At5), analysis of shared TE and sequence variants included in their confidence intervals showed that the effect of DNA-based markers was greater than the effect of epigenetic markers, suggesting these QTL are not *bona fide* epigenetic. Among the four phenotypic traits evaluated, no QTL^{epi} was detected for DI despite a significant epigenotype effect in the two growth rooms (p -value = 1.57e-06 and 1.65e-04). Results of the reverse genetics approach showed that *ddm1* was less susceptible to *P. Brassicae* compared to Col-0 suggesting that *ddm1* epialleles confer a reduction in symptoms. In this context, the absence of QTL^{epi} detected for the DI trait may be explained by i) a sampling effect, since only a subset (123 lines out of 505) of the epiRIL population was phenotyped in this study and/or ii) a low proportion of *ddm1* epi-haplotypes in the population subset (27 %), which may have been insufficient to detect small QTL effects (Holland, 2007). Most of the QTL^{epi} confidence intervals detected in this study co-localized with the pericentromeric regions. This may be explained by a more stable loss of methylation in those regions due to the loss of methylation maintenance on repeats and TE in *ddm1* (Kooke *et al.* 2015). Concerning the effect of the epialleles, the majority of the wild-type epialleles were associated with an increase in the morphological trait values (length of leaves and root biomass). However, the *ddm1*-derived epiallele led to a decrease in the amount of pathogen DNA for two of the three QTL^{epi} detected on chromosomes 1 and 4. This finding highlighted the positive effect of this epiallele on plant resistance, in agreement with the results obtained in the mutant test of this study and by Downen *et al.* (2012) who showed a modest increase in *ddm1* resistance to *P. syringae*.

1.4.2. Stability and pleiotropy of clubroot epigenetic QTL

Several QTL detected were stable across growth rooms. Indeed, despite the temperature variations between the two growth rooms, two overlapping leaf length QTL^{epi} (growth room-1: Lfi1^{epi}-At1 and Lfi1^{epi}-At3, growth room-2: Lfi2^{epi}-At1 and Lfi2^{epi}-At3) were detected in both growth rooms. Furthermore, for Pb, the QTL^{epi} Pb1^{epi}-At4, detected in the growth room-1 overlapped with the QTL^{epi} Pb^{epi}GT-At4a detected in the growth room-2 (**Figure III.3 and Table III.2**). The co-localization on chromosome 1 of four QTL^{epi} controlling 3 different traits (Lfi, Pb and Rbi) and on chromosome 4 of three QTL^{epi} controlling two traits (Pb and Rbi) may suggest the presence of pleiotropic QTL^{epi} (**Figure III.3 and Table III.2**). However, analysis of the correlation (Spearman rho correlation) between traits showed only a moderate correlation ($Rho = 0.52$, $p\text{-value} < 2.2e-16$) between Lfi and Rbi. Fine mapping is necessary to overcome the bias imposed by the large size of the QTL^{epi} confidence intervals and make further conclusions on the possibility of a pleiotropic gene or an effect due to linked genes.

1.4.3. Temperature modulates *Arabidopsis* clubroot responses

Taking into consideration the temperature variations in growth room-2, we tested the hypothesis that temperature plays a potential role in epiRIL and Col-0 clubroot symptom variations. These observations are in agreement with the variations in clubroot severity observed on *B. rapa* subsp. *chinensis* (Sharma *et al.*, 2011) and in Col-0 and the *Arabidopsis* mutant *tu8* (Siemens *et al.*, 2002) according to the temperature used for growing the plants. Similar environmental effects on clubroot response have also been shown with nitrogen supply variations in *Brassica napus* (Laperche *et al.*, 2017; Aigu *et al.*, 2018) and with flooding in *Arabidopsis* (Gravot *et al.*, 2016). Here, our analyses suggested that the temperature effect was partly triggered by the interaction with the plant epigenome for the traits root biomass and pathogen DNA quantity. The identification of temperature-dependent QTL^{epi} controlling pathogen quantity (Pb^{epi}GT-At4a and Pb^{epi}GT-At4b) suggests the presence of QTL^{epi} involved in the control of the pathogen development under temperature dependence. Similar observation concerning the detection of QTL temperature dependant was carried out by Aoun *et al.* (2017) using the model *Arabidopsis* - *Ralstonia solanacearum*. In their study, Aoun *et al.* (2017) showed that an increase of the temperature of 3 °C during the interaction between *Arabidopsis* and *R. solanacearum* lead to an increase of the sensibility of most accessions and the loss of detection of one major QTL associated with the resistance. The influence of the temperature on the epiRIL response to *P. Brassicae* could be explained by the increase of the environmental sensitivity triggered by the DNA hypomethylation suggested by Kooke *et al.* (2015).

1.4.4. Clubroot resistance, a sophisticated system of regulation involving genetics and epigenetics

Interestingly, the comparison of the clubroot genetic QTL previously identified (Jubault *et al.*, 2008; unpublished) in *Arabidopsis* and the clubroot epigenetic QTL identified in this study highlighted co-localization of some confidence intervals. Indeed, the confidence intervals of two QTL^{epi} (Rbi^{epi}-At1 and Pb2^{epi}-At4) overlapped with two QTL (Pb-At1 and Pb-At4) found in the Bur-0 x Col-0 RIL population by linkage analysis (Jubault *et al.*, 2008). Moreover, significant SNP identified by GWAS were localized in the confidence intervals of 13 QTL^{epi} (**Annexe I File S2**, unpublished data). These multiple co-localizations suggest that quantitative resistance to clubroot is modulated by a system involving both nucleotidic and epigenetic variations. These results illustrate the fact that in classical populations used for QTL detection, dissociation between causal genetic and epigenetic variations, both in linkage disequilibrium with markers, is nearly impossible (Schmitz *et al.*, 2013a, b). In addition, our findings are consistent with the suggestion that during plant pathogen interactions, plant transgenerational changes in genome structure and in DNA methylation patterns are possible (Boyko & Kovalchuk, 2011). The identification of two QTL^{epi} (Rbi^{epi}-At2 and Pb1^{epi}-At4) that did not show any overlap with previously reported QTL for these traits suggests that at these loci only epigenetic variation may be responsible for plant response variation. However, an absence of co-localization could also be due to the absence of nucleotidic and/or epigenetic variations at these loci in the previously studied populations but does not exclude their existence in other genotypes.

This first study on the role of epialleles in the *Arabidopsis* - *P. Brassicae* interaction brings to light the possibility of a complex model of quantitative resistance where alleles and epi-alleles act in concert. Furthermore, our study has shown that the temperature variations could influence epiRIL response to *P. Brassicae*. In order to confirm whether epialleles are involved in plant response to clubroot infection, the QTL^{epi} confidence intervals must be reduced to find causal epialleles, notably through a fine-mapping approach. The assessment of epiRIL in pathological tests in contrasting controlled temperature conditions is also needed to validate the possible role of temperature in modulating the epigenetic plant response to clubroot infection.

1.5. References

- Aigu Y, Laperche A, Mendes J, Lariagon C, Guichard S, Gravot A, Manzanares-Dauleux MJ. 2018.** Nitrogen supply exerts a major/minor switch between two QTLs controlling *Plasmodiophora brassicae* spore content in rapeseed. *Plant pathology*.
- Alix K., Lariagon C., Delourme R., Manzanares-Dauleux M.J. 2007.** Exploiting natural genetic diversity and mutant resources of *Arabidopsis thaliana* to study the *A. thaliana*-*Plasmodiophora brassicae* interaction. *Plant Breed* **126**: 218–221

- Aller EST, Jagd LM, Kliebenstein DJ, Burow M. 2018.** Comparison of the Relative Potential for Epigenetic and Genetic Variation To Contribute to Trait Stability. *G3* **8**: 1733–1746.
- Aoun N, Tauleigne L, Lonjon F, Deslandes L, Vaillau F, Roux F, Berthomé R. 2017.** Quantitative disease resistance under elevated temperature: genetic basis of new resistance mechanisms to *Ralstonia solanacearum*. *Frontiers in plant science* **8**: 1387.
- Becker C, Haggmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. 2011.** Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480**: 245–249.
- Bourc'his D, Bestor TH. 2002.** Helicase homologues maintain cytosine methylation in plants and mammals. *BioEssays: news and reviews in molecular, cellular and developmental biology* **24**: 297–299.
- Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR, Görlach J. 2001.** Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *The Plant cell* **13**: 1499–1510.
- Boyko A, Kovalchuk I. 2011.** Genetic and epigenetic effects of plant–pathogen interactions: an evolutionary perspective. *Molecular plant* **4**: 1014–1023.
- Broman KW, Wu H, Sen S, Churchill GA. 2003.** R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**:889-890
- Broman KW, Sen S. 2009.** Fit and exploration of multiple-QTL models. *In: Statistics for Biology and Health*. 241–282.
- Colomé-Tatché M, Cortijo S, Wardenaar R, Morgado L, Lahouze B, Sarazin A, Etcheverry M, Martin A, Feng S, Duvernois-Berthet E, et al. 2012.** Features of the *Arabidopsis* recombination landscape resulting from the combined loss of sequence variation and DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 16240–16245.
- Cortijo S, Wardenaar R, Colomé-Tatché M, Gilly A, Etcheverry M, Labadie K, Caillieux E, Hospital F, Aury J-M, Wincker P, et al. 2014.** Mapping the epigenetic basis of complex traits. *Science* **343**: 1145–1148.
- Diederichsen E, Frauen M, Linders EGA, Hatakeyama K, Hirai M. 2009.** Status and perspectives of clubroot resistance breeding in Crucifer crops. *Journal of Plant Growth Regulation* **28**: 265–281.
- Dixon GR. 2009.** The Occurrence and Economic Impact of *Plasmodiophora brassicae* and Clubroot Disease. *Journal of plant growth regulation* **28**: 194–202.

- Downen RH, Pelizzola M, Schmitz RJ, Lister R, Downen JM, Nery JR, Dixon JE, Ecker JR. 2012.** Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences* **109**: E2183–E2191.
- Fahling M, Graf H, Siemens J. 2003.** Pathotype Separation of *Plasmodiophora brassicae* by the Host Plant. *Journal of Phytopathology* **151**: 425–430.
- Gallusci P, Dai Z, Génard M, Gauffretau A, Leblanc-Fournier N, Richard-Molard C, Vile D, Brunel-Muguet S. 2017.** Epigenetics for Plant Improvement: Current Knowledge and Modeling Avenues. *Trends in plant science* **22**: 610–623.
- Gilly A, Etcheverry M, Madoui M-A, Guy J, Quadrana L, Alberti A, Martin A, Heitkam T, Engelen S, Labadie K, et al. 2014.** TE-Tracker: systematic identification of transposition events through whole-genome resequencing. *BMC bioinformatics* **15**: 377.
- Gravot A, Richard G, Lime T, Lemarié S, Jubault M, Lariagon C, Lemoine J, Vicente J, Robert-Seilaniantz A, Holdsworth MJ, et al. 2016.** Hypoxia response in *Arabidopsis* roots infected by *Plasmodiophora brassicae* supports the development of clubroot. *BMC plant biology* **16**: 251.
- Hewezi T, Lane T, Piya S, Rambani A, Rice JH, Staton M. 2017.** Cyst Nematode Parasitism Induces Dynamic Changes in the Root Epigenome. *Plant physiology* **174**: 405–420.
- Holland J. 2007.** Genetic architecture of complex traits in plants. *Current opinion in plant biology* **10**: 156–161.
- Ingram DS, Tommerup IC. 1972.** The Life History of *Plasmodiophora brassicae* Woron. *Proceedings of the Royal Society B: Biological Sciences* **180**: 103–112.
- Johannes F, Colot V, Jansen RC. 2008.** Epigenome dynamics: a quantitative genetics perspective. *Nature reviews. Genetics* **9**: 883–890.
- 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. *PLoS genetics* **5**: e1000530.
- Jubault M, Lariagon C, Simon M, Delourme R, Manzanares-Dauleux MJ. 2008.** Identification of quantitative trait loci controlling partial clubroot resistance in new mapping populations of *Arabidopsis thaliana*. *Theoretical and applied genetics*. **117**: 191–202.
- Kellenberger RT, Schlüter PM, Schiestl FP. 2016.** Herbivore-induced DNA demethylation changes floral signalling and attractiveness to pollinators in *Brassica rapa*. *Plos One* **11**: e0166646.

Koch E, Cox R, Williams PH. 1991. Infection of *Arabidopsis thaliana* by *Plasmodiophora brassicae*. *Journal of Phytopathology* **132**: 99–104.

Kooke R, Johannes F, Wardenaar R, Becker F, Etcheverry M, Colot V, Vreugdenhil D, Keurentjes JJB. 2015. Epigenetic basis of morphological variation and phenotypic plasticity in *Arabidopsis thaliana*. *The Plant cell* **27**: 337–348.

Laperche A, Aigu Y, Jubault M, Ollier M, Guichard S, Glory P, Strelkov SE, Gravot A, Manzanares-Dauleux MJ. 2017. Clubroot resistance QTL are modulated by nitrogen input in *Brassica napus*. *Theoretical and applied genetics*. **130**: 669–684.

Lee J, Izzah NK, Choi B-S, Joh HJ, Lee S-C, Perumal S, Seo J, Ahn K, Jo EJ, Choi GJ, et al. 2016. Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA research: an international journal for rapid publication of reports on genes and genomes* **23**: 29–41.

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.

Liu J, Feng L, Li J, He Z. 2015. Genetic and epigenetic control of plant heat responses. *Frontiers in plant science* **06**.

López Sánchez A, Stassen JHM, Furci L, Smith LM, Ton J. 2016. The role of DNA (de)methylation in immune responsiveness of *Arabidopsis*. *The Plant journal: for cell and molecular biology* **88**: 361–374.

Luna E, Bruce TJA, Roberts MR, Flors V, Ton J. 2012. Next-generation systemic acquired resistance. *Plant physiology* **158**: 844–853.

Manzanares-Dauleux MJ, Delourme R, Baron F, Thomas G. 2000a. Mapping of one major gene and of QTLs involved in resistance to clubroot in *Brassica napus*. *Theoretical and applied genetics*. **101**: 885–891.

Manzanares-Dauleux MJ, Divaret I, Baron F, Thomas G. 2000b. Evaluation of French *Brassica oleracea* landraces for resistance to *Plasmodiophora brassicae*. *Euphytica* **113**(3): 211-218.

Marschner IC. 2011. glm2: Fitting generalized linear models with convergence problems. *The R Journal* **3**(2): 12-15

- Piao Z, Ramchiary N, Lim YP. 2009.** Genetics of Clubroot Resistance in *Brassica* Species. *Journal of plant growth regulation* **28**: 252–264.
- Quadrana L, Colot V. 2016.** Plant Transgenerational Epigenetics. *Annual review of genetics* **50**: 467–491.
- R Core Team 2015.** *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reinders J, Wulff BBH, Mirouze M, Mari-Ordóñez 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.
- Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, et al. 2017.** Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology letters* **20**: 1576–1590.
- Rocherieux J, Glory P, Giboulot A, Boury S, Barbeyron G, Thomas G, Manzanares-Dauleux MJ. 2004.** Isolate-specific and broad-spectrum QTLs are involved in the control of clubroot in *Brassica oleracea*. *Theoretical and applied genetics*. **108**: 1555–1563.
- Russell V. Lenth 2016.** Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software* **69**(1): 1–33.
- 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* **334**: 369–373.
- Schmitz RJ, He Y, Valdes-Lopez O, Khan SM, Joshi T, Urich MA, Nery JR, Diers B, Xu D, Stacey G, et al. 2013a.** Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome research* **23**: 1663–1674.
- Schmitz RJ, Schultz MD, Urich MA, Nery JR, Pelizzola M, Libiger O, Alix A, McCosh RB, Chen H, Schork NJ, et al. 2013b.** Patterns of population epigenomic diversity. *Nature* **495**: 193–198.
- Sharma K, Gossen BD, McDonald MR. 2011.** Effect of temperature on cortical infection by *Plasmodiophora brassicae* and clubroot severity. *Phytopathology* **101**: 1424–1432.
- Sharma R, Vishal P, Kaul S, Dhar MK. 2017.** Epiallelic changes in known stress-responsive genes under extreme drought conditions in *Brassica juncea* (L.) Czern. *Plant cell reports* **36**: 203–217.

- Siemens J, Nagel M, Ludwig-Muller J, Sacristan MD. 2002.** The interaction of *Plasmodiophora brassicae* and *Arabidopsis thaliana*: Parameters for disease quantification and screening of mutant lines. *Journal of Phytopathology* **150**: 592–605.
- Some A, Manzanares MJ, Laurens F, Baron F, Thomas G, Rouxel F. 1996.** Variation for virulence on *Brassica napus* L. amongst *Plasmodiophora brassicae* Collections from France and derived single-spore isolates. *Plant pathology* **45**: 432–439.
- Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, Ciaudo C, Cruaud C, Poulain J, Berdasco M, Fraga MF, et al. 2009.** A role for RNAi in the selective correction of DNA methylation defects. *Science* **323**: 1600–1604.
- Tierney L, Rossini AJ, Li Na, Sevcikova H. 2015.** *Snow: Simple Network of Workstations*. R package version 0.4-1. <http://CRAN.R-project.org/package=snow>
- Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. 2009.** Bursts of retrotransposition reproduced in *Arabidopsis*. *Nature* **461**: 423–426.
- Vongs A, Kakutani T, Martienssen RA, Richards EJ. 1993.** *Arabidopsis thaliana* DNA methylation mutants. *Science* **260**: 1926–1928.
- Weigel D, Colot V. 2012.** Epialleles in plant evolution. *Genome biology* **13**: 249.
- Yu A, Lepère G, Jay F, Wang J, Bapaume L, Wang Y, Abraham A-L, Penterman J, Fischer RL, Voinnet O, et al. 2013.** Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *Proceedings of the National Academy of Sciences* **110**: 2389–2394.
- Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. 2013.** The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**: 193–205.
- Zhang Y-Y, Fischer M, Colot V, Bossdorf O. 2013.** Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *The New phytologist* **197**: 314–322.
- Zheng X, Chen L, Xia H, Wei H, Lou Q, Li M, Li T, Luo L. 2017.** Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition. *Scientific reports* **7**: 39843.

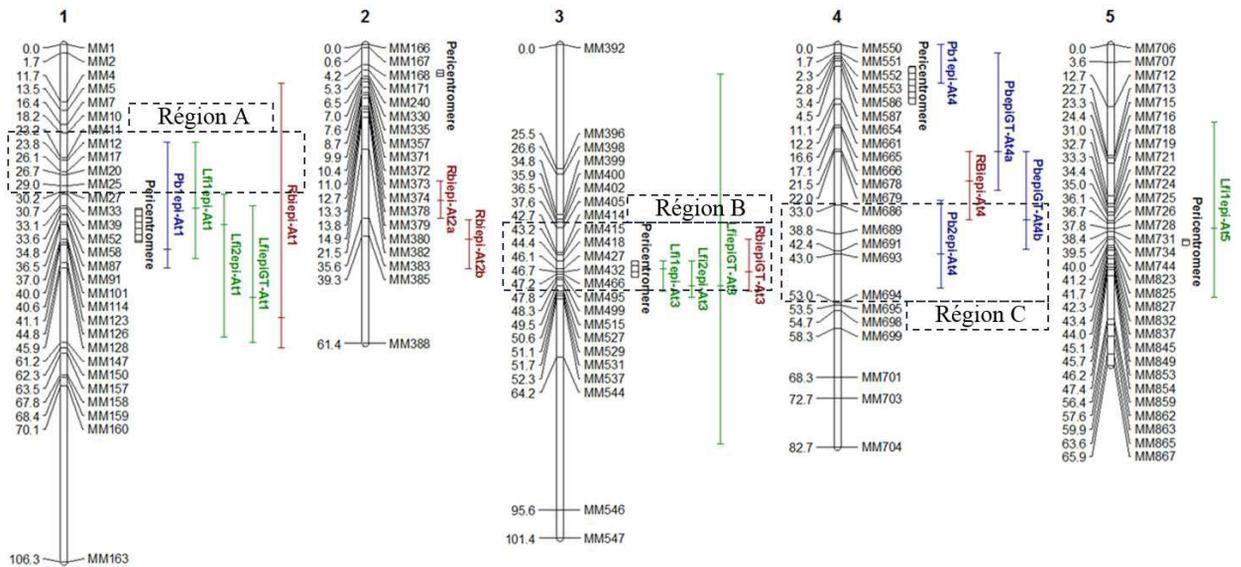


Figure III.4. Localisation des régions 3 sélectionnées. Chaque région sélectionnée est délimitée par un rectangle noir en pointillés. Figure issue de Liégard *et al.* (2018).

Tableau III.3. Intervalles des régions étudiées et marqueurs utilisés. (A). Délimitation des régions étudiées. Les marqueurs de début et de fin de région correspondent aux DMR (marqueurs) extérieurs de la région utilisées pour l'épigénotypage des 505 épiRIL. (B). Liste des 4 DMR au niveau de chaque région utilisées pour identifier les épiRIL recombinantes au niveau de chacune d'elles. Chaque DMR utilisée provient de la liste des DMR identifiées par Colomé-Tatché *et al.* (2012).

A

Région	Position	Marqueur	Position physique (pb)	Largeur de la région en pb
Région A	Début	MM10	8931527	3478393
	Fin	MM27	12409920	
Région B	Début	MM402	9435169	6745021
	Fin	MM527	16180190	
Région C	Début	MM686	8313712	2215367
	Fin	MM694	10529079	

B

	Région A	Région B	Région C
DMR utilisées comme marqueurs	MM10	MM402	MM686
	MM17	MM415	MM689
	MM20	MM466	MM691
	MM27	MM527	MM694

2. Réduction de l'intervalle de confiance de 3 régions génomiques contenant plusieurs QTL^{épi}

Grâce à l'approche d'épigénétique quantitative nous avons montré que 14 QTL sous contrôle épigénétique (QTL^{épi}) sont impliqués dans la réponse d'*Arabidopsis* à la hernie. Ces derniers sont localisés au niveau de 6 régions génomiques et, pour certains, leur effet est dépendant de la température. Cependant les QTL^{épi} détectés présentent des intervalles de confiance de taille importante (de 1,9 Mb à 22 Mb) rendant l'identification des gènes causaux et des DMR associées difficile. Nous avons donc initié une approche de cartographie fine afin de réduire les intervalles de confiance de QTL^{épi} présents au niveau de 3 des 6 régions détectées, soit 8 QTL^{épi}. Les analyses réalisées dans cette approche ont fait l'objet au sein de notre équipe du stage de M2 de Juliette Bénéjam que j'ai co-encadré.

2.1. Choix des QTL^{épi}

Les trois régions sélectionnées pour la cartographie fine sont (**Figure III.4** et **Tableau III.3A**) :

- La région A, allant du marqueur MM10 au marqueur MM27 (3,48 Mb) regroupant une partie des intervalles de confiance des QTL^{épi} *Pb1^{épi}-At1*, *Lfi1^{épi}-At1* et *Rbi^{épi}-At1* sur le chromosome 1. Cette région a été sélectionnée car elle co-localise également avec un facteur majeur de résistance à la hernie appelé *RPBI* (Arbeiter *et al.*, 2002) identifié vis-à-vis des isolats e et e3 de *P. Brassicae* et un QTL de résistance à la hernie identifié par GWAS dans le laboratoire vis-à-vis de l'isolat eH de *P. Brassicae* (Jubault *et al.*, in prep).
- La région B, allant du marqueur MM402 au marqueur MM527 (6,74 Mb) regroupant une partie ou la totalité des intervalles de confiance des QTL^{épi} *Lfi1^{épi}-At3*, *Lfi2^{épi}-At3*, *Lfi1^{épi}GT-At3* et *Rbi^{épi}GT-At3*, sur le chromosome 3. Les pics aux QTL de l'ensemble des QTL^{épi} sélectionnés étant proches mais non identiques, l'analyse de cette région s'avère intéressante afin de savoir si la causalité est commune aux différents QTL^{épi} (locus pléiotropique) ou non.
- La région C, allant du marqueur MM686 au marqueur MM694 (2,21 Mb) contient l'intervalle de confiance du QTL^{épi} *Pb2^{épi}-At4* et une petite partie de l'intervalle de confiance du QTL^{épi} *Rbi^{épi}-At4*, sur le chromosome 4. Le QTL^{épi} *Pb2^{épi}-At4* situé en région non péricentromérique était intéressant à étudier car il présente un épiallèle favorable (augmentation du niveau de résistance) de type *ddm1-2* alors que les autres QTL^{épi} sélectionnés présentent des épiallèles favorables de type WT.

Afin de diminuer les intervalles de confiance de ces QTL^{épi} et d'identifier les gènes candidats/loci sous contrôle épigénétique sous-jacents aux QTL^{épi} deux approches ont été effectuées i) la recherche d'individus recombinants supplémentaires dans les intervalles de confiance et ii) l'augmentation du nombre de marqueurs dans les régions sélectionnées.

Tableau III.4. Sélection des épiRIL recombinants. Liste des 97 épiRIL identifiées comme recombinantes parmi les 505 lignées à au moins l'une des trois régions sélectionnées. En vert les individus présentant une recombinaison dans au moins deux des trois régions étudiées.

EpiRILs recombinants													
epi4	epi37	epi58	epi71	epi107	epi122	epi152	epi174	epi202	epi229	epi305	epi371	epi443	epi523
epi8	epi45	epi60	epi74	epi112	epi129	epi163	epi177	epi206	epi230	epi323	epi373	epi454	epi556
epi9	epi46	epi62	epi79	epi114	epi137	epi165	epi183	epi208	epi247	epi328	epi391	epi466	epi558
epi17	epi48	epi64	epi80	epi115	epi147	epi166	epi187	epi211	epi262	epi333	epi419	epi471	epi561
epi23	epi49	epi65	epi92	epi118	epi148	epi167	epi195	epi212	epi274	epi340	epi424	epi492	epi570
epi30	epi51	epi68	epi97	epi119	epi150	epi169	epi197	epi213	epi277	epi363	epi432	epi497	epi573
epi36	epi55	epi70	epi99	epi121	epi151	epi172	epi201	epi222	epi292	epi369	epi437	epi499	

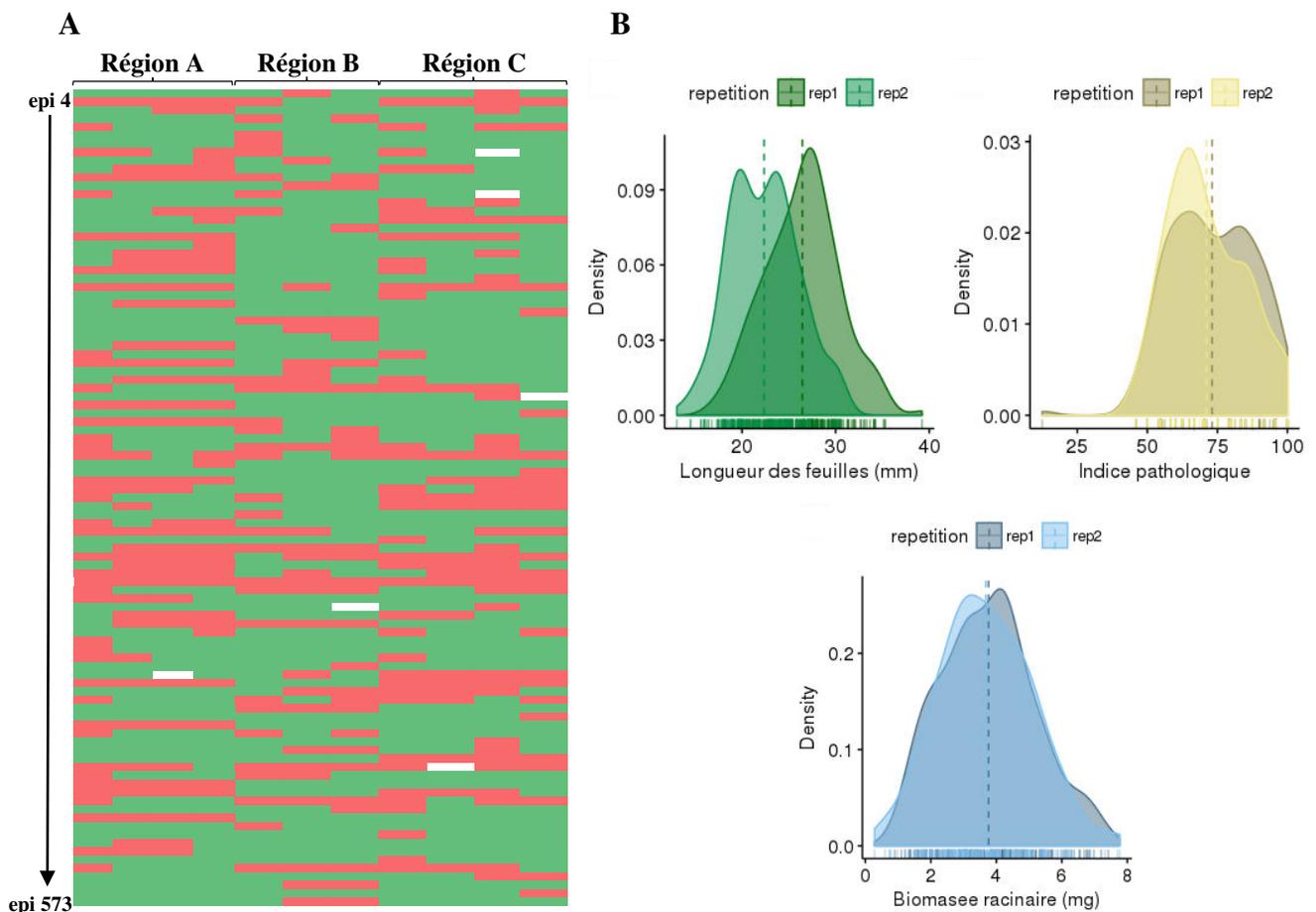


Figure III.5. Sélection des épiRIL recombinants, épihaplotypes et distribution des traits phénotypiques étudiés. (A). Épihaplotype des 97 épiRIL recombinantes au niveau des 12 DMR utilisées comme marqueur dans les trois régions sélectionnées. Les 97 épiRIL sont représentées sur l'axe y. L'épigénotype déterminé au niveau de chaque marqueur le long des régions sélectionnées est représenté selon l'axe x. En vert l'épigénotype de type sauvage. En rouge l'épigénotype de type *ddm1-2*. Les données manquantes sont représentées en blanc. (B). Distribution des traits phénotypiques mesurés dans les 2 répétitions biologiques lors des tests pathologiques. Pour chaque trait étudié, les données sont colorées en fonction de la répétition biologique. La ligne verticale pointillée indique la moyenne du trait dans chaque répétition. La distribution des données est montrée pour la longueur des feuilles (vert), l'indice pathologique (marron) et la biomasse racinaire (bleu). rep1 correspond à la répétition biologique numéro 1. rep 2 correspond à la répétition biologique 2.

2.2. Identification d'individus recombinants supplémentaires au niveau des QTL^{épi} sélectionnés

Notre étude initiale était basée sur le sous-ensemble de 123 épiRIL utilisé pour la construction de la carte épigénétique par Colomé-Tatché *et al.* (2012). Afin d'affiner la localisation de QTL^{épi} présents au niveau de 3 des 6 régions détectées, nous avons recherché de nouveaux recombinants dans ces régions en utilisant la population épiRIL entière créée par Johannes *et al.* (2009), soit 505 individus.

Afin d'identifier des lignées recombinantes supplémentaires au niveau des régions sélectionnées quatre DMR par région ont été sélectionnées à partir de la liste des 126 DMR utilisées pour épigénotyper les 123 épiRIL par Colomé-Tatché *et al.* (2012) en tenant compte de leur distribution régulière dans la région étudiée. La liste des 12 DMR utilisées comme marqueur pour la recherche de nouveaux recombinants au niveau des régions étudiées est répertoriée dans le Tableau III.3.B.

Pour chaque DMR, la détection de l'état de méthylation repose sur un différentiel de signal obtenu entre un même échantillon traité et non traité par la McrBC issu d'une amplification par PCR utilisant des amorces marquées spécifiques du locus à caractériser réalisée par une approche par multiplexage en collaboration avec la plateforme GENTYANE de Clermont-Ferrand (cf. Chapitre II, paragraphe 3.1.5.1.). Au total, 97 individus recombinants (incluant les recombinants présents initialement dans la carte des 123 épiRIL réalisée par Colomé-Tatché *et al.* (2012)) au niveau d'au moins une des 3 régions sélectionnées ont été identifiés au sein des 505 individus (Tableau III.4) puis validés au sein du laboratoire par McrBC/PCR. Un total de 42 individus recombinants a été identifié pour la région A en utilisant 4 DMR, 38 pour la région B en utilisant 3 DMR (les résultats de fluorescence obtenus ne permettant pas de distinguer les épihaplotypes WT et *ddm1-2* au niveau de la DMR MM415) et 47 pour la région C en utilisant 4 DMR (**Figure III.5A**). De plus, 32 individus présentant une recombinaison au niveau d'au moins 2 des 3 régions sélectionnées ont été identifiés (**Tableau III.4**). L'épigénotypage partiel des 505 épiRIL a ainsi permis d'augmenter le nombre d'individus recombinants pour chaque région de 1,8 à 2,3 fois par rapport au sous-ensemble de 123 lignées. Au final, l'épigénotype des 97 épiRIL a été obtenu à chaque DMR étudiée. L'analyse de ces données montre pour chaque épiRIL au moins une recombinaison au niveau de l'un des 3 QTL^{épi} sélectionnés (**Figure III.5A**). Afin de réaliser une nouvelle détection de QTL^{épi} dans les intervalles étudiés, le comportement des 97 épiRIL recombinantes en réponse à l'infection par l'isolat eH de *P. Brassicae* a été évalué dans un test pathologique comportant 2 répétitions biologiques composées chacune de 2 blocs (soit au total 24 individus testés par épigénotype). L'ensemble des traits phénotypés précédemment (Liégard *et al.*, 2018) - *i.e.* la longueur des feuilles, la quantification de l'ADN du pathogène, la biomasse racinaire et l'indice pathologique - ont été mesurés. Les valeurs obtenues pour chacun des caractères mesurés montrent une distribution continue suggérant un contrôle polygénique de ces traits comme observé précédemment (**Figure III.5B**).

Tableau III.5. Liste des 30 marqueurs utilisés dans la carte épigénétique des recombinants. En noir les marqueurs qui ont été épigénotypés à la McrBC pour les 97 épiRIL. En vert les marqueurs dont les données sont issues de la publication de Colomé-Tatché *et al.* (2012) et disponibles pour 51 % des 97 épiRIL utilisées.

	Région A			Région B				Région C
DMR utilisées comme marqueur	MM10	MM17	MM25	MM402	MM432	MM505	MM527	MM686
	MM11	MM18	MM27	MM414	MM466	MM510	MM531	MM689
	MM12	MM20		MM415	MM495	MM515	MM537	MM691
	MM13	MM22		MM418	MM499	MM520	MM544	MM694

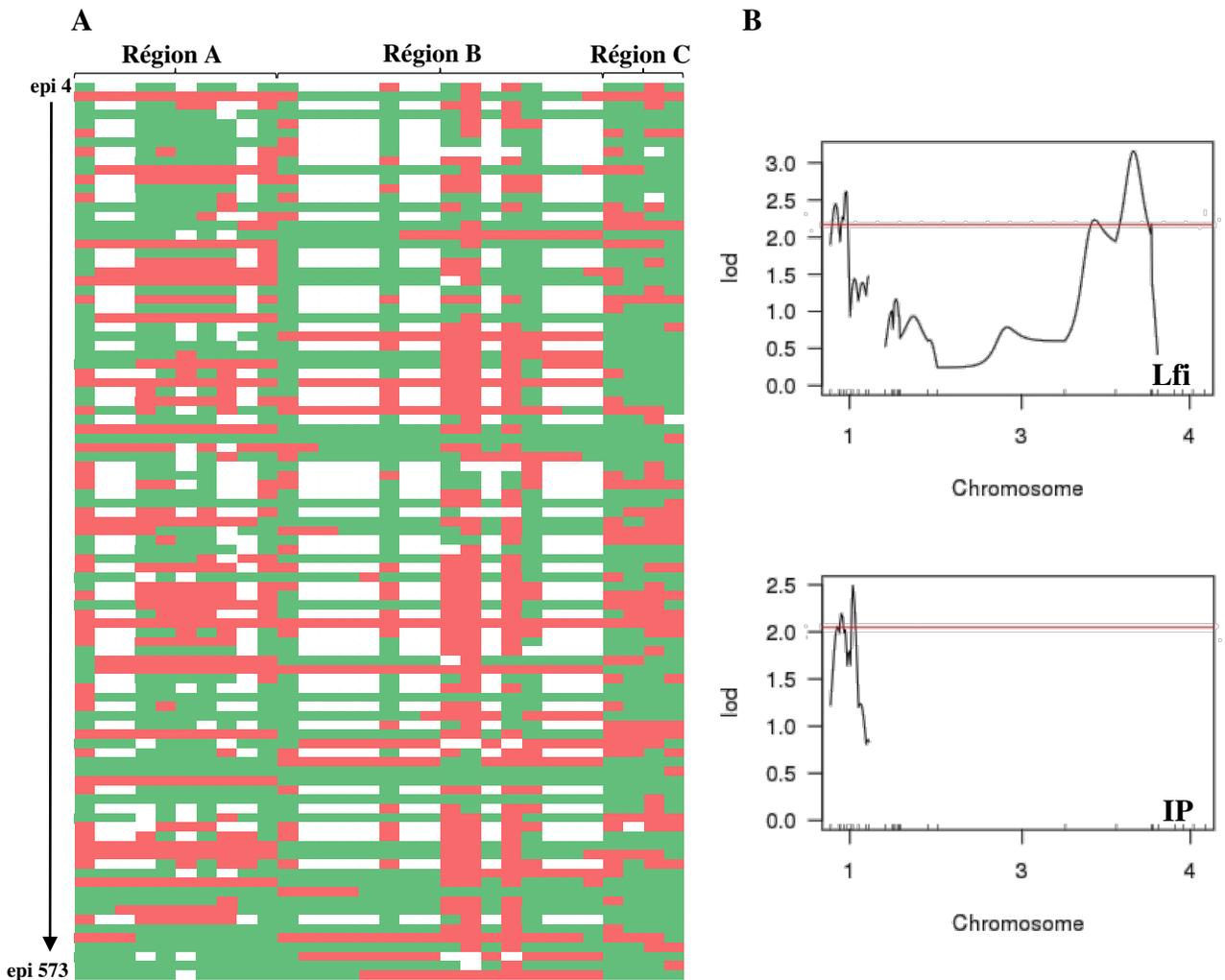


Figure III.6. Epigénotype des épiRIL sur 30 DMR et profils des QTL^{épi} détectés. (A) Epigénotype des 97 épiRIL recombinantes au niveau de 30 DMR utilisées comme marqueurs dans les trois régions sélectionnées. Les 97 épiRIL sont représentées sur l'axe y. L'épigénotype déterminé au niveau de chaque marqueur le long des régions sélectionnées est représenté selon l'axe x. En vert l'épigénotype de type sauvage. En rouge l'épigénotype de type *ddm1-2*. Les données manquantes sont représentées en blanc. (B) QTL^{épi} détectés en utilisant les traits phénotypiques mesurés sur les 97 épiRIL recombinantes suite à l'infection par *P. brassicae* et la carte contenant l'épigénotypage de 30 DMR. La ligne rouge indique le seuil de significativité à 5 %. Lfi correspond à la longueur des feuilles. IP correspond à l'indice pathologique.

Afin de réduire la taille des régions sélectionnées, en plus de l'augmentation du nombre d'individus recombinants, nous avons également augmenté le nombre de marqueurs épigénotypés au sein des régions sélectionnées pour les 97 recombinants.

2.3. Densification en marqueurs des régions étudiées chez les épiRIL recombinantes

En plus des DMR utilisées pour sélectionner les 97 individus recombinants, 6 nouveaux marqueurs (3 marqueurs au niveau de la région A et 3 au niveau de la B) et 10 marqueurs (**Tableau III.5**) dont les données d'épigénotypage sont issues de la carte épigénétique créée par Colomé-Tatché *et al.* (2012) (disponible pour 51,5 % des recombinants) ont été utilisés pour affiner l'épihaplotype des recombinants (**Figure III.6A**). Avec ces données, une carte épigénétique des recombinants a été réalisée et les régions étudiées présentaient une distance génétique de 57,1 cM (3,48 Mb) avec 10 marqueurs pour la région A, 406,7 cM (8,93 Mb) avec 16 marqueurs pour la région B et 45,3 cM (2,21 Mb) avec 4 marqueurs pour la région C. A noter que pour la région C aucun marqueur supplémentaire n'a été ajouté pour le moment car les données de quantification de l'ADN du pathogène nécessaire pour sa réduction sont toujours en cours d'acquisition. L'épigénotypage a été réalisé comme précédemment en utilisant des étapes de digestion d'ADN à la McrBC avec une lecture sur gel d'agarose et un passage par de la qPCR pour des marqueurs tendancieux (cf. Chapitre II, paragraphe 3.1.5.1).

A partir de cette carte et des données phénotypiques obtenues, la détection des QTL^{épi} a été réalisée comme décrit au Chapitre II, paragraphe 3.1.3. Les données de quantification de l'ADN de l'agent pathogène étant encore en cours d'acquisition, seules les réductions des régions A et B sont présentées. La nouvelle détection de QTL^{épi} réalisée a permis de re-détecter des QTL^{épi} à l'intérieur des régions A et B pour le caractère longueur de feuilles (**Figure III.6B**). Aucun QTL^{épi} n'a été détecté dans les régions ciblées pour la biomasse racinaire. L'intervalle de confiance du QTL^{épi} *Rbt^{épi}-At1* issu de la primo détection n'étant que partiellement inclus dans la région sélectionnée pour la cartographie fine, cette absence de détection n'est pas surprenante (pic du QTL^{épi} non inclus dans la région sélectionnée). De plus, pour le caractère indice pathologique qui n'avait pas permis de détecter de QTL^{épi} dépassant les seuils de détection lors de la primo détection, nous avons décelé un QTL^{épi} à l'intérieur de la région A (**Figure III.3B**). Les effets des marqueurs au pic des QTL^{épi} détectés dans les régions ciblées ont gardé les mêmes origines que celles décrites dans la publication Liégard *et al.* (2018) avec l'épihallèle WT qui augmente les valeurs de longueur de feuille. Pour l'indice pathologique, l'épihallèle WT amène à une diminution de l'indice pathologique (**Figure III.7**). Les intervalles de confiance des QTL^{épi} détectés avec ce set d'individus recombinants présentent des tailles inférieures à celles des régions sélectionnées pour la réduction à partir de la primo-détection (**Figure III.8**). Au final les intervalles de confiance des QTL^{épi} des régions sélectionnées présentent des tailles allant de 41,2 à 130,7 cM ou de 1,2 à 2,55 Mb englobant de 217 à 669 gènes.

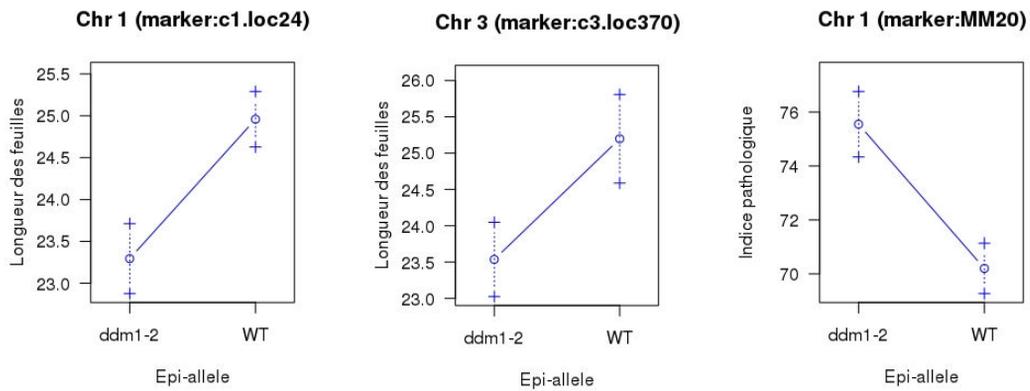


Figure III.7. Effet de l'épi-allèle au marqueur au pic des QTL^{épi}. Tous les graphiques montrent l'effet de chaque épi-allèle sur la valeur du trait au marqueur au pic du QTL. Pour les QTL^{épi} de longueur de feuilles l'épi-allèle *ddm1-2* est associé à une valeur inférieure à l'épi-allèle WT. Pour le QTL détecté pour l'indice pathologique, l'épi-allèle *ddm1-2* est associé à une valeur plus élevée que l'épi-allèle WT. WT : individu Columbia sauvage.

Tableau III.6. Tableau récapitulant les intervalles de confiance des QTL^{épi} détectés dans les régions sélectionnées.

Trait phénotypique	Région	Intervalle de confiance		
		Positionnement	Marqueur	Position physique (pb)
IP	Région A	Début	MM10	8931527
		Fin	MM22	11512649
	Région B	Début		
		Fin		
Lfi	Région A	Début	MM10	8931527
		Fin	MM18	11102716
	Région B	Début	MM515	15621495
		Fin	MM537	16825096

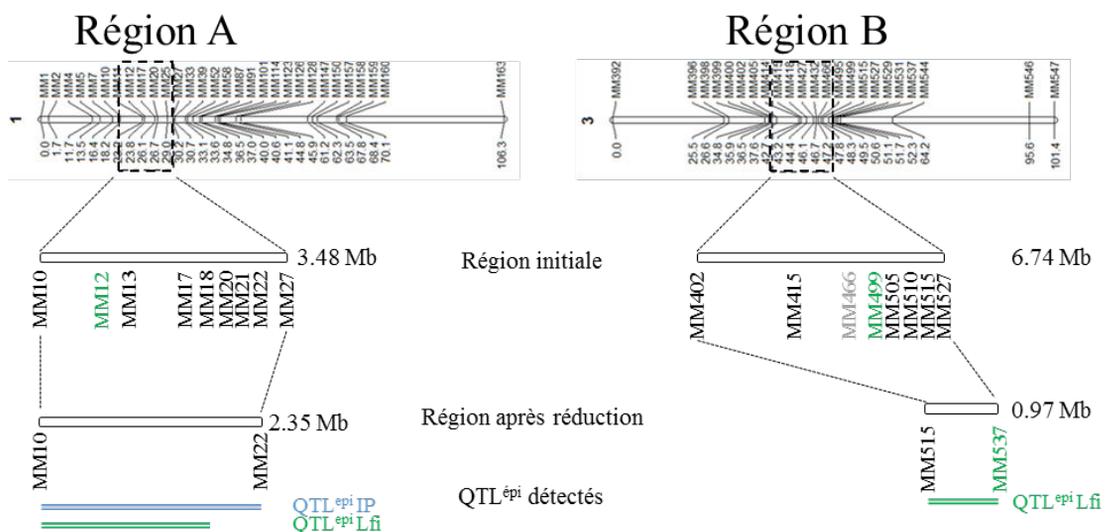


Figure III.8. Schéma récapitulatif de la réduction des régions A et B. Les numéros à gauche des chromosomes indiquent le numéro du chromosome où est positionnée la région étudiée. Les valeurs en Mb indique la taille maximale de la région étudiée en fonction de l'étape de la réduction. Les marqueurs en vert sont issus des données de la publication de Colomé-Tatché *et al.* (2012). Les marqueurs en noirs sont issus de l'épigénotypage à la McrBC des 97 épiRIL recombinantes.

3. Discussion et perspectives

La recherche de gènes ou régions candidates sous certains des QTL^{épi} identifiés par Liégard *et al.* (2018) a été réalisée *via* une approche de réduction des intervalles de confiance de QTL^{épi} avec l'identification de lignées recombinantes supplémentaires à partir des 505 individus constituant la population épiRIL (Johannes *et al.*, 2009), ainsi que par l'ajout d'épimarqueurs en utilisant la McrBC. Cette approche a permis de réduire de manière plus ou moins importante les intervalles de confiance des QTL^{épi} dans les régions A et B qui ont été sélectionnées (**Tableau III.4** et **Figure III.5**). Cependant les intervalles de confiance des QTL^{épi} détectés présentent encore des distances génomiques importantes allant de 1,2 Mb à 2,55 Mb (soit de 217 à 669 gènes). Au vu de leur taille, un nombre encore important de gènes est présent rendant difficile la recherche de gènes/loci candidats ou de régions candidates. Ainsi, malgré la réduction opérée par rapport à la taille des régions d'origine, des étapes supplémentaires de réduction sont encore nécessaires afin d'arriver à un nombre de gènes raisonnable pour poursuivre avec une liste de gènes candidats. Ainsi et par la suite, en plus des 30 marqueurs utilisés pour le moment dans la carte (**Tableau III.5**), un certain nombre de marqueurs non épigénotypés au niveau des 97 recombinants restent disponibles. Pour la région A, 3 marqueurs, les DMR MM14, MM15 et MM16 identifiées par Colomé-tatché *et al.* (2011) peuvent ainsi être utilisés pour épigénotyper et réduire encore si possible l'intervalle de confiance des QTL^{épi} de la région. Pour la région B, le décalage de l'intervalle de confiance du QTL^{épi} détecté nous a amené à utiliser des marqueurs supplémentaires de la carte de Colomé-tatché *et al.* (2011) au niveau de la borne droite de l'intervalle de confiance (**Figure III.8**). Au niveau de cette région, la réduction de l'intervalle de confiance du QTL^{épi} pourra être réalisée par l'épigénotypage de 14 marqueurs supplémentaires (MM521 à MM526 et MM 528 à MM536), mais aussi par l'épigénotypage des marqueurs MM531 et MM537 utilisés dans la carte initiale mais pour lesquels l'information n'est connue que pour 51 % des recombinants (données de Colomé-tatché *et al.* (2011)). Pour la région C, bien qu'aucune analyse n'a encore été réalisée car l'acquisition des données de quantification de l'ADN du pathogène est encore en cours, 10 marqueurs sont disponibles dans la région afin d'essayer de réduire la taille de cette dernière.

Même si l'identification de gènes candidats ou de régions candidates sous les régions réduites n'est pas encore possible sans travail supplémentaire, cette approche de réduction de QTL^{épi} a permis de mettre en avant un nouvel élément par rapport à la publication d'origine qui est la présence d'un QTL^{épi} en utilisant le trait phénotypique indice pathologique au niveau de la région A. En effet, initialement aucun QTL^{épi} dépassant le seuil de significativité n'avait été identifié en utilisant ce trait phénotypique. L'augmentation du nombre de marqueurs et d'individus recombinants au niveau de la région d'intérêt ainsi qu'un meilleur contrôle des conditions de culture permettant une plus grande stabilité de la réponse des plantes ont ainsi permis de détecter un QTL^{épi} avec ce trait phénotypique.

4. Conclusion partielle

Le travail présenté dans ce chapitre de thèse apporte des premiers éléments de réponse à la question de l'implication de l'épigénome des *Brassicaceae*, plus précisément dans notre cas du méthylome d'*A. thaliana*, dans la réponse à l'infection par *P. Brassicae* l'agent causal de la hernie. Avec cette approche non ciblée, les résultats obtenus ont permis de mettre en avant différents éléments :

- L'infection par l'isolat eH de *P. Brassicae* de différents mutants impliqués dans différentes voies épigénétiques a conduit à différentes réponses phénotypiques chez ces mutants, mettant en exergue le rôle de la méthylation de l'ADN dans la réponse de moindre sensibilité à la hernie via les résultats obtenus chez le mutant *ddm1*.
- L'infection par le même isolat de *P. Brassicae* (eH) des individus de la population épiRIL conduit à une ségrégation de réponses phénotypiques, montrant une implication du méthylome d'*Arabidopsis* dans la réponse de cette dernière à la hernie.
- La distribution continue des valeurs observées pour les différents caractères étudiés dans la descendance épiRIL suggère un contrôle polygénique (polyépigénétique) de ces traits.
- 14 QTL^{épi} additifs ont été détectés sur les chromosomes 1 à 4 d'*Arabidopsis* traduisant la régulation des gènes impliqués dans la résistance à la hernie par des variations de méthylation.
- Les effets de 5 QTL^{épi} sont dépendants de la température.
- L'augmentation du nombre d'épiRIL recombinantes, ainsi que la densification en marqueurs au niveau de régions de QTL^{épi} sélectionnées a permis de :
 - Réduire l'intervalle de confiance des 3 régions génomiques étudiées où co-localisent plusieurs QTL^{épi}. Pour la région A la réduction de l'intervalle a permis de passer de 3,48 Mb à 2,17 Mb pour la longueur des feuilles et à 2,55 Mb pour l'indice pathologique. Pour la région B l'intervalle est réduit de 6,74 Mb à 1,2 Mb.
 - Détecter des QTL^{épi} non significativement détectés lors de la première analyse. En effet, alors qu'aucun QTL^{épi} n'avait été détecté pour le trait indice pathologique dans Liégard *et al.* (2018), lors de la réduction de l'intervalle de confiance un QTL^{épi} a été détecté significativement avec ce trait dans la région A. Ce nouveau QTL^{épi} présente un intervalle de confiance de 1,1 Mb (de 10 Mb à 11,1 Mb).

Même si le travail de cartographie fine a permis de réduire les intervalles de confiance, un nombre encore important de gènes candidats, allant de 217 à 669, est encore présent sous les QTL^{épi} réduits. Pour le moment, seules des listes de gènes ont pu être dressées et une réduction supplémentaire des intervalles est encore nécessaire afin d'obtenir des listes restreintes de gènes/régions candidates à étudier.

CHAPITRE IV

**APPROCHE CIBLEE DU ROLE POTENTIEL
DES EPIMUTATIONS DANS LA RESISTANCE
PARTIELLE D'*ARABIDOPSIS THALIANA* A
*PLASMODIOPHORA BRASSICAE***

CHAPITRE IV. APPROCHE CIBLEE DU ROLE POTENTIEL DES EPIMUTATIONS DANS LA RESISTANCE PARTIELLE D'ARABIDOPSIS THALIANA A PLASMIDIOPHORA BRASSICAE

Comme présenté dans le chapitre bibliographique, différentes études ont démontré le rôle d'épimutations dans le contrôle de traits phénotypiques chez différentes espèces végétales, aussi bien dans le cadre d'épimutations naturelles (cf. chapitre I, paragraphe 2.4.5.1) que dans le cadre d'épimutations induites (cf. chapitre I, paragraphe 2.4.5.2). Les résultats que nous avons obtenus, présentés dans le chapitre précédent, mettent clairement en évidence que la résistance à la hernie peut être sous contrôle génétique et/ou épigénétique et que dans certaines régions les QTL génétiques et épigénétiques co-localisent. Dans cette deuxième partie de la thèse, nous nous sommes donc attelés à déchiffrer si des QTL identifiés préalablement par l'équipe d'accueil à partir de variations de séquence pourraient posséder une causalité épigénétique. Nous nous sommes intéressés à deux régions, localisées respectivement sur les chromosomes 1 et 5 d'*A. thaliana*.

La première partie de ce chapitre présente l'étude de la région *RPB1* sur le bras court du chromosome 1 d'*A. thaliana*. Cette région qui co-localise avec trois QTL^{épi} détectés par analyse épigénétique de liaison (Liégard *et al.*, 2018) avait déjà été identifiée comme un facteur majeur de résistance à la hernie par Arbeiter *et al.* (2002) ainsi que comme un QTL majeur de résistance à la hernie par l'équipe d'accueil (Jubault *et al.*, submitted). Dans cette région Jubault *et al.* (submitted) ont identifié le gène At1g32030 par analyse d'association faisant de ce gène et de la région adjacente des candidats potentiels pour une analyse de DMR causales de QTL^{épi} dans la population épiRIL. Des analyses bio-informatiques préliminaires réalisées sur le contexte génomique de ce gène ont montré divers éléments dans le fonds génétique Columbia : le gène At1g32030 ne présente pas d'intron (pour une taille de 1001 pb) et est encadré par des éléments transposables. Afin de déterminer si ce gène/cette région est l'élément causal des QTL^{épi} détectés dans cette région, des analyses de la variation de la méthylation et de l'expression des gènes de cette région dans différentes épiRIL, ainsi que dans des accessions naturelles d'*Arabidopsis* ont été réalisées. Les résultats de ce travail sont développés dans la première partie de ce chapitre.

La deuxième partie de ce chapitre de thèse s'intéresse à étudier la causalité épigénétique potentielle d'un autre QTL de résistance à la hernie précédemment identifié au sein du laboratoire (Jubault *et al.*, 2008) et étudié par l'équipe d'accueil. Ce QTL, nommé *Pb-At5.2*, a été initialement détecté par analyse de liaison dans la population RIL Bur-0 (partiellement résistant) x Col-0 (sensible) comme un QTL génétique impliqué dans la résistance à la hernie. Ce dernier expliquait 22,6 % de la variation

phénotypique observable. En utilisant une population HIF (Heterogenous Inbred Family) dérivée du croisement Bur-0 x Col-0, une approche de clonage positionnel a été réalisée au laboratoire. La région *Pb At-5.2* a ainsi été réduite à 7 gènes candidats potentiellement impliqués dans la résistance partielle à la hernie. Cette région présente une très faible variation de séquence entre les deux parents posant ainsi la question d'une variabilité potentielle d'origine épigénétique. Pour répondre à cette question, le statut de méthylation et l'expression des 7 gènes candidats de la région ont été analysés chez Bur-0 et Col-0 en utilisant notamment les données générées par le consortium 1001 génomes. A partir de ces données, deux des 7 gènes candidats ont montré des profils de méthylation et d'expression contrastés entre les deux accessions. Différentes analyses menées sur les accessions Col-0, Bur-0 mais aussi sur des lignées HIF et sur différentes accessions naturelles d'*Arabidopsis* ont permis d'identifier une épimutation contrôlant une partie de la résistance quantitative à la hernie chez *Arabidopsis*. Ce travail développé dans la deuxième partie de ce chapitre est présenté sous la forme d'un article en cours de rédaction pour soumission à Nature communication. La difficulté de relier l'effet du QTL *Pb-At5.2* à la résistance à la hernie chez certaines des accessions d'*A. thaliana* utilisées nous a amené à nous poser la question de la présence d'une interaction épistatique potentielle entre ce QTL et une autre région du génome. Le travail effectué sur la recherche d'une interaction épistatique, ainsi que de régions/gènes causaux potentiels est présenté dans la troisième partie de ce chapitre. Afin de valider le rôle d'un ou des deux gènes sous l'épimutation dans la résistance à la hernie, des approches de clonage positionnel et de validation par CRISPR sont en cours au sein du laboratoire. La mise en place de ce travail est présentée dans la quatrième partie de ce chapitre.

1. Analyse de la région du gène *RPB1*

Le gène de résistance *RPB1* a été précédemment identifié comme un facteur majeur de résistance à la hernie (Arbeiter *et al.*, 2002) contre les isolats e et e3 et comme un QTL de résistance contre l'isolat eH (Jubault *et al.*, submitted). Cette région a été sélectionnée comme une région candidate sous contrôle épigénétique du fait de sa co-localisation avec les QTL^{épi} *PbI^{épi}-At1*, *LfiI^{épi}-At1* et *Rbi^{épi}-At1* détectés dans la population épiRIL (Liégard *et al.*, 2018). L'étude de la structure génomique de la région entourant les SNP significatifs identifiés en analyse GWAS a montré que celle-ci était composée de trois principaux gènes/pseudo-gènes flanqués d'éléments transposables (**Figure IV.1A**). Du fait de la présence de ces éléments transposables en plus de la co-localisation avec les QTL^{épi}, nous avons suivi la piste d'une régulation d'origine épigénétique de la région. L'analyse de l'état de méthylation de cette région allant du gène At1g32010 au gène de type élément transposable At1g32040 (10 kb) en utilisant les données publiques de Stroud *et al.* (2013) a mis en avant dans le génotype Col-0 la présence d'un contexte de méthylation particulier avec un état de méthylation contrasté entre le gène At1g32030 et les éléments transposables qui l'entourent (qui sont hypométhylés) par rapport aux gènes et aux éléments transposables adjacents (qui sont eux hyperméthylés, **Figure IV.1A**). A la vue de ces premiers résultats nous avons décidé d'analyser le taux de méthylation au niveau de la région au sein des 123

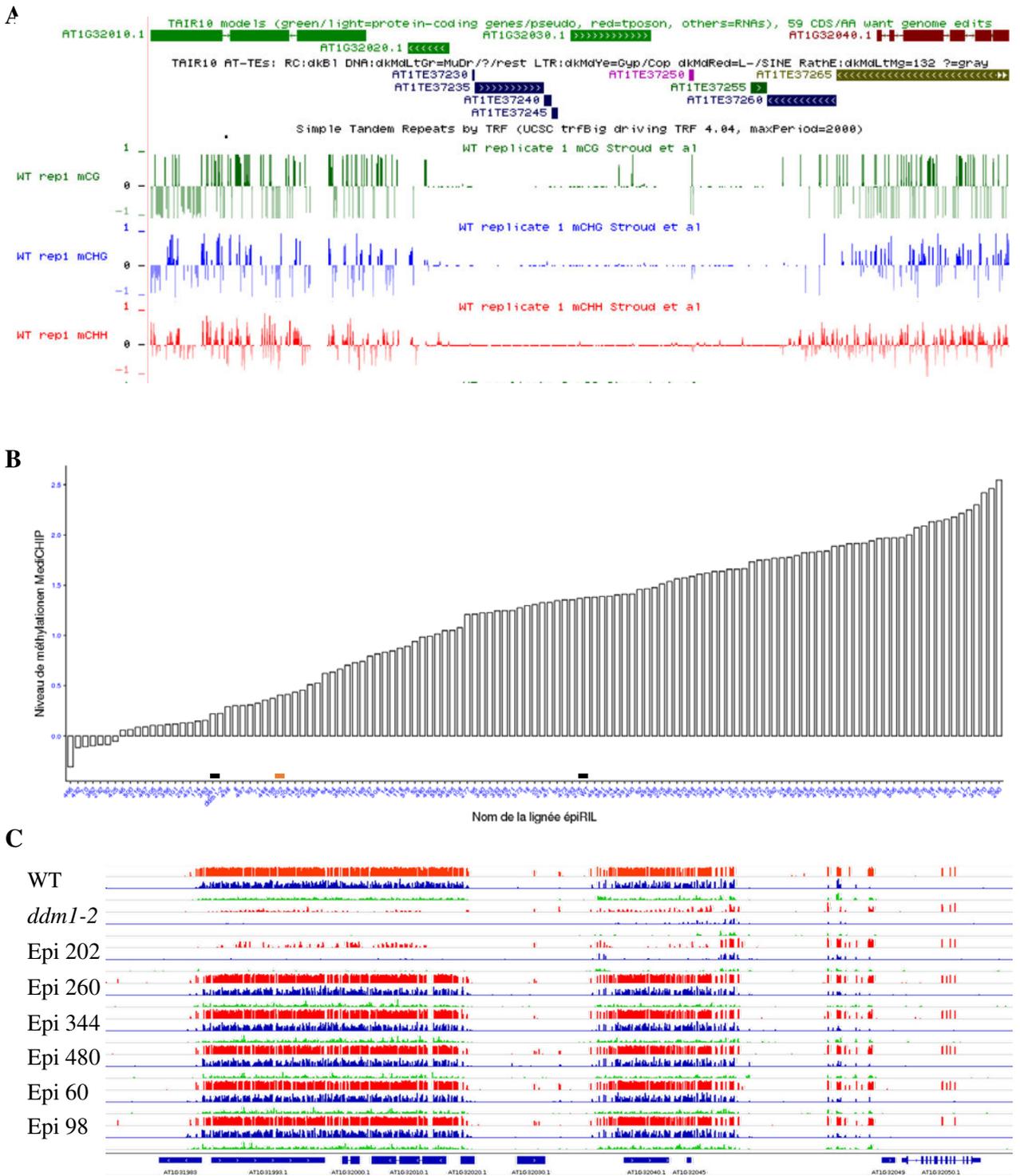


Figure IV.1. Données génomiques et de méthylation de la région de *RPBI*. (A). Contexte génomique et de méthylation de la région *RPBI* (issu de Stroud *et al.*, 2013). WT : méthylation sur les îlots CG, CHG ou CHH de la région du gène chez l'écotype Columbia d'*Arabidopsis*. (Source : www.mcdb.ucla.edu/Research/Jacobsen/LabWebSite). CG : rouge, CH : bleu, CHH : vert. (B). Etat de méthylation moyen des épiRIL dans la région de *RPBI* obtenu à partir des données de MeDIP-CHIP. Le calcul du niveau moyen de méthylation a été réalisé dans l'intervalle de la région du gène *RPBI* allant de 11 509 341 pb à 11 520 070 pb soit environ 10 kb). En ordonnées : le pourcentage moyen de méthylation. En abscisses : nom des épiRIL correspondantes. Les lignes noires en abscisses indiquent de la gauche vers la droite la position du niveau de méthylation de *ddm1-2* et de WT. La ligne orange en abscisse indique la position de l'épiRIL 202. (C). Profils de méthylation de la région de *RPBI* des 6 épiRIL obtenus par séquençage après un traitement au bisulfite. CG : rouge, CH : bleu, CHH : vert.

épiRIL ainsi que la conservation de la structure génomique et l'état de méthylation de la région dans différentes accessions naturelles d'*Arabidopsis*.

1.1. Analyse de la région au niveau des épiRIL

Afin d'analyser les niveaux de méthylation de la région dans les 123 épiRIL, les données de méthylation acquises par l'approche MeDIP-CHIP par Colomé-Tatché *et al.* (2012) ont été utilisées et mises à notre disposition par l'équipe de Vincent Colot (IBENS).

Dans un premier temps nous avons étudié la répartition des niveaux de méthylation des épiRIL au niveau de la région du gène *RPB1*. Les résultats obtenus montrent que la répartition de l'état de méthylation global de la région entre les types WT et *ddm1-2* dans la population épiRIL dévie du pourcentage attendu de 25 % de méthylation pour le type *ddm1-2* et de 75 % pour le type WT (**Figure IV.1B**). En effet, seulement 19 épiRIL montrent une méthylation globale égale ou inférieure à celle de *ddm1-2* sur 123 lignées, soit environ 15,4 %. Cette observation semble aller dans le sens de ce qui avait été démontré par Colomé-Tatché *et al.* (2012) avec la présence de phénomènes de réversion partielle ou complète de la méthylation dans certaines régions de certaines lignées épiRIL. Toutefois, le niveau de méthylation moyen de la région de type *ddm1-2* étant très faible (0,23) comparé à celui du WT (1,38) et la région couverte étant étendue (10 kb), il est possible que des épiRIL présentent un profil de méthylation de type *ddm1-2* sur seulement une partie de la région mais possèdent ainsi un taux de méthylation global supérieur à *ddm1-2*. Afin de vérifier cette hypothèse l'analyse de données de bisulfite, mises à notre disposition par le laboratoire de Vincent Colot, obtenues à partir de jeunes plantules pour 6 épiRIL (les épiRIL 60, 98, 202, 260, 344 et 480) cultivées en condition contrôle a été réalisée. Le résultat obtenu montre que bien qu'ayant un taux de méthylation global supérieur à celui de *ddm1-2* (0,41) pour la région via l'approche de MeDIP-CHIP, l'épiRIL 202 possède un profil de méthylation au bisulfite de type *ddm1-2* (**Figure IV.1C**).

Dans un deuxième temps nous avons étudié de manière plus précise le profil de méthylation de la région *RPB1* en recherchant des DMR à partir des données de Colomé-Tatché *et al.* (2012). Ainsi nous avons trouvé trois DMR candidates entre le parent WT et *ddm1-2* : les DMR58 (de 11 511 756 à 11 512 571 pb, en TAIR8), 59 (de 11 517 887 à 11 518 652 pb, en TAIR8) et 60 (de 11 519 343 à 11 519 828 pb, en TAIR8) qui sont de part et d'autre du gène *At1g32030* (**Figure IV.2A**). La DMR 58 se situe au niveau du gène *At1g32010* (de 11 509 321 à 11 512 020 pb, en TAIR8) et de l'extrémité gauche du gène *At1g32020* (de 11 512 532 à 11 513 047 pb, en TAIR8). Les DMR 59 et 60 se situent toutes les deux au niveau du gène *At1g32040* (de 11 518 398 à 11 520 050 pb, en TAIR8). Les trois DMR identifiées possèdent toutes un statut hyperméthylé dans la lignée parentale Col-0 WT et hypométhylé dans la lignée parentale *ddm1-2*. Le gène *At1g32030* quant à lui est à l'état hypométhylé dans les 2 lignées parentales. Afin de valider l'impact potentiel de la variation de méthylation sur l'expression des gènes de la région, l'analyse de données de RNA-seq, mises à notre disposition par le laboratoire de

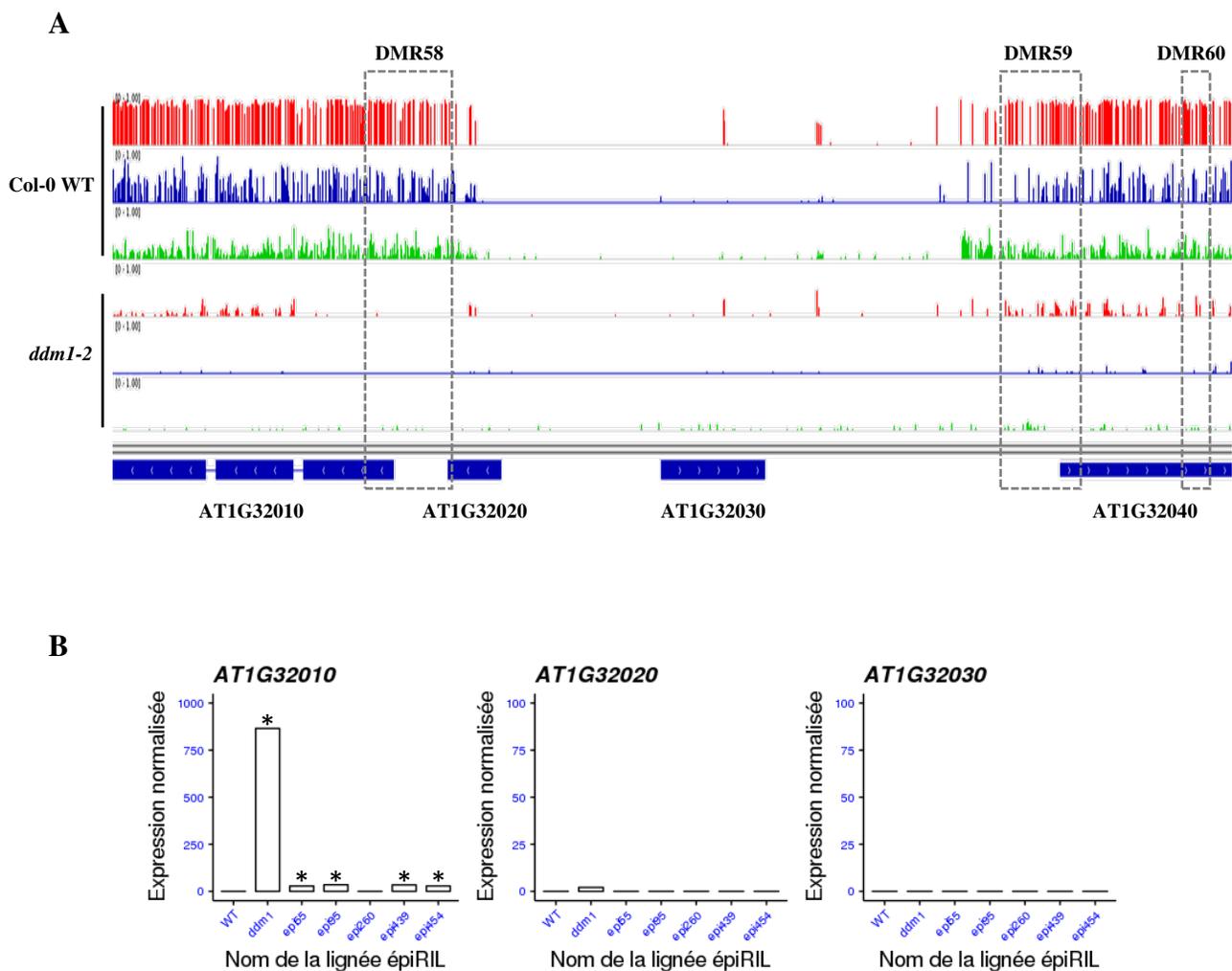


Figure IV.2. Etat de méthylation de trois DMR candidates au niveau de la région *RPB1* et niveau d'expression des gènes de la région. (A). Etat de méthylation de la région *RPB1* chez Col-0 et *ddm1-2*. Les données de méthylation ont été obtenues à partir de plantules par une approche de séquençage après traitement au bisulfite (données fournies par le laboratoire de Vincent Colot). Les trois régions différemment méthylées (DMR 58, 59 et 60) identifiées par Colomé-Tatché *et al.* (2012) sont encadrées par une ligne pointillée grise. Les taux de méthylation en contexte CG sont en rouge, en bleu pour le contexte CHG et en vert pour le contexte CHH. **(B).** Histogrammes des expressions des gènes de la région *RPB1* issues de données RNA-seq obtenues sur 5 épiRIL (55, 95, 260, 439 et 454) à partir de plantules. * indique des différences significatives d'expression entre chaque individu et WT (FDR, p -values < 0,05).

Vincent Colot, obtenues à partir de jeunes plantules de 5 épiRIL cultivées en condition contrôle a été réalisée.

L'analyse de l'expression des gènes de la région entre *ddm1-2* et WT a mis en avant la présence de variations d'expression significatives pour le gène At1g32010. En ce qui concerne les 5 épiRIL séquencées par RNA-seq (55, 95, 260, 439 et 454), 4 épiRIL (55, 95, 454 et 439) montrent une variation d'expression significative ($p\text{-value} \leq 0,001$) pour le At1g32010 par rapport au WT. Concernant le gène candidat At1g32030 aucune expression n'a été identifiée pour les épiRIL et les lignées parentales (**Figure IV.2B**). Cependant, les analyses de méthylation et d'expression ayant été réalisées sur des plantes non inoculées la question du niveau d'expression des gènes de la région et du gène At1g32030 dans les racines des plantes en réponse à l'infection se pose. De ce fait il serait intéressant de réaliser une cinétique d'infection (14 et 21 jours après infection soit en phase secondaire de l'infection), couplée à des plantes non infectées aux mêmes temps cinétiques, afin d'étudier au niveau d'épiRIL présentant des réponses contrastées à la hernie, l'expression des 4 gènes (à partir d'ARN extrait de racines) présents dans la région et de vérifier le statut de méthylation des DMR lors de l'infection par *P. Brassicae*. Pour cette approche, 6 plantes par épi-génotype/lignée parentale et par bloc pourraient être utilisées dans 2 répétitions biologiques composées de 2 blocs, soit 24 plantes au total par lignée. De plus, des marqueurs CHOP-PCR pourraient être développés au niveau des DMR58, DMR59 et DMR60 afin d'être implémentés dans la stratégie de cartographie fine de la région du chromosome 1 (épigénotypage des lignées recombinantes au niveau de la région du chromosome 1).

Ces résultats ayant été obtenus dans une population issue d'un mutant EMS, l'existence de tels profils de méthylation au niveau de la région a également été analysée dans des écotypes d'*A. thaliana*.

1.2. Analyse de la région au niveau des accessions naturelles d'*A. thaliana*

Afin d'analyser le profil de méthylation de cette région au niveau de différentes accessions d'*Arabidopsis*, l'analyse de la conservation de la structure génomique de type Col-0 de la région au sein des accessions a d'abord été réalisée en collaboration avec Léandro Quadrana au sein du laboratoire de Vincent Colot. Pour cela, les données de DNaseq des 1001 accessions d'*Arabidopsis* du projet 1001 génomes ont été utilisées. Les lectures de séquençage des 1087 accessions d'*Arabidopsis* ont été analysées et mappées sur la région afin de déterminer les couvertures verticale et horizontale présentes (**Figure IV.3A**). Ceci a été réalisé car une couverture homogène de lecture de séquençage de génome sur Col-0 est une forte indication que la structure de la région est similaire à la référence. Sur les 1087 accessions d'*Arabidopsis* analysées, seules 82 accessions présentent des couvertures verticale et horizontale permettant d'établir que leur structure génomique est similaire à celle de Col-0 (**Figure IV.3B**).

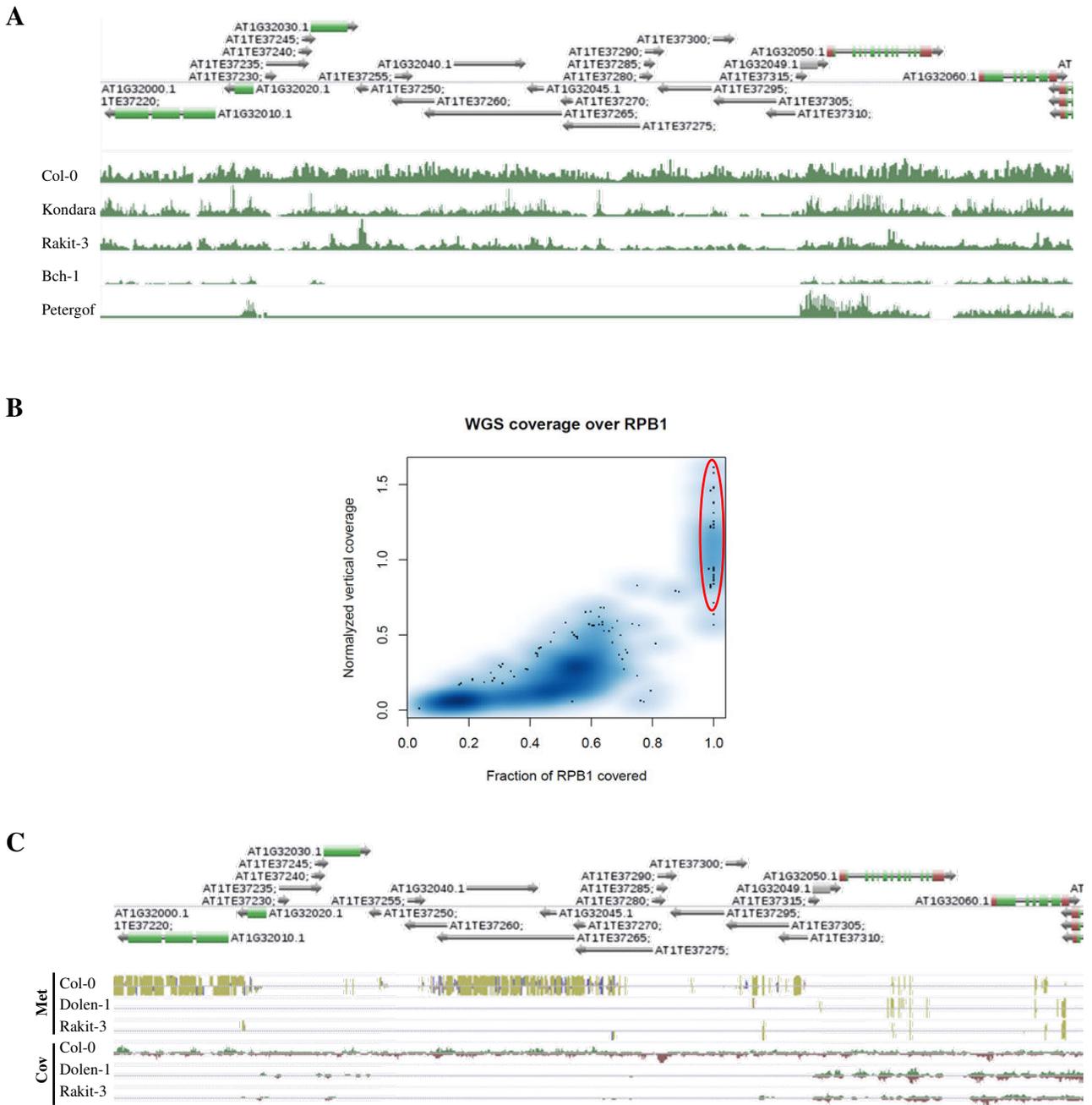


Figure IV.3. Couverture génomique et état de méthylation de la région *RPB1* pour des accessions naturelles d'*Arabidopsis*. (A). Exemple du résultat de mapping des lectures de séquençage sur la région *RPB1* de type Col-0 pour différentes accessions. Col-0 : profil de l'alignement des reads de Col-0 sur le génome Col-0. Kondara : profil de l'alignement des reads de Kondara sur le génome Col-0, montrant une région de type Col-0. Rakit-3 : profil de l'alignement des reads de Rakit-3 sur le génome Col-0, montrant une région de type Col-0. Bch-1 : profil de l'alignement des reads de Bch-1 sur le génome Col-0, montrant une région différente du type Col-0. Petergof : profil de l'alignement des reads de Petergof sur le génome Col-0, montrant une région différente du type Col-0. (B). Couverture verticale en pourcentage de reads. Couverture horizontale : région couverte par au moins un read. i) Une couverture verticale normalisée $< 0,25$ et horizontale $< 0,3$ indique l'absence totale du locus. ii) Une couverture verticale normalisée $[0,25 - 0,8]$ et horizontale $[0,3 - 0,8]$ indique l'absence partielle du locus. iii) Une couverture verticale normalisée $> 0,8$ et horizontale $> 0,5$ indique la présence du locus. *RPB1* région (Chr1 : 11502466-11520824). Dans le cercle rouge les 82 accessions présentant une structure génomique similaire à Col-0. (C). Profil de méthylation de la région *RPB1* pour les accessions Col-0, Dolen-1 et Rakit-3. Met : niveau de méthylation présent représenté par les barres verticales de couleur. Cov : couverture des lectures de bisulfite représentés en rouge vert.

L'analyse des données de méthylation mises à disposition par le projet 1001 génomes a montré que 73 accessions des 82 identifiées présentent des données de méthylation. L'analyse des données de méthylation de ces 73 accessions montre que le profil de méthylation pour la quasi-totalité des accessions est similaire à celui de Col-0 sur la région. En effet, deux accessions Dolen-1 et Rakit-3 semblaient présenter un état hypométhylé de la région, mais celui-ci est dû à une absence de lecture de bisulfite au niveau de la région (**Figure IV.3C**). Les résultats obtenus montrent que les profils de méthylation de type Col-0 identifiés au niveau de plantules d'épiRIL sont retrouvés dans les feuilles des écotypes étudiés.

Au vu du faible nombre d'accessions de type Col-0 au niveau de la région *RPBI*, la comparaison de la structure génomique de cette région chez Col-0 avec celle des écotypes Ler-1 et Cvi, dont les génomes ont été assemblés *de novo*, a été réalisée. La comparaison effectuée, au sein du laboratoire de Vincent Colot par Léandro Quadrana, a mis en évidence qu'aucune des trois accessions ne présente la même structuration génomique dans la région, indiquant une forte variabilité structurale de la région.

L'ensemble des résultats obtenus sur les accessions pour cette région montre que la seule variation identifiée au niveau de la région du gène *RPBI* pour les accessions présentant une structure génomique de type Col-0 est une variation de séquence. En effet, des variations majeures de structure, telles que la perte complète du locus (Chr1:11502466-11520824), sont présentes dans la collection d'accessions. L'haplotype Col-0 représente un allèle rare (seulement 83 des 1087 accessions soit ~7,6 % MAF (Fréquence de l'Allèle Mineur)). Parmi les accessions de type Col-0, 71 accessions identifiées présentent un état hyperméthylé de la région *RPBI*, indiquant qu'*a priori* la structure de type Col-0 est associée à une hyperméthylation d'une partie de la région. Ainsi, aucune accession naturelle ne semble posséder une structure génomique de la région de type Col-0 associé à un état de méthylation de type *ddm1-2*.

En perspective, il serait intéressant de déterminer la séquence de cette région pour les 1014 accessions restantes et de déterminer si la variation de séquence est due à une présence différentielle d'éléments transposables dans cette région. Cette hypothèse irait dans le sens de la présence d'un état hyperméthylé pour les accessions présentant une structure génomique de type Col-0 du fait de la présence des éléments transposables afin de maintenir la répression de l'expression de ces derniers. De ce fait, il serait intéressant d'essayer de reconstruire la séquence génomique de cette région au niveau des accessions présentant une structure génomique différente de Col-0. Ceci permettrait de déterminer si la structure de la région est modifiée, si les TE sont toujours présents ou non et de déterminer le niveau de méthylation au niveau de ces régions structurellement différentes. Les séquences DNA-seq générées pour les différentes accessions pourraient être mappées sur les génomes assemblés *de novo* des accessions Ler-1, Cvi et Bur-0 afin de déterminer si différentes accessions possèdent la même structure. Puis, l'annotation de ces régions pourrait être réalisée afin de voir quels éléments sont présents/absents

de la région par rapport à Col-0. Enfin les lectures de RNA-seq et de bisulfite pourraient être mappées sur les régions précédemment définies afin de déterminer l'expression et le niveau de méthylation des éléments présents dans la région. Finalement ce travail pourrait mettre en évidence le statut de méthylation et d'expression du ou des gènes causaux expliquant la différence de sensibilité d'accessions d'*Arabidopsis* à la hernie.

2. Analyse de la région du QTL^{épi} *Pb-At5.2*

Cette partie de thèse est présentée sous la forme d'un article en préparation pour soumission dans la revue *Nature communication*. Des données supplémentaires sont fournies dans les paragraphes qui suivent l'article.

NATURAL EPIALLELIC VARIATION ASSOCIATED WITH QUANTITATIVE RESISTANCE TO PLANT PATHOGEN

Benjamin Liégard¹, Antoine Grivot¹, Leandro Quadrana², Yoann Aigu¹, Juliette Benejam¹, Christine Lariagon¹, Jocelyne Lemoine¹, Vincent Colot², Maria Manzanarès-Dauleux¹ and Mélanie Jubault¹

AGROCAMPUS OUEST – INRA – Université de Rennes 1 UMR 1349 IGEPPP, F-35000 Rennes, France

Abstract

Plant phenotypic variations are generally triggered by genomic sequence variations. However, some plant traits were recently described as controlled by epialleles. Clubroot caused by the protist *Plasmodiophora brassicae* is a major disease of *Brassicaceae* including crops and the model plant *Arabidopsis thaliana*. Both qualitative (major genes) and quantitative (QTL) clubroot resistances have been identified and studied in *Brassicaceae*. Here we describe a natural epimutation controlling one moderate-effect QTL involved in *Arabidopsis* clubroot resistance. Fine mapping of the QTL, previously identified in a biparental population, was first carried out using 2400 progenies from Heterogeneous Inbred Families. We thus identified a 25 kb-region containing eight genes with a reduced number of SNP between the parental accessions. Transcriptomic and methylation data analyses highlighted in this region a causal 6 kb epimutation spreading two NLR genes. These results revealed the implication of a natural epimutation in the variation of plant response to clubroot infection. Combination of favorable epiallelic variations to allelic ones could therefore constitute new opportunities to breed plant resistant varieties to biotic stress.

2.1. Introduction

Intraspecific diversity of plant immune interactions is connected to the high level of sequence variations at hundreds of NLR, one of the largest and most rapidly diversifying plant gene families¹⁻⁶. Each NLR protein can be involved in the direct or indirect recognition of a small range of effector proteins secreted by specific strains of plant pathogens, potentially triggering the induction of strong plant defense responses that can rapidly stop pathogen invasion^{7,8}. Then, with only few exceptions (including non-NLR driven resistances⁹⁻¹¹ and NLR-driven broad-spectrum resistances^{12,13}, the catalogue of NLR encoding R-genes expressed in a given plant genotype shapes the range of isolate-specific full resistances (incompatible interactions).

Contrasting with R-gene driven resistances, quantitative resistances involve combinations of allelic variations at several loci (QTL), collectively contributing to post-invasive partial resistance in compatible plant-pathogen interactions. An emerging list of cloned resistance QTL supports that Quantitative Resistance Genes (QRG) are molecularly more diverse than R genes^{14,15}. This short list however still includes genes encoding NLR¹⁶⁻¹⁹ and other receptors^{20,21} or co-receptors²². Thus, variations on NLR genes (or other non-self-recognition loci) can contribute to the variations of basal resistance levels in compatible interactions.

The expression of NLR is limited by autoimmunity drawbacks and is potentially deleterious in plant genetic backgrounds expressing conformational variants of their corresponding guard proteins. NLR expression levels are thus tightly downregulated by a variety of cellular mechanisms^{23,24}. Those include regulation at transcriptional, post-transcriptional²⁵ (*i.e.* alternative splicing) and post-translational levels (*i.e.* Ub-dependent proteolytic regulation). NLR regulation also involves a wealth of epigenetic-related cellular processes, including redundant networks of small RNA²⁶⁻²⁸, histone modifications²⁹⁻³¹, histone-mark dependent-alternative splicing³², regulation of chromatin structure and DNA methylation³³. There are increasing evidences that epigenetic processes can play roles in the transitory imprinting of some plant biotic stress responses, for at least a few generations^{34,35}. It is however not clear yet to which extent the stable transgenerational inheritance of epigenetically regulated gene expression contributes to the natural intraspecific diversity of plant-pathogen interactions.

Most of the few available examples of transgenerational epigenetically controlled traits come from various plant models³⁶, where differentially methylated regions (DMR) have been associated with phenotypic traits with stable or (most often) metastable inheritance across the generations. Such regions, designed as epialleles, can affect agronomically relevant traits: compatibility, accumulation of vitamin E and fruit ripening in tomato, sex determination in melon, and fruit productivity in oil palm³⁶.

In plants, DNA methylation is monitored at three different sequence contexts, CG, CHG and CHH³⁷ (where H could be A, C or T) and its impact varies in function of the targeted genomic region

(transposable elements, gene promoters or gene bodies). DNA methylation state at a genomic level is the result of a dynamic regulation of *de novo* methylation, maintenance of the methylation and DNA demethylation. The DNA *de novo* methylation is triggered by canonical and non-canonical RNA-directed DNA methylation (RdDM) pathway involving several proteins and siRNA³⁸. Maintenance of the DNA methylation across the mitosis division, meiosis division and generations involves an RNA-dependent pathway (involving siRNA and DRM proteins) and RNA-independent pathways (involving DDM1, MET1 and VIM proteins in the CG context, and DDM1, KYP, CMT2/3 and histone mark HK9me2 in the CHG and CHH contexts^{39,40}). One first report of DNA methylation as a possible inherited epigenetic mechanism controlling plant NLR function came from the dwarfing variant *bal*, derived from the inbred progeny of the *ddm1* mutant, following a multigenerational accumulation of epigenetic and genetic alterations⁴¹. This artificial epimutant harbours a metastable phenotype of short stature/curled leaves/constitutively activated defence, that was however later attributed to a 55 kb duplication event responsible for the overexpression of the NLR-gene *SNCI*⁴². Then, this overexpression was not associated with a specific methylation state of the genomic sequence of *SNCI*, as it was initially assumed.

In the present work, we report the fine-mapping based identification of a naturally occurring stable epivariation underlying a quantitative trait locus involved in partial resistance to clubroot in *Arabidopsis*. Clubroot is a root gall disease caused by the telluric biotrophic pathogen *Plasmodiophora brassicae* (Rhizaria), affecting all *Brassicaceae* crops. The infection process involves a primary infection in root hairs for only few days, followed by a secondary infection where secondary plasmodia develop in root cortical cells, provoking hyperplasia and hypertrophy that ultimately impair plant water and nutrient uptake. Partial resistance to eH isolate in the *Arabidopsis* accession Bur-0 is controlled by four additive QTL⁴³. Fine mapping of the resistance QTL *PbAt5.2* led to the identification of a strong association between partial resistance and the methylation and expression patterns of two adjacent NLR genes At5g47260 and At5g47260. We show that both hypomethylated and hypermethylated epivariants are frequent in natural *Arabidopsis* accessions. An analysis of 127 natural accessions assessed that the highly methylation state at this locus was negatively correlated with the expression of the two NLR-genes and with quantitative resistance against *P. Brassicae*, independently of nucleotide variations at this locus. We further show that the hypermethylated epiallele was maintained in Col-0 by the RNA independent pathway involving DDM1, MET1, VIM and CMT2/3. The small region including At5g47260 and At5g47280 was found massively targeted by siRNA in Col-0, with a majority of 24 nt siRNA. By contrast, the Bur-0 epiallele was weakly targeted by siRNA. Collectively, the presented data demonstrate that the quantitative resistance to a major root disease affecting *Brassicaceae* is associated with the stable inheritance of a natural epivariation involved in the control of the constitutive expression of a pair of NLR-genes.

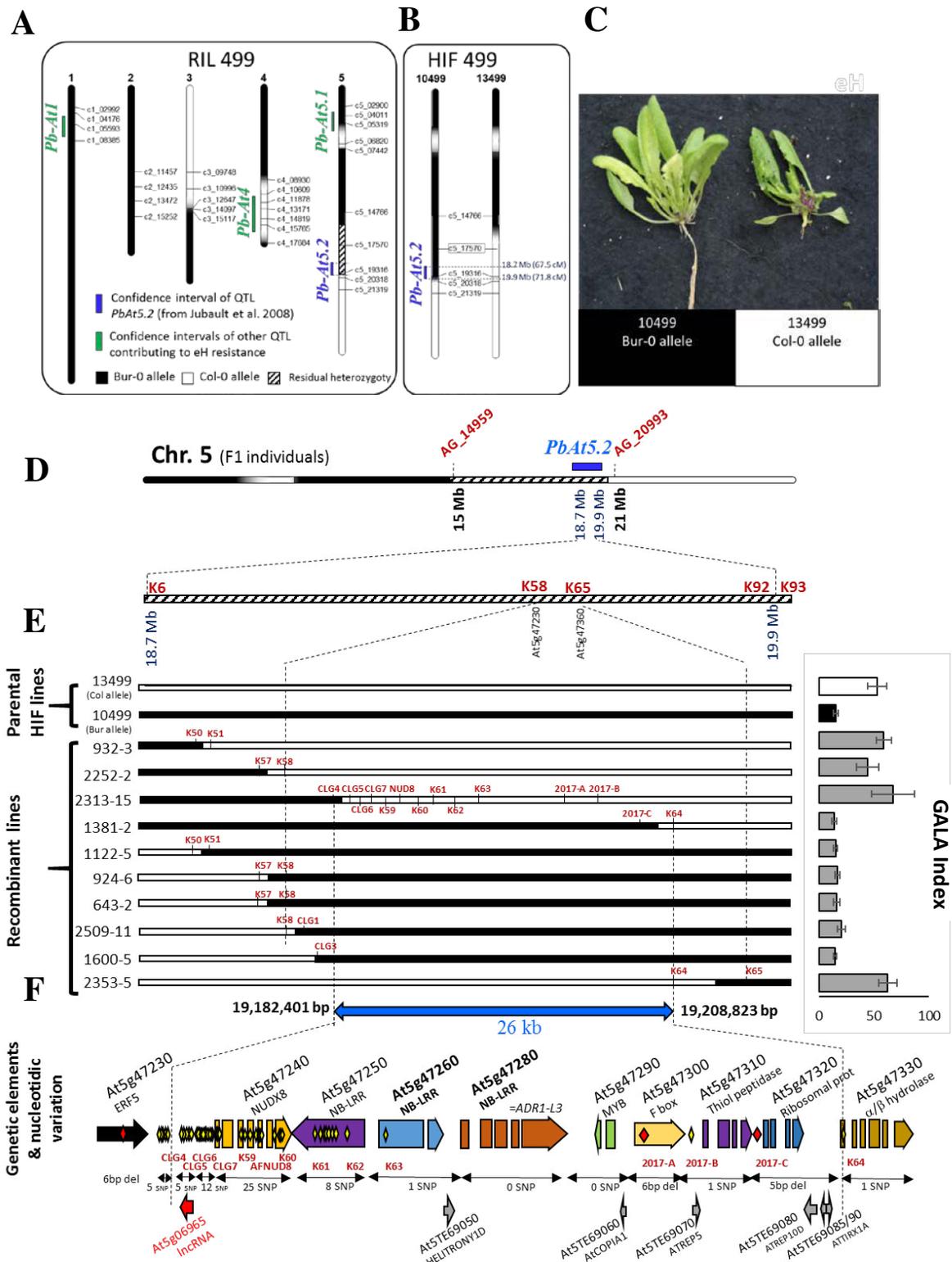


Figure IV.4. Fine mapping of *PbAt5.2*. (A). Genetic map and residual heterozygosity in the Recombinant Inbred Line (RIL) 499 derived from Bur-0 and Col-0 (from⁴⁶) and position of clubroot resistance QTL (from⁴³). Black: Bur-0 allele, White: Col-0 allele (B). Allele configuration at *PbAt5.2* in the two derived HIF lines 10499 and 13499 (C). Illustration of the partial resistance to eH isolate conferred by *PbAt5.2* in the RIL499 genetic background (D). First round of fine mapping: allelic structure in the F1 lines derived from reciprocal crosses between 10499 and 13499. 554 F3 lines with recombination in the confidence interval were screened using 10 SNP markers between AG_14959 and AG_20993. High density genotyping (94 SNP from K1 to K94), in a series of 88 recombinant F3 lines, and clubroot phenotyping in their bulked segregating F4 progenies, led to a new interval between markers K58 (At5g47230) and K65 (At5g47360). (E). Second round of fine mapping: Recombination positions in homozygous individuals obtained from selected recombinant lines. For each line, GALA index is indicated on the right panel. Genetic markers are indicated for each recombination position. Markers between CLG5 and 2017-C were used in every line, but indicated only for 2313-15 (F). New 26 kb interval of *PbAt5.2* between markers CL4 (excluded) and K64 (excluded), containing eight annotated ORF, six transposons and one lncRNA. Yellow and red diamonds indicate SNP and nucleotide deletions, respectively.

2.2. Results

Fine mapping of the locus *PbAt5.2* responsible for clubroot resistance. In a previous screen among *Arabidopsis thaliana* accessions, Bur-0 had been identified as partially resistant to eH isolate of *P. Brassicae*, when the canonical accession Col-0 is fully susceptible (illustrated in **Annexe III Figure S1**; first report in⁴⁴). Partial resistance in Bur-0 was associated with a slow development of root gall, a lower amount of pathogen in infected root tissues, and a reduced impact on leaf development compared to Col-0 (**Annexe III Figure S1**; see also⁴⁵). QTL controlling resistance to eH had been thereafter analysed⁴³ in a RIL progeny derived from an initial cross Bur-0 (NASC accession N1028) x Col-0 (NASC accession N1092) followed by 6 generation SSD (Recombinant Inbred Line set 20RV described in⁴⁶). Those previous works had led to the identification of four additive QTL (illustrated in the **Figure IV.4A**). Among them the resistance QTL *PbAt5.2*, on the chromosome 5, displayed a medium effect ($R^2=20\%$, Bur-0 allele providing partial resistance), with a rough position at 19.3 Mb, and a confidence interval containing 157 annotated sequences between At5g46260 and At5g47690. Among the RIL set 20RV from Simon *et al.* (2008), one RIL line (RIL499) displayed a single residual heterozygous region including markers c5_17570, c5_19313 and c5_20318, framed by the homozygous loci c5_14766 and c5_21319 (**Figure IV.4A**). Heterogeneous Inbred Family (HIF) lines 10499 and 13499 have been derived from RIL499 by the Institute Jean Pierre Bourgin (INRA Versailles, France). Those two near-isogenic lines then harbour identical combinations of Bur-0 and Col-0 homozygous genome sequences at every loci, except in the region between c5_14766 and c5_21319, where 10499 and 13499 are homozygous for Bur-0 or Col-0 alleles, respectively (**Figure IV.4B**). HIF line 10499 displayed a higher level of partial resistance to eH isolate compared to 13499 (**Figure IV.4C**, **Annexe III Figure S2**) thus corroborating the position of *PbAt5.2* in this interval (previously reported in⁴⁵).

The initial aim of the present work was the fine mapping of *PbAt5.2*, starting by the reciprocal crosses between HIF lines 10499 and 13499. Clubroot symptoms in the individuals of the F1 progeny were found identically high as those found in the susceptible parental line HIF 13499, suggesting that the resistance allele Bur-0 was recessive (**Annexe III Figure S3**). This conclusion was thereafter confirmed by the phenotyping of the segregating progeny (See details in **Annexe II**). The *PbAt5.2* resistance locus was thereafter fine mapped through several rounds of genotyping and clubroot phenotyping (**Figure IV.4D-F**, details in **Annexe II**), revealing a final confidence interval of 26 kb between the marker CLG4 (19,182,401 bp, in the promoter region of At5g47240), and the marker K64 (on SNP at position 19,208,823 bp, in At5g47330). This region contained eight annotated ORF (including the three NLR-encoding genes At5g47250/At5g47260/At5g47280), six transposons and one lncRNA (**Figure IV.4F**). The two F5 progeny lines 1381-2 and 2313-15 (harbouring the closest recombination events from both sides of the 26 kb interval, see **Figure IV.4E**) were further challenged with a series of additional *P. Brassicae* isolates (from pathotypes 1, 4 and 7 following the classification of⁴⁷). This

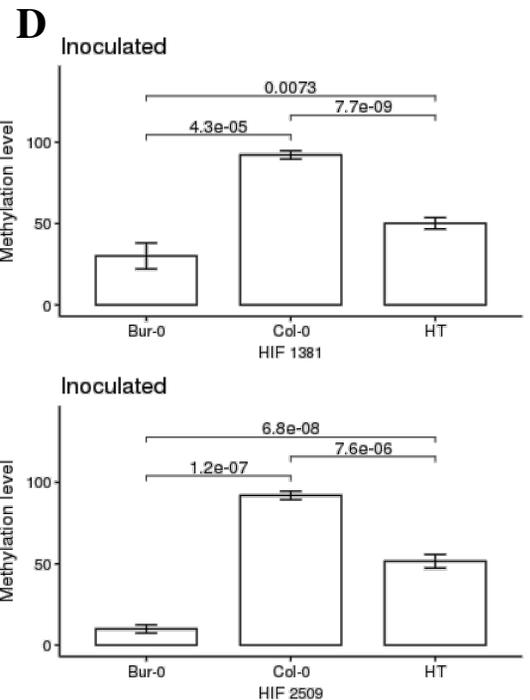
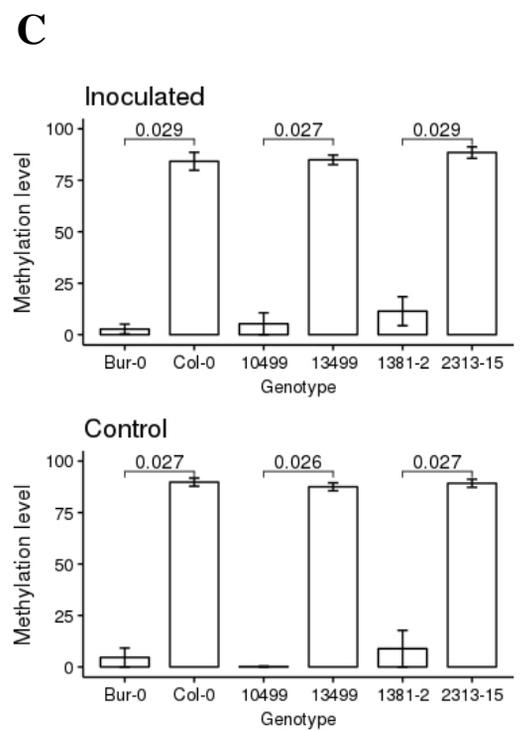
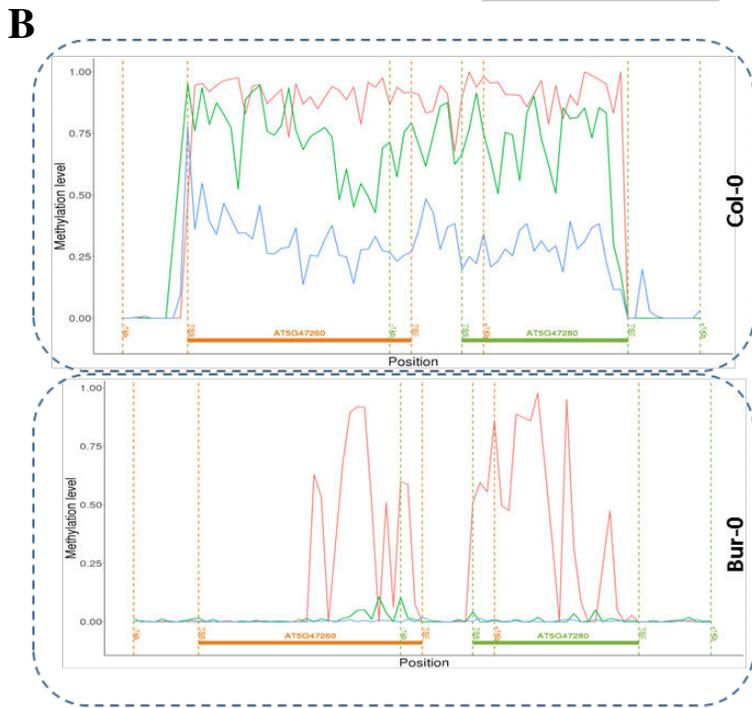
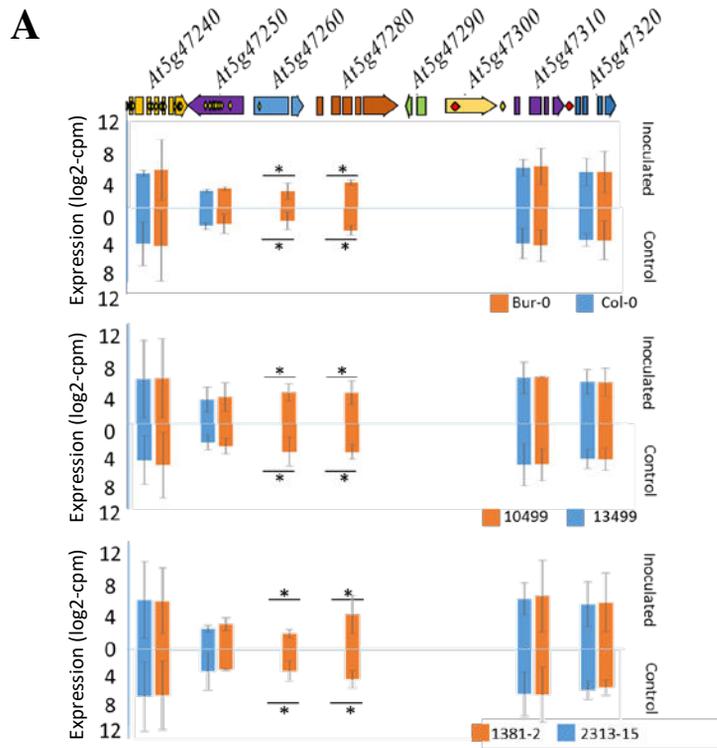


Figure IV.5. Identification of the two candidates under the *Pb-At5.2* region. (A). RNAseq expression (log₂ normalized cpm) of genes on the *PbAt5.2* region for the Col-0/Bur-0 accessions, the 13499/10499 HIF lines and 2313-15/1381-2 HIF lines in inoculated and control conditions. Statistical significant differences between genes expression represented by the stars traduce the FDR adjusted *p*-values. **(B).** Distribution of the methylation in the Col-0 and Bur-0 accessions using the bisulfite data from⁴⁸. Average methylation level was calculated within non overlapping 100-bp windows starting 1 kb before the TSS site of At5g47260 and stopping 1 kb after the TSE site of At5g47280. In red the methylation in CG context. In green the methylation in CHG context. In blue the methylation in CHH context. **(C).** Barplot of At5g47260 methylation profiles in Bur-0, Col-0 and the HIF 10499, 13499, 1381-2 and 2313-15. Methylation level was obtained using CHOP-qPCR. Vertical lines indicate the standard errors (N = 4) and *p*-values are shown (Wilcoxon test) **(D).** Barplot of At5g47260 methylation profiles in HIF 1381 and 2509. Methylation level was obtained using CHOP-qPCR. Vertical lines indicate the standard errors and *p*-values are shown (Wilcoxon test). Bur-0 indicates HIF with the homozygous Bur-0 epiallele at *PbAt-5.2*. Col-0 indicates HIF with the homozygous Col-0 epiallele at *PbAt-5.2*. HT indicates HIF which are heterozygous at *PbAt-5.2*. HIF 1381 Bur-0 (n = 12), HIF 1381 Col-0 (n = 22), HIF 1381 HT (n = 42), HIF 2509 Bur-0 (n = 23), HIF 2509 Col-0 (n = 16), HIF 2509 HT (n = 47).

highlighted the broad spectrum of the partial resistance conferred by the Bur-0 allele at this locus (**Annexe III Figure S4**).

Possible causal role of a stably inherited expression/methylation polymorphism affecting two NLR genes. An RNA-seq analysis was performed in Bur-0 and Col-0 accessions, in the parental lines 10499 and 13499 and in the F5 progeny lines 1381-2 and 2313-15 (harbouring the closest recombination events from both sides of the 26 kb interval, see **Figure IV.4E**). In the accession Col-0, sequencing reads were found for only four genes of this short list: At5g47240/At5g47250/At5g47310/At5g47320. At5g472310 and At5g47320 were both devoid of nucleotide variations, to the exception of one SNP and one indel in their promoter regions, respectively. The possible causal role of these variations was discarded, as those genes were expressed at similar levels whatever the presence of Bur or Col allele in the *PbAt5.2* region (**Figure IV.5A**). By contrast, most SNP variations in the region were associated with the *NUDX8* encoding gene At5g47240 and to a smaller extent to the adjacent NLR-gene At5g47250 (**Figure IV.4F**). Clubroot symptoms were however not different from Col-0 in the homozygous T-DNA mutant lines SALK_092325C (T-DNA in the At5g47240 gene) and WiscDsLoxHs110_09B (T-DNA in the At5g47250 gene), thus invalidating the idea that those genes may drive clubroot susceptibility in Col-0 (**Annexe III Figure S5**). The two non-TIR NLR-encoding genes At5g47260 and At5g47280 displayed one single non-synonymous SNP, and no SNP, respectively. RNA-seq analysis indicated that those two genes were constitutively expressed in the roots of Bur-0 accession and the genotypes 10499 and 1381-2 (with the Bur-0 allele, **Figure IV.5A**), but were almost not detectable in Col-0 accession, 13499 and 2313-15 (with the Col-0 allele).

To understand why the expression of the two NLR genes was significantly differentially expressed between Bur-0 and Col-0 alleles, we analysed the methylation level of these two genes in Bur-0 and Col-0 using the public bisulfite data⁴⁸. The genomic interval between 19,188,411 and 19,196,559, which includes At47260 and At547280 and the transposon At5TE69050 in-between, was highly methylated in CG, CHG and CHH contexts in Col-0, as CHH/RdDM genes, whereas the Bur-0 region was poorly methylated with only some CG methylation, as gene body methylation (gbM) genes⁴⁹ (**Figure IV.5A, B**). This contrasted methylation state between the two accessions was experimentally confirmed using CHOP qPCR with samples from plants cultivated in our laboratory, in both uninfected and infected conditions (**Figure IV.5C**). This epivariation was also found between the progeny HIF lines 10499 and 13499 and in the pair of HIF-derived near-isogenic lines 1381-2/2313-15 (**Figure IV.5C**). The ‘Col-like’ high-methylation pattern in the interval between At5g47260 and At5g47280 is systematically associated with a low expression of the two NLR genes and a lower level of partial resistance to *P. Brassicae* infection. To further investigate the inheritance behaviour of this epivariation and its penetrance on gene expression, the SNP status, methylation state and transcriptional expression of these two genes were then assessed in two progenies of 100 plant individuals, respectively derived from the

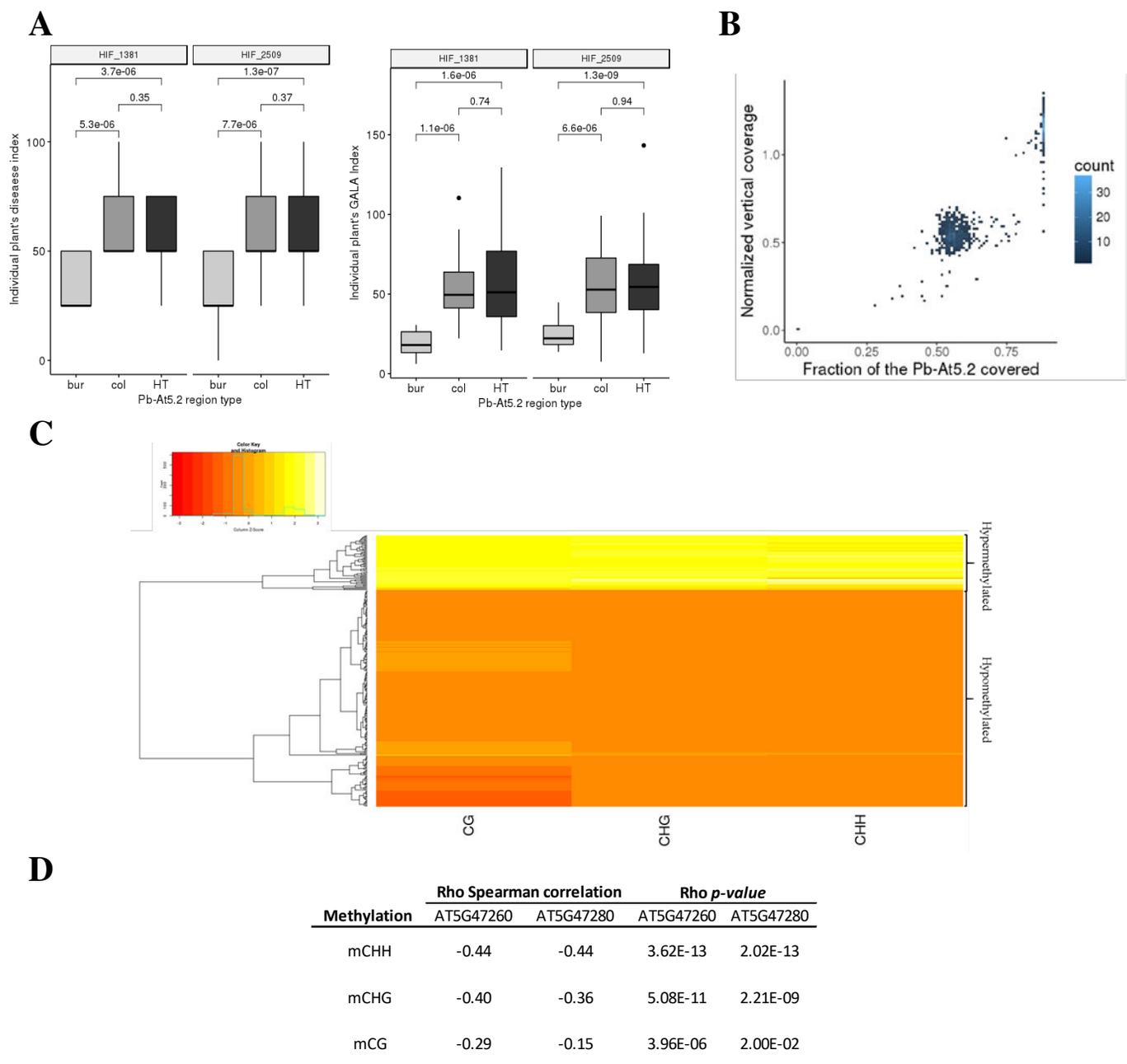


Figure IV.6. Natural variations at the *PbAt5.2* locus. (A). Boxplot of individual plant clubroot disease evaluation on the two HIF lines in segregation. Disease index and GALA index were calculated for each individual plant and their genetic state at the locus homozygote Bur-0 (bur), homozygote Col-0 (col) or heterozygote (HT) was determined by genotyping. Number of individuals in each class is as follows : HIF 1381 Bur-0 n = 15, HIF 1381 Col-0 n = 25, HIF 1381 HT n = 47; HIF 2509 Bur-0 n = 28, HIF 2509 Col-0 n = 18, HIF 2509 HT n = 52. Horizontal lines indicate the means. Vertical lines indicate the standard errors. Statistically different values (from wilcox test) are indicated in each boxplot. (B) Natural *Arabidopsis* accessions DNA-seq read coverage in the *PbAt5.2* region. DNA-seq data were derived from⁸¹. Vertical coverage in read percentage. Horizontal coverage: region covered by at least one read. i) One normalized vertical coverage < 0.25 and horizontal < 0.3 indicates the total absence of the locus. ii) One normalized vertical [0.25 - 0.8] and horizontal [0.3 - 0.8] indicates partial absence of the locus. iii) One normalized vertical > 0.8 and horizontal > 0.5 indicates the locus presence. *Pb-At5.2* region (chr5: 19185600 - 19200600). (C) Clustering of the accessions by their level of methylation on the *Pb-At5.2* region. Bisulfite data were obtained from the 1001 genome data project. Average of methylation level was calculated 1 kb before the TSS site of At5g47260 and stopping 1 kb after the TSE site of At5g47280 for each context. In total 58 accessions were hypermethylated like Col-0 and 229 were hypomethylated like Bur-0. (D) Spearman correlation between the epimutation methylation and the NLR expression in the 127 *Arabidopsis* accessions. For each NLR (At5g47260 and At5g47280) the correlation of their expression (RNAseq data of⁴⁸) and the methylation level in each context of the interval 1 kb before the TSS site of At5g47260 and stopping 1 kb after the TSE site of At5g47280.

selfing of heterozygous 1381 and 2313 lines (harbouring heterozygosity at the locus). The methylation state of the region in these two progenies was monitored by CHOP qPCR on At5g47260. The study of the methylation state in homozygous Bur-0, Col-0 and heterozygous HIF 1381 and 2313 highlighted an intermediate level of methylation in heterozygous HIF compared to homozygous HIF (**Figure IV.5D**). The evaluation of plant disease for each individual plant in the two progenies indicated a clear 3:1 mendelian segregation of the partial resistance phenotype. Levels of clubroot symptoms in individuals harbouring only one Bur-0 resistance allele were not distinguishable from individuals harbouring the two susceptible alleles Col-0, thus confirming that Bur-0 presents a recessive partial resistance allele (**Figure IV.6A**).

Altogether, those results suggested a strong link between *PbAt5.2* driven partial resistance to *P. Brassicae* and a stably inherited epivariation, which controls the expression of two NLR genes at this locus. The possible causal influence of nucleotidic variations in this small region (including one non-synonymous SNP in At5g47260) was however not fully excluded. Thus, the respective roles of epivariations and/or nucleotidic variations in the control of clubroot resistance was further investigated by relying on the natural distribution of the methylation status at At5g47260/At5g47280 across natural *Arabidopsis* accessions.

The Bur-like hypo-methylated epiallele is well represented among Arabidopsis accessions and contributes to reduce clubroot symptoms. A preliminary analysis of the genomic sequence of those two genes from the 1001 genome data highlighted genomic structural variations at this locus in a number of *Arabidopsis* accessions, since they displayed an absence of clear mapping of genomic reads on the two genes At5g47260 and At5g47280. It was thus necessary to focus our analysis on a set of *Arabidopsis* accessions showing a genomic structure similar to Col-0 and Bur-0 accessions. For that, DNA seq reads of 1135 accessions were mapped on the Col-0 reference genomic sequence of the *PbAt5.2* region, revealing 401 accessions with a normalized vertical and horizontal coverages superior of 0.6 (**Figure IV.6B; Annexe III Table S1**). This indicated that the Col-0/Bur-0-like genomic structure at this locus was present with a high frequency (Minor Allele Frequency MAF ~0.37). Frequency also varied following geographic groups, from 52.7 % in the Spain group to 17.7 % in the Asia group (**Annexe III Table S2**). Among those 401 accessions, the methylation state of the two NLR genes was analysed on the 287 accessions for which public bisulfite data are available⁴⁸. Based on methylation data, those 287 accessions split into two clearly distinct subsets of genotypes (**Figure IV.6C**): one composed of 229 accessions with a Bur-0-like methylation state at the locus, and another subset with 58 accessions displaying a Col-0-like highly methylated state (**Figure IV.6C, Annexe III Table S3**). Mean proportion of Col-like hypermethylated state varied considerably along geographic origin of accessions: *i.e.* 3.7 % among the 54 accessions from the Spain group contrasting with 54.5 % among the 55 accessions from the Central Europe group (**Annexe III Table S2**). Among this set of 287 lines,

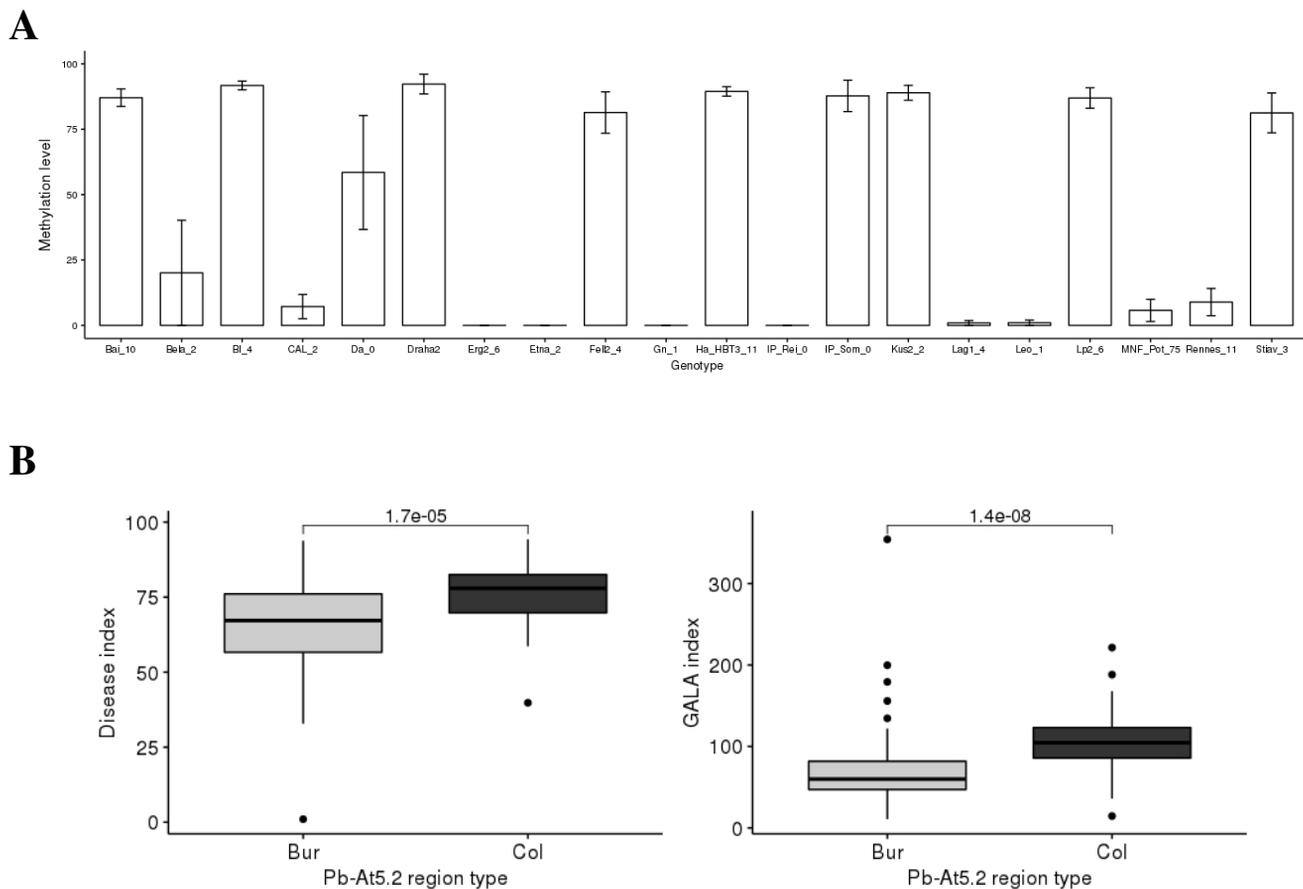


Figure IV.7. Methylation profiles and disease evaluation on *Arabidopsis thaliana* accessions. (A). Barplot of At5g47260 methylation profiles in 20 ecotypes. Methylation level was obtained using CHOP-qPCR. Vertical lines indicate the standard errors (N = 4). (B). Boxplot of disease evaluation carried out on the 127 *Arabidopsis* accessions according to their epiallele. For each trait (Disease index and GALA index) and accession the mean was obtained by modelling raw data of two biological replicates with two blocks. Horizontal lines indicates the means for all accessions with the same epiallele. Vertical lines indicate the standard errors. Statistically different values (from Wilcox test) are indicated in each boxplot. In total 43 accessions have Col-0 like epiallele and 84 have Bur-0 like epiallele.

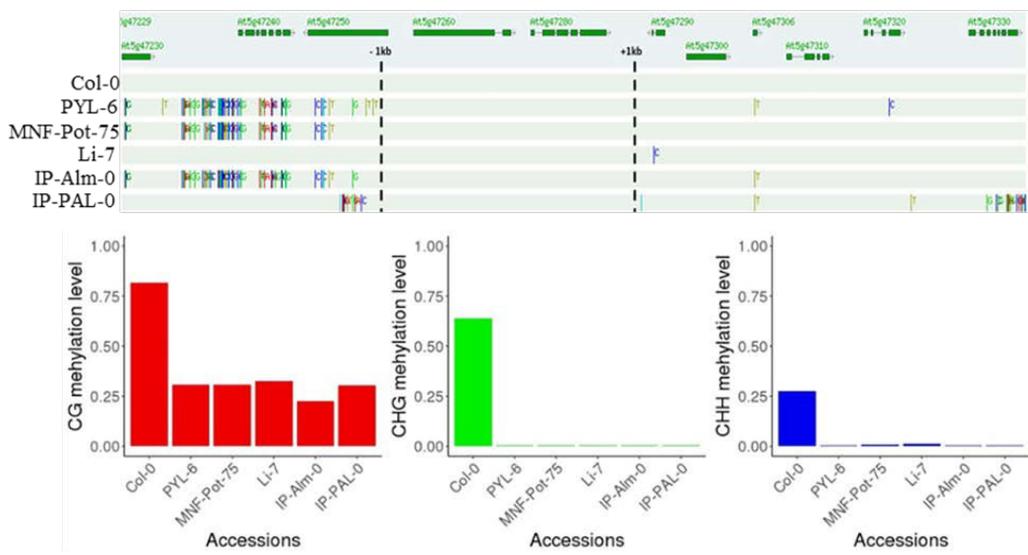
the comparison of publicly available bisulfite and RNA-seq data from 253 accessions revealed a sharp negative correlation between methylation level (in each context) and gene expression for At5g47260 and At5g47280 (**Figure IV.6D**). The methylation state of the two NLR genes was confirmed on the roots infected and not-infected, in a subset of 24 accessions, using CHOP qPCR for methylation state (**Figure IV.7A**).

The apparent balanced polymorphism between Col-like and Bur-like epialleles across *Arabidopsis* accessions (about 20 % of Col-like epialleles among the accessions with the Col/Bur-like structure at the locus) offered an interesting genetic material for further investigation of possible causal role of the epivariation in the control of clubroot partial resistance. A selection of 127 accessions, including 42 accessions with the Col-like epiallele and 84 accessions with the Bur-like epiallele, were inoculated with the *P. Brassicae* eH isolate (two biological replicates composed of two blocks, **Annexe III Table S4**). Based on two different variables used to evaluate disease susceptibility (*i.e.* Disease Index and the GALA index), this experiment showed that low methylation state type Bur-0 at the At5g47260/At5g47280 locus was associated with an enhanced partial resistance level (**Figure IV.7B**). Mean GALA disease symptoms was 69.1 in Bur-like accessions and 102.3 in Col-like accessions, a delta similar to those found between HIF lines 10499 and 13499. This result corroborated and extended the conclusions from our analysis of the Bur-0 x Col-0 progeny.

PbAt5.2 epivariation is independent of cis- genetic variations. Following the nomenclature proposed by⁵⁰, epivariations can be categorized either as 1/obligatory/facultative epialleles (*i.e.* methylation variations caused or favoured by a cis-nucleotide variation) or 2/pure epialleles. At the locus *PbAt5.2*, the transposon At5TE69050 was identically present in both parental genotypes, with no sequence variation that might have been the primary cause of the variation of methylation on the two adjacent genes. The only cis-nucleotide variation between Bur-0 and Col-0 accessions was a SNP in the coding sequence of At5g47260. This SNP was however not likely the primary driver of the methylation patterns, as five accessions with a Bur-0-like methylation pattern (among the 229 analyzed), also harboured a sequence strictly identical to Col-0 in the region (-1 kb before At5g47260 to +1 kb after At5g47280, **Figure IV.8A**). Further GWAS analysis among the 287 accessions also failed to identify any nucleotide variation in the region between At5g47240 and At5g47330, which would explain the variation of methylation. These suggested that the switch of methylation between At5g47260 and At5g47280 was independent of cis-natural sequence variations.

The hypermethylated epivariant is maintained by the RNA-independent pathway despite targeted by smRNA. Screening of the small RNA in the epiallele region using publicly available data⁵¹ has shown that in Col-0, the At5g47260/At5g47280 region is targeted by 21, 22, 23 and 24-nt small RNA. Analyses in the non-inoculated roots of Col-0 and Bur-0 accessions at a time point corresponding to 14 dpi confirmed the presence of a large range of smRNA targeting the Col-0 epiallele, but also highlighted

A



B



C

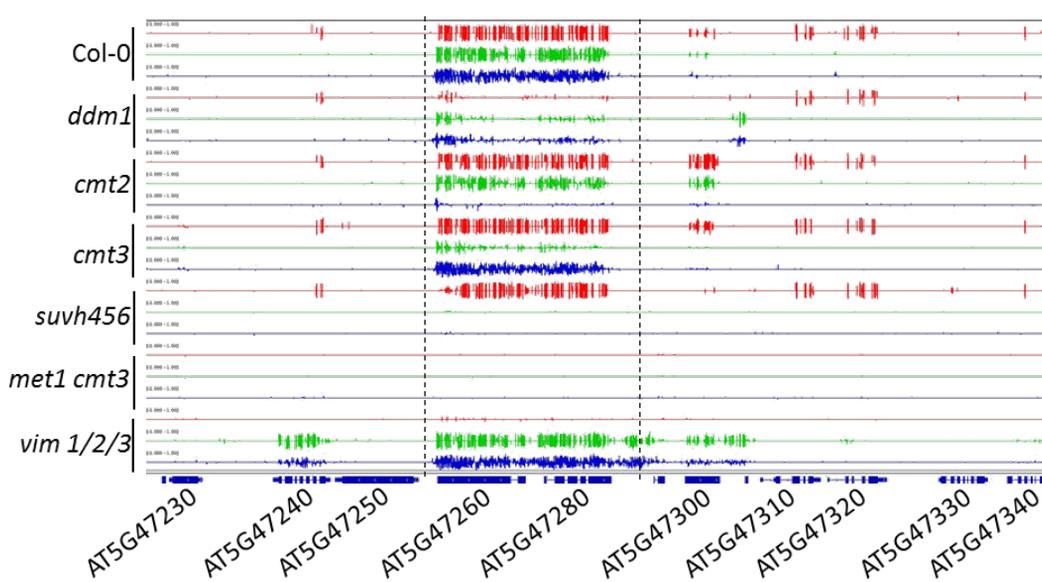


Figure IV.8. Genomic and epigenetic characterisation of *Pb-At5.2* epimutation. (A). Characterisation of the link between sequence variant and methylation state in *Pb-At5.2* region. Sequence variant data were obtained from the signal SALK genome browser based on 1001 genome data. Average of methylation level was calculated 1 kb before the TSS site of *At5g47260* and stopping 1 kb after the TSE site of *At5g47280*. (B). Mapping of smRNA-seq read. Read were obtained from roots of Col-0 and Bur-0 accessions at two time points 7 and 14 days after inoculation. (C). Methylation of *PbAt5.2* region in Col-0 and in mutants in genes involved in epigenetic pathways⁵¹. In red the methylation in CG context. In green the methylation in CHG context; in blue the methylation in CHH context.

their absence in Bur-0 (**Figure IV.8B**). Maintenance of the DNA methylation involves an RNA-dependent pathway (involving siRNA and DRM proteins) and RNA-independent pathways. An analysis of publicly available methylation profiles in a series of classical mutants affected for methylation maintenance highlighted the major role of a series of genes involved in the RNA-independent methylation pathway (*DDMI*, *MET1*, *CMT2/3* and *SUVH456*) in the maintenance of the Col-like epiallele (**Figure IV.8C**). By contrast, mutations affecting the RNA-dependent methylation pathway did not affect hypermethylation of the locus in the accession Col-0 (**Annexe III Figure S6**). Mutant methylation profiles ask thus the question of the role of the smRNA which target the Col-like hypermethylated region whereas the methylation maintenance is already ensured by the RNA-independent pathway.

2.3. Discussion

Only a limited number of resistance QTL have been molecularly characterized so far¹⁵. Detection and fine mapping of resistance QTL can be tricky not only because of the challenge of measuring small variations of partial resistance in large number of progeny individuals, but also because resistance QTL can be sensitive to environmental variations⁵²⁻⁵⁴. However, technical issues may have been only part of the problem and recent developments in the field of epigenetics also suggests that the detection of an important part of inherited resistance factors could be also missed (or fluctuant) because regular genetic approaches are based only on nucleotide variations. In the present work, we report for the first time an example where an epigenetically driven expression polymorphism substantially contributes to the natural diversity of quantitative responses to a plant pathogen.

Many examples of transgenerational inheritance of epigenetic variants, in particular ‘facultative epialleles’, are metastable, *i.e.* they can be reversed by stochastic or non-identified factors⁵⁵. Stability over multiple generations is a primary concern from both evolutionary and breeding perspectives. In the particular example studied in this work, well before we understood its epigenetic nature, we focused our fine mapping efforts on the clubroot resistance QTL *PbAt5.2* precisely because it was robustly detected in all our previous QTL investigations with *Arabidopsis*. This included two independent segregating progenies derived from Bur-0 and Col-0⁴³ and additional works with HIF lines 10499/13499^{45,56}. The high level of methylation and absence of gene expression of At5g47260 and At5g47280 in Col-0 was found in series of publicly available data obtained in different laboratories, with a diversity of plant tissues and conditions (efpBrowser, Trava...). It was also confirmed by our own data generated with inoculated roots, non-inoculated roots and leaf samples. Finally, this methylation pattern was also robustly found in multiple replicates of individual plants. Thus, *PbAt5.2* can be unreservedly classified in the category of stable epialleles.

Beyond the example of *SNCI* (discussed earlier in the introduction), there has been only one recent report of a plant disease resistance caused by an inherited methylation variant affecting the expression

of a resistance-related gene⁵⁷. In their work, a stable expression polymorphism (between Ler-0 and Ag-0 accessions) on the TIR-only encoding gene *RBA1* (At1g47370) affected ETI-responses to the *Pseudomonas syringae* effector *hopBA1*. This expression polymorphism was linked to the nearby presence/absence of a transposon and to the apparently resulting *MET1-3* dependent methylation of the promoter region of this gene. This epimutation on the promoter was reversed when the transposon was segregated away (**Figure S2** in⁵⁷), thus suggesting that structural sequence variations on this transposon was the actual causal polymorphism, probably driving the expression of the adjacent gene through a methylation spreading process (similarly to the archetypal gynoecy locus reported in⁵⁸).

In our case however, the low/high methylation polymorphism in the region between At5g47260 and At5g47280 was not linked to any nucleotide/structural variation on At5TE69050 whose location (the transposon is surrounded by the two genes) and sequence are identical in Bur-0 and Col-0. Taking into account also that the three genetic elements At5g47260/At5TE69050/At5g47280 were all abundantly targeted by siRNA in Col-0, one can hypothesize that the methylation pattern of this transposon probably follows the same process affecting its two neighbours, more than it drives them. One major remaining question to clarify was the possible role that could play other nucleotide variations in the 26 kb interval, on the transgenerational inheritance of the hyper/hypomethylation state. Our efforts through SNP-based GWAS analysis of the epiallele variation (considered here as a phenotype) failed at identifying any nucleotide variation that would correlate with the methylation state at this locus. This suggests that *PbAt5.2* behaves as a genuine ‘pure epiallele’ (as defined by⁵⁰), *i.e.* the divergence between the two methylation variants should have emerged from unknown past events, and was thereafter autonomously inherited, *causa sui*, without any additional genetic cause.

From available genomic and epigenomic data from the ‘1001 genome project’, it can be extrapolated that the Bur-like clubroot resistance epiallele is present in about half of the accessions from the ‘Relict’, ‘Spain’, and ‘Italy/Balkans/Caucasus’ groups and 39 % of the accessions from the ‘North Sweden’ group (**Annexe III Table S2**), but stays below 10 % in the ‘Germany’ group. By contrast, the clubroot susceptibility Col-like epiallele was absent from the accessions from ‘Relict’ and ‘North Sweden’ groups, but reached 22 % in the ‘Central Europe’ group. This sharp geographical structuration suggests that both epialleles confer fitness gains depending on environmental contexts, but it cannot be evidently related to the distribution of clubroot incidence on *Brassica* cultures (usually low in the warm southern European regions). Keeping in mind that NLR can detect unrelated effectors from distinct microbial species⁵⁹, and echoing to the work of⁶⁰, we can hypothesise that the maintenance of this epivariation in natural populations may reflect additional *PbAt5.2* roles against other plant pathogens (besides the control of clubroot infection).

At5g47260 and At5g47280 belong to a small heterogeneous cluster of three non-TIR-NLR located on chromosome 5, not far from the largest NLR-hot spot in the *Arabidopsis* genome². None of those three

genes has been previously involved in plant pathogen interactions before. At5g47260 belongs to the large sub-group CNL-B, which notably contains *RPS5*². At5g47280 is 66 % identical to the paradigmatic gene ACTIVATED DISEASE RESISTANCE 1 (*ADR1*, At1g33560) and belongs to the family of ADR-like genes, a small non-TIR clade which also includes in *Arabidopsis* *ADR1-L1* (At4g33300) and *ADR1-L2* (At5g04720). Present in all plant species, these evolutionary conserved genes share a RPW8 domain on N-ter and this clade has been thus referred as CCR-NBS-LRR⁶¹⁻⁶³. Consistently with their unusually low level of variations in the LRR domain, ADR1-like proteins do not play regular ‘specific sensor’ roles (that are typically played by NLR in ‘gene-for-gene’ interactions). They are now usually referred as ‘helpers NLR’ thanks to the pivotal roles they play as downstream signaling components for multiple TNL-mediated ETI, per part through the control of a regulatory loop of SA-driven responses^{61,62,64-66}. The first genome-wide analysis of NLR genes had initially classified *ADR1-L3* (At5g47280) in this clade as a pseudogene because of the lack of the 5’ part encoding the RPW8 domain² and very low expression levels in Col-0. It has since been generally considered as an outsider and was therefore not included in previous functional studies of the ADR1-like gene family^{62,65}. However, based on several examples where NLR proteins can modulate resistance levels in spite of the loss of important conserved motifs,⁶² did not fully exclude a possible role for *ADR1-L3*. The results presented in the present work would tend to corroborate this assumption.

There are few examples of tandem NLR genes coding for pairs of proteins that function as heterodimers⁶⁷⁻⁶⁹. Similarly, the proteins encoded by the two epigenetically joint-regulated genes At5g47260/80 may function together in the control of cell defence responses during clubroot infection. Interestingly, and although the underlying molecular mechanisms are not known, the canonical example of TIR-NLR heterodimer *RRS1/RPS4* corresponds to a recessive resistance locus, similarly to *PbAt5.2*. Among the 1001 genome genotypes, not a single accession was found methylated only for one of the two genes, and it was thus not possible to use such genetic resources to get any answer about the impact of the silencing of only one of the two genes. Additional molecular work (including the CRISPR-Cas9-based nucleotide deletions or demethylation, in the Bur-0 accession, and the analysis of protein-protein interactions) would be necessary to assess the individual role and possible partnership of these two genes.

From the data from⁷⁰, we found that 32 genes from the NLR family (list of²) are methylated in Col-0, when 77 are unmethylated (others being ambiguous). It would be now interesting to start a careful genome wide analysis of methylation profiles on all NLR-genes among *Arabidopsis* accessions, which would take into account the structural diversity of all these individual genes (supported by additional targeted resequencing of NLR-loci). Beyond, the intraspecific diversity of methylation patterns of NLR and RLK/RLP genes in plants, their heritability, and their consequences on plant biotic interactions, might deserve further attention in future studies.

2.1. Methods

Plant materials and growth conditions. Heterogeneous Inbred Family (HIF) lines 10499 and 13499 and their parental accessions Col-0 (186AV) and Bur-0 (162AV) were obtained from the Institute Jean Pierre Bourgin (INRA Versailles, France). *Arabidopsis thaliana* accessions were all purchased from the Nottingham Stock Center. The panel of 127 accessions was selected according to their methylation level at interest region⁴⁸.

All accessions used in this study were listed in **Annexe III Table S4**. Seed germination were synchronised by placing seeds on wet blotting paper in Petri dishes for two days at 4 °C. Seeds were sown individually in pots (four cm diameter) containing a sterilised mix (by autoclaving) composed of 2/3 compost and 1/3 vermiculite. Growth chamber controlled conditions of 16 h light ($110\mu\text{mol m}^{-2}\text{s}^{-1}$) at 20 °C and 8 h dark at 18 °C were used to grow plants. The 127 accessions of *Arabidopsis* and HIF were assessed to *P. Brassicae* in two biological replicates in randomised complete block designs (with two blocks per replicate). The *Arabidopsis* accessions Col-0, Bur-0 and the HIF 10499, 13499, 1381-2 and 2313-15 used in RNA-seq and smRNA-seq approaches were assessed to *P. Brassicae* or in uninfected condition in three randomised blocks. Almost all clubroot tests were performed with the isolate eH of *P. Brassicae* described by⁷¹ which belongs to the most virulent pathotype P1. The resistance spectrum of *PbAt5.2* was also assessed with a series of additional isolates Pb137-522, Ms6, K92-16 and P1⁺. For every isolate used in this study, pathotype was validated in every experiment using the differential host set according to⁴⁷, including also two genotypes of *B. oleracea* ssp. *acephala* C10 and CB151. One ml of resting spore suspension (10^7 spores.ml⁻¹) prepared according to⁷² was used for pathogen inoculation 10 days (d) after germination (stage 1.04⁷³). This inoculum was applied on crown of each seedling.

Phenotyping. HIF and *Arabidopsis* accessions were phenotyped three weeks after inoculation (21 days post inoculation (dpi) for their susceptibility to *P. Brassicae*. Plants were thoroughly rinsed with water and photographed. Infected roots were removed and frozen in liquid nitrogen. In order to validate the involvement of the epivariation in the plant response to the pathogen two related phenotypic disease traits (the disease index and the GALA index) previously used were measured^{43,74,75}. Disease index calculation was carried out according to⁷² and the GALA index was calculated according to⁷⁴.

Fine mapping of the loci responsible to clubroot resistance. Fine mapping of *PbAt5.2* was performed starting from crosses between HIF lines 10499 and 13499, followed by successive rounds of genotyping and clubroot phenotyping in subsequent plant generations (full details are given in **Annexe II**).

RNA isolation, mRNA Sequencing and differentially gene expression determination. Total RNA extraction was carried out on frozen and lyophilized roots (collected 14 days after inoculation) using the mirVana™ miRNA Isolation Kit (Invitrogen) according to the manufacturer's instructions. After

extraction, the RNA were quantified using the NanoDrop ND-1000 technologies and their quality was controlled using the RNA 6000 assay kit total RNA (Agilent). Samples with RIN superior or equal to seven were used for sequencing. Construction of cDNA sequencing libraries and the sequencing were performed by the NGS platform of Marie Curie Institute of Paris. For each sample, one Paired-end (100 x 100 bp) sequencing using HiSeq 2500 technology was carried out. For each sample Paired-End reads were aligned on TAIR10 genome annotation and assembly of Columbia *Arabidopsis thaliana* genome concatenated with *P. Brassicae* genome using STAR software⁷⁶ (Version 2.5.3.a). Alignment conditions were selected according to *Arabidopsis* genome. A maximum of five multiple read alignments were accepted and for each alignment no more than three mismatches were enabled. Once alignment performed, BAM files were used to carry out read counts using the function Counts of Feature Counts software (Version 1.4.6) and the TAIR10 gff of *Arabidopsis* concatenated with the gff of *P. Brassicae*. The determination of the differentially expressed genes was carried out using the package EdgeR⁷⁷ on the R software⁷⁸ (version 3.3.0). Raw counts obtained as described previously were used as enter data in EdgeR. Once CPM (Counts Per Million) determined, only genes with at least one CPM in three samples were conserved. Normalization of expression signals were carried out using the TMM method (Trimmed Mean of M-values) of the CalcNormFactors function of Edge R. Finally, the genes differentially expressed were determined using the decideTests function of Edge R using one minimum fold change between -1.5 and 1.5.

Small RNA isolation, sequencing, clusterization and differential presence determination. Small RNA extraction was carried out on frozen and lyophilized roots (collected 14 days after inoculation) using the mirVana™ miRNA Isolation Kit (Invitrogen) according to the manufacturer's instructions. After extraction, the small RNA was quantified using the NanoDrop ND-1000 technologies and their quality was controlled using the Small RNA assay kit (Agilent). Samples with RIN superior or equal to seven were used for sequencing. Construction of cDNA sequencing libraries and the sequencing were performed by the NGS platform of Marie Curie Institute of Paris. For each sample, one Single-end (50 bp) sequencing using Illumina HiSeq 2500 technology was carried out. For all sample Single-End reads were aligned on TAIR10 genome annotation and assembly of Columbia *Arabidopsis thaliana* genome concatenated with *P. Brassicae* genome using STAR software⁷⁶ (Version 2.5.3.a), counts and clustered using Shortstack software⁷⁹. The determination of the differential presence of smRNA was carried out using the package EdgeR⁷⁷ on the R software⁷⁸ (version 3.3.0). Raw counts obtained as described previously were used as enter data in EdgeR. Once CPM (Counts Per Million) determined, only genes with at least one CPM in three samples were conserved. Normalization of expression signals were carried out using the TMM method (Trimmed Mean of M-values) of the CalcNormFactors function of Edge R. Finally, the small RNA differentially present in gene were determined using the decideTests function of Edge R using one minimum fold change between -1.5 and 1.5.

RNA isolation and Rt-qPCR analysis. Accessions and HIF total RNA were extracted from the root of 21 days post infection plants previously lyophilised using the trizol extraction protocol. Samples with residual traces of DNA were treated with DNase (Promega ref M6 10A). Before reverse transcription of RNA to cDNA by the SuperScript II (Invitrogen), RNA quality was checked by electrophoresis gel migration. RT-qPCR was carried out in the thermocycler LightCycler® 480 (Roche) on cDNA obtained as described above. Gene expression was normalized using as reference two *Arabidopsis* genes defined as stable during infection using RNA-seq data (At1g54610, At5g38470). Primer sets were designed for each gene and were listed in **Annexe III Table S5**.

CHOP PCR and qPCR assays. Gene methylation profiles were investigated using McrBC enzyme (M0272L, BioLabs®). Conditions used to treat DNA with McrBC was 0.5 µL BSA (20 mg/mL), 0.5 µL GTP (20 mM), 5 µL NEB (10X) and 0.2 µL McrBC (10000 U/mL) to treat 40 ng of DNA. For CHOP PCR and qPCR 2 ng of digested and undigested DNA was used. For the CHOP PCR the temperature conditions were adjusted according to primer design and 35 cycles of amplification were used. To determine the methylation state of the targeted region, each sample is digested and undigested (control) by McrBC before amplification. For the CHOP qPCR the temperature conditions were adjusted according to primer design and 30 cycles of amplification were used. Methylation level of the target region was calculated in percentage of molecules lost through McrBC digestion according to⁸⁰ with the formula: $(1 - (2^{-(Ct\ digested\ sample - Ct\ undigested\ sample)})) * 100$. The percentage of DNA methylation for At5g13440 and At5g47400, was calculated in all CHOP qPCR as control. Primer sets designed for each gene were listed in **Annexe III Table S5**.

Published data. The DNaseq data, RNAseq data, variant sequence and bisulfite data of the natural accessions studied were obtained to the previous studies^{48,81} archived at the NCBI SRA number SRP056687 and the NCBI GEO reference: GSE43857, GSE80744. The bisulfite data and small RNA data of *Arabidopsis* mutants studied were obtained to the previous study⁵¹ archived at the NCBI GEO reference: GSE39901.

Statistical analysis. Data were statistically analysed using the R program⁷⁸.

2.2. References

1. Stahl, E. A., Dwyer, G., Mauricio, R., Kreitman, M. & Bergelson, J. Dynamics of disease resistance polymorphism at the Rpm1 locus of *Arabidopsis*. *Nature* **400**, 667 (1999).
2. Meyers, B. C., Kozik, A., Griego, A., Kuang, H. & Michelmore, R. W. Genome-wide analysis of NBS-LRR–Encoding genes in *Arabidopsis*. *Plant Cell* **15**, 809 (2003).
3. Clark, L. B. *et al.* Systematic mutagenesis of the leucine-rich repeat (LRR) domain of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in CCR4 deadenylase activity. *Journal of Biological Chemistry* **279**, 13616–13623 (2004).

4. Innes, R. W. *et al.* Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. *Plant Physiology* **148**, 1740–1759 (2008).
5. Yue, J.-X., Meyers, B. C., Chen, J.-Q., Tian, D. & Yang, S. Tracing the origin and evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes. *New Phytologist* **193**, 1049–1063 (2012).
6. Shao, Z.-Q. *et al.* Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiology* **170**, 2095–2109 (2016).
7. Maekawa, T., Kufer, T. A. & Schulze-Lefert, P. NLR functions in plant and animal immune systems: so far and yet so close. *Nat. Immunol.* **12**, 817–826 (2011).
8. Jones, J. D. G., Vance, R. E. & Dangl, J. L. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, (2016).
9. Thomas, C. M. Genetic and molecular analysis of tomato Cf genes for resistance to *Cladosporium fulvum*. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **353**, 1413–1424 (1998).
10. Xiao, S. *et al.* Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. *Science* **291**, 118–120 (2001).
11. Larkan, N. J. *et al.* The *Brassica napus* blackleg resistance gene LepR3 encodes a receptor-like protein triggered by the *Leptosphaeria maculans* effector AVRML1. *New Phytologist* **197**, 595–605 (2013).
12. Ernst, K. *et al.* The broad-spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. *The Plant Journal* **31**, 127–136 (2002).
13. Qu, S. *et al.* The broad-spectrum blast resistance gene *Pi9* encodes a nucleotide-binding site–leucine-rich repeat protein and is a member of a multigene family in rice. *Genetics* **172**, 1901 (2006).
14. Pilet-Nayel, M.-L. *et al.* Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Frontiers in plant science* **8**, 1838 (2017).
15. Nelson, R., Wiesner-Hanks, T., Wissner, R. & Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nature Reviews Genetics* **19**, 21 (2017).
16. Hayashi, N. *et al.* Durable panicle blast-resistance gene Pb1 encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through local genome duplication. *The Plant Journal* **64**, 498–510 (2010).
17. Xu, X. *et al.* Rice blast resistance gene *Pikahei-1 (t)*, a member of a resistance gene cluster on chromosome 4, encodes a nucleotide-binding site and leucine-rich repeat protein. *Molecular breeding* **34**, 691–700 (2014).
18. Fukuoka, S. *et al.* Multiple functional polymorphisms in a single disease resistance gene in rice enhance durable resistance to blast. *Scientific Reports* **4**, 4550 (2014).

19. Debieu, M., Huard-Chauveau, C., Genissel, A., Roux, F. & Roby, D. Quantitative disease resistance to the bacterial pathogen *Xanthomonas campestris* involves an *Arabidopsis* immune receptor pair and a gene of unknown function. *Molecular plant pathology* **17**, 510–520 (2016).
20. Diener, A. C. & Ausubel, F. M. Resistance to *Fusarium oxysporum*, a dominant *Arabidopsis* disease-resistance gene, is not race specific. *Genetics* **171**, 305 (2005).
21. Hurni, S. *et al.* The maize disease resistance gene *Htn1* against northern corn leaf blight encodes a wall-associated receptor-like kinase. *Proceedings of the National Academy of Sciences* **112**, 8780–8785 (2015).
22. Huard-Chauveau, C. *et al.* An atypical kinase under balancing selection confers broad-spectrum disease resistance in *Arabidopsis*. *PLoS genetics* **9**, e1003766 (2013).
23. Lai, Y. & Eulgem, T. Transcript-level expression control of plant NLR genes. *Mol. Plant Pathol.* **19**, 1267–1281 (2018).
24. Li, X., Kapos, P. & Zhang, Y. NLRs in plants. *Current Opinion in Immunology* **32**, 114–121 (2015).
25. Zhang, X.-C. & Gassmann, W. Alternative splicing and mRNA levels of the disease resistance gene RPS4 are induced during defense responses. *Plant physiology* **145**, 1577–1587 (2007).
26. Shivaprasad, P. V. *et al.* A microRNA superfamily regulates nucleotide binding site–leucine-rich repeats and other mRNAs. *The Plant Cell* **24**, 859–874 (2012).
27. Fei, Q., Xia, R. & Meyers, B. C. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* **25**, 2400 (2013).
28. Deng, Y., Liu, M., Li, X. & Li, F. microRNA-mediated R gene regulation: molecular scabbards for double-edged swords. *Science China Life Sciences* **61**, 138–147 (2018).
29. Palma, K. *et al.* Autoimmunity in *Arabidopsis* *acd11* is mediated by epigenetic regulation of an immune receptor. *PLoS pathogens* **6**, e1001137 (2010).
30. Xia, S. *et al.* Regulation of transcription of nucleotide-binding leucine-rich repeat-encoding genes *SNCI* and *RPP4* via H3K4 Trimethylation. *Plant Physiol.* **162**, 1694 (2013).
31. Zou, B., Yang, D.-L., Shi, Z., Dong, H. & Hua, J. Monoubiquitination of Histone 2B at the disease resistance gene locus regulates its expression and impacts immune responses in *Arabidopsis*. *Plant physiology* **165**, 309–318 (2014).
32. Tsuchiya, T. & Eulgem, T. An alternative polyadenylation mechanism coopted to the *Arabidopsis* *RPP7* gene through intronic retrotransposon domestication. *Proc Natl Acad Sci USA* **110**, E3535 (2013).
33. Li, Y., Tessaro, M. J., Li, X. & Zhang, Y. Regulation of the expression of plant Resistance Gene *SNCI* by a protein with a conserved BAT2 domain. *Plant Physiol.* **153**, 1425 (2010).
34. Molinier, J., Ries, G., Zipfel, C. & Hohn, B. Transgeneration memory of stress in plants. *Nature* **442**, 1046 (2006).
35. Slaughter, A. *et al.* Descendants of primed *Arabidopsis* plants exhibit resistance to biotic stress. *Plant physiology* **158**, 835–843 (2011).

36. Quadrana, L. & Colot, V. Plant transgenerational epigenetics. *Annual review of genetics* **50**, 467–491 (2016).
37. Henderson, I. R. & Jacobsen, S. E. Epigenetic inheritance in plants. *Nature* **447**, 418–424 (2007).
38. Zhang, H., Lang, Z. & Zhu, J.-K. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology* **19**, 489–506 (2018).
39. Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010).
40. Matzke, M. A. & Mosher, R. A. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15**, 394–408 (2014).
41. Stokes, T. L., Kunkel, B. N. & Richards, E. J. Epigenetic variation in *Arabidopsis* disease resistance. *Genes & Development* **16**, 171–182 (2002).
42. Yi, H. & Richards, E. J. Gene duplication and hypermutation of the pathogen Resistance gene SNC1 in the *Arabidopsis* bal variant. *Genetics* **183**, 1227–1234 (2009).
43. Jubault, M. *et al.* Identification of quantitative trait loci controlling partial clubroot resistance in new mapping populations of *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **117**: 191–202 (2008).
44. Alix, K., Lariagon, C., Delourme, R. & Manzanares-Dauleux, M. J. Exploiting natural genetic diversity and mutant resources of *Arabidopsis thaliana* to study the *A. thaliana*–*Plasmodiophora brassicae* interaction. *Plant Breeding* **126**, 218–221 (2007).
45. Lemarié, S. *et al.* Both the jasmonic acid and the salicylic acid pathways contribute to resistance to the biotrophic clubroot agent *Plasmodiophora brassicae* in *Arabidopsis*. *Plant Cell Physiol.* **56**, 2158–2168 (2015).
46. Simon, M. *et al.* Quantitative trait loci mapping in five new large recombinant inbred line populations of *Arabidopsis thaliana* genotyped with consensus single-nucleotide polymorphism markers. *Genetics* **178**, 2253–2264 (2008).
47. Some, A. *et al.* Variation for virulence on *Brassica napus* L. amongst *Plasmodiophora brassicae* collections from France and derived single-spore isolates. *Plant Pathology* **45**, 432–439 (1996).
48. Kawakatsu, T. *et al.* Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell* **166**, 492–505 (2016).
49. Bewick, A. J. & Schmitz, R. J. Gene body DNA methylation in plants. *Curr. Opin. Plant Biol.* **36**, 103–110 (2017).
50. Richards, E. J. Inherited epigenetic variation--revisiting soft inheritance. *Nat. Rev. Genet.* **7**, 395–401 (2006).
51. Stroud, H. *et al.* Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat. Struct. Mol. Biol.* **21**, 64–72 (2014).
52. Laperche, A. *et al.* Clubroot resistance QTL are modulated by nitrogen input in *Brassica napus*. *Theor Appl Genet* **130**, 669–684 (2017).

53. Aigu, Y. *et al.* Nitrogen supply exerts a major/minor switch between two QTLs controlling *Plasmodiophora brassicae* spore content in rapeseed. *Plant Pathology* **0**,
54. Aoun, N. *et al.* Quantitative disease resistance under elevated temperature: genetic basis of new resistance mechanisms to *Ralstonia solanacearum*. *Front Plant Sci* **8**, (2017).
55. Weigel, D. & Colot, V. Epialleles in plant evolution. *Genome Biol.* **13**, (2012).
56. Gravot, A. *et al.* Hypoxia response in *Arabidopsis* roots infected by *Plasmodiophora brassicae* supports the development of clubroot. *BMC Plant Biology* **16**, 251 (2016).
57. Nishimura, M. T. *et al.* TIR-only protein RBA1 recognizes a pathogen effector to regulate cell death in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **114**, E2053–E2062 (2017).
58. Martin, A. *et al.* A transposon-induced epigenetic change leads to sex determination in melon. *Nature* **461**, 1135 (2009).
59. Narusaka, M. *et al.* RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *The Plant Journal* **60**, 218–226 (2009).
60. Karasov, T. L. *et al.* The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* **512**, 436 (2014).
61. Chini, A. & Loake, G. J. Motifs specific for the ADR1 NBS–LRR protein family in *Arabidopsis* are conserved among NBS–LRR sequences from both dicotyledonous and monocotyledonous plants. *Planta* **221**, 597–601 (2005).
62. Bonardi, V. *et al.* Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16463–16468 (2011).
63. Collier, S. M., Hamel, L.-P. & Moffett, P. Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Mol. Plant Microbe Interact.* **24**, 918–931 (2011).
64. Grant, J. J., Chini, A., Basu, D. & Loake, G. J. Targeted activation tagging of the *Arabidopsis* NBS-LRR gene, ADR1, conveys resistance to virulent pathogens. *Molecular Plant-Microbe Interactions* **16**, 669–680 (2003).
65. Dong, O. X. *et al.* TNL-mediated immunity in *Arabidopsis* requires complex regulation of the redundant ADR1 gene family. *New Phytologist* **210**, 960–973 (2016).
66. Roberts, M., Tang, S., Stallmann, A., Dangl, J. L. & Bonardi, V. Genetic requirements for signaling from an autoactive plant NB-LRR intracellular innate immune receptor. *PLoS genetics* **9**, e1003465 (2013).
67. Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T. & Dodds, P. N. A novel conserved mechanism for plant NLR protein pairs: the ‘integrated decoy’ hypothesis. *Front Plant Sci* **5**, 606 (2014).
68. Williams, S. J. *et al.* Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* **344**, 299–303 (2014).
69. Saucet, S. B. *et al.* Two linked pairs of *Arabidopsis* TNL resistance genes independently confer recognition of bacterial effector AvrRps4. *Nature communications* **6**, 6338 (2015).

70. Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. & Henikoff, S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* **39**, 61–69 (2007).
71. Föhling, M., Graf, H. & Siemens, J. Pathotype separation of *Plasmodiophora brassicae* by the host plant. *Journal of Phytopathology* **151**, 425–430 (2003).
72. Manzanares-Dauleux, M. J., Divaret, I., Baron, F. & Thomas, G. Evaluation of French *Brassica oleracea* landraces for resistance to *Plasmodiophora brassicae*. *Euphytica* **113**, 211–218 (2000).
73. Boyes, D. C. *et al.* Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *The Plant Cell* **13**, 1499–1510 (2001).
74. Gravot, A. *et al.* Genetic and physiological analysis of the relationship between partial resistance to clubroot and tolerance to trehalose in *Arabidopsis thaliana*. *New Phytologist* **191**, 1083–1094 (2011).
75. Liegard, B. *et al.* *Arabidopsis* methylome pattern variations linked to clubroot infection. *The New Phytologist* (submitted).
76. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
77. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
78. Team, R. C. R: A language and environment for statistical computing. (2013).
79. Axtell, M. J. ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* **19**, 740–751 (2013).
80. Silveira, A. B. *et al.* Extensive natural epigenetic variation at a *de novo* originated gene. *PLOS Genetics* **9**, e1003437 (2013).
81. Alonso-Blanco, C. *et al.* 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491 (2016).
82. Clark, L. B. *et al.* Systematic mutagenesis of the leucine-rich repeat (LRR) domain of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in CCR4 deadenylase activity. *Journal of Biological Chemistry* **14**, 13616 (2004).

Interaction plot for *c1_05593 and *c5_19316

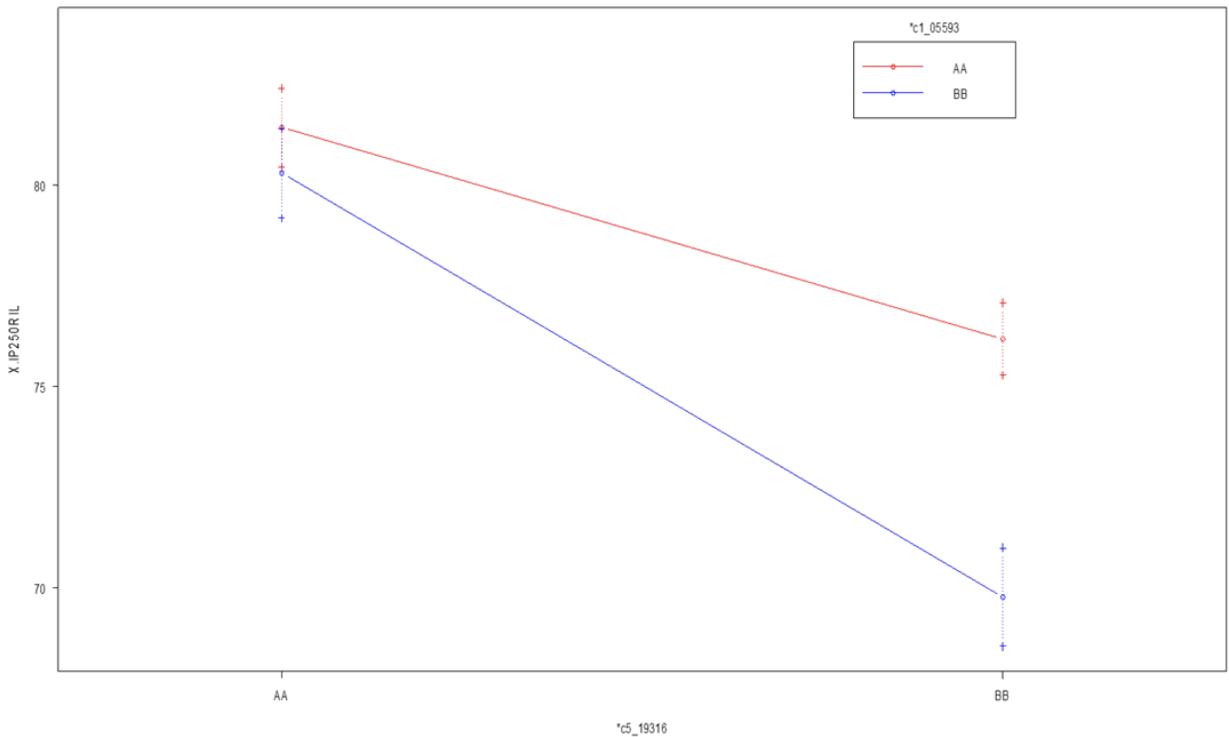


Figure IV.9. Plot de la valeur de l'indice pathologique en fonction de l'état allélique au QTL *Pb-At5.2* et de la région épistatique sur le chromosome 1. En ordonnées : valeur de l'indice pathologique. En abscisses état allélique au QTL *Pb-At5.2*. Les couleurs rouge et bleu correspondent à deux états alléliques différents au niveau de la région épistatique sur le chromosome 1. AA correspond à l'allèle Col-0. BB correspond à l'allèle Bur-0.

Tableau IV.1. Liste des gènes candidats potentiels sous la région en interaction épistatique avec le QTL *Pb-At5.2*.

Locus	Position physique (pb)	Type	Description	Processus biologique
AT1G23730	8395509 - 8398387	Code une protéine	ATBCA3	Utilisation du carbone, régulation des mouvements des stomates, réponse au dioxyde de carbone
AT1G27285	9476197 - 9480516	Gène de type élément transposable		
AT1G27565	9574550 - 9575189	Code une protéine	Protéine hypothétique	
AT1G27570	9575538 - 9582376	Code une protéine	Protéine de la famille des phosphatidylinositol 3- and 4-kinase	

Nous avons démontré qu'une épimutation contrôle la réponse d'*A. thaliana* à la hernie. Cependant la mise en évidence difficile de l'effet du QTL *Pb-At5.2* dans certaines accessions nous a mené à essayer d'en déterminer les causes. L'un des éléments souvent pris en compte dans les études de génétique quantitative et qui pourrait expliquer ce problème est la présence d'une ou d'interaction(s) épistatique(s) avec une autre ou d'autres régions du génome. En effet, une telle interaction si elle existe aurait pour effet de moduler l'effet de la résistance apporté par le QTL *Pb-At5.2* en fonction des allèles présent au locus en épistasie. Ainsi, la recherche d'une potentielle interaction épistatique entre le QTL *Pb-At5.2* et une autre région du génome d'*Arabidopsis* est présentée dans le paragraphe ci-après.

3. Recherche d'une interaction épistatique avec le QTL *Pb-At5.2*

Dans l'optique de déterminer la présence ou non d'une interaction épistatique modulant le niveau de résistance partielle apporté par le QTL *Pb-At5.2*, nous avons ré-analysé les données phénotypiques de Jubault *et al.* (2008) ayant mené à l'identification du QTL *Pb-At5.2* dans la population RIL Bur-0 x Col-0. Ainsi, les données ajustées du trait phénotypique indice pathologique ont été ré-analysées avec le package « qtl » de R. Cette nouvelle analyse a permis d'identifier une interaction épistatique entre le QTL *Pb-At5.2* et une zone du génome située sur le chromosome 1 d'*A. thaliana*. L'intervalle de confiance de la région épistatique identifiée englobe une région génomique de 1,4 Mb qui est située de 8,39 Mb à 9,78 Mb sur le chromosome 1 d'*Arabidopsis*.

L'étude de l'effet des marqueurs aux pics des deux régions (au QTL *Pb-At5.2* et à la région épistatique identifiée) a montré que la présence des allèles favorables de type Bur-0 aux 2 loci permettait une diminution de l'indice pathologique de 12 points en moyenne. En effet, l'indice pathologique passe de 82, lors de la présence de l'allèle favorable Bur-0 uniquement au niveau du QTL *Pb-At5.2*, à 69 lors de la présence de l'allèle Bur-0 aux 2 loci (**Figure IV.9**). Ce résultat met en avant l'importance de cette interaction épistatique dans la modulation de l'effet du QTL *Pb-At5.2* sur la diminution du développement de la maladie chez *Arabidopsis*. L'analyse des éléments présents dans l'intervalle de confiance de la région épistatique identifiée a montré la présence dans ce dernier de 396 gènes, 3 miRNA, 12 ncRNA, 1 snoRNA, 12 pseudogènes et 12 gènes éléments transposables (GFF TAIR10 d'*Arabidopsis*).

Au vu du nombre important de gènes présents dans la région une recherche directe de gènes candidats est difficile. Cependant, cette recherche a quand même été menée et est présentée ci-dessous en utilisant notamment des données de RNA-seq de racines de plantes infectées par *P. Brassicae* disponibles au sein du laboratoire.

Afin de réaliser cette sélection de gènes candidats différentes hypothèses ont été faites en fonction des résultats en notre possession. Ainsi les gènes candidats identifiés sont issus d'une approche avec *a priori* suggérant que les gènes proposés peuvent être différents de ceux réellement impliqués.

Les deux hypothèses utilisées pour la recherche de gènes candidats sont les suivantes :

- Le ou les gènes candidats de la région épistatique sont potentiellement exprimés chez Bur-0 mais pas chez Col-0 comme cela est le cas pour les gènes At5g47260 et At5g47280 du QTL *Pb-At5.2*.
- Le ou les gènes candidats de la région épistatique sont aussi bien exprimés dans les racines que dans les feuilles des plantes comme cela est le cas des gènes At5g47260 et At5g47280 du QTL *Pb-At5.2*

A partir de la première hypothèse, 4 gènes candidats (un élément transposable et 3 gènes) ont été identifiés dans la région épistatique située sur le chromosome 1 en analysant les données de RNA-seq disponibles au sein du laboratoire pour l'HIF 13499 sensible et l'HIF 10499 partiellement résistante ainsi que pour Col-0 (**Tableau IV.1**). Puis en se basant sur notre seconde hypothèse nous avons testé l'implication de l'un ou de plusieurs de ces gènes dans la résistance à la hernie. Pour cela nous avons analysé l'expression de ces 4 gènes dans les accessions d'*Arabidopsis* pour lesquelles les gènes At5g47260 et At5g47280 sont exprimés (84 accessions des 127 accessions testées contre *P. Brassicae* dans la publication Liégard *et al.*, in prep du chapitre IV). Nous nous sommes intéressés uniquement à ce set d'accessions car contrairement au QTL *Pb-At5.2*, la région en épistasie n'a pas été détectée en absence d'interaction avec le QTL *Pb-At5.2*. Ceci montre qu'à elle seule la région en épistasie n'est pas capable d'impacter le développement du pathogène chez *Arabidopsis*. Ainsi, pour arriver à déterminer la causalité de cette région, il est nécessaire de comparer les individus possédant les allèles favorables à l'un (seulement au QTL *Pb-At5.2*) ou aux deux loci en interaction (au QTL *Pb-At5.2* et au niveau du locus en épistasie).

Les caractères DI et GALA ont été mesurés pour ces 84 accessions. Les valeurs phénotypiques vont respectivement de 1,01 à 93,78 et 10,77 à 354,24. Nous avons ensuite vérifié si la variation d'expression d'un ou de plusieurs des gènes candidats permettait d'expliquer la variation de la résistance à la hernie. Aucune différence dans l'expression des gènes candidats n'a pu être mise en évidence en utilisant les données de RNA-seq de 1001 génomes afin d'expliquer une partie de la variabilité de réponse à la hernie visible au sein des 84 accessions d'*Arabidopsis* sélectionnées. Nous en avons donc conclu que les premières hypothèses que nous avons formulées afin d'essayer d'affiner la liste de gènes candidats étaient erronées. En effet, il est possible que le ou les gènes candidats ne soient pas exprimés de la même manière entre les feuilles et les racines, qu'ils ne soient pas exprimés de manière constitutive comme les gènes causaux du QTL *Pb-At5.2* ou que la différence ne soit pas due à une différence d'expression mais de séquence au niveau du ou des gènes candidats.

Afin de réduire la liste de gènes candidats de la région en épistasie, une approche de cartographie fine pourrait être réalisée comme pour la réduction du QTL *Pb-At5.2*. Dans un premier temps une lignée RIL Bur-0 x Col-0 possédant un état résiduel hétérozygote au niveau de la région épistatique et la structure génomique de Bur-0 au QTL *Pb-At5.2* devra être identifiée et sélectionnée. Cette lignée RIL sera par la suite utilisée pour la création de 2 lignées HIF l'une homozygote Bur-0, l'autre homozygote Col-0 au niveau de la région épistatique. Puis les deux lignées HIF seront croisées de manière réciproque afin d'obtenir des lignées F1. Ces F1 devront être autofécondées pour donner des F2 qui seront génotypées afin d'identifier les recombinants potentiels au niveau de la région épistatique. Les F2 recombinantes identifiées seront alors autopollinisées pour donner des F3 qui seront soumises à des tests pathologiques avec l'isolat eH de *P. Brassicae*. Finalement, la mise en relation des données phénotypiques obtenues avec les marqueurs génétiques devrait permettre de réduire l'intervalle de confiance de la région épistatique. Si possible, de nouveaux marqueurs pourront être utilisés pour génotyper les F3 afin de réduire davantage l'intervalle de confiance de la région épistatique. Une fois la région suffisamment réduite une approche gène candidat pourra être réalisée.

Concernant le QTL *Pb-At5.2*, l'épimutation causale identifiée englobant deux gènes NLR (At5g47260 et At5g47280), il est difficile en l'état de déterminer si les 2 NLR sont requis pour obtenir le phénotype de résistance partielle à la hernie notamment observé chez Bur-0. Afin de répondre à cette question, le clonage des 2 gènes présents au niveau de l'épimutation ainsi que leur altération au niveau génomique par une approche de CRISPR sont en cours au sein du laboratoire.

4. Validation du rôle de l'un ou des deux gènes présents sous l'épimutation dans la résistance à la hernie.

Afin de déterminer si seulement l'un des deux gènes ou si les deux gènes présents sous l'épimutation identifiée est (sont) nécessaire(s) à la résistance à la hernie, deux approches financées par le projet CIRCLE du docteur Antoine Grivot sont en cours au sein du laboratoire.

La première approche vise à exprimer de manière native ainsi qu'à surexprimer la version de Bur-0 de l'un des deux gènes ou les deux gènes dans l'accession Columbia et dans une lignée HIF de type Col-0. Ce travail a pour but d'établir un niveau de résistance partielle plus important chez le Columbia transformé par rapport au sauvage ou d'établir une résistance équivalente à celle d'une HIF résistante chez la HIF sensible transformée. Il existe toutefois une probabilité de ne voir aucune variation de résistance dans les plantes sensibles transformées. En effet, deux scénarii pourraient expliquer un tel résultat i) lors de son intégration dans le génome de la plante transformée, la séquence exogène introduite deviendrait méthylée comme la séquence endogène, l'expression du ou des gènes introduits deviendrait alors nulle et aucune différence de phénotype ne serait observable, ii) uniquement chez Columbia, on pourrait imaginer que d'autres gènes nécessaires à la réponse de résistance partielle

possédés par Bur-0 ne seraient pas présents chez Col-0 rendant la voie de résistance incomplète et aucune différence de phénotype ne serait visible. De ce fait une deuxième approche complémentaire est réalisée.

Dans la deuxième approche, le rôle de l'un ou des deux gènes dans la résistance partielle à la hernie n'est plus recherché par leur surexpression dans des plantes sensibles mais par leur altération dans les plantes partiellement résistantes Bur-0 et HIF. La technique CRISPR-Cas9 est utilisée afin d'obtenir différentes combinaisons de gènes de résistance fonctionnels, soit At5g47260, soit At5g47280 en altérant la séquence du second gène. Cette approche est réalisée en collaboration avec l'équipe de Fabien Nogué de l'INRA de Versailles dans le cadre du projet CIRCLE. En plus des constructions précédentes visant à comprendre le rôle des gènes dans la résistance partielle à la hernie, des constructions GUS visant à connaître le lieu d'expression des gènes sont en cours de création au sein du laboratoire.

En perspective, parallèlement à cette approche de validation du rôle de l'un ou des deux NLR dans la résistance partielle à la hernie, il serait intéressant de confronter les accessions Col-0, Bur-0 ainsi que les HIF sensibles et partiellement résistantes à la hernie à d'autres agents pathogènes. En effet au cours de nos études nous avons pu démontrer que le QTL *Pb-At5.2* avait un effet sur différents isolats de *P. Brassicae*, maintenant il serait intéressant de déterminer si celui-ci est également impliqué dans la résistance à d'autres agents pathogènes.

5. Conclusion partielle

Le travail présenté dans ce chapitre de thèse a permis de répondre à la question du rôle potentiel d'épimutations dans la résistance partielle d'*Arabidopsis* à la hernie. Avec cette approche ciblée, les résultats obtenus ont mis en évidence pour la région du gène *RPB1* que :

- Chez les épiRIL, une variation de méthylation existe entre le parent *ddm1-2* et Col-0 WT, ainsi qu'entre différentes lignées épiRIL.
- Contrairement à ce qui est observé dans la population épiRIL, la variation majoritaire au locus n'est pas d'ordre épigénétique mais est due à une importante variation structurale dans les accessions naturelles d'*Arabidopsis*.
- La variation structurale observée pour les accessions naturelles d'*Arabidopsis* pourrait être le résultat de la présence/absence d'éléments transposables entre les accessions dans la région.

En ce qui concerne la région du QTL *Pb-At5.2*, nous avons montré que :

- La causalité de la détection de cette région est une épimutation contrôlant l'expression de deux gènes NLR et ainsi qu'un QTL détecté en utilisant des marqueurs basés sur des variations de séquence peut avoir une causalité épigénétique.

- Au niveau génomique, aucun SNP en cis ne semble causal dans la variation de méthylation.
- La présence d'une interaction épistatique entre ce QTL et une autre région du génome sur le chromosome 1 d'*Arabidopsis* possédant un intervalle de confiance de 1,4 Mb.
- Ce dernier présente un effet protecteur de la plante face à différents isolats de *P. Brassicae* (eH, Pb137-522, Ms6, K92-16 et P1+)

CHAPITRE V

DISCUSSION GENERALE ET PERSPECTIVES

CHAPITRE V. DISCUSSION GENERALE ET PERSPECTIVES

A l'heure où agriculture rime avec respect de l'environnement, la mise en œuvre de stratégies durables de protection des cultures contre les bioagresseurs constitue l'un des principaux défis sociétaux et économiques mais aussi scientifiques actuels. Dans ce contexte, le développement de méthodes complémentaires ou synergiques de lutte contre les bioagresseurs des espèces végétales d'intérêt agronomique est fondamental pour une agriculture plus respectueuse de l'environnement. L'utilisation de variétés résistantes est l'une des composantes majeures de ces systèmes de protection intégrés. Ainsi l'identification de sources de résistance génétique pour la construction de variétés résistantes est devenue une priorité notamment devant la recrudescence de maladies telles que la hernie des crucifères, l'une des principales maladies des *Brassicaceae* dans le monde (Terre Inovia, 2008). Du fait de la faible efficacité ou de la difficulté de mise en œuvre des solutions de type traitement chimique ou pratiques culturales sur le développement de la maladie, la lutte contre la hernie des crucifères passe majoritairement par la lutte génétique. La plupart des variétés résistantes à la hernie commercialisées jusqu'à présent présentent des résistances de type qualitatif basées sur l'introduction (ou introgression) de gènes majeurs de type R. Ce type de résistance, à effet presque total, est principalement isolat spécifique, c'est à dire que la protection conférée n'est efficace que vis-à-vis de certains isolats/souches de l'agent pathogène. Par ailleurs, l'utilisation de gènes de type R impose une très forte pression de sélection sur les populations de pathogènes, favorisant ainsi l'émergence d'isolats plus virulents qui vont réussir à contourner la résistance et à la rendre inefficace. Cela a par exemple été le cas de la résistance à la hernie portée par la variété de colza Mendel qui après une période de culture relativement courte a été contournée par *P. Brassicae* (Diederichsen *et al.*, 2009).

Pour remédier à ce problème de contournement et ainsi augmenter la durabilité des résistances, d'autres sources de résistance d'origine génétique, les résistances dites quantitatives, ont été étudiées pour lutter contre les maladies. La résistance quantitative confère une protection partielle contre l'agent pathogène, c'est à dire qu'elle ne stoppe pas totalement le développement de l'agent pathogène mais réduit sa multiplication, sa colonisation ou la gravité des symptômes observés sur la plante hôte. L'un des avantages de ce type de résistance est de réduire ainsi la pression de sélection imposée sur l'agent pathogène et donc l'apparition de nouvelles souches virulentes. De plus, ce type de résistance repose non pas sur un gène unique mais sur l'effet de multiples gènes permettant ainsi des combinaisons de mécanismes de résistance divers et agissant à des moments différents du cycle de vie de l'agent pathogène et exerçant des pressions de sélection contradictoires sur l'évolution de l'agent pathogène. L'ensemble de ces données suggère ainsi que la résistance quantitative serait plus durable que la résistance qualitative (French *et al.*, 2016; Pilet-Nayel *et al.*, 2017). Des résistances quantitatives ont

été décrites chez de nombreuses espèces végétales vis-à-vis de différents agents pathogènes (Kim & Diers, 2000; Manzanares-Dauleux *et al.*, 2000a ; Foolad *et al.*, 2002 ; Zhao *et al.*, 2006 ; Darvishzadeh R. *et al.*, 2007 ; Jubault *et al.*, 2008b ; Kump *et al.*, 2011 ; Benson *et al.*, 2015 ; Tanweer *et al.*, 2015 ; Corwin *et al.*, 2016 ; Aoun *et al.*, 2017 ; Kumar *et al.*, 2018). Chez les *Brassicaceae* de nombreux facteurs génétiques contrôlant la résistance quantitative à la hernie ont été identifiés chez *B. napus*, *B. rapa*, *B. oleracea* et *A. thaliana* (Aigu *et al.*; Landry *et al.*, 1992 ; Figdore *et al.*, 1993 ; Fuchs & Sacristán, 1996 ; Grandclément & Thomas, 1996 ; Voorrips *et al.*, 1997, p. 199 ; Matsumoto *et al.*, 1998 ; Moriguchi *et al.*, 1999 ; Manzanares-Dauleux *et al.*, 2000a, 2003 ; Piao *et al.*, 2002, 2004 ; Hirai *et al.*, 2004 ; Rocherieux *et al.*, 2004 ; Saito *et al.*, 2006 ; Hayashida *et al.*, 2008 ; Jubault *et al.*, 2008b ; Nomura *et al.*, 2008 ; Sakamoto *et al.*, 2008 ; Diederichsen *et al.*, 2009 ; Kato *et al.*, 2012, 2013 ; Ueno *et al.*, 2012 ; Li *et al.*, 2016b ; Strelkov *et al.*, 2016 ; Hatakeyama *et al.*, 2017 ; Laperche *et al.*, 2017 ; Yu *et al.*, 2017). Cependant, lors des analyses génétiques effectuées, la variance phénotypique observée ne peut être expliquée en totalité par l'effet des différents QTL détectés et/ou par l'effet de l'environnement. Ce phénomène est appelé l'héritabilité manquante (« missing heritability » ; Brachi *et al.*, 2011 ; Caballero *et al.*, 2015). L'un des enjeux actuels des recherches menées sur les résistances quantitatives est de comprendre qu'elle(s) est (sont) la (les) source(s) de variabilité permettant d'expliquer une partie de cette héritabilité perdue. Une des hypothèses explicatives est basée sur l'existence d'une variabilité d'origine épigénétique. En effet, des études menées ces dernières années ont mis en évidence la modulation et l'implication des marques épigénétique dans la réponse des végétaux à des stress d'origine abiotique et biotique (Stokes *et al.*, 2002 ; Steward *et al.*, 2002 ; Pavet *et al.*, 2006 ; Akimoto *et al.*, 2007 ; Agorio & Vera, 2007 ; Boyko *et al.*, 2010 ; López *et al.*, 2011 ; Luna *et al.*, 2012 ; Downen *et al.*, 2012 ; Yu *et al.*, 2013 ; Jiang *et al.*, 2014 ; Bond & Baulcombe, 2015 ; Rambani *et al.*, 2015 ; López Sánchez Ana *et al.*, 2016 ; Zhang *et al.*, 2016 ; Li *et al.*, 2016a ; Zheng *et al.*, 2017 ; Liu *et al.*, 2017 ; Hewezi *et al.*, 2017). De plus, il a été montré que des modifications du méthylome des végétaux, sous la forme de cytosines différenciellement méthylées et/ou de régions différenciellement méthylées (DMR), pouvaient apparaître dans le génome d'*A. thaliana* avec des fréquences respectivement plus importantes ou similaires que celle des mutations nucléotidiques (Becker *et al.*, 2011 ; Schmitz *et al.*, 2011 ; van der Graaf *et al.*, 2015). Certaines de ces modifications sont transmissibles à travers les générations (Richards, 2006) et impliquées dans la variabilité phénotypique chez différentes espèces végétales (Cubas *et al.*, 1999 ; Manning *et al.*, 2006 ; Miura *et al.*, 2009 ; Martin *et al.*, 2009 ; Durand *et al.*, 2012 ; Silveira *et al.*, 2013 ; Ong-Abdullah *et al.*, 2015 ; Blevins *et al.*, 2017 ; He *et al.*, 2018). A partir de ces données de la littérature nous avons posé l'hypothèse qu'une partie de la réponse de résistance quantitative à la hernie pourrait résulter d'une ségrégation de différents épiallèles.

Les travaux que j'ai effectués durant ces trois années de thèse sont ancrés dans cet axe de recherche et ont pour objectif de déterminer si des variations de méthylation de l'ADN transmissibles au cours des

générations sont impliquées dans la résistance partielle à la hernie des crucifères chez *A. thaliana*. Afin d'apporter des éléments de réponse deux approches différentes ont été utilisées, une approche non ciblée et une approche ciblée. L'approche non ciblée repose sur l'étude de mutants affectés dans la structure de la chromatine et sur une analyse d'(épi)généétique quantitative à partir d'une population épiRIL présentant un fonds génétique quasi isogénique mais une mosaïque des épigénomes parentaux. L'approche ciblée s'est intéressée à l'étude de deux régions génomiques spécifiques, l'une issue des résultats de l'approche non ciblée et l'autre concernant un QTL de résistance détecté précédemment au laboratoire par une approche de génétique quantitative classique basée sur du polymorphisme de séquence et pour lequel la causalité sous-jacente ne semblait pas reposer sur un polymorphisme de séquence.

Les résultats majeurs de mes travaux de thèse sont les suivants :

- La méthylation de l'ADN est impliquée dans la résistance quantitative à la hernie des crucifères chez *Arabidopsis*. En effet le mutant *ddm1* impliqué dans le maintien de la méthylation de l'ADN présente moins de symptômes vis-à-vis de l'isolat eH de *P. Brassicae* que l'écotype sauvage Col-0.
- Des régions différenciellement méthylées transmissibles aux générations suivantes (actuellement génération 10) sont impliquées dans la résistance quantitative à la hernie chez *A. thaliana* : 14 QTL^{épi} additifs ont été identifiés au sein de la population épiRIL.
- Cinq QTL^{épi} ont un effet dépendant de la température.
- Le QTL de résistance *Pb-At5.2* étudié lors de l'approche ciblée présente une causalité sous-jacente d'origine épigénétique : une épimutation transmissible à travers les générations qui englobe deux gènes NLR. Ce résultat suggère qu'un QTL détecté en utilisant des marqueurs basés sur des variations de séquence peut avoir une causalité épigénétique.
- Ce QTL confère une résistance à différents isolats de *P. Brassicae* (eH, Pb137-522, Ms6, K92-16 et P1⁺).

Dans la suite de ce chapitre, je discuterai dans un premier temps de l'intérêt d'étudier la variabilité épigénétique et des méthodologies disponibles pour étudier l'implication de l'épigénétique dans la résistance quantitative chez le pathosystème *A. thaliana/P. Brassicae*. Par la suite j'exposerai la pertinence, les avantages et les inconvénients liés à l'utilisation des approches ciblée et non ciblée dans l'étude de l'implication de l'épigénétique dans la résistance quantitative. Je continuerai par une discussion autour de la stabilité des QTL^{épi} par rapport à l'environnement et j'aborderai la question de la causalité d'épimutations dans la détection de QTL. Finalement je terminerai ce chapitre par la question du transfert des résultats obtenus chez *A. thaliana* à d'autres *Brassicaceae* et de l'utilisation de la variation d'origine épigénétique et des épiallèles en sélection.

1. Résistance quantitative et épigénétique

1.1. Pourquoi prendre en compte ce type de variabilité ?

La première question relative à ma thèse a été de comprendre pourquoi l'hypothèse épigénétique était un élément cohérent à prendre en considération pour expliquer les mécanismes de résistance quantitative. Actuellement, la résistance quantitative est majoritairement étudiée en utilisant des approches de génétique quantitative de liaison ou d'association (GWAS) basées sur des marqueurs de polymorphisme de séquence. En effet, de manière classique les caractères complexes héréditaires sont considérés comme issus de variants de séquence de l'ADN transmis à travers les allèles. De nombreux facteurs génétiques de résistance ont ainsi été identifiés chez des nombreuses plantes hôtes vis-à-vis différents agents pathogènes y compris chez les *Brassicaceae* contre *P. Brassicae*. Cependant, bien qu'un nombre important de facteurs génétiques ait été décelé, l'ensemble de l'effet de ceux-ci ne permettait pas d'expliquer la totalité de la variabilité de la réponse de résistance quantitative.

En parallèle de ces études, de nombreux variants de la structure de la chromatine (Kakutani *et al.*, 1999 ; Riddle & Richards, 2002, 2005 ; Vaughn *et al.*, 2007 ; Zhang *et al.*, 2008) potentiellement impliqués dans des variations phénotypiques et transmis entre les générations ont été mis en évidence dans des populations naturelles et expérimentales. Différents chercheurs se sont alors questionnés sur le rôle de l'épigénétique dans la mise en place de caractères complexes, *i.e.* contrôlés par un grand nombre de gènes en interaction, dans la diversité des phénotypes observables (Johannes *et al.*, 2008 ; Richards, 2008). Le comportement de certains mutants affectés dans la mise en place de la structure de la chromatine vis-à-vis de différents agents pathogènes suggérait par ailleurs que la variabilité d'origine épigénétique pourrait être impliquée dans la réponse de résistance.

Toutefois, aborder l'étude de l'implication épigénétique dans la réponse aux stress biotiques s'est avéré très compliqué. En effet, il est complexe de dissocier l'effet génétique de l'effet épigénétique, les deux pouvant être connectés comme dans le cas des épimutations obligatoires ou facilitées (cf chapitre I paragraphe 2.4.5). Les résultats que nous avons obtenus au cours de ma thèse sont une bonne illustration de cette difficulté. En effet nous avons démontré que certains QTL détectés par analyse génétique classique - cas du QTL *Pb-At5.2* - étaient en fait des QTL^{épi}. Dans notre cas ceci a été possible car l'épimutation causale était en déséquilibre de liaison avec des SNP dans les lignées parentales Bur-0 et Col-0 à l'origine de la population en ségrégation utilisée pour l'analyse génétique. Cependant cette détection n'aurait pas été possible dans le cas d'un QTL^{épi} dont la causalité sous-jacente aurait été un épiallèle pur (Richards, 2006) en équilibre de liaison avec les SNP connus. Une des méthodes alternatives développées pour résoudre ce problème de déséquilibre de liaison entre variations épigénétiques et variations génétiques est l'analyse EWAS (Epigenome Wide Association Study; Rakyan *et al.*, 2011). L'EWAS repose sur l'utilisation de marqueurs issus de polymorphisme épigénétique (tel que des DMR) afin trouver des associations entre des variations épigénétiques et des

traits phénotypiques d'intérêt. En perspective, il serait intéressant de réaliser une approche d'EWAS en utilisant les données de méthylation de 1001 génomes sur les 127 écotypes pour lesquels nous avons obtenu des données phénotypiques de réponse à l'infection par la hernie.

1.2. Pourquoi le modèle *Arabidopsis thaliana* - *Plasmodiophora brassicae*

L'une des questions de recherche menée au sein de l'équipe résistance et adaptation (RA) de l'UMR IGEPP où mon travail de thèse a été effectué est de comprendre, exploiter et combiner différents leviers afin de contrôler de manière durable les épidémies. Parmi eux se trouve la lutte génétique qui est étudiée dans différents pathosystèmes (*Brassica napus/Leptosphaeria maculans*, *Brassicaceae/P. Brassicae*, *Pisum/Aphanomyces euteiches*...) notamment par des approches de génétique/génomique quantitative. A l'heure actuelle beaucoup de données ont été accumulées sur le pathosystème *Brassicaceae/P. Brassicae* du fait de son importance socioéconomique (Terres Inovia, 2008), de sa distribution à l'échelle mondiale (Dixon, 2009) et ceci pour différents isolats de l'agent pathogène et différentes espèces hôtes telles que *B. napus*, *B. oleracea*, *B. rapa* et *A. thaliana* (Aigu *et al.*; Fuchs & Sacristán, 1996 ; Manzanares-Dauleux *et al.*, 2000a, 2003 ; Alix *et al.*, 2007 ; Werner *et al.*, 2007 ; Jubault *et al.*, 2008a,b ; Gravot Antoine *et al.*, 2011 ; Gravot *et al.*, 2012, 2016a,b ; Lemarié *et al.*, 2015b,a ; Li *et al.*, 2016b ; Laperche *et al.*, 2017). Un des avantages de l'étude de ce pathosystème est que *P. Brassicae* ne possède pas de spécificité d'hôte ; un même isolat peut infecter et être agressif sur différentes espèces et ainsi par exemple des données acquises sur *Arabidopsis* lors de l'interaction avec un isolat donné peuvent être transférables chez *B. napus*. De plus l'isolat privilégié pour notre étude, l'isolat eH, est un isolat monospore (stable et homogène génétiquement), représentatif du pathotype P1, l'un des pathotypes les plus fréquents en France (Orgeur *et al.*, 2015), dont la séquence a été obtenue au sein du laboratoire (Daval *et al.*, 2018) et de nombreuses données issues de l'infection de *Brassicaceae* par cet isolat sont disponibles. *A. thaliana*, hôte naturel de *P. Brassicae*, est une espèce modèle comportant d'importantes ressources génétiques (telles que des mutants, y compris épigénétiques ; des lignées HIF, NIL, RIL, y compris des lignées épiRIL...) et génomiques (données de RNA-seq, DNaseq, BS-seq, ChIP-seq ...) qui ont été générées pour différents écotypes (voire même pour des milliers avec le projet 1001 génomes). Dans le cadre d'une approche de d'épigénétique quantitative, il est nécessaire de posséder une bonne résolution au niveau génomique qui soit compatible avec les méthodes d'identification des marques épigénétiques (méthylation de l'ADN : MeDIP-CHIP, séquençage après traitement au bisulfite (BS-seq) ; marques histones : ChIP-seq) ce qui est le cas chez *A. thaliana*. Ce choix important d'outils et de données disponibles couplé à la non spécificité d'hôte de *P. Brassicae* font du pathosystème *A. thaliana/P. Brassicae* un modèle idéal pour étudier l'implication de l'épigénétique dans la résistance quantitative à la hernie.

1.3. Quel matériel végétal est disponible pour étudier l'impact de l'épigénétique dans le cadre de la résistance quantitative ?

Le premier matériel végétal disponible est la collection de mutants altérés au niveau des gènes impliqués dans la structure de la chromatine. Initialement, l'impact de l'épigénétique sur des caractères complexes a été découvert en étudiant ces mutants. Certains de ces mutants montraient des altérations du phénotype notamment au niveau du phénotype floral (Jacobsen *et al.*, 2000 ; Soppe *et al.*, 2000). Ces mutants ou des lignées dérivées ont été utilisés récemment afin de mettre en relation la variation épigénétique et la réponse à l'infection par différents agents pathogènes tels que *P. syringae* DC3000, *A. tumefaciens*, *B. cinerea* ou *P. cucumerina* (Stokes *et al.*, 2002 ; López *et al.*, 2011 ; Yu *et al.*, 2013 ; Bond & Baulcombe, 2015 ; López Sánchez Ana *et al.*, 2016). Ainsi, nous avons criblé certains de ces mutants pour la résistance à la hernie. Les premiers résultats obtenus dans ma thèse ont mis en avant une moindre sensibilité à la hernie du mutant *ddm1*, mutant qui présente une diminution globale de la méthylation de l'ADN (58 % en CG, 57 % en CHG et 32 % en CHH ; Zemach *et al.*, 2013). De manière intéressante, nos résultats vont à l'inverse des observations réalisées chez *Arabidopsis* en interaction avec un autre agent pathogène à galle, *A. tumefaciens*, où la méthylation en CHG et CHH chez *Arabidopsis* empêcherait potentiellement la formation des galles. En effet, les mutants *ddc* (triple mutant *cmt3 drm1 drm2*) et *ago4* présentaient une forte augmentation de la taille des galles induites par cet agent pathogène. Néanmoins, bien que le mutant *ddm1* nous ait fourni une information importante sur l'implication de la méthylation de l'ADN dans la résistance à la hernie, cela ne permettait pas d'aller plus loin dans la connaissance des mécanismes sous-jacents. En effet, la mutation *ddm1* impacte d'autres régions que celles impliquées dans la réponse à la hernie.

Un autre type de matériel végétal permettant ce type d'études sont les lignées accumulatrices de mutations d'*A. thaliana* chez lesquelles il a été évalué la fréquence d'apparition des épimutations en étudiant le méthylome (Becker *et al.*, 2011 ; Schmitz *et al.*, 2011 ; van der Graaf *et al.*, 2015). Ces lignées accumulatrices de mutations et d'épimutations ont été utilisées afin de caractériser les changements de l'état de méthylation de l'ADN induits après un stress abiotique (Jiang *et al.*, 2014 ; Zheng *et al.*, 2017). Cependant, nous n'avons pas utilisé ces lignées dans notre étude du fait du faible nombre de lignées caractérisées disponibles par rapport à celui nécessaire pour une approche de génétique quantitative (environ 100 lignées au minimum).

Enfin une dernière possibilité était d'utiliser une population epiRIL. Ce type de population est issue du croisement de deux individus possédant le même fonds génétique mais avec des fonds épigénétiques différents. Deux populations epiRIL quasi isogéniques mais avec une mosaïque des méthylomes parentaux sont disponibles chez *Arabidopsis*. Une population de 505 individus résultant du croisement entre le mutant *ddm1-2* et un individu de type sauvage Col-0 (Johannes *et al.*, 2009) et une autre de 68 individus issus du croisement entre Col-0 et le mutant *met 1-3* (Reinders *et al.*, 2009). Ces populations

ont été utilisées afin d'évaluer l'impact des variations de méthylation sur l'expression de caractères complexes tels que le temps de floraison, la croissance des plantes, la taille des plantes en condition contrôle et la germination en condition stress salin (Johannes *et al.*, 2009 ; Reinders *et al.*, 2009). Une seule étude a concerné l'étude de la réponse de la population epiRIL Col-0 x *met1-3* à un agent pathogène, *P. syringae pv tomato* où une augmentation de la réponse de résistance a été observée pour 34 % des epiRIL et une plus grande sensibilité pour 4 % d'entre elles (Reinders *et al.*, 2009). La population *ddm1-2* x Col-0 a été utilisée dans différentes analyses d'épigénétique quantitative afin de détecter des QTL^{épi} associés à différents traits phénotypiques mais uniquement en condition contrôle ou en présence d'un stress abiotique (Cortijo *et al.*, 2014 ; Kooke *et al.*, 2015 ; Aller *et al.*, 2018).

Finalement parmi tous les matériels disponibles, la population epiRIL de Johannes *et al.* (2009) semblait donc être la plus adaptée pour étudier la contribution épigénétique dans la réponse à la hernie car celle-ci regroupait toutes les caractéristiques nécessaires : i) un nombre de lignées suffisant pour des analyses statistiques robustes, ii) peu de variabilité génétique et une forte variabilité épigénétique, permettant de dissocier l'effet génétique de l'effet épigénétique et iii) l'existence d'une carte épigénétique regroupant de manière précise l'état de méthylation de chaque lignée à chaque marqueur épigénétique stable (Colomé-Tatché *et al.*, 2012).

2. Pertinence, avantages et inconvénients des approches ciblées et non ciblées utilisées

Dans notre étude, nous avons évalué les symptômes provoqués par l'isolat eH de *P. Brassicae* via l'indice pathologique (Manzanares-Dauleux *et al.*, 2000b) chez des mutants d'*A. thaliana* affectés dans la structure de la chromatine puis sur la population epiRIL de Johannes *et al.* (2009). L'utilisation de certains mutants dans le cadre de ma thèse a ainsi permis de mettre en évidence le rôle de *DDMI* dans la résistance à la hernie et donc de l'implication de la méthylation de l'ADN. Ce premier élément de réponse nous a permis de valider l'utilisation de la population epiRIL de Johannes *et al.* (2009) afin de mieux caractériser l'implication de la méthylation de l'ADN dans la réponse à la hernie via l'identification de QTL^{épi}. Un des avantages de cette population était que celle-ci avait déjà été utilisée dans différentes études afin d'identifier des QTL^{épi} impliqués dans le contrôle de caractères complexes (Cortijo *et al.*, 2014 ; Kooke *et al.*, 2015 ; Aller *et al.*, 2018). Ces trois études ont permis de mettre en évidence différents QTL^{épi} contrôlant une part plus ou moins importante des caractères complexes étudiés montrant ainsi i) la solidité de l'approche de détection de QTL^{épi} en utilisant cette population et ceci même lors d'un stress abiotique, ii) l'implication de régions sous dépendance épigénétique dans le contrôle de divers traits phénotypiques et iii) l'influence des stress abiotiques sur la détection de QTL^{épi} pour une même caractéristique complexe avec la détection de QTL^{épi} communs et spécifiques d'une condition. Toutefois, dans aucune de ces études des DMR candidates ou gènes candidats ont été identifiés, les intervalles de confiance des QTL^{épi} détectés étant très importants.

Les résultats que j'ai obtenus grâce à cette analyse ont permis, comme pour les autres détections de QTL^{épi} réalisées avec cette population en condition contrôle ou lors d'un stress abiotique, de mettre en évidence l'implication de QTL^{épi} dans la résistance à la hernie chez *A. thaliana*. Dans notre étude nous avons en plus pris en considération l'effet d'un facteur abiotique sur l'interaction épigénome *Arabidopsis*/hernie, l'influence de la température. La prise en compte de variations de température présentes dans la chambre de culture utilisée nous a permis de déceler pour la première fois des QTL sous dépendance épigénétique influencés par la température (l'interaction de QTL^{épi} avec l'environnement est discutée dans le chapitre V.3).

L'approche non ciblée s'est relevée très pertinente car avec peu de résultats préliminaires sur le sujet mais avec une bonne connaissance de notre pathosystème nous avons pu mettre en évidence la présence de QTL^{épi} liés à la résistance à la hernie. Toutefois la population épiRIL utilisée possède un fonds génétique de type Columbia, accession sensible à l'isolat eH de *P. Brassicae* (Alix *et al.*, 2007), contrairement à la population RIL utilisée précédemment pour les analyses classiques, constituée d'individus présentant différentes combinaisons du génome (et de l'épigénome) de Col-0 et de Bur-0, accession partiellement résistante à la hernie (Jubault *et al.*, 2008b). Or d'après ce que nous avons démontré avec la région *Pb-At5.2*, des relations d'épistasie concernant des QTL^{épi} existent et de ce fait le fonds génétique dans lequel est réalisée la population épiRIL conditionne fortement la détection de QTL^{épi}. Il serait alors intéressant de développer une nouvelle population épiRIL en utilisant une accession partiellement résistante à la hernie telle que Bur-0 (Alix *et al.*, 2007). Une telle population permettrait de réaliser la détection de QTL^{épi} dans un autre fonds génétique et d'augmenter ainsi le nombre de QTL^{épi} détectés, avec potentiellement la présence de QTL^{épi} montrant des effets plus importants que ceux identifiés dans mon travail de thèse.

Une fois l'implication de l'épigénétique dans la réponse à la hernie mise en évidence nous avons essayé de caractériser plus finement l'origine épigénétique potentielle de deux régions impliquées dans la résistance à la hernie, la région du gène *RPB1* et du QTL *Pb-At5.2*. Afin d'arriver à notre objectif, nous avons utilisé des données de méthylation de l'ADN, de transcriptomique et de génomique de différents écotypes d'*Arabidopsis* (régions *RPB1* et QTL *Pb-At5.2*), de lignées HIF (QTL *Pb-At5.2*) et de lignées épiRIL (région du gène *RPB1*). Des lignées HIF et épiRIL ont été utilisées afin d'identifier des régions différenciellement méthylées causales au niveau des régions ciblées. L'avantage de ces deux populations est qu'elles présentent peu de variations dues à du polymorphisme de séquence entre les lignées utilisées facilitant la recherche de DMR.

Afin de déterminer si la structure génétique/épigénétique des deux régions ciblées est fréquente dans la nature, nous avons étudié ces régions dans une collection d'écotypes d'*Arabidopsis*. En effet, Cortijo *et al.* (2014) avaient observé que seuls 30 % des DMR identifiées comme stables au cours des générations dans la population épiRIL étaient retrouvées au sein d'un panel de 138 accessions naturelles

d'*Arabidopsis*. Dans le cas des accessions Col-0, Bur-0 et des lignées HIF dérivées, l'identification d'une DMR causale aurait pu être un évènement naturel mais rare qui n'existerait que dans une proportion infime de la collection des écotypes d'*A. thaliana*. De ce fait, afin de pouvoir valider le rôle de nos DMR candidates dans une population non expérimentale et ainsi de valider leur impact au niveau de la variation de résistance à la hernie *in natura*, la validation de ces dernières au sein d'accessions naturelles d'*Arabidopsis* était nécessaire.

Les résultats obtenus dans le cadre de ma thèse sur la région du gène *RPBI* ont permis de montrer que les variations de l'état de méthylation présentes dans la population épiRIL ne sont pas retrouvées dans les écotypes d'*Arabidopsis* possédant une structure de la région de type Col-0 (80 écotypes sur 1035 testés). Nos résultats suggèrent que, dans la nature, la variation au niveau de la région est principalement une variation de structure génomique induite potentiellement par des mouvements d'éléments transposables. Cependant, les variations de méthylation peuvent potentiellement exister au niveau de cette région chez des accessions avec des structures génomiques différentes de Col-0. Afin d'éclaircir ce point il serait intéressant de déterminer la structure génomique des autres 955 écotypes et de réaliser une nouvelle cartographie des lectures de bisulfite sur la région. Ce résultat illustre une autre limite de l'utilisation de populations expérimentales dans le cadre de l'interaction plante pathogène et la nécessité de valider les résultats obtenus en utilisant des accessions naturelles.

Concernant la région du QTL *Pb-At5.2* les résultats obtenus ont permis de mettre en évidence l'implication d'une épimutation dans la résistance à la hernie grâce à l'utilisation de couples de lignées HIF et d'écotypes d'*Arabidopsis*. Les lignées HIF ont permis de réduire la région causale à 8 gènes candidats (travail effectué par le docteur Antoine Gravot) et d'étudier au sein de deux lignées HIF homozygotes (1381 et 2509) et dans les lignées hétérozygotes issues de leurs autofécondations, le taux de méthylation au niveau de l'épimutation et l'expression des deux gènes NLR sous-jacents. Bien que les HIF hétérozygotes soient sensibles à la hernie comme les HIF homozygotes de type Col-0, le niveau de méthylation des hétérozygotes, intermédiaire de celui des homozygotes, a permis d'écarter l'existence d'un phénomène de paramutation. Ce résultat pose la question de l'origine de la sensibilité des hétérozygotes. Une des hypothèses possibles pour expliquer cette sensibilité est que la résistance ne s'établit que s'il existe un niveau suffisant de transcrits/protéines des deux NLR : l'hyperméthylation des deux brins ou d'un seul ne permettrait pas ainsi d'avoir un niveau de transcrits suffisant pour l'expression de la résistance. Afin de vérifier cette hypothèse le niveau d'expression des deux gènes NLR sous l'épimutation dans les HIF homozygotes de type Col-0 et Bur-0 et dans les hétérozygotes est en cours d'étude au sein du laboratoire. La variation de méthylation au niveau des deux NLR a été validée dans une proportion non négligeable d'écotypes (401 sur 1087 écotypes) chez lesquels cette variation est associée à la réponse de résistance. Nous avons aussi montré dans la collection d'écotypes l'absence de polymorphisme en *cis* responsable de l'épimutation. Cependant, la présence d'un

polymorphisme de séquence ayant un effet en *trans* sur l'épimutation n'a pas pu être écartée pour le moment.

Finalement les deux approches ciblée et non ciblée utilisées dans mon travail de thèse ont été très complémentaires permettant de moduler le risque d'absence de résultats ainsi que la finesse des résultats obtenus.

3. Interaction des QTL^{épi} par rapport à l'environnement

L'identification de multiples épimutations stables mais aussi métastables au sein de plusieurs espèces végétales (Cubas *et al.*, 1999 ; Bender, 2004 ; Manning *et al.*, 2006 ; Miura *et al.*, 2009 ; Martin *et al.*, 2009 ; Durand *et al.*, 2012 ; Silveira *et al.*, 2013 ; Quadrana *et al.*, 2014 ; Ong-Abdullah *et al.*, 2015 ; Blevins *et al.*, 2017 ; He *et al.*, 2018) et la modulation du méthylome des végétaux en réponse à un stress (Steward *et al.*, 2002 ; Boyko *et al.*, 2010 ; Jiang *et al.*, 2014 ; Zhang *et al.*, 2016 ; Li *et al.*, 2016a ; Zheng *et al.*, 2017 ; Liu *et al.*, 2017) pose la stabilité des épimutations dans des environnements fluctuants. Nous avons identifié cinq QTL^{épi} dépendant de la température dans le cadre de l'interaction *A. thaliana*/*P. Brassicae*. Par ailleurs, des travaux réalisés dans le laboratoire d'accueil en utilisant différents génotypes d'*A. thaliana* et de *B. napus* ont permis de montrer l'effet de l'irrigation et de l'apport azoté dans la modulation de la réponse de la plante à la hernie (Gravot *et al.*, 2016a ; Laperche *et al.*, 2017 ; Aigu Y. *et al.*, 2018). L'influence de l'environnement sur l'effet des résistances a été décrit chez d'autres pathosystèmes et le nombre de QTL de résistance détectés varie en fonction des années et donc des conditions climatiques (Calenge & Durel, 2006 ; Coyne *et al.*, 2015 ; Huang *et al.*, 2016) mais aussi de l'apport azoté (Talukder *et al.*, 2005 ; Ballini *et al.*, 2013).

Mes résultats préliminaires suggèrent que la modulation de la réponse des plantes à un stress biotique par l'environnement (ici un stress abiotique : la température) ne touche pas simplement leur architecture génétique mais aussi leur architecture épigénétique. Maintenant, la question est de comprendre comment les variations environnementales telles que la température modulent le méthylome et l'architecture épigénétique de la réponse de la plante à un autre stress. Une hypothèse pourrait être que la température influencerait sur le fonctionnement de protéines impliquées dans la mise en place et le maintien des marques épigénétiques.

Afin de valider les QTL^{épi} dépendant de la température, des tests pathologiques à deux températures distinctes (par exemple 20 °C et 25 °C) peuvent être envisagés ce qui permettrait d'une part de valider les QTL^{épi} détectés dans Liégard *et al.* (2018) et de réduire potentiellement leurs intervalles de confiance par un phénotypage plus précis et d'autre part de détecter d'autres QTL^{épi} sous la dépendance de la température. Une approche similaire pour l'étude de l'impact de la température sur la résistance a déjà été réalisée dans le cadre du pathosystème *A. thaliana*/*Ralstonia solanacearum* via une approche d'analyse d'association basée sur des polymorphismes de séquence de 176 accessions d'*A. thaliana*. La

sensibilité des plantes augmente lors d'une augmentation de la température de 3 °C et l'architecture génétique de la réponse d'*A. thaliana* à *R. solanacearum* est différente. Cette variation de l'architecture génétique de la réponse se traduit notamment par l'absence de détection des deux QTL majeurs associés à la résistance à l'agent pathogène à 30 °C par rapport à 27 °C (Aoun *et al.*, 2017).

4. Transfert des résultats obtenus chez *Arabidopsis* à d'autres *Brassicaceae*

Une question importante lors de l'utilisation d'une espèce modèle est de savoir si les résultats obtenus sur cette dernière sont transférables à des espèces d'intérêt agronomique. Dans le cas du QTL *Pb-At5.2* il est ainsi important de savoir dans un premier temps si les gènes NLR identifiés chez *A. thaliana* existent chez *B. napus*, *B. oleracea* et/ou *B. rapa*. Une première analyse par BLAST sur le génome de *B. napus* suggère la présence d'une copie orthologue du gène At5g47260. Pour le gène At5g47280, nous avons identifié un gène orthologue mais qui a été attribué à un autre membre de la famille des ADR1-like à laquelle il appartient. Cependant, comme les gènes de cette famille sont proches et que le gène At5g47280 est annoté dans le génome d'*A. thaliana* comme un pseudogène, la question d'une mauvaise attribution du membre de la famille ADR1-like au gène orthologue chez *B. napus* se pose. Ces résultats étant issus d'une analyse préliminaire, il serait intéressant de réaliser une analyse plus poussée afin de déterminer si plusieurs copies orthologues existent, si elles sont exprimées de manière constitutive chez *B. napus* comme chez Bur-0 et si une différence d'expression et de méthylation existe entre des génotypes sensibles et partiellement résistants à la hernie chez *B. napus*.

Par ailleurs une approche d'épi/génomique comparative (recherche de régions chromosomiques conservées entre plusieurs espèces) pourrait être réalisée chez les *Brassicaceae* afin de déceler des régions génomiques conservées. Des résultats récents du laboratoire soulignent la présence de 3 QTL conservés entre différentes *Brassicaceae* (*Arabidopsis*, chou et colza ; Evrard *et al.*, en préparation). Dans ce contexte, l'hypothèse que des régulations épigénétiques différentielles seraient conservées entre les *Brassicaceae* semble très prometteuse. En perspective, la recherche des régions orthologues aux régions détectées chez *A. thaliana* (par détection de QTL^{épi}) dans le génome de *B. napus* pourra être possible du fait de la disponibilité du séquençage et de l'annotation du génome du génotype Darmorbzh (Chalhoub *et al.*, 2014). Afin de faciliter cette étude, le logiciel Askomics développé au sein de l'INRIA en collaboration avec l'UMR IGEPP pourra être utilisé pour mettre en relation les différentes régions identifiées chez les *Brassicaceae* au niveau génétique et épigénétique (Evrard *et al.*, 2016). Finalement, une comparaison des gènes présents dans des régions communes ou spécifiques à une espèce et de leur état épigénétique pourrait être effectuée afin de valider ou non leur maintien entre les différentes *Brassicaceae*. Si les résultats obtenus chez *Arabidopsis* sont maintenus chez les différentes *Brassicaceae*, cela permettrait d'obtenir une vision plus large du contrôle génétique et épigénétique de la résistance quantitative à la hernie mais aussi d'ouvrir de nouvelles pistes en sélection pour le développement de variétés résistantes à la hernie.

L'intégration des données épigénétiques, acquises au cours de la thèse, aux données génétiques, déjà disponibles au laboratoire, sur les deux espèces permettra d'établir une vision intégrative des régions génomiques impliquées au niveau génétique et épigénétique dans le contrôle de la résistance quantitative à la hernie chez les *Brassicaceae* et devrait apporter des éléments de réponse sur l'impact des duplications et des réarrangements structuraux sur l'organisation, l'effet et la diversité des facteurs génétiques et épigénétiques impliqués dans la résistance quantitative à la hernie des Crucifères.

5. Utilisation de la variabilité épigénétique et des épiallèles en sélection

Le travail effectué durant ma thèse a permis de mettre en évidence le rôle de variations de méthylation de l'ADN dans la résistance quantitative à la hernie et ceci notamment par l'identification de QTL^{épi} et par l'identification d'une épimutation comme causale d'un QTL génétique précédemment identifié (Jubault *et al.*, 2008b). Ces travaux additionnés à ceux déjà réalisés au sein du laboratoire mettent en avant que la résistance quantitative repose à la fois sur une variabilité génétique et sur une variabilité épigénétique ceci augmentant la complexité des mécanismes mis en jeu dans la résistance à la hernie. Cela suggère que le transfert d'un trait d'intérêt agronomique en sélection repose à la fois sur une transmission d'une séquence d'ADN mais aussi de son état épigénétique qui pour produire le phénotype attendu ne doit pas être modifié entre la plante donneuse et la plante receveuse. Les résultats obtenus sur le QTL *Pb-At5.2* illustrent le fait que cette stabilité des épimutations au cours des processus de sélection est possible. En effet dans le cadre de ce QTL nous avons montré que l'épimutation sous-jacente a été conservée depuis les accessions parentales Bur-0 et Col-0 et jusqu'aux HIF après plusieurs générations.

L'intégration de sources de résistance quantitative d'origine épigénétique en sélection pourrait permettre d'augmenter la durabilité des résistances en mettant ainsi en place un niveau supplémentaire de résistance à contourner par l'agent pathogène. Ainsi dans l'approche de pyramidage de gènes envisagé pour augmenter la durabilité de la résistance (Brun *et al.*, 2010 ; Delourme *et al.*, 2014 ; Pilet-Nayel *et al.*, 2017), un pyramidage de résistances basées sur des allèles et des épiallèles pourrait être une bonne perspective. De plus comme nous l'avons montré avec l'épimutation au QTL *Pb-At5.2*, les épiallèles comme les allèles peuvent être source de variabilité au niveau de gènes possédant un effet sur de multiples isolats (comme c'est le cas ici pour la hernie) mais également potentiellement sur de multiples agents pathogènes. Afin de répondre à ce dernier point il serait intéressant pour le QTL *Pb-At5.2* de tester si son effet est efficace vis-à-vis d'autres agents pathogènes en étudiant le comportement des accessions Col-0, Bur-0 et plus particulièrement des lignées HIF 1381 et 2509 vis-à-vis d'autres agents pathogènes, à galle comme les nématodes à galles du genre *Meloidogyne* ou *Agrobacterium tumefaciens* et/ou même sur des agents pathogènes ne produisant pas de galles.

Cependant, même si des résultats intéressants ont été montrés sur l'implication d'épiallèles dans la mise en place de caractères complexes, l'utilisation d'épiallèles en sélection est encore un challenge. En effet,

l'utilisation d'épiallèles en sélection impose leur identification dans des espèces d'intérêt agronomique. Or pour le moment, un nombre restreint d'épiallèles a été mis en évidence dans des plantes d'intérêt agronomique (Manning *et al.*, 2006 ; Quadrana *et al.*, 2014 ; Ong-Abdullah *et al.*, 2015). Une des difficultés liées à leur identification passe notamment par l'application des techniques d'identification de marques épigénétiques (Bisulfite, CHIP-seq...) largement éprouvées sur des espèces modèles mais qui le sont largement moins sur des espèces de grande culture. En effet les génomes des espèces d'intérêt agronomique sont dans la majorité des cas complexes : i) polyploïde, ii) de grande taille, iii) avec un nombre de séquences répétées important, iv) avec des variations génétiques importantes entre les génotypes et v) souvent assemblé de façon partielle, autant de difficultés techniques à dépasser pour les études d'épigénomique. En effet là où la cartographie de lecture d'approches omiques (RNAseq, bisulfite ...) en des localisations spécifiques d'un génome de référence est une routine sur des espèces modèles, réaliser la même opération sur des génomes présentant des gènes paralogues, orthologues et homéologues sur des génomes de grande taille partiellement annotés est nettement plus complexe. De plus certaines techniques comme le bisulfite nécessitent des niveaux importants de couverture afin d'identifier des régions différenciellement méthylées ce qui est plus difficile à obtenir sur des génomes de grande taille. Enfin, l'utilisation d'une telle source de variabilité en sélection nécessite de bien connaître sa stabilité à travers de multiples générations et d'étudier l'impact potentiel que pourraient avoir les techniques de sélection classiques pour l'introduire dans une espèce d'intérêt.

En conclusion les résultats obtenus au cours de ma thèse ouvrent la voie quant à la prise en compte de variations épigénétiques dans la résistance quantitative basée notamment sur la ségrégation de multiples épiallèles au sein des génomes des végétaux. De manière plus large, une des questions majeures qui reste concerne l'origine des épimutations qui peut être un élément important en sélection. En effet, contrairement à un QTL issu d'un polymorphisme de séquence, un QTL^{épi} pourrait perdre son effet en fonction du fonds génétique. On peut imaginer que suivant le fond génétique, la présence d'un polymorphisme de séquence en cis ou en trans mènerait au passage d'un épiallèle associé à la résistance vers un épiallèle associé à la sensibilité.

REFERENCES
BIBLIOGRAPHIQUES

RÉFÉRENCES BIBLIOGRAPHIQUES

Agarwal A, Kaul V, Faggian R, Rookes J, Ludwig-Müller J, Cahill D. 2011. Analysis of global host gene expression during the primary phase of the *Arabidopsis thaliana*–*Plasmodiophora brassicae* interaction. *Functional Plant Biology* **38**: 462–478.

Agorio A, Vera P. 2007. ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in *Arabidopsis*. *The Plant Cell* **19**: 3778.

Ahmed HU, Hwang SF, Strelkov SE, Gossen BD, Peng G, Howard RJ, Turnbull GD. 2011a. Assessment of bait crops to reduce inoculum of clubroot (*Plasmodiophora brassicae*) of canola. *Canadian Journal of Plant Science* **91**: 545–551.

Ahmed I, Sarazin A, Bowler C, Colot V, Quesneville H. 2011b. Genome-wide evidence for local DNA methylation spreading from small RNA-targeted sequences in *Arabidopsis*. *Nucleic Acids Research* **39**: 6919–6931.

Aigu Y., Laperche A., Mendes J., Lariagon C., Guichard S., Gravot A., Manzanares-Dauleux M. J. 2018. Nitrogen supply exerts a major/minor switch between two QTLs controlling *Plasmodiophora brassicae* spore content in rapeseed. *Plant Pathology*.

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.

Alix K, Lariagon C, Delourme R, Manzanares-Dauleux MJ. 2007. Exploiting natural genetic diversity and mutant resources of *Arabidopsis thaliana* to study the *A. thaliana*–*Plasmodiophora brassicae* interaction. *Plant Breeding* **126**: 218–221.

Aller EST, Jagd LM, Kliebenstein DJ, Burow M. 2018. Comparison of the relative potential for epigenetic and genetic variation to contribute to trait stability. *G3 (Bethesda, Md.)* **8**: 1733–1746.

Allfrey VG, Faulkner R, Mirsky AE. 1964. Acetylation and methylation of histones and their possible role in the regulation of rna synthesis. *Proceedings of the National Academy of Sciences* **51**: 786–794.

Alonso-Blanco C, Andrade J, Becker C, Bemm F, Bergelson J, Borgwardt KM, Cao J, Chae E, Dezaan TM, Ding W. 2016. 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**: 481–491.

Aoun N, Tauleigne L, Lonjon F, Deslandes L, Vaillau F, Roux F, Berthomé R. 2017. Quantitative disease resistance under elevated temperature: genetic basis of new resistance mechanisms to *Ralstonia solanacearum*. *Frontiers in Plant Science* **8**.

***Arabidopsis* Genome I. 2000.** Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature (London)* **408**: 796–815.

Aruga D, Ueno H, Matsumura H, Matsumoto E, Hayashida N. 2013. Distribution of CRa in clubroot resistance (CR) cultivars of Chinese cabbage. *Plant Biotechnology* **30**: 393–397.

Ascenzi R, Gantt JS. 1997. A drought-stress-inducible histone gene in *Arabidopsis thaliana* is a member of a distinct class of plant linker histone variants. *Plant Molecular Biology* **34**: 629–641.

Auer S, Ludwig-Müller J. 2014. Effects of the endophyte *Acremonium alternatum* on oilseed rape (*Brassica napus*) development and clubroot progression. *Albanian Journal of Agricultural Sciences*: **15**.

- Avramova Z. 2009.** Evolution and pleiotropy of TRITHORAX function in *Arabidopsis*. *International Journal of Developmental Biology* **53**: 371–381.
- Axtell MJ. 2013.** ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* **19**: 740–751.
- Ayers GW. 1944.** Studies on the life history of the club root organism, *Plasmodiophora brassicae*. *Canadian Journal of Research* **22c**: 143–149.
- Baggs E, Dagdas G, Krasileva KV. 2017.** NLR diversity, helpers and integrated domains: making sense of the NLR Identity. *Current opinion in plant biology* **38**: 59–67.
- Ballini E, Nguyen TT, Morel J-B. 2013.** Diversity and genetics of nitrogen-induced susceptibility to the blast fungus in rice and wheat. *Rice* **6**: 32.
- Baroux C, Pien S, Grossniklaus U. 2007.** Chromatin modification and remodeling during early seed development. *Genomes and evolution* **17**: 473–479.
- Bashir Z, Ahmad A, Shafique S, Anjum T, Shafique S, Akram W. 2013.** Hypersensitive response — A biophysical phenomenon of producers. *European journal of microbiology & immunology* **3**: 105–110.
- Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R, Schulz I, Reuter G, Aalen RB. 2001.** The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Research* **29**: 4319–4333.
- Baute J, Depicker A. 2008.** Base excision repair and its role in maintaining genome stability. *Critical Reviews in Biochemistry and Molecular Biology* **43**: 239–276.
- Becker C, Hagemann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. 2011.** Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480**: 245.
- BENDER J. 2004.** DNA Methylation of the endogenous PAI genes in *Arabidopsis*. *Cold Spring Harbor Symposia on Quantitative Biology* **69**: 145–154.
- Benson JM, Poland JA, Benson BM, Stromberg EL, Nelson RJ. 2015.** Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. *PLOS Genetics* **11**: e1005045.
- Berger SL. 2007.** The complex language of chromatin regulation during transcription. *Nature* **447**: 407–412.
- Bernatavichute YV, Zhang X, Cokus S, Pellegrini M, Jacobsen SE. 2008.** Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *Plos One* **3**: e3156.
- Bewick AJ, Ji L, Niederhuth CE, Willing E-M, Hofmeister BT, Shi X, Wang L, Lu Z, Rohr NA, Hartwig B, et al. 2016.** On the origin and evolutionary consequences of gene body DNA methylation. *Proceedings of the National Academy of Sciences* **113**: 9111–9116.
- Bewick AJ, Schmitz RJ. 2017.** Gene body DNA methylation in plants. *Current Opinion in Plant Biology* **36**: 103–110.

- Blázquez MA, Santos E, Flores CL, Martínez-Zapater JM, Salinas J, Gancedo C. 1998.** Isolation and molecular characterization of the *Arabidopsis* TPS1 gene, encoding trehalose-6-phosphate synthase. *The Plant Journal: For Cell and Molecular Biology* **13**: 685–689.
- Blevins T, Wang J, Pflieger D, Pontvianne F, Pikaard CS. 2017.** Hybrid incompatibility caused by an epiallele. *Proceedings of the National Academy of Sciences* **114**: 3702.
- Bohnsack CW, Albert LS. 1977.** Early effects of boron deficiency on indoleacetic acid oxidase levels of squash root tips. *Plant Physiology* **59**: 1047–1050.
- Boller T, Felix G. 2009.** A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology* **60**: 379–406.
- Bonardi V, Tang S, Stallmann A, Roberts M, Cherkis K, Dangl JL. 2011.** Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. *Proceedings of the National Academy of Sciences* **108**: 16463–16468.
- Bond DM, Baulcombe DC. 2015.** Epigenetic transitions leading to heritable, RNA-mediated de novo silencing in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **112**: 917.
- Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR, Görlach J. 2001.** Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *The Plant Cell* **13**: 1499–1510.
- Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, Ilnytskyy Y, Hollander J, Meins F Jr, Kovalchuk I. 2010.** Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of Dicer-like proteins. *Plos One* **5**: e9514.
- Brachi B, Morris GP, Borevitz JO. 2011.** Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol* **12**: 232.
- Braselton JP. 1982.** Karyotypic analysis of *Plasmodiophora brassicae* based on serial thin sections of pachytene nuclei. *Canadian Journal of Botany* **60**: 403–408.
- Braselton JP. 1995.** Current Status of the Plasmodiophorids. *Critical Reviews in Microbiology* **21**: 263–275.
- Brodmann A, Schuller A, Ludwig-Müller J, Aeschbacher RA, Wiemken A, Boller T, Wingler A. 2002.** Induction of trehalase in *Arabidopsis* plants infected with the trehalose-producing pathogen *Plasmodiophora brassicae*. *Molecular plant-microbe interactions* **15**: 693–700.
- Brun H, Chèvre A-M, Fitt BDL, Powers S, Besnard A-L, Ermel M, Huteau V, Marquer B, Eber F, Renard M, et al. 2010.** Quantitative resistance increases the durability of qualitative resistance to *Leptosphaeria maculans* in *Brassica napus*. *The New Phytologist* **185**: 285–299.
- Buczacki ST, Cadd SE. 1976.** Size distribution of resting spores of *Plasmodiophora brassicae*. *Transactions of the British Mycological Society* **67**: 133–136.
- Buczacki ST, Ockendon JG. 1979.** Preliminary observations on variation in susceptibility to clubroot among collections of some wild crucifers. *Annals of Applied Biology* **92**: 113–118.
- Buisine N, Quesneville H, Colot V. 2008.** Improved detection and annotation of transposable elements in sequenced genomes using multiple reference sequence sets. *Genomics* **91**: 467–475.

- Burki F, Kudryavtsev A, Matz MV, Aglyamova GV, Bulman S, Fiers M, Keeling PJ, Pawlowski J. 2010.** Evolution of Rhizaria: new insights from phylogenomic analysis of uncultivated protists. *BMC evolutionary biology* **10**: 377.
- Butcher DN, El-Tigani S, Ingram DS. 1974.** The rôle of indole glucosinolates in the club root disease of the *Cruciferae*. *Physiological Plant Pathology* **4**: 127–140.
- Caballero A, Tenesa A, Keightley PD. 2015.** The nature of genetic variation for complex traits revealed by GWAS and regional heritability mapping analyses. *Genetics: genetics*. 115.177220.
- Cai C, Wang X, Liu B, Wu J, Liang J, Cui Y, Cheng F, Wang X. 2017.** *Brassica rapa* genome 2.0: a reference upgrade through sequence re-assembly and gene re-annotation. *Molecular Plant* **10**: 649–651.
- 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.
- Calenge F, Durel C-E. 2006.** Both stable and unstable QTLs for resistance to powdery mildew are detected in apple after four years of field assessments. *Molecular Breeding* **17**: 329–339.
- Calenge F, Faure A, Goerre M, Gebhardt C, Van de Weg WE, Parisi L, Durel C-E. 2004.** Quantitative trait loci (QTL) analysis reveals both broad-spectrum and isolate-specific QTL for scab resistance in an apple progeny challenged with eight isolates of *Venturia inaequalis*. *Phytopathology* **94**: 370–379.
- Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, et al. 2011.** Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nature Genetics* **43**: 956–963.
- Caranta C, Lefebvre V, Palloix A. 1997.** Polygenic resistance of pepper to potyviruses consists of a combination of isolate-specific and broad-spectrum quantitative trait loci. *Molecular plant-microbe interactions* **10**: 872–878.
- Cavalier-Smith T, Chao EE-Y. 2003.** Phylogeny and classification of *Phylum cercozoa* (Protozoa). *Protist* **154**: 341–358.
- Cesari S, Bernoux M, Moncuquet P, Kroj T, Dodds PN. 2014.** A novel conserved mechanism for plant NLR protein pairs: the ‘integrated decoy’ hypothesis. *Frontiers in Plant Science* **5**: 606.
- Chai AL, Xie XW, Shi YX, Li BJ. 2014.** Research status of clubroot (*Plasmodiophora brassicae*) on cruciferous crops in China. *Canadian Journal of Plant Pathology* **36**: 142–153.
- Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, et al. 2014.** Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* **345**: 950–953.
- Chan SW-L, Henderson IR, Jacobsen SE. 2005.** Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Reviews. Genetics* **6**: 351–360.
- Chiang MS, Crête R. 1985.** Acadie: a clubroot resistant cabbage cultivar. *Canadian journal of plant science* **65**: 233–235.

- Chini A, Loake GJ. 2005.** Motifs specific for the ADR1 NBS–LRR protein family in *Arabidopsis* are conserved among NBS–LRR sequences from both dicotyledonous and monocotyledonous plants. *Planta* **221**: 597–601.
- Chisholm ST, Coaker G, Day B, Staskawicz BJ. 2006.** Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**: 803–814.
- Cho SH, Addo-Quaye C, Coruh C, Arif MA, Ma Z, Frank W, Axtell MJ. 2008.** *Physcomitrella patens* DCL3 is required for 22-24 nt siRNA accumulation, suppression of retrotransposon-derived transcripts, and normal development. *PLoS genetics* **4**: e1000314.
- Chong C, Chiang MS, Crete R. 1985.** Studies on glucosinolates in clubroot resistant selections and susceptible commercial cultivars of cabbages. *Euphytica* **34**: 65–73.
- Clark LB, Viswanathan P, Quigley G, Chiang Y-C, McMahon JS, Yao G, Chen J, Nelsbach A, Denis CL. 2004.** Systematic mutagenesis of the leucine-rich repeat (LRR) domain of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in CCR4 deadenylase activity. *Journal of Biological Chemistry*.
- 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.
- Coleman-Derr D, Zilberman D. 2012.** Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. *PLOS Genetics* **8**: e1002988.
- Colhoun J. 1952.** Factors affecting the incidence of club root disease of *Brassicaceae*. *Nature* **169**: 21–22.
- Colhoun J. 1953.** A study of the epidemiology of club-root disease of *Brassicaceae*. *Annals of Applied Biology* **40**: 262–283.
- Collier SM, Hamel L-P, Moffett P. 2011.** Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Molecular plant-microbe interactions* **24**: 918–931.
- Colomé-Tatché M, Cortijo S, Wardenaar R, Morgado L, Lahouze B, Sarazin A, Etcheverry M, Martin A, Feng S, Duvernois-Berthet E, et al. 2012a.** Features of the *Arabidopsis* recombination landscape resulting from the combined loss of sequence variation and DNA methylation. *Proceedings of the National Academy of Sciences* **109**: 16240.
- Cortijo S, Wardenaar R, Colomé-Tatché M, Gilly A, Etcheverry M, Labadie K, Caillieux E, Hospital F, Aury J-M, Wincker P, et al. 2014.** Mapping the epigenetic basis of complex traits. *Science* **343**: 1145–1148.
- Corwin JA, Copeland D, Feusier J, Subedy A, Eshbaugh R, Palmer C, Maloof J, Kliebenstein DJ. 2016.** The quantitative basis of the *Arabidopsis* innate immune system to endemic pathogens depends on pathogen genetics. *PLOS Genetics* **12**: e1005789.
- Corwin JA, Kliebenstein DJ. 2017.** Quantitative resistance: more than just perception of a pathogen. *The Plant Cell* **29**: 655–665.
- Crute IR, Gray AR, Crisp P, Buczacki ST. 1980.** Variation in *Plasmodiophora brassicae* and resistance to clubroot disease in *Brassicaceae* and allied crops - a critical review. *Plant Breeding Abstracts* **50**: 91–104.

- Cubas P, Vincent C, Coen E. 1999.** An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**: 157.
- Dangl JL, Jones JD. 2001.** Plant pathogens and integrated defence responses to infection. *Nature* **411**: 826–833.
- Darvishzadeh R., Kiani S. Poormohammad, Dechamp-Guillaume G., Gentzbittel L., Sarrafi A. 2007.** Quantitative trait loci associated with isolate specific and isolate nonspecific partial resistance to *Phoma macdonaldii* in sunflower. *Plant Pathology* **56**: 855–861.
- Daval S, Belcour A, Gazengel K, Legrand L, Gouzy J, Cottret L, Lebreton L, Aigu Y, Mougel C, Manzanares-Dauleux M. 2018.** Computational analysis of the *Plasmodiophora brassicae* genome: mitochondrial sequence description and metabolic pathway database design. *bioRxiv*: 335406.
- Daxinger L, Kanno T, Bucher E, van der Winden J, Naumann U, Matzke AJM, Matzke M. 2009.** A stepwise pathway for biogenesis of 24-nt secondary siRNAs and spreading of DNA methylation. *The EMBO journal* **28**: 48–57.
- Deal RB, Topp CN, McKinney EC, Meagher RB. 2007.** Repression of flowering in *Arabidopsis* requires activation of FLOWERING LOCUS C expression by the histone variant H2A. *Z. The Plant Cell* **19**: 74–83.
- Debieu M, Huard-Chauveau C, Genissel A, Roux F, Roby D. 2016.** Quantitative disease resistance to the bacterial pathogen *Xanthomonas campestris* involves an *Arabidopsis* immune receptor pair and a gene of unknown function. *Molecular plant pathology* **17**: 510–520.
- Dekhuijzen HM. 1981.** The occurrence of free and bound cytokinins in plasmodia of *Plasmodiophora brassicae* isolated from tissue cultures of clubroots. *Plant Cell Reports* **1**: 18–20.
- Dekhuijzen HM, Overeem JC. 1971.** The role of cytokinins in clubroot formation. *Physiological Plant Pathology* **1**: 151–161.
- Deleris A, Halter T, Navarro L. 2016.** DNA methylation and demethylation in plant immunity. *Annual review of phytopathology* **54**: 579–603.
- Delourme R, Bousset L, Ermel M, Duffe P, Besnard A-L, Marquer B, Fudal I, Linglin J, Chadoeuf J, Brun H. 2014a.** Quantitative resistance affects the speed of frequency increase but not the diversity of the virulence alleles overcoming a major resistance gene to *Leptosphaeria maculans* in oilseed rape. *Infection, genetics and evolution* **27**: 490–499.
- Delourme R, Chevre AM, Brun H, Rouxel T, Balesdent MH, Dias JS, Salisbury P, Renard M, Rimmer SR. 2006.** Major gene and polygenic resistance to *Leptosphaeria maculans* in oilseed rape (*Brassica napus*). *European Journal of Plant Pathology* **114**: 41–52.
- Deng Y, Liu M, Li X, Li F. 2018.** microRNA-mediated R gene regulation: molecular scabbards for double-edged swords. *Science China Life Sciences*: 1–10.
- Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M, Boucher C, Somssich I, Genin S, Marco Y. 2003.** Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proceedings of the National Academy of Sciences* **100**: 8024–8029.
- Devos S, Laukens K, Deckers P, Van Der Straeten D, Beeckman T, Inzé D, Van Onckelen H, Witters E, Prinsen E. 2006.** A hormone and proteome approach to picturing the initial metabolic events

during *Plasmodiophora brassicae* infection on *Arabidopsis*. *Molecular Plant-Microbe Interactions* **19**: 1431–1443.

Diederichsen E., Sacristan M. D. 1996. Disease response of resynthesized *Brassica napus* L. lines carrying different combinations of resistance to *Plasmodiophora brassicae* Wor. *Plant Breeding* **115**: 5–10.

Diederichsen E, Frauen M, Linders EGA, Hatakeyama K, Hirai M. 2009. Status and perspectives of clubroot resistance breeding in crucifer crops. *Journal of Plant Growth Regulation* **28**: 265–281.

Diener AC, Ausubel FM. 2005. Resistance to *Fusarium oxysporum* 1, a dominant *Arabidopsis* disease-resistance gene, is not race specific. *Genetics* **171**: 305.

Dixon GR. 2009a. *Plasmodiophora brassicae* in its environment. *Journal of Plant Growth Regulation* **28**: 212–228.

Dixon GR. 2009b. The Occurrence and Economic Impact of *Plasmodiophora brassicae* and Clubroot Disease. *Journal of Plant Growth Regulation* **28**: 194–202.

Dixon GR, Page LV. 1998. Calcium and nitrogen eliciting alterations to growth and reproduction of *Plasmodiophora brassicae* (Clubroot). *Acta Horticulturae* **459**: 343–349.

Doan TT, Jäschke D, Ludwig-Müller J. 2010. An endophytic fungus induces tolerance against the clubroot pathogen *Plasmodiophora brassicae* in *Arabidopsis thaliana* and *Brassica rapa* roots. *Acta Horticulturae* **867**: 173–180.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

Dodds PN, Lawrence GJ, Catanzariti A-M, Teh T, Wang C-IA, Ayliffe MA, Kobe B, Ellis JG. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences* **103**: 8888–8893.

Dodds PN, Rathjen JP. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews. Genetics* **11**: 539–548.

Donald EC, Cross SJ, Lawrence JM, Porter IJ. 2006. Pathotypes of *Plasmodiophora brassicae*, the cause of clubroot, in Australia. *Annals of Applied Biology* **148**: 239–244.

Donald C, Porter I. 2009. Integrated Control of Clubroot. *Journal of Plant Growth Regulation* **28**: 289–303.

Dong OX, Tong M, Bonardi V, Kasmi FE, Woloshen V, Wunsch LK, Dangl JL, Li X. 2016. TNL-mediated immunity in *Arabidopsis* requires complex regulation of the redundant ADR1 gene family. *New Phytologist* **210**: 960–973.

Downen RH, Pelizzola M, Schmitz RJ, Lister R, Downen JM, Nery JR, Dixon JE, Ecker JR. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences* **109**: E2183.

Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, et al. 2012. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* **151**: 167–180.

- Dunn DB, Smith JD. 1958.** The occurrence of 6-methylaminopurine in deoxyribonucleic acids. *Biochemical Journal* **68**: 627–636.
- Dunoyer P, Brosnan CA, Schott G, Wang Y, Jay F, Alioua A, Himber C, Voinnet O. 2010.** An endogenous, systemic RNAi pathway in plants. *The EMBO journal* **29**: 1699–1712.
- Durand S, Bouché N, Perez Strand E, Loudet O, Camilleri C. 2012.** Rapid establishment of genetic incompatibility through natural epigenetic variation. *Current Biology* **22**: 326–331.
- Ebbs ML, Bartee L, Bender J. 2005.** H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Molecular and Cellular Biology* **25**: 10507–10515.
- Ebbs ML, Bender J. 2006.** Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *The Plant Cell* **18**: 1166–1176.
- Ellis JG, Lagudah ES, Spielmeier W, Dodds PN. 2014.** The past, present and future of breeding rust resistant wheat. *Frontiers in plant science* **5**: 641.
- Ernst K, Kumar A, Kriseleit D, Kloos D-U, Phillips MS, Ganai MW. 2002.** The broad-spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. *The Plant Journal* **31**: 127–136.
- Evans JL, Scholes JD. 1995.** How does clubroot alter the regulation of carbon metabolism in its host? *Aspects of Applied Biology*: 125–132.
- Evrard A, Bettembourg C, Jubault M, Dameron O, Filangi O, Bretaudeau A, Legeai F. 2016.** Integration and query of biological datasets with Semantic Web technologies: AskOmics. *Journées Ouvertes Biologie, Informatique et Mathématiques*.
- Fähling M, Graf H, Siemens J. 2003.** Pathotype separation of *Plasmodiophora brassicae* by the host plant. *Journal of Phytopathology* **151**: 425–430.
- Fähling M, Graf H, Siemens J. 2004.** Characterization of a single-spore isolate population of *Plasmodiophora brassicae* resulting from a single club. *Journal of Phytopathology* **152**: 438–444.
- Fei Q, Xia R, Meyers BC. 2013.** Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *The Plant Cell* **25**: 2400.
- 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 J, Hwang RU, HWANG S-F, Strelkov SE, Gossen BD, ZHOU Q-X, Peng G. 2010b.** Molecular characterization of a serine protease Pro1 from *Plasmodiophora brassicae* that stimulates resting spore germination. *Molecular plant pathology* **11**: 503–512.
- Feng S, Jacobsen SE, Reik W. 2010c.** Epigenetic reprogramming in plant and animal development. *Science* **330**: 622–627.
- Figdore SS, Ferreira ME, Slocum MK, Williams PH. 1993.** Association of RFLP markers with trait loci affecting clubroot resistance and morphological characters in *Brassica oleracea* L. *Euphytica* **69**: 33–44.
- Fink GR. 1998.** Anatomy of a revolution. *Genetics* **149**: 473.

- Flor HH. 1971.** Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9**: 275–296.
- Foolad M, Zhang L, Khan AA, Niño-Liu D, Lin G. 2002.** Identification of QTLs for early blight (*Alternaria solani*) resistance in tomato using backcross populations of a *Lycopersicon esculentum* × *L. hirsutum* cross. *Theoretical and Applied Genetics* **104**: 945–958.
- French E, Kim B-S, Iyer-Pascuzzi AS. 2016.** Mechanisms of quantitative disease resistance in plants. *Seminars in Cell & Developmental Biology* **56**: 201–208.
- Friberg H, Lagerlöf J, Rämert B. 2006.** Usefulness of nonhost plants in managing *Plasmodiophora brassicae*. *Plant Pathology* **55**: 690–695.
- Fu ZQ, Dong X. 2013.** Systemic acquired resistance: turning local infection into global defense. *Annual Review of Plant Biology* **64**: 839–863.
- Fuchs H, Sacristán MD. 1996.** Identification of a gene in *Arabidopsis thaliana* controlling resistance to clubroot (*Plasmodiophora brassicae*) and characterization of the resistance response.
- Fukagawa T, Earnshaw WC. 2014.** The centromere: chromatin foundation for the kinetochore machinery. *Developmental Cell* **30**: 496–508.
- Fukuoka S, Yamamoto S-I, Mizobuchi R, Yamanouchi U, Ono K, Kitazawa N, Yasuda N, Fujita Y, Nguyen TTT, Koizumi S. 2014.** Multiple functional polymorphisms in a single disease resistance gene in rice enhance durable resistance to blast. *Scientific Reports* **4**: 4550.
- Fultz D, Choudury SG, Slotkin RK. 2015.** Silencing of active transposable elements in plants. *Current Opinion in Plant Biology* **27**: 67–76.
- Fusco G, Carrer R, Serrelli E. 2014.** The landscape metaphor in development. *Towards a Theory of Development*. 114–128.
- Gabriëls SHEJ, Vossen JH, Ekengren SK, van Ooijen G, Abd-El-Haliem AM, van den Berg GCM, Rainey DY, Martin GB, Takken FLW, de Wit PJGM, et al. 2007.** An NB-LRR protein required for HR signalling mediated by both extra- and intracellular resistance proteins. *The Plant Journal: For Cell and Molecular Biology* **50**: 14–28.
- Ganguly DR, Crisp PA, Eichten SR, Pogson BJ. 2017.** The *Arabidopsis* DNA methylome is stable under transgenerational drought stress. *Plant Physiology* **175**: 1893.
- Gehring M, Reik W, Henikoff S. 2009.** DNA demethylation by DNA repair. *Trends in genetics: TIG* **25**: 82–90.
- Gendler K, Paulsen T, Napoli C. 2008.** ChromDB: the chromatin database. *Nucleic Acids Research* **36**: D298–D302.
- Giolai M, Paaajanen P, Verweij W, Witek K, Jones JDG, Clark MD. 2017.** Comparative analysis of targeted long read sequencing approaches for characterization of a plant's immune receptor repertoire. *BMC genomics* **18**: 564.
- Glazebrook J. 2005.** Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology* **43**: 205–227.
- Gong Z, Morales-Ruiz T, Ariza RR, Roldán-Arjona T, David L, Zhu JK. 2002.** ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell* **111**: 803–814.

- Gorovsky MA, Hattman S, Pleger GL. 1973.** [6n]methyl adenine in the nuclear DNA of a eucaryote, *Tetrahymena pyriformis*. *The Journal of Cell Biology* **56**: 697–701.
- Gossen BD, Deora A, Peng G, Hwang S-F, McDonald MR. 2014.** Effect of environmental parameters on clubroot development and the risk of pathogen spread. *Canadian Journal of Plant Pathology* **36**: 37–48.
- Gossen BD, Kasinathan H, Cao T, Manolii VP, Strelkov SE, Hwang S-F, McDonald MR. 2013.** Interaction of pH and temperature affect infection and symptom development of *Plasmodiophora brassicae* in canola. *Canadian Journal of Plant Pathology* **35**: 294–303.
- van der Graaf A, Wardenaar R, Neumann DA, Taudt A, Shaw RG, Jansen RC, Schmitz RJ, Colomé-Tatché M, Johannes F. 2015.** Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proceedings of the National Academy of Sciences* **112**: 6676.
- Graf H, Föhling M, Siemens J. 2004.** Chromosome polymorphism of the obligate biotrophic parasite *Plasmodiophora brassicae*. *Journal of Phytopathology* **152**: 86–91.
- Graf H, Sokolowski F, Klewer A, Diederichsen E, Luerßen H, Siemens J. 2001.** Electrophoretic karyotype of the obligate biotrophic parasite *Plasmodiophora brassicae* Wor. *Journal of Phytopathology* **149**: 313–318.
- Grandclément C, Thomas G. 1996.** Detection and analysis of QTLs based on RAPD markers for polygenic resistance to *Plasmodiophora brassicae* Woron in *Brassica oleracea* L. *Theoretical and Applied Genetics* **93**: 86–90.
- Grandperret V, Nicolas-Francès V, Wendehenne D, Bourque S. 2013.** Type-II histone deacetylases: elusive plant nuclear signal transducers. *Plant, Cell & Environment* **37**: 1259–1269.
- Grant JJ, Chini A, Basu D, Loake GJ. 2003.** Targeted activation tagging of the *Arabidopsis* NBS-LRR gene, ADR1, conveys resistance to virulent pathogens. *Molecular Plant-Microbe Interactions* **16**: 669–680.
- Gravot A, Grillet L, Wagner G, Jubault M, Lariagon C, Baron C, Deleu C, Delourme R, Bouchereau A, Manzanares-Dauleux MJ. 2011.** Genetic and physiological analysis of the relationship between partial resistance to clubroot and tolerance to trehalose in *Arabidopsis thaliana*. *New Phytologist* **191**: 1083–1094.
- Gravot A, Deleu C, Wagner G, Lariagon C, Lugan R, Todd C, Wendehenne D, Delourme R, Bouchereau A, Manzanares-Dauleux MJ. 2012.** Arginase induction represses gall development during clubroot infection in *Arabidopsis*. *Plant and Cell Physiology* **53**: 901–911.
- Gravot A, Lemarié S, Richard G, Lime T, Lariagon C, Manzanares-Dauleux MJ. 2016a.** Flooding affects the development of *Plasmodiophora brassicae* in *Arabidopsis* roots during the secondary phase of infection. *Plant Pathology* **65**: 1153–1160.
- Gravot A, Richard G, Lime T, Lemarié S, Jubault M, Lariagon C, Lemoine J, Vicente J, Robert-Seilaniantz A, Holdsworth MJ, et al. 2016b.** Hypoxia response in *Arabidopsis* roots infected by *Plasmodiophora brassicae* supports the development of clubroot. *BMC Plant Biology* **16**: 251.
- Greer EL, Blanco MA, Gu L, Sendinc E, Liu J, Aristizábal-Corrales D, Hsu C-H, Aravind L, He C, Shi Y. 2015.** DNA Methylation on N6-Adenine in *C. elegans*. *Cell* **161**: 868–878.
- Gu Lee J, Soon Kim J. 2010.** Variation of glucosinolate content in the root of susceptible and resistant chinese cabbage cultivars during development of clubroot disease.

- Gurard-Levin ZA, Quivy J-P, Almouzni G. 2014.** Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annual Review of Biochemistry* **83**: 487–517.
- Ha M. 2013.** Understanding the chromatin remodeling code. *Plant Science: An International Journal of Experimental Plant Biology* **211**: 137–145.
- Halkier BA, Gershenzon J. 2006.** Biology and biochemistry of glucosinolates. *Annual Review of Plant Biology* **57**: 303–333.
- Hatakeyama K, Niwa T, Kato T, Ohara T, Kakizaki T, Matsumoto S. 2017.** The tandem repeated organization of NB-LRR genes in the clubroot-resistant CRb locus in *Brassica rapa* L. *Molecular Genetics and Genomics* **292**: 397–405.
- Hatakeyama K, Suwabe K, Tomita RN, Kato T, Nunome T, Fukuoka H, Matsumoto S. 2013.** Identification and Characterization of Crr1a, a Gene for Resistance to Clubroot Disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *Plos One* **8**: e54745.
- Hattman S, Kenny C, Berger L, Pratt K. 1978.** Comparative study of DNA methylation in three unicellular eucaryotes. *Journal of Bacteriology* **135**: 1156–1157.
- Hayashi N, Inoue H, Kato T, Funao T, Shiota M, Shimizu T, Kanamori H, Yamane H, Hayano-Saito Y, Matsumoto T. 2010.** Durable panicle blast-resistance gene Pb1 encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through local genome duplication. *The Plant Journal* **64**: 498–510.
- Hayashi Keiko, Yoshida Hitoshi. 2009.** Refunctionalization of the ancient rice blast disease resistance gene Pit by the recruitment of a retrotransposon as a promoter. *The Plant Journal* **57**: 413–425.
- Hayashida N, Takabatake Y, Nakazawa N, Aruga D, Nakanishi H, Taguchi G, Sakamoto K, Matsumoto E. 2008.** Construction of a practical SCAR marker linked to clubroot resistance in Chinese cabbage, with intensive analysis of HC352b genes. *Journal of the Japanese Society for Horticultural Science* **77**: 150–154.
- He D, Fiz-Palacios O, Fu C-J, Fehling J, Tsai C-C, Baldauf SL. 2014.** An alternative root for the eukaryote tree of life. *Current biology: CB* **24**: 465–470.
- He L, Wu W, Zinta G, Yang L, Wang D, Liu R, Zhang H, Zheng Z, Huang H, Zhang Q, et al. 2018.** A naturally occurring epiallele associates with leaf senescence and local climate adaptation in *Arabidopsis* accessions. *Nature Communications* **9**: 460.
- Heard E, Martienssen RA. 2014.** Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* **157**: 95–109.
- Henderson IR, Jacobsen SE. 2007.** Epigenetic inheritance in plants. *Nature* **447**: 418–424.
- Hewezi T, Lane T, Piya S, Rambani A, Rice JH, Staton M. 2017.** Cyst nematode parasitism induces dynamic changes in the root epigenome. *Plant Physiology* **174**: 405.
- Hirai M, Harada T, Kubo N, Tsukada M, Suwabe K, Matsumoto S. 2004.** A novel locus for clubroot resistance in *Brassica rapa* and its linkage markers. *Theoretical and applied genetics*. **108**: 639–643.
- HONIG F. 1931.** The pathogene of club root (*Plasmodiophora brassicae* Wor.). A monograph. *Gartenbauwissenschaft* **5**: 116–225.

- van der Hoorn RAL, Kamoun S. 2008.** From guard to decoy: a new model for perception of plant pathogen effectors. *The Plant Cell* **20**: 2009–2017.
- Howard RJ, Strelkov SE, Harding MW. 2010.** Clubroot of cruciferous crops – new perspectives on an old disease. *Canadian Journal of Plant Pathology* **32**: 43–57.
- Huang Y, Kendall T, Mosher RA. 2013.** Pol IV-Dependent siRNA production is reduced in *Brassica rapa*. *Biology* **2**: 1210–1223.
- Huard-Chauveau C, Perchepied L, Debieu M, Rivas S, Kroj T, Kars I, Bergelson J, Roux F, Roby D. 2013.** An atypical kinase under balancing selection confers broad-spectrum disease resistance in *Arabidopsis*. *PLoS genetics* **9**: e1003766.
- Huh JH, Bauer MJ, Hsieh T-F, Fischer RL. 2008.** Cellular programming of plant gene imprinting. *Cell* **132**: 735–744.
- Hurni S, Scheuermann D, Krattinger SG, Kessel B, Wicker T, Herren G, Fitze MN, Breen J, Presterl T, Ouzunova M. 2015.** The maize disease resistance gene Htn1 against northern corn leaf blight encodes a wall-associated receptor-like kinase. *Proceedings of the National Academy of Sciences*: 201502522.
- Hwang S-F, Strelkov SE, Feng J, Gossen BD, Howard RJ. 2012.** *Plasmodiophora brassicae*: a review of an emerging pathogen of the Canadian canola (*Brassica napus*) crop. *Molecular Plant Pathology* **13**: 105–113.
- Hwang SF, Strelkov SE, Gossen BD, Turnbull GD, Ahmed HU, Manolii VP. 2011.** Soil treatments and amendments for amelioration of clubroot of canola. *Canadian Journal of Plant Science* **91**: 999–1010.
- Ikegami H. 1981.** *Growth of Plasmodiophora brassicae in the root and callus of Chinese cabbage. Chinese cabbage.*
- Ikegami H. 1985.** Decrease of clubroot fungus by cultivation of different crops in heavily infested soil. *Bull. Fac. Agric. Gifu Univ.* **50**: 19–32.
- Ikegami H, Mukobata H, Naiki T. 1978.** Scanning electron microscopy of *Plasmodiophora brassicae* in diseased root cells of turnip and Chinese cabbage. *Japanese Journal of Phytopathology* **44**: 456–464.
- Ingram DS, Tommerup IC. 1972.** The life history of *Plasmodiophora brassicae* Woron. *Proc. R. Soc. Lond. B* **180**: 103–112.
- Innes RW, Ameline-Torregrosa C, Ashfield T, Cannon E, Cannon SB, Chacko B, Chen NW, Couloux A, Dalwani A, Denny R. 2008.** Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. *Plant Physiology*.
- Ito S, Yano S, Tanaka S, Kameya-Iwaki M. 1994.** The use of resting spore spheroplasts in the DNA analysis of *Plasmodiophora brassicae*. *Japanese Journal of Phytopathology* **60**: 491–495.
- J. Bryan R, T. Trese A, P. Braselton J. 1996.** Molecular karyotypes for the obligate, intracellular, plant pathogens, *Plasmodiophora brassicae* and *Spongospora subterranea*. *Mycologia* **88**: 358.
- Jackson JP, Lindroth AM, Cao X, Jacobsen SE. 2002.** Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**: 556–560.

- Jacob Y, Feng S, LeBlanc CA, Bernatavichute YV, Stroud H, Cokus S, Johnson LM, Pellegrini M, Jacobsen SE, Michaels SD. 2009.** ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nature Structural & Molecular Biology* **16**: 763.
- Jacob F, Monod J. 1961.** Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**: 318–356.
- Jacobsen SE, Meyerowitz EM. 1997.** Hypermethylated SUPERMAN Epigenetic Alleles in *Arabidopsis*. *Science* **277**: 1100.
- Jacobsen SE, Sakai H, Finnegan EJ, Cao X, Meyerowitz EM. 2000.** Ectopic hypermethylation of flower-specific genes in *Arabidopsis*. *Current Biology* **10**: 179–186.
- Jäschke D, Dugassa-Gobena D, Karlovsky P, Vidal S, Ludwig-Müller J. 2010.** Suppression of clubroot (*Plasmodiophora brassicae*) development in *Arabidopsis thaliana* by the endophytic fungus *Acremonium alternatum*. *Plant pathology* **59**: 100–111.
- Jenuwein T, Allis CD. 2001.** Translating the histone code. *Science* **293**: 1074.
- Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B. 2000.** Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO journal* **19**: 4004–4014.
- Jiang D, Berger F. 2017.** Histone variants in plant transcriptional regulation. *Biochimica et biophysica acta* **1860**: 123–130.
- Jiang C, Mithani A, Belfield EJ, Mott R, Hurst LD, Harberd NP. 2014.** Environmentally responsive genome-wide accumulation of de novo *Arabidopsis thaliana* mutations and epimutations. *Genome Research*: gr.177659.114.
- Johannes F, Colot V, Jansen RC. 2008.** Epigenome dynamics: a quantitative genetics perspective. *Nature Reviews. Genetics* **9**: 883–890.
- Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuissou J, Heredia F, Audigier P. 2009.** Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genetics* **5**: e1000530.
- Jones JDG, Dangl JL. 2006.** The plant immune system. *Nature* **444**: 323–329.
- Jones JDR, Ingram DS, Dixon GR. 1982a.** Characterization of isolates derived from single resting spores of *Plasmodiophora brassicae* and studies of their interaction. *Plant Pathology* **31**: 239–246.
- Jones JDR, Ingram DS, Dixon GR. 1982b.** Factors affecting tests for differential pathogenicity in populations of *Plasmodiophora brassicae*. *Plant Pathology* **31**: 229–238.
- Jones JDG, Vance RE, Dangl JL. 2016.** Intracellular innate immune surveillance devices in plants and animals. *Science (New York, N.Y.)* **354**.
- Jubault M, Hamon C, Gravot A, Lariagon C, Delourme R, Bouchereau A, Manzanares-Dauleux MJ. 2008a.** Differential regulation of root arginine catabolism and polyamine metabolism in clubroot-susceptible and partially resistant *Arabidopsis* genotypes. *Plant Physiology* **146**: 2008–2019.
- Jubault M, Lariagon C, Simon M, Delourme R, Manzanares-Dauleux MJ. 2008b.** Identification of quantitative trait loci controlling partial clubroot resistance in new mapping populations of *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **117**: 191–202.

- Jubault M, Lariagon C, Tacconat L, Renou J-P, Gravot A, Delourme R, Manzanares-Dauleux MJ. 2013.** Partial resistance to clubroot in *Arabidopsis* is based on changes in the host primary metabolism and targeted cell division and expansion capacity. *Functional & Integrative Genomics* **13**: 191–205.
- Kageyama K, Asano T. 2009.** Life cycle of *Plasmodiophora brassicae*. *Journal of Plant Growth Regulation* **28**: 203–211.
- Kakutani T, Jeddelloh JA, Flowers SK, Munakata K, Richards EJ. 1996.** Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proceedings of the National Academy of Sciences* **93**: 12406–12411.
- Kakutani T, Munakata K, Richards EJ, Hirochika H. 1999.** Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of *Arabidopsis thaliana*. *Genetics* **151**: 831–838.
- Kankel MW, Ramsey DE, Stokes TL, Flowers SK, Haag JR, Jeddelloh JA, Riddle NC, Verbsky ML, Richards EJ. 2003.** *Arabidopsis* MET1 cytosine methyltransferase mutants. *Genetics* **163**: 1109.
- Karasov TL, Kniskern JM, Gao L, DeYoung BJ, Ding J, Dubiella U, Lastra RO, Nallu S, Roux F, Innes RW. 2014.** The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* **512**: 436.
- Kato T, Hatakeyama K, Fukino N, Matsumoto S. 2012.** Identification of a clubroot resistance locus conferring resistance to a *Plasmodiophora brassicae* classified into pathotype group 3 in Chinese cabbage (*Brassica rapa* L.). *Breeding science* **62**: 282–287.
- Kato T, Hatakeyama K, Fukino N, Matsumoto S. 2013.** Fine mapping of the clubroot resistance gene CRb and development of a useful selectable marker in *Brassica rapa*. *Breeding science* **63**: 116–124.
- Kato M, Miura A, Bender J, Jacobsen SE, Kakutani T. 2003.** Role of CG and non-CG methylation in immobilization of transposons in *Arabidopsis*. *Current Biology* **13**: 421–426.
- Kavanagh JA, Williams PH. 1981.** Indole auxins in *Plasmodiophora* infected cabbage roots and hypocotyls. *Transactions of the British Mycological Society* **77**: 125–129.
- Kawakatsu T, Huang S-SC, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R, Nery JR, Barragan C, He Y, et al. 2016.** Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell* **166**: 492–505.
- Kawashima T, Lorković ZJ, Nishihama R, Ishizaki K, Axelsson E, Yelagandula R, Kohchi T, Berger F. 2015.** Diversification of histone H2A variants during plant evolution. *Trends in Plant Science* **20**: 419–425.
- Keen NT. 1990.** Gene-for-gene complementarity in plant-pathogen interactions. *Annual Review of Genetics* **24**: 447–463.
- Keen NT, Williams PH. 1969.** Translocation of sugars into infected cabbage tissues during clubroot development 1. *Plant Physiology* **44**: 748–754.
- Kim HS, Diers BW. 2000.** Inheritance of partial resistance to *Sclerotinia* stem rot in soybean research supported by the Michigan agricultural experiment station and grants from the Michigan soybean promotion committee. *Crop Science* **40**: 55–61.

- Kolmer JA, Leonard KJ. 1986.** Genetic selection and adaptation of *Cochliobolus heterostrophus* to corn hosts with partial resistance. *Phytopathology (USA)*.
- Kooke R, Johannes F, Wardenaar R, Becker F, Etcheverry M, Colot V, Vreugdenhil D, Keurentjes JJ. 2015a.** Epigenetic basis of morphological variation and phenotypic plasticity in *Arabidopsis thaliana*. *The Plant Cell Online*: tpc. 114.133025.
- Kooke R, Johannes F, Wardenaar R, Becker F, Etcheverry M, Colot V, Vreugdenhil D, Keurentjes JJB. 2015b.** Epigenetic basis of morphological variation and phenotypic plasticity in *Arabidopsis thaliana*. *The Plant Cell* **27**: 337.
- Kopsell DA, Kopsell DE. 2006.** Accumulation and bioavailability of dietary carotenoids in vegetable crops. *Trends in Plant Science* **11**: 499–507.
- Kouzarides T. 2007.** Chromatin modifications and their function. *Cell* **128**: 693–705.
- Krasileva KV, Dahlbeck D, Staskawicz BJ. 2010.** Activation of an *Arabidopsis* resistance protein is specified by the in-planta association of its leucine-rich repeat domain with the cognate oomycete effector. *The Plant Cell* **22**: 2444–2458.
- Kroj T, Chanclud E, Michel-Romiti C, Grand X, Morel J-B. 2016.** Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *The New Phytologist* **210**: 618–626.
- Kumar V, Paillard S, Fopa-Fomeju B, Falentin C, Deniot G, Baron C, Vallée P, Manzanares-Dauleux MJ, Delourme R. 2018.** Multi-year linkage and association mapping confirm the high number of genomic regions involved in oilseed rape quantitative resistance to blackleg. *Theoretical and Applied Genetics*.
- Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Zwonitzer JC, Kresovich S, McMullen MD, Ware D, et al. 2011.** Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nature Genetics* **43**: 163.
- Kushalappa AC, Yogendra KN, Karre S. 2016.** Plant innate immune response: qualitative and quantitative resistance. *Critical Reviews in Plant Sciences* **35**: 38–55.
- Lahlali R, Peng G, Gossen BD, McGregor L, Yu FQ, Hynes RK, Hwang SF, McDonald MR, Boyetchko SM. 2013.** Evidence that the biofungicide Serenade (*Bacillus subtilis*) suppresses clubroot on canola via antibiosis and induced host resistance. *Phytopathology* **103**: 245–254.
- Lahlali R, Peng G, McGregor L, Gossen BD, Hwang SF, McDonald M. 2011.** Mechanisms of the biofungicide Serenade (*Bacillus subtilis* QST713) in suppressing clubroot. *Biocontrol Science and Technology* **21**: 1351–1362.
- Lai Y, Eulgem T. 2018.** Transcript-level expression control of plant NLR genes. *Molecular Plant Pathology* **19**: 1267–1281.
- Landry BS, Hubert N, Crete R, Chang MS, Lincoln SE, Etoh T. 1992.** A genetic map for *Brassica oleracea* based on RFLP markers detected with expressed DNA sequences and mapping of resistance genes to race 2 of *Plasmodiophora brassicae* (Woronin). *Genome* **35**: 409–420.
- Laperche A, Aigu Y, Jubault M, Ollier M, Guichard S, Glory P, Strelkov S, Gravot A, Manzanares-Dauleux M. 2017.** Clubroot resistance QTL are modulated by nitrogen input in *Brassica napus*. *Theoretical and Applied Genetics* **130**: 669–684.

Larkan NJ, Lydiate DJ, Parkin IAP, Nelson MN, Epp DJ, Cowling WA, Rimmer SR, Borhan MH. 2013. The *Brassica napus* blackleg resistance gene LepR3 encodes a receptor-like protein triggered by the *Leptosphaeria maculans* effector AVR1M1. *New Phytologist* **197**: 595–605.

Larson RH, Walker JC. 1934. Soil treatment in relation to clubroot of cabbage. *J Agric Res* **48**: 749–759.

Law CW, Alhamdoosh M, Su S, Smyth GK, Ritchie ME. 2016. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research* **5**.

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

Le T-N, Schumann U, Smith NA, Tiwari S, Au PCK, Zhu Q-H, Taylor JM, Kazan K, Llewellyn DJ, Zhang R, et al. 2014. DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. *Genome Biology* **15**: 458.

Lee M-L, Hsieh W-H. 1992. Influence of pH, sodium and potassium salts on viability of resting spores of *Plasmodiophora brassicae*. *Plant Pathology* **1**: 1.

Lee J, Izzah NK, Choi B-S, Joh HJ, Lee S-C, Perumal S, Seo J, Ahn K, Jo EJ, Choi GJ, et al. 2015. Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA Research*: dsv034.

Lee J-S, Shilatifard A. 2007. A site to remember: H3K36 methylation a mark for histone deacetylation. *Chromatin: Repair, Remodeling and Regulation* **618**: 130–134.

Lee J-S, Smith E, Shilatifard A. 2010. The Language of Histone Crosstalk. *Cell* **142**: 682–685.

Lefebvre V, Palloix A. 1996. Both epistatic and additive effects of QTLs are involved in polygenic induced resistance to disease: a case study, the interaction pepper—*Phytophthora capsici* Leonian. *Theoretical and Applied Genetics* **93**: 503–511.

Lehman JS, Shaner G. 1997. Selection of populations of *Puccinia recondita* f. sp. tritici for shortened latent period on a partially resistant wheat cultivar. *Phytopathology* **87**: 170–176.

Lemarié S, Robert-Seilantantz A, Lariagon C, Lemoine J, Marnet N, Jubault M, Manzanares-Dauleux MJ, Gravot A. 2015a. Both the jasmonic acid and the salicylic acid pathways contribute to resistance to the biotrophic clubroot agent *Plasmodiophora brassicae* in *Arabidopsis*. *Plant & Cell Physiology* **56**: 2158–2168.

Lemarié S, Robert-Seilantantz A, Lariagon C, Lemoine J, Marnet N, Levrel A, Jubault M, Manzanares-Dauleux M, Gravot A. 2015b. Camalexin contributes to the partial resistance of *Arabidopsis thaliana* to the biotrophic soilborne protist *Plasmodiophora brassicae*. *Frontiers in Plant Science* **6**: 539.

Li J, Huang Q, Sun M, Zhang T, Li H, Chen B, Xu K, Gao G, Li F, Yan G, et al. 2016a. Global DNA methylation variations after short-term heat shock treatment in cultured microspores of *Brassica napus* cv. Topas. *Scientific Reports* **6**: 38401.

Li X, Kapos P, Zhang Y. 2015. NLRs in plants. *Current Opinion in Immunology* **32**: 114–121.

Li L, Luo Y, Chen B, Xu K, Zhang F, Li H, Huang Q, Xiao X, Zhang T, Hu J, et al. 2016b. A genome-wide association study reveals new loci for resistance to clubroot disease in *Brassica napus*. *Frontiers in Plant Science* **7**: 1483.

- Li Y, Tessaro MJ, Li X, Zhang Y. 2010.** Regulation of the expression of plant resistance gene SNC1 by a protein with a conserved BAT2 domain. *Plant Physiology* **153**: 1425.
- Liegard B, Baillet V, Joseph E, Lariagon C, Lemoine J, Evrard A, Colot V, Gravot A, Manzanares-Dauleux M, Jubault M. submitted.** *Arabidopsis* methylome pattern variations linked to clubroot infection. *The New Phytologist*.
- Lindhout P. 2002.** The perspectives of polygenic resistance in breeding for durable disease resistance. *Euphytica* **124**: 217–226.
- Lippert K, Galinski EA, Trüper HG. 1993.** Biosynthesis and function of trehalose in *Ectothiorhodospira halochloris*. *Antonie Van Leeuwenhoek* **63**: 85–91.
- 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.
- Liu T, Li Y, Duan W, Huang F, Hou X. 2017.** Cold acclimation alters DNA methylation patterns and confers tolerance to heat and increases growth rate in *Brassica rapa*. *Journal of Experimental Botany* **68**: 1213–1224.
- Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IAP, Zhao M, Ma J, Yu J, Huang S, et al. 2014.** The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications* **5**: 3930.
- Lombard V, Delourme R. 2001.** A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. *Theoretical and applied genetics* **103**: 491–507.
- Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, Zhang Q, Vilhjálmsson BJ, Korte A, Nizhynska V, et al. 2013.** Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. *Nature Genetics* **45**: 884–890.
- López A, Ramírez V, García-Andrade J, Flors V, Vera P. 2011.** The RNA silencing enzyme RNA polymerase V is required for plant immunity. *PLoS Genetics* **7**: e1002434.
- López Sánchez Ana, Stassen Joost H.M., Furci Leonardo, Smith Lisa M., Ton Jurriaan. 2016.** The role of DNA (de)methylation in immune responsiveness of *Arabidopsis*. *The Plant Journal* **88**: 361–374.
- Lovelock DA, Donald CE, Conlan XA, Cahill DM. 2013.** Salicylic acid suppression of clubroot in broccoli (var. Italica) caused by the obligate biotroph *Plasmodiophora brassicae*. *Australasian Plant Pathology* **42**: 141–153.
- Ludwig-Muller J, Bendel U, Thermann P, Ruppel M, Epstein E, Hilgenberg W. 1993.** Concentrations of indole-3-acetic acid in plants of tolerant and susceptible varieties of Chinese cabbage infected with *Plasmodiophora brassicae* Woron. *New Phytologist* **125**: 763–769.
- Ludwig-Müller J, Jülke S, Geiß K, Richter F, Mithöfer A, Šola I, Rusak G, Keenan S, Bulman S. 2014.** A novel methyltransferase from the intracellular pathogen *Plasmodiophora brassicae* methylates salicylic acid. *Molecular Plant Pathology* **16**: 349–364.
- Ludwig-Müller J, Prinsen E, A. Rolfe S, Scholes J. 2009.** Metabolism and plant hormone action during clubroot disease. *J. Plant Growth Regul.* **28**: 229–244.

- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997.** Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260.
- Luna E, Bruce TJA, Roberts MR, Flors V, Ton J. 2012.** Next-generation systemic acquired resistance. *Plant Physiology* **158**: 844.
- Luna E, Ton J. 2012.** The epigenetic machinery controlling transgenerational systemic acquired resistance. *Plant signaling & behavior* **7**: 615–618.
- Macho AP, Zipfel C. 2014.** Plant PRRs and the activation of innate immune signaling. *Molecular Cell* **54**: 263–272.
- Maekawa T, Kufer TA, Schulze-Lefert P. 2011.** NLR functions in plant and animal immune systems: so far and yet so close. *Nature Immunology* **12**: 817–826.
- Malinowski R, Novák O, Borhan MH, Spíchal L, Strnad M, Rolfe SA. 2016.** The role of cytokinins in clubroot disease. *European Journal of Plant Pathology* **145**: 543–557.
- Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB. 2006.** A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics* **38**: 948.
- Manzanares MJ, Baron F, Thomas G. 1996.** Host-pathogen interactions between *Brassica oleracea* genotypes and single spore derived isolates of *Plasmodiophora brassicae*. *Acta Horticulturae: International Society for Horticultural Science (ISHS)*, Leuven, Belgium, 411–418.
- Manzanares-Dauleux MJ, Delourme R, Baron F, Thomas G. 2000a.** Mapping of one major gene and of QTLs involved in resistance to clubroot in *Brassica napus*. *Theoretical and Applied Genetics* **101**: 885–891.
- Manzanares-Dauleux MJ, Delourme R, Glory P, Giboulot A, Thomas G. 2003.** Mapping QTLs and major resistance genes to clubroot (*Plasmodiophora brassicae*) in *Brassica napus*. In: 13th Crucifer Genetics Workshop, UC Davis, California, March. 23–26.
- Manzanares-Dauleux MJ, Divaret I, Baron F, Thomas G. 2000b.** Evaluation of french *Brassica oleracea* landraces for resistance to *Plasmodiophora brassicae*. *Euphytica* **113**: 211–218.
- Mari-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. 2013.** Reconstructing de novo silencing of an active plant retrotransposon. *Nature Genetics* **45**: 1029.
- Marschner H. 1995.** *Mineral Nutrition of Higher Plants*. Gulf Professional Publishing.
- Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A. 2009.** A transposon-induced epigenetic change leads to sex determination in melon. *Nature* **461**: 1135.
- Matsumoto E, Yasui C, Ohi M, Tsukada M. 1998.** Linkage analysis of RFLP markers for clubroot resistance and pigmentation in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Euphytica* **104**: 79.
- Matzke MA, Moshier RA. 2014.** RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics*.
- Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, et al. 2010.** Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**: 253–257.

- McCLINTOCK B. 1951.** Chromosome organization and genic expression. *Cold Spring Harbor Symposia on Quantitative Biology* **16**: 13–47.
- McClintock B. 1961.** Some parallels between gene control systems in maize and in bacteria. *The American Naturalist* **95**: 265–277.
- McCue AD, Panda K, Nuthikattu S, Choudury SG, Thomas EN, Slotkin RK. 2015.** ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *The EMBO Journal* **34**: 20–35.
- Meyer P. 2015.** Epigenetic variation and environmental change. *Journal of Experimental Botany* **66**: 3541–3548.
- Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. 2003.** Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *The Plant Cell* **15**: 809.
- Mitani S, Araki S, Takii Y, Ohshima T, Matsuo N, Miyoshi H. 2001.** The biochemical mode of action of the novel selective fungicide Cyazofamid: specific inhibition of mitochondrial complex III in *Phythium spinosum*. *Pesticide Biochemistry and Physiology* **71**: 107–115.
- Mitani S, Sugimoto K, Hayashi H, Takii Y, Ohshima T, Matsuo N. 2003.** Effects of cyazofamid against *Plasmodiophora brassicae* Woronin on Chinese cabbage. *Pest Management Science* **59**: 287–293.
- Mithen R, Magrath R. 1992.** A contribution to the life history of *Plasmodiophora brassicae*: secondary plasmodia development in root galls of *Arabidopsis thaliana*. *Mycological Research* **96**: 877–885.
- Miura K, Agetsuma M, Kitano H, Yoshimura A, Matsuoka M, Jacobsen SE, Ashikari M. 2009.** A metastable DWARF1 epigenetic mutant affecting plant stature in rice. *Proceedings of the National Academy of Sciences* **106**: 11218.
- 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.
- Molinier J, Ries G, Zipfel C, Hohn B. 2006.** Transgeneration memory of stress in plants. *Nature* **442**: 1046.
- Montarry J, Cartier E, Jacquemond M, Palloix A, Moury B. 2012.** Virus adaptation to quantitative plant resistance: erosion or breakdown ? *Journal of evolutionary biology* **25**: 2242–2252.
- Monteith J. 1924.** *Relation of soil temperature and soil moisture to infection by Plasmodiophora brassicae.*
- Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marín MI, Martínez-Macías MI, Ariza RR, Roldán-Arjona T. 2006.** DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proceedings of the National Academy of Sciences* **103**: 6853–6858.
- Moriguchi K, Kimizuka-Takagi C, Ishii K, Nomura K. 1999.** A genetic map based on RAPD, RFLP, isozyme, morphological markers and QTL analysis for clubroot resistance in *Brassica oleracea*. *Breeding science* **49**: 257–265.
- Mousdale DMA. 1981.** Endogenous indolyl-3-acetic acid and pathogen-induced plant growth disorders: distinction between hyperplasia and neoplastic development. *Experientia* **37**: 972–973.

- Müller J, Aeschbacher RA, Wingler A, Boller T, Wiemken A. 2001.** Trehalose and Trehalase in *Arabidopsis*. *Plant Physiology* **125**: 1086–1093.
- Müller P, Hilgenberg W. 1986.** Isomers of zeatin and zeatin riboside in clubroot tissue: evidence for trans-zeatin biosynthesis by *Plasmodiophora brassicae*. *Physiologia Plantarum* **66**: 245–250.
- Mullin WJ, Collins MJ, Proudfoot KG. 1980.** Glucosinolate content and clubroot of rutabaga and turnip. *Canadian Journal of Plant Science* **60**: 605–612.
- Mundt CC. 2014.** Durable resistance: a key to sustainable management of pathogens and pests. *Infection, Genetics and Evolution* **27**: 446–455.
- Murakami H., Tsushima S., Akimoto T., Murakami K., Goto I., Shishido Y. 2001.** Effects of growing leafy daikon (*Raphanus sativus*) on populations of *Plasmodiophora brassicae* (clubroot). *Plant Pathology* **49**: 584–589.
- Murakami H, Tsushima S, Kuroyanagi Y, Shishido Y. 2002.** Reduction of resting spore density of *Plasmodiophora brassicae* and clubroot disease severity by liming. *Soil Science and Plant Nutrition* **48**: 685–691.
- Myers DF, Campbell RN. 1985.** Lime and the control of clubroot of crucifers: effects of pH, calcium, magnesium, and their interactions. *Phytopathology (USA)*.
- Mysore KS, Ryu C-M. 2004.** Nonhost resistance: how much do we know? *Trends in Plant Science* **9**: 97–104.
- Nagaoka T, Doullah MAU, Matsumoto S, Kawasaki S, Ishikawa T, Hori H, Okazaki K. 2010.** Identification of QTLs that control clubroot resistance in *Brassica oleracea* and comparative analysis of clubroot resistance genes between *B. rapa* and *B. oleracea*. *Theoretical and Applied Genetics* **120**: 1335–1346.
- Naiki T. 1987.** Life cycle and control of *Plasmodiophora brassicae*, causing clubroot disease of cruciferous plants. *Soil Microorganisms* **29**: 23–37.
- Narisawa K, Shimura M, Usuki F, Fukuhara S, Hashiba T. 2005.** Effects of pathogen density, soil moisture, and soil pH on biological control of clubroot in Chinese cabbage by *Heteroconium chaetospora*. *Plant Disease* **89**: 285–290.
- Narisawa K, Tokumasu S, Hashiba T. 1998.** Suppression of clubroot formation in Chinese cabbage by the root endophytic fungus, *Heteroconium chaetospora*. *Plant Pathology* **47**: 206–210.
- Narusaka M, Shirasu K, Noutoshi Y, Kubo Y, Shiraishi T, Iwabuchi M, Narusaka Y. 2009.** RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *The Plant Journal* **60**: 218–226.
- Naumann K, Fischer A, Hofmann I, Krauss V, Phalke S, Irmeler K, Hause G, Aurich A, Dorn R, Jenuwein T. 2005.** Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *The EMBO journal* **24**: 1418–1429.
- Naumova NA. 1933.** Contribution to the knowledge of the influence of soil factors on the development of club root in the *Cruciferae*. *Bull Plant Prot Leningrad Ser II Phytopathol* **3**: 32–52.
- Nelson R, Wiesner-Hanks T, Wissner R, Balint-Kurti P. 2017.** Navigating complexity to breed disease-resistant crops. *Nature Reviews Genetics* **19**: 21.

- Neuhauser S, Kirchmair M, Gleason FH. 2011.** Ecological roles of the parasitic phytomyxids (plasmodiophorids) in marine ecosystems – a review. *Marine & freshwater research* **62**: 365–371.
- Ng DW-K, Wang T, Chandrasekharan MB, Aramayo R, Kertbundit S, Hall TC. 2007.** Plant SET domain-containing proteins: Structure, function and regulation. *Plant Chromatin: Structure and Expression* **1769**: 316–329.
- Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, et al. 2016.** Widespread natural variation of DNA methylation within angiosperms. *Genome Biology* **17**: 194.
- Niks RE, Qi X, Marcel TC. 2015.** Quantitative resistance to biotrophic filamentous plant pathogens: concepts, misconceptions, and mechanisms. *Annual Review of Phytopathology* **53**: 445–470.
- Niks RE, Rubiales D. 2002.** Potentially durable resistance mechanisms in plants to specialised fungal pathogens. *Euphytica* **124**: 201–216.
- Nishimura MT, Anderson RG, Cherkis KA, Law TF, Liu QL, Machius M, Nimchuk ZL, Yang L, Chung E-H, El Kasmi F. 2017.** TIR-only protein RBA1 recognizes a pathogen effector to regulate cell death in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **114**: E2053–E2062.
- Nobuta K, Venu RC, Lu C, Beló A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang G, et al. 2007.** An expression atlas of rice mRNAs and small RNAs. *Nature Biotechnology* **25**: 473–477.
- Nomura K., Minegishi Y., Kimizuka-Takagi C., Fujioka T., Moriguchi K., Shishido R., Ikehashi H. 2008.** Evaluation of F2 and F3 plants introgressed with QTLs for clubroot resistance in cabbage developed by using SCAR markers. *Plant Breeding* **124**: 371–375.
- Ong-Abdullah M, Ordway JM, Jiang N, Ooi S-E, Kok S-Y, Sarpan N, Azimi N, Hashim AT, Ishak Z, Rosli SK, et al. 2015.** Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature* **525**: 533.
- Orgeur G, Jestin C, Manzanares-Dauleux M, Grimault V. 2015.** Hernie sur colza: la résistance variétale peut enfin être évaluée. *Phytoma* **684**: 14–17.
- Ortega-Galisteo AP, Morales-Ruiz T, Ariza RR, Roldán-Arjona T. 2008.** *Arabidopsis* DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Molecular Biology* **67**: 671–681.
- Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. 2010.** The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* **327**: 92.
- Palloix A, Ayme V, Moury B. 2009.** Durability of plant major resistance genes to pathogens depends on the genetic background, experimental evidence and consequences for breeding strategies. *New Phytologist* **183**: 190–199.
- Palm ET. 1963.** Effect of mineral nutrition on the invasion and response of turnip tissue to *Plasmodiophora brassicae* Wor. *Contrib Boyce Thompson Inst* **22**: 91–112.
- Palma K, Thorgrimsen S, Malinovsky FG, Fiil BK, Nielsen HB, Brodersen P, Hofius D, Petersen M, Mundy J. 2010.** Autoimmunity in *Arabidopsis* *acd11* is mediated by epigenetic regulation of an immune receptor. *PLoS pathogens* **6**: e1001137.

- Pandey R, Müller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, Bender J, Mount DW, Jorgensen RA. 2002.** Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Research* **30**: 5036–5055.
- Parkin IA, Koh C, Tang H, Robinson SJ, Kagale S, Clarke WE, Town CD, Nixon J, Krishnakumar V, Bidwell SL, et al. 2014.** Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biology* **15**: R77.
- Pavet V, Quintero C, Cecchini NM, Rosa AL, Alvarez ME. 2006.** *Arabidopsis* displays centromeric DNA hypomethylation and cytological alterations of heterochromatin upon attack by *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions* **19**: 577–587.
- Peng G, McGregor L, Lahlali R, Gossen BD, Hwang SF, Adhikari KK, Strelkov SE, McDonald MR. 2011.** Potential biological control of clubroot on canola and crucifer vegetable crops. *Plant Pathology* **60**: 566–574.
- 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.
- Pfluger J, Wagner D. 2007.** Histone modifications and dynamic regulation of genome accessibility in plants. *Current Opinion in Plant Biology* **10**: 645–652.
- Piao ZY, Deng YQ, Choi SR, Park YJ, Lim YP. 2004.** SCAR and CAPS mapping of CRb, a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Theoretical and Applied Genetics* **108**: 1458–1465.
- Piao ZY, Park YJ, Choi SR, Hong CP, Park JY, Choi YS, Lim YP. 2002.** Conversion of an AFLP marker linked to clubroot resistance gene in Chinese cabbage into a SCAR marker. *Journal Korean Society for Horticultural Science* **43**: 653–659.
- Pien S, Fleury D, Mylne JS, Crevillen P, Inzé D, Avramova Z, Dean C, Grossniklaus U. 2008.** *ARABIDOPSIS* TRITHORAX1 dynamically regulates FLOWERING LOCUS C activation via histone 3 lysine 4 trimethylation. *The Plant Cell* **20**: 580.
- Pilet-Nayel M-L, Moury B, Caffier V, Montarry J, Kerlan M-C, Fournet S, Durel C-E, Delourme R. 2017.** Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Frontiers in plant science* **8**: 1838.
- Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ. 2009.** Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science* **14**: 21–29.
- Pontvianne F, Blevins T, Pikaard CS. 2010.** *Arabidopsis* histone lysine methyltransferases. *Advances in Botanical Research*. 1–22.
- Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, Browse J, Chapple C, Colot V, Cutler S, et al. 2016.** 50 years of *Arabidopsis* research: highlights and future directions. *The New Phytologist* **209**: 921–944.
- Qu S, Liu G, Zhou B, Bellizzi M, Zeng L, Dai L, Han B, Wang G-L. 2006.** The broad-spectrum blast resistance gene Pi9 encodes a nucleotide-binding site–leucine-rich repeat protein and is a member of a multigene family in rice. *Genetics* **172**: 1901.

- Quadrana L, Almeida J, Asís R, Duffy T, Dominguez PG, Bermúdez L, Conti G, Corrêa da Silva JV, Peralta IE, Colot V, et al. 2014.** Natural occurring epialleles determine vitamin E accumulation in tomato fruits. *Nature Communications* **5**: 4027.
- Quadrana L, Bortolini Silveira A, Mayhew GF, LeBlanc C, Martienssen RA, Jeddloh JA, Colot V. 2016.** The *Arabidopsis thaliana* mobilome and its impact at the species level (D Zilberman, Ed.). *eLife* **5**: e15716.
- Quadrana L, Colot V. 2016.** Plant transgenerational epigenetics. *Annual review of genetics* **50**: 467–491.
- Raa J. 1971.** Indole-3-acetic acid levels and the role of indole-3-acetic acid oxidase in the normal root and clubroot of cabbage. *Physiologia Plantarum* **25**: 130–134.
- Raiber E-A, Hardisty R, van Delft P, Balasubramanian S. 2017.** Mapping and elucidating the function of modified bases in DNA. *Nature Reviews Chemistry* **1**: s41570-017.
- Rakyan VK, Down TA, Balding DJ, Beck S. 2011.** Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics* **12**: 529.
- Rambani A, Rice JH, Liu J, Lane T, Ranjan P, Mazarei M, Pantalone V, Stewart CN, Staton M, Hewezi T. 2015.** The methylome of soybean roots during the compatible interaction with the soybean cyst nematode. *Plant Physiology* **168**: 1364.
- Rasmann S, De Vos M, Casteel CL, Tian D, Halitschke R, Sun JY, Agrawal AA, Felton GW, Jander G. 2011.** Herbivory in the previous generation primes plants for enhanced insect resistance. *Plant physiology*: 111.187831.
- Ravi M, Shibata F, Ramahi JS, Nagaki K, Chen C, Murata M, Chan SWL. 2011.** Meiosis-specific loading of the centromere-specific histone CENH3 in *Arabidopsis thaliana*. *PLOS Genetics* **7**: e1002121.
- Regulski M, Lu Z, Kendall J, Donoghue MTA, Reinders J, Llaca V, Deschamps S, Smith A, Levy D, McCombie WR, et al. 2013.** The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Research* **23**: 1651–1662.
- Reinders J, Wulff BBH, Mirouze M, Marí-Ordóñez 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.
- Řičařová V, Kaczmarek J, Strelkov SE, Kazda J, Lueders W, Rysanek P, Manolii V, Jedryczka M. 2016.** Pathotypes of *Plasmodiophora brassicae* causing damage to oilseed rape in the Czech Republic and Poland. *European Journal of Plant Pathology* **145**: 559–572.
- Richards EJ. 2006.** Inherited epigenetic variation--revisiting soft inheritance. *Nature Reviews. Genetics* **7**: 395–401.
- Richards EJ. 2008.** Population epigenetics. *Current opinion in genetics & development* **18**: 221–226.
- Riddle NC, Richards EJ. 2002.** The control of natural variation in cytosine methylation in *Arabidopsis*. *Genetics* **162**: 355–363.
- Riddle NC, Richards EJ. 2005.** Genetic variation in epigenetic inheritance of ribosomal RNA gene methylation in *Arabidopsis*. *The Plant Journal* **41**: 524–532.

- Roberts M, Tang S, Stallmann A, Dangl JL, Bonardi V. 2013.** Genetic requirements for signaling from an autoactive plant NB-LRR intracellular innate immune receptor. *PLoS genetics* **9**: e1003465.
- Robinson MD, McCarthy DJ, Smyth GK. 2010.** edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
- Rocherieux J, Glory P, Giboulot A, Boury S, Barbeyron G, Thomas G, Manzanares-Dauleux MJ. 2004.** Isolate-specific and broad-spectrum QTLs are involved in the control of clubroot in *Brassica oleracea*. *Theoretical and Applied Genetics* **108**: 1555–1563.
- Rolfe SA, Strelkov SE, Links MG, Clarke WE, Robinson SJ, Djavaheri M, Malinowski R, Haddadi P, Kagale S, Parkin IAP, et al. 2016.** The compact genome of the plant pathogen *Plasmodiophora brassicae* is adapted to intracellular interactions with host *Brassica* spp. *BMC genomics* **17**: 272.
- Roudier F, Ahmed I, Berard 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*. *Embo Journal* **30**: 1928–1938.
- Roux F, Voisin D, Badet T, Balagué C, Barlet X, Huard-Chauveau C, Roby D, Raffaele S. 2014.** Resistance to phytopathogens e tutti quanti: placing plant quantitative disease resistance on the map. *Molecular Plant Pathology* **15**: 427–432.
- Saito M, Kubo N, Matsumoto S, Suwabe K, Tsukada M, Hirai M. 2006.** Fine mapping of the clubroot resistance gene, *Crr3*, in *Brassica rapa*. *Theoretical and Applied Genetics* **114**: 81.
- Sakamoto K, Saito A, Hayashida N, Taguchi G, Matsumoto E. 2008.** Mapping of isolate-specific QTLs for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Theoretical and Applied Genetics* **117**: 759–767.
- Saleh A, Alvarez-Venegas R, Yilmaz M, Le O, Hou G, Sadler M, Al-Abdallat A, Xia Y, Lu G, Ladunga I, et al. 2008.** The highly similar *Arabidopsis* homologs of *Trithorax* ATX1 and ATX2 encode proteins with divergent biochemical functions. *The Plant Cell* **20**: 568.
- Saucet SB, Ma Y, Sarris PF, Furzer OJ, Sohn KH, Jones JD. 2015.** Two linked pairs of *Arabidopsis* TNL resistance genes independently confer recognition of bacterial effector AvrRps4. *Nature communications* **6**: 6338.
- 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* **334**: 369.
- Schmitz RJ, Schultz MD, Urich MA, Nery JR, Pelizzola M, Libiger O, Alix A, McCosh RB, Chen H, Schork NJ, et al. 2013.** Patterns of population epigenomic diversity. *Nature* **495**: 193–198.
- 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 (New York, N.Y.)* **326**: 1112–1115.
- Schulze-Lefert P, Panstruga R. 2011.** A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in Plant Science* **16**: 117–125.
- Schwelm A, Fogelqvist J, Knaust A, Jülke S, Lilja T, Bonilla-Rosso G, Karlsson M, Shevchenko A, Dhandapani V, Choi SR, et al. 2015.** The *Plasmodiophora brassicae* genome reveals insights in its life cycle and ancestry of chitin synthases. *Scientific Reports* **5**: 11153.

- Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. 2012.** The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Molecular Cell* **48**: 849–862.
- Sequeira-Mendes J, Aragüez I, Peiró R, Mendez-Giraldez R, Zhang X, Jacobsen SE, Bastolla U, Gutierrez C. 2014.** The functional topography of the *Arabidopsis* genome is organized in a reduced number of linear motifs of chromatin states. *The Plant Cell Online* **26**: 2351–2366.
- Shao Z-Q, Xue J-Y, Wu P, Zhang Y-M, Wu Y, Hang Y-Y, Wang B, Chen J-Q. 2016.** Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiology*: 01487.2015.
- Sharma K, Gossen BD, McDonald MR. 2011.** Effect of temperature on primary infection by *Plasmodiophora brassicae* and initiation of clubroot symptoms. *Plant Pathology* **60**: 830–838.
- Shaw RG, Byers DL, Darmo E. 2000.** Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana*. *Genetics* **155**: 369–378.
- Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES. 1999.** The FLF MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. *The Plant Cell* **11**: 445.
- Shivaprasad PV, Chen H-M, Patel K, Bond DM, Santos BA, Baulcombe DC. 2012.** A microRNA superfamily regulates nucleotide binding site–leucine-rich repeats and other mRNAs. *The Plant Cell* **24**: 859–874.
- Shu H, Nakamura M, Siretskiy A, Borghi L, Moraes I, Wildhaber T, Grissem W, Hennig L. 2014.** *Arabidopsis* replacement histone variant H3. 3 occupies promoters of regulated genes. *Genome Biology* **15**: R62.
- Siemens J, Graf H, Bulman S, In O, Ludwig-Müller J. 2009.** Monitoring expression of selected *Plasmodiophora brassicae* genes during clubroot development in *Arabidopsis thaliana*. *Plant Pathology* **58**: 130–136.
- Siemens J, González M-C, Wolf S, Hofmann C, Greiner S, Du Y, Rausch T, Roitsch T, Ludwig-Müller J. 2010.** Extracellular invertase is involved in the regulation of clubroot disease in *Arabidopsis thaliana*. *Molecular Plant Pathology* **12**: 247–262.
- Siemens J, Keller I, Sarx J, Kunz S, Schuller A, Nagel W, Schmülling T, Parniske M, Ludwig-Müller J. 2006.** Transcriptome analysis of *Arabidopsis* clubroots indicate a key role for cytokinins in disease development. *Molecular Plant-Microbe Interactions* **19**: 480–494.
- Siemens J, Nagel M, Ludwig-Müller J, Sacristán MD. 2003.** The interaction of *Plasmodiophora brassicae* and *Arabidopsis thaliana*: parameters for disease quantification and screening of mutant lines. *Journal of Phytopathology* **150**: 592–605.
- Silveira AB, Trontin C, Cortijo S, Barau J, Del Bem LEV, Loudet O, Colot V, Vincentz M. 2013.** Extensive natural epigenetic variation at a *de novo* originated gene. *PLOS Genetics* **9**: e1003437.
- Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal F-X, Durand-Tardif M, Pelletier G, Camilleri C. 2008.** Quantitative trait loci mapping in five new large recombinant inbred line populations of *Arabidopsis thaliana* genotyped with consensus single-nucleotide polymorphism markers. *Genetics* **178**: 2253–2264.

- 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.
- Slaughter A, Daniel X, Flors V, Luna E, Hohn B, Mauch-Mani B. 2011.** Descendants of primed *Arabidopsis* plants exhibit resistance to biotic stress. *Plant physiology*: 111.191593.
- Some A, Manzanares MJ, Laurens F, Baron F, Thomas G, Rouxel F. 1996.** Variation for virulence on *Brassica napus* L. amongst *Plasmodiophora brassicae* collections from France and derived single-spore isolates. *Plant Pathology* **45**: 432–439.
- Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters AJ. 2000.** The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Molecular Cell* **6**: 791–802.
- Springer NM, Napoli CA, Selinger DA, Pandey R, Cone KC, Chandler VL, Kaeppler HF, Kaeppler SM. 2003.** Comparative analysis of SET domain proteins in maize and *Arabidopsis* reveals multiple duplications preceding the divergence of monocots and dicots. *Plant Physiology* **132**: 907.
- St Clair DA. 2010.** Quantitative disease resistance and quantitative resistance Loci in breeding. *Annual Review of Phytopathology* **48**: 247–268.
- Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J. 1999.** Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature* **400**: 667.
- Steward N, Ito M, Yamaguchi Y, Koizumi N, Sano H. 2002.** Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *Journal of Biological Chemistry* **277**: 37741–37746.
- Stokes TL, Kunkel BN, Richards EJ. 2002.** Epigenetic variation in *Arabidopsis* disease resistance. *Genes & Development* **16**: 171–182.
- Strelkov SE, Hwang S-F. 2014.** Clubroot in the Canadian canola crop: 10 years into the outbreak. *Canadian Journal of Plant Pathology* **36**: 27–36.
- Strelkov SE, Hwang S-F, Manolii VP, Cao T, Feindel D. 2016.** Emergence of new virulence phenotypes of *Plasmodiophora brassicae* on canola (*Brassica napus*) in Alberta, Canada. *European Journal of Plant Pathology* **145**: 517–529.
- Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. 2014.** Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nature Structural & Molecular Biology* **21**: 64–72.
- Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE. 2013.** Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **152**: 352–364.
- Stroud H, Otero S, Desvoyes B, Ramírez-Parra E, Jacobsen SE, Gutierrez C. 2012.** Genome-wide analysis of histone H3. 1 and H3. 3 variants in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **109**: 5370–5375.
- Stuthman DD, Leonard KJ, Miller-Garvin J. 2007.** Breeding crops for durable resistance to disease. *Advances in agronomy* **95**: 319–367.
- Suzanne Dorn E, Cook J. 2011.** *Nucleosomes in the neighborhood*.

Suzuki K, Matsumiya E, Ueno Y, Mizutani J. 1992. Some properties of germination-stimulating factor from plants for resting spores of *Plasmodiophora brassicae*. *Japanese Journal of Phytopathology* **58**: 699–705.

Talukder ZI, McDonald AJS, Price AH. 2005. Loci controlling partial resistance to rice blast do not show marked QTL × environment interaction when plant nitrogen status alters disease severity. *New Phytologist* **168**: 455–464.

Tanaka S, Kochi S, Kunita H, Ito S, Kameya-Iwaki M. 1999. Biological mode of action of the fungicide, flusulfamide, against *Plasmodiophora brassicae* (clubroot). *European journal of plant pathology* **105**: 577–584.

Tanweer FA, Rafii MY, Sijam K, Rahim HA, Ahmed F, Latif MA. 2015. Current advance methods for the identification of blast resistance genes in rice. *Comptes Rendus Biologies* **338**: 321–334.

Tariq M, Saze H, Probst AV, Lichota J, Habu Y, Paszkowski J. 2003. Erasure of CpG methylation in *Arabidopsis* alters patterns of histone H3 methylation in heterochromatin. *Proceedings of the National Academy of Sciences* **100**: 8823–8827.

Team RC. 2013. R: A language and environment for statistical computing.

Thevelein JM. 1984. Regulation of trehalose mobilization in fungi. *Microbiological Reviews* **48**: 42–59.

Thomas CM. 1998. Genetic and molecular analysis of tomato Cf genes for resistance to *Cladosporium fulvum*. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **353**: 1413–1424.

Thorstensen T, Fischer A, Sandvik SV, Johnsen SS, Grini PE, Reuter G, Aalen RB. 2006. The *Arabidopsis* SUVH4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. *Nucleic Acids Research* **34**: 5461–5470.

Tsuchiya T, Eulgem T. 2013. An alternative polyadenylation mechanism coopted to the *Arabidopsis* RPP7 gene through intronic retrotransposon domestication. *Proceedings of the National Academy of Sciences* **110**: E3535.

Ueno H, Matsumoto E, Aruga D, Kitagawa S, Matsumura H, Hayashida N. 2012. Molecular characterization of the CRa gene conferring clubroot resistance in *Brassica rapa*. *Plant Molecular Biology* **80**: 621–629.

Van der Biezen EA, Jones JD. 1998. Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences* **23**: 454–456.

Vanyushin B, N. Belozersky A, A. Kokurina N, X. Kadirova D. 1968. 5-methylcytosine and 6-methylaminopurine in bacterial DNA. *Nature* **218**: 1066–1067.

Vanyushin BF, Tkacheva SG, Belozersky AN. 1970. Rare bases in animal DNA. *Nature* **225**: 948–949.

Vaughn MW, Tanurdžić M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, Dedhia N, McCombie WR, Agier N, Bulski A. 2007. Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS biology* **5**: e174.

Vogel G, Fiehn O, Jean-Richard-dit-Bressel L, Boller T, Wiemken A, Aeschbacher RA, Wingler A. 2001. Trehalose metabolism in *Arabidopsis*: occurrence of trehalose and molecular cloning and

characterization of trehalose-6-phosphate synthase homologues. *Journal of Experimental Botany* **52**: 1817–1826.

Vongs A, Kakutani T, Martienssen R, Richards E. 1993. *Arabidopsis thaliana* DNA methylation mutants. *Science* **260**: 1926.

Voorrips RE, Jongerius MC, Kanne HJ. 1997. Mapping of two genes for resistance to clubroot (*Plasmodiophora brassicae*) in a population of doubled haploid lines of *Brassica oleracea* by means of RFLP and AFLP markers. *Theoretical and Applied Genetics* **94**: 75–82.

Waddington CH. 1942. Canalization of development and the inheritance of acquired characters. *Nature* **150**: 563–565.

Wagner G, Charton S, Lariagon C, Laperche A, Lugan R, Hopkins J, Frendo P, Bouchereau A, Delourme R, Gravot A, et al. 2012. Metabotyping: a new approach to investigate rapeseed (*Brassica napus* L.) genetic diversity in the metabolic response to clubroot infection. *Molecular Plant-Microbe Interactions* **25**: 1478–1491.

Walker JC, Hooker WJ. 1945. Plant nutrition in relation to disease development. Cabbage clubroot. *American Journal of Botany* **32**: 487–490.

Wallenhammar A-C. 1996. Prevalence of *Plasmodiophora brassicae* in a spring oilseed rape growing area in central Sweden and factors influencing soil infestation levels. *Plant Pathology* **45**: 710–719.

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun J-H, Bancroft I, Cheng F, et al. 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics* **43**: 1035.

Webster MA. 1986. pH and nutritional effects on infection by *Plasmodiophora brassicae* Wor. and on clubroot symptoms.

Webster MA, Dixon GR. 1991. Boron, pH and inoculum concentration influencing colonization by *Plasmodiophora brassicae*. *Mycological Research* **95**: 74–79.

Weigel D, Colot V. 2012. Epialleles in plant evolution. *Genome Biology* **13**.

Wellman FL. 1930. *Clubroot of crucifers. US Dept. of Agriculture.*

Werner S, Diederichsen E, Frauen M, Schondelmaier J, Jung C. 2007. Genetic mapping of clubroot resistance genes in oilseed rape. *Theoretical and Applied Genetics* **116**: 363–372.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, et al. 2007. A unified classification system for eukaryotic transposable elements. *Nature Reviews. Genetics* **8**: 973–982.

Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. 2009. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genetics* **41**: 630–634.

Wiesner-Hanks T, Nelson R. 2016. Multiple disease resistance in plants. *Annual review of phytopathology* **54**: 229–252.

Williams PH. 1966. A system for the determination of races of *Plasmodiophora brassicae* that infect cabbage and rutabaga. *Phytopathology* **56**: 624–626.

Williams PH, McNabola SS. 1967. Fine structure of *Plasmodiophora brassicae* in sporogenesis. *Canadian Journal of Botany* **45**: 1665–1669.

Williams SJ, Sohn KH, Wan L, Bernoux M, Sarris PF, Segonzac C, Ve T, Ma Y, Saucet SB, Ericsson DJ. 2014. Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* **344**: 299–303.

Wollmann H, Holec S, Alden K, Clarke ND, Jacques P-É, Berger F. 2012. Dynamic deposition of histone variant H3.3 accompanies developmental remodeling of the *Arabidopsis* transcriptome. *PLOS Genetics* **8**: e1002658.

Woo HR, Dittmer TA, Richards EJ. 2008. Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in *Arabidopsis*. *PLOS Genetics* **4**: e1000156.

Woronin MS. 1878. *Plasmodiophora brassicae* - The cause of cabbage Hernia. *The American Phytopathological Society*, 9–32.

Wu C-H, Krasileva KV, Banfield MJ, Terauchi R, Kamoun S. 2015. The “sensor domains” of plant NLR proteins: more than decoys? *Frontiers in Plant Science* **6**.

Xia S, Cheng YT, Huang S, Win J, Soards A, Jinn T-L, Jones JDG, Kamoun S, Chen S, Zhang Y, et al. 2013. Regulation of transcription of nucleotide-binding leucine-rich repeat-encoding genes SNC1 and RPP4 via H3K4 trimethylation. *Plant Physiology* **162**: 1694.

Xiang-Hua XC-GG. 2002. Biological characteristic of *Plasmodiophora brassicae* [J]. *Mycosystema* **4**: 026.

Xiao S, Ellwood S, Calis O, Patrick E, Li T, Coleman M, Turner JG. 2001. Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. *Science* **291**: 118–120.

Xu X, Hayashi N, Wang C-T, Fukuoka S, Kawasaki S, Takatsuji H, Jiang C-J. 2014. Rice blast resistance gene Pikahei-1 (t), a member of a resistance gene cluster on chromosome 4, encodes a nucleotide-binding site and leucine-rich repeat protein. *Molecular breeding* **34**: 691–700.

Yang Hao, Yang Ning, Wang Tai. 2016. Proteomic analysis reveals the differential histone programs between male germline cells and vegetative cells in *Lilium davidii*. *The Plant Journal* **85**: 660–674.

Yang X-J, Seto E. 2008. Lysine acetylation: codified crosstalk with other posttranslational modifications. *Molecular Cell* **31**: 449–461.

Yelagandula R, Stroud H, Holec S, Zhou K, Feng S, Zhong X, Muthurajan UM, Nie X, Kawashima T, Groth M, et al. 2014. The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. *Cell* **158**: 98–109.

Yi H, Richards EJ. 2009. Gene duplication and hypermutation of the pathogen Resistance gene SNC1 in the *Arabidopsis* bal variant. *Genetics* **183**: 1227–1234.

Yoshikawa H. 1993. Studies on breeding of clubroot resistance in cole [*Cruciferae*] crops. *Bulletin of the National Research Institute of Vegetables, Ornamental Plants and Tea. Series A.(Japan)*.

Yu A, Lepère G, Jay F, Wang J, Bapaume L, Wang Y, Abraham A-L, Penterman J, Fischer RL, Voinnet O, et al. 2013. Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *Proceedings of the National Academy of Sciences* **110**: 2389.

Yu F, Zhang X, Peng G, Falk KC, Strelkov SE, Gossen BD. 2017. Genotyping-by-sequencing reveals three QTL for clubroot resistance to six pathotypes of *Plasmodiophora brassicae* in *Brassica rapa*. *Scientific Reports* **7**: 4516.

- Yue J-X, Meyers BC, Chen J-Q, Tian D, Yang S. 2012.** Tracing the origin and evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes. *New Phytologist* **193**: 1049–1063.
- Zamani-Noor N.** Variation in pathotypes and virulence of *Plasmodiophora brassicae* populations in Germany. *Plant Pathology* **66**: 316–324.
- Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. 2013.** The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**: 193–205.
- Zemach A, McDaniel IE, Silva P, Zilberman D. 2010.** Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**: 916–919.
- Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. 2007a.** Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLOS Biology* **5**: e129.
- Zhang X-C, Gassmann W. 2007.** Alternative splicing and mRNA levels of the disease resistance gene RPS4 are induced during defense responses. *Plant physiology* **145**: 1577–1587.
- Zhang G, Huang H, Liu D, Cheng Y, Liu X, Zhang W, Yin R, Zhang D, Zhang P, Liu J, et al. 2015.** N6-Methyladenine DNA modification in *Drosophila*. *Cell* **161**: 893–906.
- Zhang H, Lang Z, Zhu J-K. 2018.** Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology* **19**: 489–506.
- 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. *PLoS Genetics* **4**: e1000032.
- Zhang K, Sridhar VV, Zhu J, Kapoor A, Zhu J-K. 2007b.** Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *Plos One* **2**: e1210.
- Zhang B, Tieman DM, Jiao C, Xu Y, Chen K, Fei Z, Giovannoni JJ, Klee HJ. 2016.** Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. *Proceedings of the National Academy of Sciences* **113**: 12580.
- Zhang Yuan-Ye, Fischer Markus, Colot Vincent, Bossdorf Oliver. 2012.** Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist* **197**: 314–322.
- Zhao J, Udall JA, Quijada PA, Grau CR, Meng J, Osborn TC. 2006.** Quantitative trait loci for resistance to *Sclerotinia sclerotiorum* and its association with a homeologous non-reciprocal transposition in *Brassica napus* L. *Theoretical and Applied Genetics* **112**: 509–516.
- Zheng X, Chen L, Xia H, Wei H, Lou Q, Li M, Li T, Luo L. 2017.** Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition. *Scientific Reports* **7**: 39843.
- Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ, Jacobsen SE. 2014.** Molecular mechanism of action of plant DRM *de novo* DNA methyltransferases. *Cell* **157**: 1050–1060.
- Zhong X, Hale CJ, Law JA, Johnson LM, Feng S, Tu A, Jacobsen SE. 2012.** DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nature Structural & Molecular Biology* **19**: 870–875.

Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. 2008. Histone H2A. Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456**: 125–129.

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.

Zou B, Yang D-L, Shi Z, Dong H, Hua J. 2014. Monoubiquitination of histone 2B at the disease resistance gene locus regulates its expression and impacts immune responses in *Arabidopsis*. *Plant physiology*: 113.227801.

ANNEXES

SOMMAIRE DES ANNEXES

Annexe I : New Phytologist Article Supporting Information..... 162

Annexe II : Nature Communication Supplementary Methods 233

Annexe III : Nature Communication Supporting Information 237

ANNEXE I : *New Phytologist* Article Supporting Information

Article title: Quantitative resistance to clubroot infection mediated by transgenerational epigenetic variation in *Arabidopsis*

Authors: Benjamin Liégard, Victoire Baillet, Mathilde Etcheverry, Evens Joseph, Christine Lariagon, Jocelyne Lemoine, Aurélie Evrard, Vincent Colot, Antoine Gravot, Maria J. Manzanares-Dauleux, Mélanie Jubault.

The following Supporting Information is available for this article:

Figure S1. Epi-allele effects at the closest QTL ^{epi} peak marker.	163
Figure S2. Boxplot of temperature data recorded in each growth room.....	164
Table S1. List of primer sets used to confirm homozygosity of the T-DNA insertion in mutants.	165
Table S2. (A) Disease index for Col-0 and mutants after infection by <i>Plasmodiophora brassicae</i> . (B) Dunnett post hoc test results.....	166
Table S3. Phenotypic responses of epiRIL and their parent lines to infection by <i>Plasmodiophora brassicae</i>	167
Table S4. Dunnett test comparison of temperatures recorded by each temperature sensor in growth room-2.	168
File S1. Location of segregating DNA sequence variants (SNP, indel, TE, SV) among the clubroot QTL intervals.	170
File S2. Co-localization of QTL ^{epi} confidence intervals with candidate SNP identified by GWAS after clubroot infection.....	215
File S3. Statistic analysis of TE and sequence variant effect compared to methylated markers on QTL ^{epi} detection for each trait study.	217

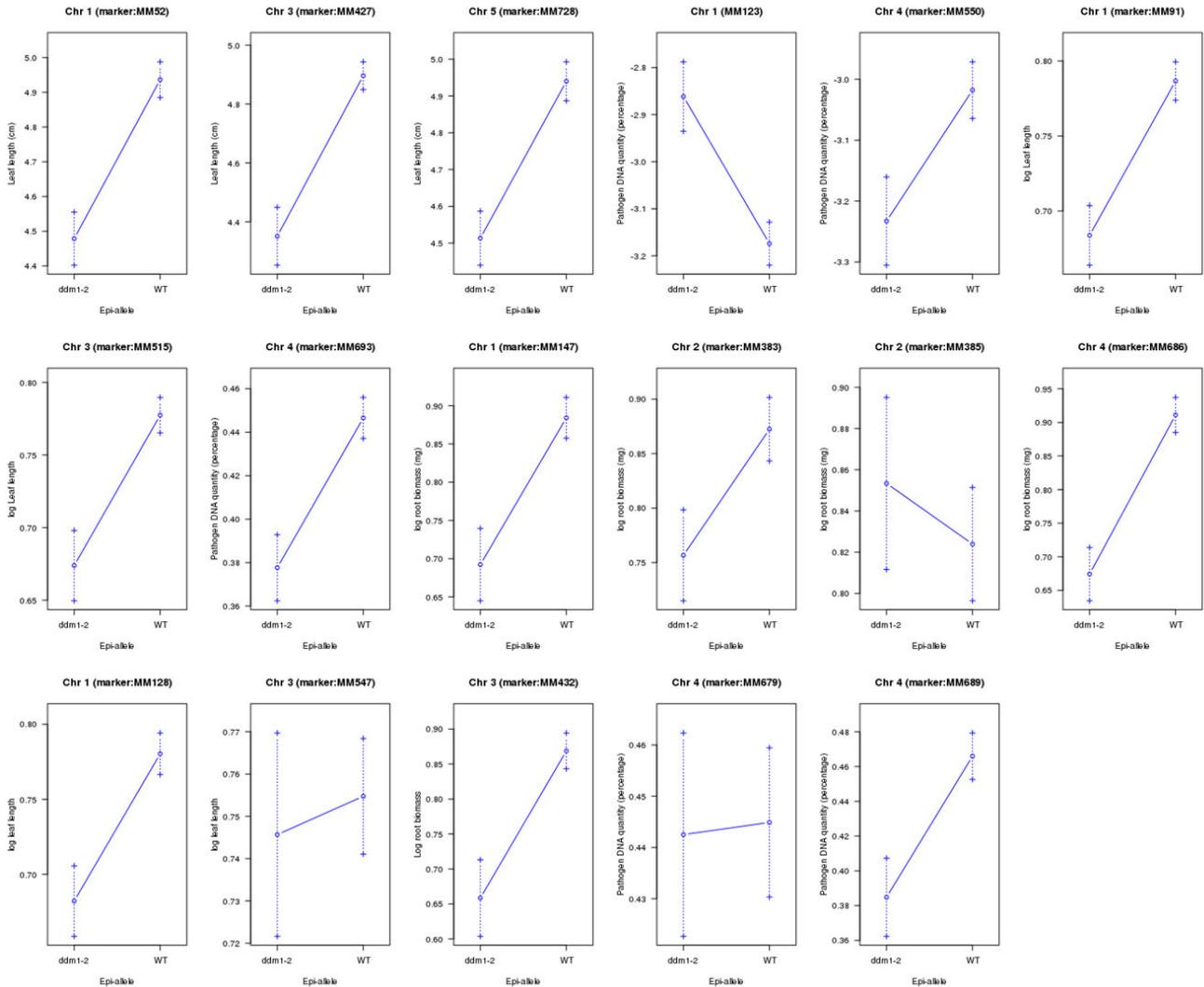


Figure S1. Epi-allele effects at the closest QTL^{epi} peak marker. All graphics show the effect of each epi-allele on the trait value at the closest QTL peak marker. For all QTL^{epi} detected for root biomass, leaf length and pathogen quantity in growth room-2, the *ddm1-2* epi-allele is associated with a lower value than the WT epi-allele. For all QTL detected for the pathogen quantity in growth room-1, the *ddm1-2* epi-allele is associated with a higher value than the WT epi-allele.

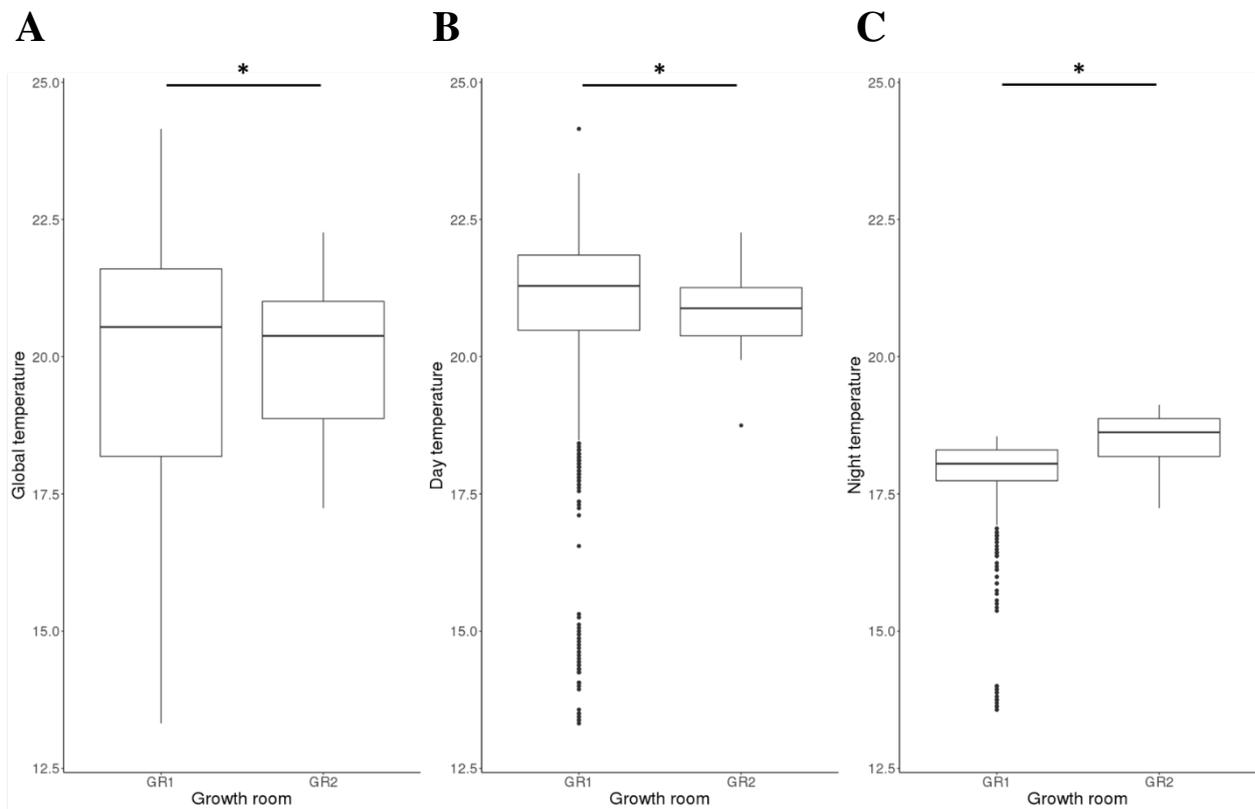


Figure S2. Boxplot of temperature data recorded in each growth room. (A) Global temperature in each growth room. Stars indicate the significant difference (Fisher permutation test, $F = 2.67$, $p\text{-value} = 0.002$) of the variance of the global temperature between the two growth rooms. (B) Day temperature in each growth room. Stars indicate the significant difference (Fisher permutation test, $F = 9.05$, $p\text{-value} = 0.002$) of the variance of the day temperature between the two growth rooms. (C) Night temperature in each growth room. Stars indicate the significant difference (Fisher permutation test, $F = 3.67$, $p\text{-value} = 0.002$) of the variance of the day temperature between the two growth rooms. Boxplot showed the median (inner box black horizontal line), the 25 and the 75 percentiles (box), the 1.5 times interquartile (vertical line) and the extreme values (black dots).

Table S1. List of primer sets used to confirm homozygosity of the T-DNA insertion in mutants.

NASC number	Mutant	Gene	LP primer (5' > 3')	RP primer (5' > 3')
N650863	SALK_150863	At5g14620	AGATCGCTTCCAGAGTTAGCC	TTGTCGCAAAAAGCAAAAGAG
N655230	SALK_004027C	At3g18520	CTTCTCTGTTCATCGTTTCGC	AGCAACATTCTCTCGTCGAAC
N667337	SALK_130607C	At5g09790	TTTCTCTTGTCCGGTGAAATG	CCTGCAACAATCAGTGTGATG
N681376	SALK_082118C	At1g79000	GACTGACAGGTCTTTGCTTGG	ATGAACCTGGAATGGAGAACC
N681865	SALK_149295C	At5g09230	CGCAGAGAGAGAACAAAATCG	TTCCACATTCTGTGCTAACCC
N681885	SALK_000590C	At5g66750	GAAGACAGGTTTCACCTGTGC	CGTGAGAAATAGCTCAATGCC

Table S2. (A) Disease index for Col-0 and mutants after infection by *Plasmodiophora brassicae*. **(B)** Dunnett post hoc test results.

A

Lines	Means	Sd
<i>ddm1</i> *	46.56	14.63
<i>srt2</i>	61.67	12.58
<i>drm2</i>	66.67	14.43
<i>atxr5</i>	70.28	15.28
<i>hac1</i>	70.56	18.96
Col-0	77.09	8.23
<i>hdc15</i>	84.72	19.7

Disease index calculation was carried out according to Manzanares-Dauleux *et al.* (2000b). Means were obtained by averaging the data from three blocks with six plants by block. Standard deviation (Sd) was obtained from the difference in DI between the three blocks for each mutant and Col-0. The star next to *ddm1* indicates significant differences for the disease index compared to Col-0.

B

DI genotype comparison	Estimate	Std.Error	t value	Pr(> t)
<i>atxr5</i> - Col-0	-6.8083	6.880	-0.990	0.86361
<i>ddm1</i> - Col-0	-30.5283	6.880	-4.437	0.00269
<i>drm2</i> - Col-0	-10.4183	6.880	-1.514	0.53771
<i>hac1</i> - Col-0	-6.5283	6.880	-0.949	0.88306
<i>hdc15</i> - Col-0	7.6383	6.880	1.110	0.79812
<i>srt2</i> - Col-0	-15.4183	6.880	-2.241	0.18262

Test was carried using multcomp R package.

Table S3. Phenotypic responses of epiRIL and their parent lines to infection by *Plasmodiophora brassicae*.

Lines	Growth room	Leaf length (cm)	Root biomass (mg per plant)	Pathogen DNA quantity (%)	Disease index
Col-0	1	5.52 ± 0.59	NA	0.036 ± 0.01	53.25 ± 2.36
	2	1.98 ± 0.34	1.95 ± 0.23	0.45 ± 0.05	90.75 ± 9.64
<i>ddm1-2</i>	1	NA	NA	NA	NA
	2	2.12 ± 0.54	2.16 ± 0.72	0.31 ± 0.13	71.75 ± 17.11
Epirils	1	4.80 ± 0.75	NA	0.052 ± 0.04	51.17 ± 11.42
	2	2.17 ± 0.47	2.41 ± 0.95	0.43 ± 0.17	86.27 ± 10.74

Disease index calculation was carried out according to Manzanares-Dauleux *et al.* (2000b). Pathogen DNA quantification was determined by quantitative PCR. For each trait and growth room, means were obtained by averaging data of two biological replicates, each one designed with two blocks and six plants per block. Standard deviation (Sd) was obtained from the difference in each trait between the four blocks.

Table S4. Dunnett test comparison of temperatures recorded by each temperature sensor in growth room-2.

	Ts1	Ts2	Ts3	Ts4	Ts5	Ts6	Ts7	Ts8	Ts9	Ts10	Ts11	Ts12	Ts13	Ts14	Ts15
Ts2	< 2.00E-16	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ts3	< 2.00E-16	6.50E-06	-	-	-	-	-	-	-	-	-	-	-	-	-
Ts4	< 2.00E-16	1.10E-04	< 2.00E-16	-	-	-	-	-	-	-	-	-	-	-	-
Ts5	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	-	-	-	-	-	-	-	-	-	-	-
Ts6	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	1.00	-	-	-	-	-	-	-	-	-	-
Ts7	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	5.90E-02	2.90E-06	-	-	-	-	-	-	-	-	-
Ts8	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	1.77E-02	1.00	2.90E-06	-	-	-	-	-	-	-	-
Ts9	2.6 E-16	< 2.50E-06	1.00	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	-	-	-	-	-	-	-
Ts10	< 2.00E-16	< 2.00E-16	< 2.00E-16	3.50E-05	6.40E-13	< 2.00E-16	2.30E-04	< 2.00E-16	< 2.00E-16	-	-	-	-	-	-
Ts11	< 2.00E-16	1.00	3.60E-08	2.54E-02	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	1.50E-08	1.70E-14	-	-	-	-	-
Ts12	< 2.00E-16	< 2.00E-16	< 2.00E-16	3.90E-12	2.50E-06	2.20E-12	1.00	< 2.00E-16	< 2.00E-16	1.00	< 2.00E-16	-	-	-	-
Ts13	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	2.5E-16	1.70E-09	< 2.00E-16	7.40E-06	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	-	-	-
Ts14	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	4.80E-12	3.10E-06	< 2.00E-16	2.68E-03	< 2.00E-16	< 2.00E-16	< 2.00E-16	< 2.00E-16	1.00	-	-
Ts15	< 2.00E-16	< 2.00E-16	< 2.00E-16	2.50E-14	5.80E-06	3.90E-12	1.00	< 2.00E-16	< 2.00E-16	4.29E-01	< 2.00E-16	1.00	< 2.00E-16	< 2.00E-16	-
Ts16	< 2.00E-16	< 2.00E-16	< 2.00E-16	8.60E-13	4.60E-07	1.20E-13	7.89E-01	< 2.00E-16	< 2.00E-16	1.00	< 2.00E-16	1.00	< 2.00E-16	< 2.00E-16	1.00

Adjusted *p-values* (Bonferroni method) calculated with the Dunnett test between each temperature sensor are presented in the bottom left of the table.

Ts corresponds to the temperature sensor number. Non-significant *p-values* of test were highlighted in bold.

Table S5. Median temperature monitored by each temperature sensor in growth room-2.

Temperature sensor	Median temperature	MAD
Ts1	21.40	0.74
Ts2	22.90	0.74
Ts3	22.50	0.74
Ts4	22.90	0.74
Ts5	23.20	0.74
Ts6	23.60	0.74
Ts7	23.10	0.74
Ts8	23.20	1.11
Ts9	22.10	0.74
Ts10	23.10	0.74
Ts11	22.60	0.74
Ts12	23.20	0.74
Ts13	24.10	0.74
Ts14	23.70	0.74
Ts15	23.10	0.74
Ts16	23.10	0.74

MAD: Median absolute deviation. The median temperatures are in °C. Ts corresponds to the temperature sensor number.

File S1. Location of segregating DNA sequence variants (SNP, indel, TE, SV) among the clubroot QTL intervals. 1 and 0 denote respectively the presence or absence of the variant in the epiRIL under scrutiny; het denotes heterozygosity, in black lack of evidence. * in the epiRIL name indicates that this line was not considered for TE detection.

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 8	epi RIL 11*	epi RIL 14	epi RIL 18	epi RIL 20*	epi RIL 24	epi RIL 36	epi RIL 46	epi RIL 52
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0		1	0		0	0	1	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	0	1	1	1	0	0	1	0
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	1	het	0	0	0	het	1	1	1
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	1	1	1	1	het	0	0	0
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0		0	0		0	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0		0	0		0	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	0	1	1		het	0	1	1
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	1	1	0	het	0	1	1
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	0	0	1	0	1	0	0
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	0	0	1	0	1	0	0
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	1	0	0	0	1	0
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	0	0	1	0	1		0
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	1	1	0	1	0	1	1
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	0	0	1	0	1	0	0

TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-AT1, Lfi2epi-GTAt1	102	1		0	0		0	0	1	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	1	1	0	1	0	1	1
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	0	0	1	0	1	0	0
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	0	0	1	0	1	0	0
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	102	1		0	0		0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	0	0	1	1	1	0	0
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	0	het	0	0	1	1	1	0	0
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0		1	0		0	0	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	1	het	0	0	1	0	1	1	0
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	1		0	0		0	0	0	0
SNP	Chr3	3666840	A	G	Lfi2epi-GTAt3	122	1	het	0	0	1	1	0	0	0
SNP	Chr3	4740011	C	T	Lfi2epi-GTAt3	122	1	het	0	0	1	1	0	0	0
TE	Chr3	5595169	.	ATENSPM3	Lfi2epi-GTAt3	102	0		0	0		0	0	0	0
Footprint	Chr3	5595169	C	CTTAA	Lfi2epi-GTAt3	122	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfi2epi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfi2epi-GTAt3	122	1	0	0	0	1	0	1	0	0
TE	Chr3	8759468	.	ATCOPIA78	Lfi2epi-GTAt3	102	0		1	1	0	1	1	1	1
indel	Chr3	9589063	TA	T	Lfi2epi-GTAt3	122	0	1	1	1	1	1	0	1	1
SNP	Chr3	9858537	T	C	Lfi2epi-GTAt3, Rbiepi-GTAt3	122	0	1	1	1	1	1	0	1	1
SNP	Chr3	10202051	A	G	Lfi2epi-GTAt3, Rbiepi-GTAt3	122	1	0	0	0	0	1	1	0	0
SNP	Chr3	10743038	G	A	Lfi2epi-GTAt3, Rbiepi-GTAt3	122	0	0	1	1	1	0	0	1	1
SNP	Chr3	13092809	C	T	Lfi2epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	het	0	0	0
SNP	Chr3	14288480	A	T	Lfi2epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	0	1	het	0	1	1
SNP	Chr3	14509867	G	T	Lfi2epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	0	het	1	0	0
SNP	Chr3	14518339	C	G	Lfi2epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	0	het	1	0	0
SNP	Chr3	14530903	G	A	Lfi2epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	0	het	1	0	0

SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	0	het	1	0	0
SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	0	1	het	0	1	1
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	1	1	0	1	het	0	1	1
SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	0	1	0	0	1	0	0	1	1
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	0	1	0	0	1	0	0	1	1
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	1	0	1	1	0	0	1	0	0
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	1	1	1	0	0	0
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	0	0	1	1	1	1
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	0	0	1	1	1	1
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	1	█	1	0	0	0
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	0	0	0	1	het	1
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	1	0	1	1	0	0	0
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	1	0	█	1	0	0	0
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	0	1	0	0	1	1	1
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	0	1	0	0	1	1	1
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	1	0	█	1	1	0	0
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	█	0	1	█	0	0	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	1	0	0	1	0	0	0	0	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	1	1	0	1	1	1	1	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	1	0	0	1	0	0	1	0	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0	█	1	0	█	0	0	1	1
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	0	1	0	het	1	1	0	0	0
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	0	1	0	het	1	1	0	0	0
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	1	0	1	het	0	0	1	1	1
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	0	1	0	het	1	1	0	0	0

SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	1	0	1	het	0	0	1	1	1
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	1	0	1	het	0	0	1	1	1
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	1	0	1	het	0	0	1	1	1
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	0	1	0	het	1	1	0	0	0
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	0	1	0	het	1	1	0	0	0
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	0	1	0	het	1	1	0	0	0
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	1	0	1	het	0	0	1	1	1
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	1	0	1	0	0	0	1	1	1

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 53	epi RIL 54	epi RIL 55	epi RIL 60	epi RIL 62	epi RIL 64*	epi RIL 69*	epi RIL 70	epi RIL 71
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	1	0	0	1	1			0	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	1	1	1	1	1	0	1	0	0
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	1	0	1	0	0	1	0	1	1
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	1	0	1	1	0	1	0	0
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0	1	0			0	1
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0	0	0			0	1
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	1	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	0	1	1	1	1	0	1	1
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	1	1	1	1	0	1	1
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0		1	0	0
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0	0	1	0	0
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	0	1	0	1	0	1	1

SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0	0	1	0	0
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	1	1	1	1	0	1	1
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0	0	1	0	0
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0	0			0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	1	1	1	1	0	1	1
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0	0	1	0	0
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0	0	1	0	0
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0	0			0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0	0	1	0	0
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	0	0	0	0	0	0	1	0	0
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0	0	0	0			0	1
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	1	0	0	1	0		0	0	1
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0	0	0	0			0	0
SNP	Chr3	3666840	A	G	Lfi1epi-GTAt3	122	het	1	0	0	1	1	1	1	0
SNP	Chr3	4740011	C	T	Lfi1epi-GTAt3	122	het	1	0	0	1	1	1	1	0
TE	Chr3	5595169	.	ATENSPM3	Lfi1epi-GTAt3	102	0	0	0	0	0			0	0
Footprint	Chr3	5595169	C	CTTAA	Lfi1epi-GTAt3	122	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfi1epi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfi1epi-GTAt3	122	het	1	0	0	1	0	0	0	0
TE	Chr3	8759468	.	ATCOPIA78	Lfi1epi-GTAt3	102	1	0	1	1	0			0	1
indel	Chr3	9589063	TA	T	Lfi1epi-GTAt3	122	het	0	0	0	0		1	1	1
SNP	Chr3	9858537	T	C	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	het	0	1	0	0	1	1	1	1
SNP	Chr3	10202051	A	G	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	het	1	0	1	1	0	0	0	0
SNP	Chr3	10743038	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	1	0	1	1	0	1	1	0	1

SNP	Chr3	13092809	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	0	0	0	0	0	
SNP	Chr3	14288480	A	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	1	0	1	1	0	1	0	1	0	1	
SNP	Chr3	14509867	G	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	0	1	0	0	1	0	0	1	0	0	
SNP	Chr3	14518339	C	G	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	0	1	0	0	1	0	0	1	0	0	
SNP	Chr3	14530903	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	0	1	0	0	1	0	0	1	0	0	
SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	0	1			0	0	1	0	0	0	
SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	1	0	1	1	0	1	1	0	0	1	
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	1	0	0	1	0			1	1	0	0	1	0	1
SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	1	0	0	1	0	1	1	0	1	1	0	0	1	1
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	1	0	0	1	0	1	1	0	1	1	0	0	1	1
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	0	0	1	0	1	0	0	1	0	0	1	1	1	1
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	0	0	0	0	1	0	1	0	0	0	0
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	1	1			0	0	1	1	0	1	1
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	1	1	1	0	1	0	1	1	0	1	1
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	0	0	0	1	0	1	0	0	0	0	0
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	1	1			0	0	1	1	0	1	1
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	0	0	1			1	0	1	0	0	0	0
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	0	0	1	0	1	0	1	0	0	0	0	0
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	0	1	0	1	0	1	1	0	1	1
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	1	1	0			0	0	1	0	1	1	1
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	0	0	1	0	1	0	1	0	0	0	0	0
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0	1	1	0					0	0	0	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	0	0	1	1	0	0	0	0	0	0	0	0	0	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	1	1	1	0	1	1	1	1	1	1	1	1	1	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	het	1	1	0	1	1	1	1	1	0	0	0	0	1

TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	1	0	0	0	0	0	0	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	het	0	0	1	1	1	0	0	0
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	het	0	0	1	1	1	0	0	0
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	het	1	1	1	0	0	1	1	1
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	0	0	0	0	1	1	0	0	0
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	1	0	1	1	0	0	1	1	1
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	1	0	1	1	0	0	1	1	1
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	1	0	1	1	0	0	1	1	1
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	0	1	0	0	1	1	0	0	0
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	0	1	0	0	1	1	0	0	1
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	0	1	0	0	1	1	0	0	1
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	0	1	1	0	0	1	1	0
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	0	1	1	0	0	1	1	0

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 73	epi RIL 92	epi RIL 93	epi RIL 94	epi RIL 95	epi RIL 98	epi RIL 99*	epi RIL 101	epi RIL 108
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	0	0	0	0	0	0	0	1
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	1	0	0	0	0	1	0	1
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	0	1	1	1	1	0	0	1	0
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	1	0	0	0	0	1	1	0	1
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	1	0	0	0	het	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	1	0	0	0	het	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	0	1	1	1	1	het	0	1	0
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	1	1	1	het	0	1	0

SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	0	0	het	1	0	1
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	0	0	het	1	0	1
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	1	0	0	het	0	1	0
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	0	0		1	0	1
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	1	1	1	het	0	1	0
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	0	0		1	0	1
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0	0	0	0	0	0		0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	1	1	1	het	1	1	0
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	0	0	0	0	0	1
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	0	0	0	0	0	1
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0	0	0	0	0	0		0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	0	0	0	0	0	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	0	0	1	0	0	0	1	1
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0	0	1	0	1		0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	0	1	1	1	1	1	0	1	0
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0	0	0	0	0		0	0
SNP	Chr3	3666840	A	G	Lfiepi-GTAt3	122	1	1	1	0	1	het	1	1	1
SNP	Chr3	4740011	C	T	Lfiepi-GTAt3	122	1	1	1	0	1	het	1	1	1
TE	Chr3	5595169	.	ATENSPM3	Lfiepi-GTAt3	102	0	0	0	0	0	0		0	1
Footprint	Chr3	5595169	C	CTTAA	Lfiepi-GTAt3	122	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfiepi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfiepi-GTAt3	122	1	1	1	0	1	1	1	1	1
TE	Chr3	8759468	.	ATCOPIA78	Lfiepi-GTAt3	102	0	0	0	0	0	1		0	1

indel	Chr3	9589063	TA	T	Lfi1epi-GTAt3	122	0	0	0	1	0	1	0	0	0
SNP	Chr3	9858537	T	C	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	0	0	0	1	0	1	0	0	0
SNP	Chr3	10202051	A	G	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	1	1	1	0	1	0	1	1	1
SNP	Chr3	10743038	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	0	0	0	1	0	1	0	0	0
SNP	Chr3	13092809	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	1	0	0
SNP	Chr3	14288480	A	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	1	0	0	0
SNP	Chr3	14509867	G	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	0	1	1	1
SNP	Chr3	14518339	C	G	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	0	1	1	1
SNP	Chr3	14530903	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	0	1	1	1
SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	0	1	1	1
SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	1	0	0	0
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	0	0	1	0	1	0	0	0
SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	1	0	0	1	0	1	0	0	0
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	1	0	0	1	0	1	0	0	0
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	0	1	1	0	1	0	1	0	0
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	0	0	1	0	0	1	1	0
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	0	1	1	0	0	1
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	0	1	1	0	0	1
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	0	0	1	0	0	1	1	0
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	0	1	1	0	0	1
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	0	0	1	0	1
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	0	0	1	0	1
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	het	het	0	1	1	0	0	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	het	het	0	1	1	0	1	0
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0		0	1	0		1	0	1

TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	1	1	0	0	1	1		0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	1	het	1	0	1	1	0	1	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	1	0	1	0	0	1	0	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	0	1	1	0	1	1	0	1	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	1	0	0	0	0	0		0	1
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	1	0	0	1	0	0	0	0	1
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	0	0	1	0	0	het	0	1
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	1	1	0	1	1	1	1	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	1	0	0	1	0	0	het	0	1
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	0	1	1	1	1	1	0	1	0
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	0	1	1	1	1	1	0	1	0
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	0	1	1	1	1	1	0	1	0
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	1	0	0	0	0	0	1	0	1
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	1	0	0	0	0	0	1	0	1
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	1	0	0	0	0	0	1	0	1
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	1	1	1	1	1	0	1	0
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	1	1	1	1	1	0	1	0

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 112	epi RIL 114	epi RIL 118*	epi RIL 122	epi RIL 137	epi RIL 144	epi RIL 147	epi RIL 148	epi RIL 150
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	0		0	0	0	0	0	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	0	0	0	0	1	het	0	1
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	0	1	1	1	1	0	0	1	0
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	1	0	0	0	0	1	het	0	1
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0		0	0	0	0	0	0

TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0		0	0	0	0	0	0	0	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	0	1	1	1	1	0	het	1	0				
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	1	1	1	0	het	1	0				
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	0	0	0	1	het	0	1				
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	0	0	0	1	het	0	1				
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	1	0	het	0	het	1	1				
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	0	0	0	1	het	0	0				
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	1	1	1	1	het	1	1				
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	0	0	0	0	het	0	0				
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	102	0	0		0	0	0	0	0	0	0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	1	1	1	1	0	1	1				
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	0	0	0	0	1	0	0				
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	0	0	0	0	1	0	0				
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	102	0	0		0	0	0	0	0	0	0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	0	0	0	0	1	0	0				
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	0	0	0	0	0	0	0	0				
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0		0	0	0	0	0	0	0	0	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	1	1	1	1	1	0	het	1	1				
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0		0	0	0	0	0	0	0	0	0	0
SNP	Chr3	3666840	A	G	Lfi2epi-GTAt3	122	1	1	1	1	0	0	het	1	1				
SNP	Chr3	4740011	C	T	Lfi2epi-GTAt3	122	1	1	1	1	0	1	het	1	1				

TE	Chr3	5595169	.	ATENSPM3	Lfiempi-GTAt3	102	0	0	0	0	0	0	het	0	1
Footprint	Chr3	5595169	C	CTTAA	Lfiempi-GTAt3	122	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfiempi-GTAt3	122	0	0	0	0	0	1	0	0	0
SNP	Chr3	8547929	G	A	Lfiempi-GTAt3	122	1	1	1	1	1	1	het	1	1
TE	Chr3	8759468	.	ATCOPIA78	Lfiempi-GTAt3	102	0	0	0	0	0	0	1	0	0
indel	Chr3	9589063	TA	T	Lfiempi-GTAt3	122	1	0	0	0	0	0	1	0	1
SNP	Chr3	9858537	T	C	Lfiempi-GTAt3, Rbiepi-GTAt3	122	1	0	0	0	0	0	1	0	1
SNP	Chr3	10202051	A	G	Lfiempi-GTAt3, Rbiepi-GTAt3	122	0	1	1	1	1	1	0	1	0
SNP	Chr3	10743038	G	A	Lfiempi-GTAt3, Rbiepi-GTAt3	122	1	0	0	0	0	0	1	0	1
SNP	Chr3	13092809	C	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	1	1	0	1	0	0	het
SNP	Chr3	14288480	A	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	0	0	1	0	het
SNP	Chr3	14509867	G	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	1	1	1	0	1	het
SNP	Chr3	14518339	C	G	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	1	1	1	0	1	het
SNP	Chr3	14530903	G	A	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	1	1	1	0	1	het
SNP	Chr3	14545591	C	A	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	1	1	1	0	1	het
SNP	Chr3	15861145	C	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	0	0	1	0	0
SNP	Chr3	16205156	C	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	0	0	0	0	0	1	0	0
SNP	Chr3	18112512	G	A	Lfiempi-GTAt3	122	0	0	0	0	0	0	1	0	1
SNP	Chr3	18253358	C	G	Lfiempi-GTAt3	122	0	0	0	0	0	0	1	0	1
SNP	Chr3	20306786	C	T	Lfiempi-GTAt3	122	1	1	1	1	1	0	0	1	0
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	0	0	1	het	1	0
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	1	1	0	het	0	1
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	1	1	0	het	0	1
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	0	0	1	het	1	0
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	0	1	1	0	het	0	1
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	1	het	het	1	0

SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	1	het	het	1	0
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	0	0	het	0	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	0	1	0	het	het	0	0
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	1	het	het	1	1
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0		1	0	0	0	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	0	1	0	1	0	0	het	0	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	1	0	1	0	1	1	het	0	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	1	1	1	1	1	1	het	het	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0	0		0	0	0	0	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	1	0		0		0	het	1	1
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	0	0	0	1	0	het	1	1
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	1	1	1	0	1	het	0	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	1	0	0	0	1	0	het	1	0
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	0	1	1	1	0	1	het	0	1
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	0	1	1	1	0	1	het	0	1
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	0	1	1	1	0	1	het	0	1
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	1	0	0	0	1	0	het	1	0
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	1	0	0	0	1	0	0	1	0
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	1	0	0	0	1	0	0	1	0
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	1	1	1	0	1	1	0	1
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	1	1	1	0	1	1	0	1

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 159	epi RIL 164	epi RIL 166	epi RIL 169	epi RIL 170	epi RIL 172	epi RIL 183*	epi RIL 193	epi RIL 195
--------------	------------	----------	-----	-----	--------------	--------------------------------	-------------	-------------	-------------	-------------	-------------	-------------	--------------	-------------	-------------

TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	0	0	0	0	0	0	0	0	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	0	1	1	1	0	0	1	1	1
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	1	1	0	1	0	1	0	0	0	0
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	1	0	1	0	1	1	1	1
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0	0	0	0	0	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	1	0	0	0	0	0	0	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	1	0	0	0	0	1	0	0	0
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	0	0	0	0	0	0	0
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	1	1	1	1	1	1	1
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	1	1	1	1	1	1	1
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	0	0	0	0	0	0	0	0
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	1	1	1	1	1	1	1
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	0	0	0	0	0	0	0
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	1	1	1	1	1	1	1
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0	0	0	0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	0	0	0	0	0	0	0
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	1	1	1	1	1	1	1
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	1	1	1	1	1	1	1
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0	0	0	0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	1	1	1	1	1	1	1	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	0	0	1	0	1	1	1	1	1	1

TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	het	1	1	0	0	0	0	1	0	0	0	0	0
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	1	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr3	3666840	A	G	Lfiapi-GTAt3	122	1	0	0	0	1	1	1	1	0	0	0	0	0
SNP	Chr3	4740011	C	T	Lfiapi-GTAt3	122	1	0	0	0	1	1	1	0	0	0	0	0	0
TE	Chr3	5595169	.	ATENSPM3	Lfiapi-GTAt3	102	0	0	0	0	0	1	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTAA	Lfiapi-GTAt3	122	1	0	0	0	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfiapi-GTAt3	122	0	0	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfiapi-GTAt3	122	1	0	0	0	1	1	0	0	0	0	0	0	0
TE	Chr3	8759468	.	ATCOPIA78	Lfiapi-GTAt3	102	0	0	1	0	0	0	0	0	0	0	0	0	0
indel	Chr3	9589063	TA	T	Lfiapi-GTAt3	122	1	1	1	1	0	0	1	0	1	0	1	0	1
SNP	Chr3	9858537	T	C	Lfiapi-GTAt3, Rbiepi-GTAt3	122	1	1	1	1	0	0	1	0	1	0	1	0	1
SNP	Chr3	10202051	A	G	Lfiapi-GTAt3, Rbiepi-GTAt3	122	0	0	0	0	1	1	0	1	0	1	0	1	0
SNP	Chr3	10743038	G	A	Lfiapi-GTAt3, Rbiepi-GTAt3	122	1	1	1	1	0	0	1	0	1	0	1	0	1
SNP	Chr3	13092809	C	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	1	1	0	1	1
SNP	Chr3	14288480	A	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	1	1	1	1	1	0	0	0	0	0
SNP	Chr3	14509867	G	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	1	1	0	1	1
SNP	Chr3	14518339	C	G	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	1	1	0	1	1
SNP	Chr3	14530903	G	A	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	1	1	0	1	1
SNP	Chr3	14545591	C	A	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	1	1	0	1	1
SNP	Chr3	15861145	C	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	1	1	1	1	1	0	0	0	0	0
SNP	Chr3	16205156	C	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	1	1	1	1	1	1	1	1	0	0	0	0	0
SNP	Chr3	18112512	G	A	Lfiapi-GTAt3	122	1	1	1	1	1	1	1	1	0	0	0	0	0
SNP	Chr3	18253358	C	G	Lfiapi-GTAt3	122	1	1	1	1	1	1	1	1	0	0	0	0	0
SNP	Chr3	20306786	C	T	Lfiapi-GTAt3	122	0	1	1	0	0	0	0	1	1	0	1	1	1
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	0	0	1	1	0	0	0	0	0	0	0	het
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	0	0	1	1	1	1	1	1	1	het

SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	0	0	1	1	1	het
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	0	0	1	1	0	0	0	het
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	0	0	1	1	1	het
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	0	1	1	0	0	0	0
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	0	1	1	0	0	0	0
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	0	1	0	1	1
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	1	0	0	1	0	1	1
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	0	0	1	0	1	0	0
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0	0	0	0	1		1	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	0	1	1	1	0	1	0	1	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	1	0	0	0	1	0	1	0	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	0	0	0	1	1	1	1	1	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0	0	1	0	0	0		0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	1	het	1	0	0	0	het	0	het
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	1	1	0	0	0	het	1	1
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	0	0	1	1	1	het	0	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	1	1	1	0	0	0	het	1	1
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	0	0	0	1	1	1	het	0	0
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	0	0	0	1	1	1	het	0	0
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	0	0	0	1	1	1	het	0	0
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	1	1	1	0	0	0	het	1	1
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	1	1	1	0	0	0	1	1	1
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	1	1	1	0	0	0	1	1	1
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	0	0	1	1	1	0	0	0
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	0	0	1	1	1	0	0	0

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 202	epi RIL 208	epi RIL 215	epi RIL 216*	epi RIL 218	epi RIL 222	epi RIL 225	epi RIL 229	epi RIL 232
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	0	0		0	0	0	0	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	1	0	0	0	het	1	1	0	0
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	0	1	1	1	0	1	0	1	1
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	1	0	0	0	1	0	1	0	0
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0		0	0	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0		0	0	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	1	0	0	0	1	0	1	1
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	0	0	0	1	0	1	1
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	1	1	1	0	1	0	0
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	1	1	1	0	1	0	0
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	0	0	0	1	0	1	1
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	1	1	1	0	1	0	0
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	0	0	0	1	0	1	1
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0		1	1	1	0	1	0	0
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0	1	0		0	1	0	0	1
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	1	0	1	1	0	1	1
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	1	0	0	1	0	0

SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	0	1	0	0	1	0	0
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0		0	1	0	0	1
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	0	het	0	0	1	0	0
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	0	0	het	1	0	1	0	0
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0	1		het	0	0	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	0	1	1	1	1	0	0	1	1
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0	0		0	0	0	0	1
SNP	Chr3	3666840	A	G	Lfi1epi-GTAt3	122	0	1	het	1	0	0	0	1	1
SNP	Chr3	4740011	C	T	Lfi1epi-GTAt3	122	0	1	het	1	0	0	1	1	1
TE	Chr3	5595169	.	ATENSPM3	Lfi1epi-GTAt3	102	0	0	0		0	0	1	0	0
Footprint	Chr3	5595169	C	CTTAA	Lfi1epi-GTAt3	122	0	0	het	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfi1epi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfi1epi-GTAt3	122	0	1	0	1	0	0	1	1	1
TE	Chr3	8759468	.	ATCOPIA78	Lfi1epi-GTAt3	102	0	0	0		1	0	0	0	0
indel	Chr3	9589063	TA	T	Lfi1epi-GTAt3	122	1	0	1	0	1	1	0	0	0
SNP	Chr3	9858537	T	C	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	1	0	1	0	1	1	0	0	0
SNP	Chr3	10202051	A	G	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	0	1	0	1	0	0	1	1	1
SNP	Chr3	10743038	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	0	0	1	0	1	1	0	0	0
SNP	Chr3	13092809	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	1	0	0	1	1	1
SNP	Chr3	14288480	A	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	0	1	1	0	0	0
SNP	Chr3	14509867	G	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	1	0	0	1	1	1
SNP	Chr3	14518339	C	G	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	1	0	0	1	1	1
SNP	Chr3	14530903	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	1	0	0	1	1	1
SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	1	0	0	1	1	1
SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	0	1	1	0	0	0
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	1	0	1	0	1	1	0	0	0

SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	1	0	1	0	1	1	0	0	0
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	1	0	1	0	1	1	0	0	0
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	0	1	0	1	0	0	1	1	1
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	1	0	0	1	0	0
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	1	1	0	1	1
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	1	1	0	1	1
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	1	0	0	1	0	0
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	1	1	0	1	1
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	0	0	1	0	0
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	0	0	1	0	0
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	0	0	1	1	0	0	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	0	1	1	0	1	1
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	1	0	1	0	0
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0	0		0	0	0	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	1	1	0	0	0	1	0	1	1
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	0	1	1	1	0	1	0	0
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	1	1	0	0	0	1	1	het	1
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0	0	1		1	0	0	1	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	0	0	1	0	1	1	0	het	0
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	0	1	0	1	1	0	het	0
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	1	0	1	0	0	1	1	1
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	1	0	1	0	1	1	0	0	0
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	0	1	0	1	0	0	0	1	1
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	0	1	0	1	0	0	0	1	1
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	0	1	0	1	0	0	0	1	1

SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	1	0	1	0	1	1	1	0	0
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	1	0	1	0	1	1	1	0	0
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	1	0	1	0	1	1	1	0	0
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	1	0	1	0	0	0	1	1
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	1	0	1	0	0	0	1	1

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 238	epi RIL 244	epi RIL 252	epi RIL 257	epi RIL 258*	epi RIL 260	epi RIL 262	epi RIL 275	epi RIL 276
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	1	0	0		1	0	0	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	1	1	0	1	1	1	het	1	1
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	0	0	1	0	0	0	1	0	0
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	1	1	0	1	1	1	0	1	1
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0	0		0	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0	0		0	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	0	0	0	0	0	0	1	0	0
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	0	0	1	0	0
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	1	1	1	1	0	1	1
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	1	1	1	1	0	1	1
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	1	1	1	1	0	1	1
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	0	0	1	0	0

indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	1	1	1	1	0	1	1
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0		0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	0	0	0	0	1	0	0
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	het	0	1	1	1	1	0	1	1
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	1	1	1	1	0	1	1
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0		0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	1	1	1	1	0	1	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	0	1	1	0	1	0	1	1
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0	0	0		0	0	0	het
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	0	0	0	1	1	0	1	0	1
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0	0	0		0	0	0	0
SNP	Chr3	3666840	A	G	Lfi1epi-GTAt3	122	1	0	1	0	1	1	1	1	1
SNP	Chr3	4740011	C	T	Lfi1epi-GTAt3	122	1	0	1	0	1	1	1	1	1
TE	Chr3	5595169	.	ATENSPM3	Lfi1epi-GTAt3	102	0	0	0	0		0	0	1	0
Footprint	Chr3	5595169	C	CTTAA	Lfi1epi-GTAt3	122	0	0	0	0		0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfi1epi-GTAt3	122	0	0	0	0		0	0	0	0
SNP	Chr3	8547929	G	A	Lfi1epi-GTAt3	122	1	0	het	0	1	1	0	1	1
TE	Chr3	8759468	.	ATCOPIA78	Lfi1epi-GTAt3	102	0	1	1	0		1	0	0	0
indel	Chr3	9589063	TA	T	Lfi1epi-GTAt3	122	0	1	1	1	0	0	1	0	0
SNP	Chr3	9858537	T	C	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	0	1	1	1	0	0	1	0	0
SNP	Chr3	10202051	A	G	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	1	0	0	0	1	1	0	1	1
SNP	Chr3	10743038	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	0	1	1	1	0	0	1	0	0
SNP	Chr3	13092809	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	1	0	0	1	0
SNP	Chr3	14288480	A	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	1	0	0	1	0	0
SNP	Chr3	14509867	G	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	1	1	0	1	1

SNP	Chr3	14518339	C	G	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	1	1	0	1	1
SNP	Chr3	14530903	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	1	1	0	1	1
SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	1	1	0	1	1
SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	0	0	0	1	0	0
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	1	1	0	0	0	1	0	0
SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	0	1	1	0	1	0	1	0	0
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	0	1	1	0	1	0	1	0	0
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	1	0	1	0	0	1	0	1	1
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	0	1	1	0	1
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	1	1	0	0	1	0
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	1	1	0	0	1	0
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	0	1	1	0	1
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	1	1	0	0	1	0
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	0	1	1	0	1
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	0	1	1	0	1
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	1	0	0	0	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	0	1	1	0	0	1	0
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	0	1	1	0	1
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0	0	1		0	0	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	0	1	0	1	1	0	0	1	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	1	0	1	0	0	1	1	0	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	1	1	0	0	0	1	1	0	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0	0	1	0		0	0	1	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	0	0	1	0		0	0	1	1
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	0	0	1	0	1	0	0	1	1

indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	1	1	0	1	0	0	1	0	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	0	0	1	0	1	1	0	1	1
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	1	1	0	1	0	0	1	0	0
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	1	1	0	1	0	0	1	0	0
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	1	1	0	1	0	0	1	0	0
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	0	0	1	0	1	1	0	1	1
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	0	0	1	0	1	1	0	1	1
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	0	0	1	0	1	1	0	1	1
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	1	1	0	1	0	0	1	0	0
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122		1	0	1	0	0	1	0	0

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 277	epi RIL 297	epi RIL 305*	epi RIL 315*	epi RIL 323	epi RIL 326	epi RIL 333	epi RIL 340	epi RIL 344
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	0			1	0	0	0	1
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	0	1	0	1	0	0	0	1
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	1	1	1	1	0	1	0	0	1
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	1	0	1	1	0
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	1			0	0	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	1			0	0	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	1	0	0	0	0	0	1
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	1	1	1	0	1	0	0	1
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	1	1	0	1	0	0	1
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	1	0	1	1	0

SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	1	0	1	1	0
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	1	0	0	0	0	0	1
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	1	0	1	1	0
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	1	1	0	1	0	0	1
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	1	0	1	1	0
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0	0			0	0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	het	1	1	0	het	1	0	het
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	het	0	0	1	het	0	1	het
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	het	0	0	1	het	0	1	het
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0	0			0	0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	het	0	0	1	het	0	1	het
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	0	het	0	0	0	het	0	1	het
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	1	0			0	0	0	1	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	1	1		0	1	0	1	1	0
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0			0	0	0	0	0
SNP	Chr3	3666840	A	G	Lfiepi-GTAt3	122	0	0	1	0	0	1	1	0	0
SNP	Chr3	4740011	C	T	Lfiepi-GTAt3	122	0	0		0	0	1	1	0	0
TE	Chr3	5595169	.	ATENSPM3	Lfiepi-GTAt3	102	0	0			0	1	0	0	0
Footprint	Chr3	5595169	C	CTTAA	Lfiepi-GTAt3	122	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfiepi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfiepi-GTAt3	122	0	1	1	0	0	1	1	0	0
TE	Chr3	8759468	.	ATCOPIA78	Lfiepi-GTAt3	102	1	0			0	0	0	0	1
indel	Chr3	9589063	TA	T	Lfiepi-GTAt3	122	1	0	0	1	1	0	0	1	1

SNP	Chr3	9858537	T	C	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	1	0	0	1	0	0	0	1	1
SNP	Chr3	10202051	A	G	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	het	1	1	0	1	1	1	0	0
SNP	Chr3	10743038	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	het	0	0	1	0	0	0	1	1
SNP	Chr3	13092809	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	het	0	0	0	1	0	1	0	0
SNP	Chr3	14288480	A	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	het	0	0	1	0	0	0	1	1
SNP	Chr3	14509867	G	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	het	1	1	0	1	1	1	0	0
SNP	Chr3	14518339	C	G	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	het	1	1	0	1	1	1	0	0
SNP	Chr3	14530903	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	het	1	1	0	1	1	1	0	0
SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	het	1	1	0	1	1	1	0	0
SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	0	0	1	1
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	0	0	1	0	0	0	1	1
SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	0	0	0	1	0	0	0	1	1
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	0	0	0	1	0	0	0	1	1
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	1	1	0	1	1	1	1	0	0
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	0	0	1	0	1	1
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	1	0	1	0	0
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	1	0	1	0	0
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	0	0	1	0	1	1
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	1	0	1	0	0
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	0	1	0	het	1	1	1
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	0	1	0	0	1	1	1
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	1	0	1	1	0	0	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	1	1	0	0	0
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	0	1	0	0	1	1	1
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0			1	1	0	0	0

SNP	Chr4	10012411	A	G	Pb2epi-At4	122	1	0		0	1	1	0	0	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	1	0	1	0	1	1	1	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	0	1	1	0	0	0	1	1	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	1	0			1	1	0	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	1	0	0	1	1	1	1	1	1
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	0	0	1	1	1	1	1	1
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	1	1	0	0	0	1	0	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	1	0	0	1	1	1	0	1	1
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	0	1	1	0	0	0	1	0	0
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	0	1	1	0	0	0	1	0	0
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	0	1	1	0	0	0	1	0	0
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	1	0	0	1	1	1	0	1	1
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	1	0	0	1	1	1	0	1	0
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	1	0	0	1	1	1	0	1	0
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	1	1	0	0	0	1	0	1
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	1	1	0	0	0	1	0	1

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 350*	epi RIL 356	epi RIL 361	epi RIL 362	epi RIL 363	epi RIL 366	epi RIL 368	epi RIL 371	epi RIL 375
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102		0	0	0	0	0	0	0	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	0	0	0	0	0	1	1	0
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	1	1	1	1	1	1	0	1	1
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	1	0	0
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102		0	0	0	0	0	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102		0	0	0	0	0	0	0	0

Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	1	1	0	1	0	0	0	0	1	0	0	1
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	1	1	0	1	0	0	0	0	0	0	0	1
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	0	0	1	0	1	1	1	1	0	0	0	0
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	0	0	1	0	1	1	1	1	0	0	0	0
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	1	1	0	1	0	0	0	0	0	0	0	0
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	0	0	1	0	1	1	1	1	0	0	0	0
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	1	1	0	1	0	0	0	0	0	0	0	1
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	0	0	1	0	1	1	1	1	0	0	0	0
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	102		1	0	0	1	0	0	0	0	0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	1	1	0	1	0	0	0	0	0	0	0	1
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	0	0	1	0	1	1	1	1	0	0	0	0
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	0	0	1	0	het	1	1	1	0	0	0	0
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	102		0	0	0	1	0	0	0	0	0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	0	0	1	0	het	1	1	1	0	0	0	0
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	0	0	1	1	0	0	1	1	0	0	0	0	0
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106		0	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	0	1	0	1	1	1	0	het	1	0	0	0	1
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102		0	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr3	3666840	A	G	Lfi2epi-GTAt3	122	0	1	0	1	1	0	0	het	het	0	0	0	0
SNP	Chr3	4740011	C	T	Lfi2epi-GTAt3	122	0	1	0	1	1	0	0	het	het	0	0	0	1
TE	Chr3	5595169	.	ATENSPM3	Lfi2epi-GTAt3	102		0	0	0	0	0	0	0	0	0	0	het	0

Footprint	Chr3	5595169	C	CTTAA	LfiEpi-GTAt3	122	0	0	0	0	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	LfiEpi-GTAt3	122	0	0	0	0	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	LfiEpi-GTAt3	122	1	1	0	0	1	0	0	1	1	1	1	1	1
TE	Chr3	8759468	.	ATCOPIA78	LfiEpi-GTAt3	102		0	0	1	0	0	0	0	0	0	0	0	0
indel	Chr3	9589063	TA	T	LfiEpi-GTAt3	122	0	0	1	1	0	1	1	0	1	0	0	0	0
SNP	Chr3	9858537	T	C	LfiEpi-GTAt3, Rbiepi-GTAt3	122	0	0	1	1	0	1	1	0	1	0	0	0	0
SNP	Chr3	10202051	A	G	LfiEpi-GTAt3, Rbiepi-GTAt3	122	1	1	0	0	1	0	0	1	0	1	1	1	1
SNP	Chr3	10743038	G	A	LfiEpi-GTAt3, Rbiepi-GTAt3	122	0	0	1	1	0	0	1	0	1	0	0	0	0
SNP	Chr3	13092809	C	T	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3, Lfi1Epi-At3	122	1	0	0	0	0	0	0	0	0	0	0	1	1
SNP	Chr3	14288480	A	T	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3, Lfi1Epi-At3	122	0	0	1	1	0	0	1	1	0	1	1	0	0
SNP	Chr3	14509867	G	T	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3, Lfi1Epi-At3	122	1	1	0	0	1	1	0	0	1	0	0	1	1
SNP	Chr3	14518339	C	G	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3, Lfi1Epi-At3	122	1	1	0	0	1	1	0	0	1	0	0	1	1
SNP	Chr3	14530903	G	A	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3, Lfi1Epi-At3	122	1	1	0	0	1	1	0	0	1	0	0	1	1
SNP	Chr3	14545591	C	A	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3, Lfi1Epi-At3	122	1	1	0	0	1	1	0	0	1	0	0	1	1
SNP	Chr3	15861145	C	T	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3, Lfi1Epi-At3	122	0	0	1	1	0	0	1	1	0	1	1	0	0
SNP	Chr3	16205156	C	T	LfiEpi-GTAt3, Rbiepi-GTAt3, Lfi2Epi-At3	122	0	0	1	1	0	0	1	1	0	1	1	0	0
SNP	Chr3	18112512	G	A	LfiEpi-GTAt3	122	0	0	1	1	0	0	1	1	0	1	1	0	0
SNP	Chr3	18253358	C	G	LfiEpi-GTAt3	122	0	0	1	1	0	0	1	1	0	1	1	0	0
SNP	Chr3	20306786	C	T	LfiEpi-GTAt3	122	1	1	0	0	1	1	0	0	1	0	0	0	0
SNP	Chr4	2834559	G	A	Pb1Epi-At4, Pbepi-GTAt4a	122	1	0	1	0	0	0	0	1	het	1	1	1	1
SNP	Chr4	2920544	G	A	Pb1Epi-At4, Pbepi-GTAt4a	122	0	1	0	1	1	1	0	0	het	0	0	0	0
SNP	Chr4	3318778	C	G	Pb1Epi-At4, Pbepi-GTAt4a	122	0	1	0	1	1	1	0	0	het	0	0	0	0
SNP	Chr4	4474111	ACCAGAT	A	Pb1Epi-At4, Pbepi-GTAt4a	122	1	0	1	0	0	0	1	0	het	1	1	1	1
SNP	Chr4	5115842	T	A	Pb1Epi-At4, Pbepi-GTAt4a	122	0	1	0	1	1	1	0	0	het	0	0	0	0
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2Epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	0	0	1	1	0	1	1	0	0
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2Epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	0	0	1	1	0	1	1	0	0

indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	0	0	0	0	1
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	1	0	1	1	1	0	0	1
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	0	0	1	1	0
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102		0	0	1	0	0	0	1	1
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	0	1	0	1	1	1	0	1	1
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	0	1	0	0	0	1	0	0
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	het	1	0	1	1	1	0	0	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102		0	1	0	0	0	0	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	0	0	1	0	0	0	1	1	1
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	0	0	1	0	0	0	1	1	0
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	1	1	0	1	1	1	0	0	1
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	0	0	1	0	0	0	1	1	0
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	1	1	0	1	1	1	0	0	1
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	1	1	0	1	1	1	0	0	1
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	1	1	0	1	1	1	0	0	1
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	0	0	1	0	0	0	1	1	0
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	0	0	1	0	0	0	1	1	0
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	0	0	1	0	0	0	1	1	0
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	1	1	het	1	1	1	0	0	1
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	1	1	het	1	1	1	0	0	0

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 391*	epi RIL 393	epi RIL 394	epi RIL 400	epi RIL 408	epi RIL 410	epi RIL 425	epi RIL 432	epi RIL 434
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102		0	1	0	1	0	0	0	1

SNP	Chr1	7241579	G	A	Rbiepi-At1	122		1	1	0	0	1	0	0	1
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	0	0	0	0	1	0	1	1	0
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122		1	1	1	0	1	0	0	1
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102		0	0	0	0	0	1	1	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102		0	0	0	0	0	1	1	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	0	0	0	1	1	0	1	1	0
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	0	1	1	0	1	1	0
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	1	0	0	1	0	0	1
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	1	0	0	1	0	0	1
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	0	0	0	0	1	1	0
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	1	0	0	1	0	0	1
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	0	1	1	0	1	1	0
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	1	0	0	1	0	0	1
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102		0	0	0	0	0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	1	1	0	1	1	0
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	0	0	1	0	0	1
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	0	0	1	0	0	1
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102		0	0	0	0	0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	0	0	1	0	0	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	1	0	0	0	1	0	0	het
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106		0	0	0	0	0	0	0	0

SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	1	0	1	0	0	0	0	0	1
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102		0	0	0	0	0	0	0	0
SNP	Chr3	3666840	A	G	Lfiapi-GTAt3	122	1	1	1	0	0	0	0	1	1
SNP	Chr3	4740011	C	T	Lfiapi-GTAt3	122	1	1	1	0	0	0	0	1	0
TE	Chr3	5595169	.	ATENSPM3	Lfiapi-GTAt3	102		0	0	0	0	0	0	1	0
Footprint	Chr3	5595169	C	CTTAA	Lfiapi-GTAt3	122	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfiapi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfiapi-GTAt3	122	1	0	1	0	0	0	0	1	0
TE	Chr3	8759468	.	ATCOPIA78	Lfiapi-GTAt3	102		1	0	1	0	0	0	0	0
indel	Chr3	9589063	TA	T	Lfiapi-GTAt3	122		1	1	1	1	0	1	0	1
SNP	Chr3	9858537	T	C	Lfiapi-GTAt3, Rbiepi-GTAt3	122		1	1	1	1	0	0	0	1
SNP	Chr3	10202051	A	G	Lfiapi-GTAt3, Rbiepi-GTAt3	122	1	0	0	0	0	1	1	1	0
SNP	Chr3	10743038	G	A	Lfiapi-GTAt3, Rbiepi-GTAt3	122	0	1	1	1	1	0	0	0	1
SNP	Chr3	13092809	C	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	14288480	A	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	1	1	0	0	0	1
SNP	Chr3	14509867	G	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	0	0	1	1	1	0
SNP	Chr3	14518339	C	G	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	0	0	1	1	1	0
SNP	Chr3	14530903	G	A	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	0	0	1	1	1	0
SNP	Chr3	14545591	C	A	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	1	0	0	1	1	1	0
SNP	Chr3	15861145	C	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	0	1	1	1	0	0	1
SNP	Chr3	16205156	C	T	Lfiapi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	1	0	1	1	1	0	0	1
SNP	Chr3	18112512	G	A	Lfiapi-GTAt3	122	1	0	0	1	1	1	0	0	1
SNP	Chr3	18253358	C	G	Lfiapi-GTAt3	122	1	0	0	1	1	1	0	0	1
SNP	Chr3	20306786	C	T	Lfiapi-GTAt3	122	0	0	1	0	0	0	1	1	0
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	0	1	0	1	0	1
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	0	1	0	1	0
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	0	1	0	1	0

SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	0	1	0	1	0	1
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	0	1	0	1	0
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	0	0	1	0	1
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	0	0	1	0	1
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	0	1	0	0	0	0	1	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	1	1	1	0	1	0
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	0	0	1	0	1
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102		0	1	0	0	0	0	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	1	1	1	1	1	1	0	0	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	0	0	0	0	0	1	1	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	1	1	1	1	1	0	1	0	1
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102		0	0	0	0	0	0	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	0	1	0	0	0	0	0	1	0
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	0	1	0	0	0	0	0	1	1
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	1	0	1	1	1	1	1	0	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	0	1	0	0	0	0	0	1	1
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	1	0	1	1	1	1	1	0	0
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	1	0	1	1	1	1	1	0	0
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	1	0	1	1	1	1	1	0	0
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	0	1	0	0	0	0	0	1	1
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	0	1	0	0	0	0	0	1	1
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	0	1	0	0	0	0	0	1	1
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	1	0	1	1	1	1	1	0	0
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122		0	1	1	1	1	1	0	0

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 437	epi RIL 438	epi RIL 439	epi RIL 454	epi RIL 458	epi RIL 466	epi RIL 467	epi RIL 473	epi RIL 477
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	0	1	0	0	0	0	0	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	0	1	1	1	0	0	het	0
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	1	1	0	0	0	1	1	1	0
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	het	1	1	0	0	0	1
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	foot	0	0	0	1	1	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	foot	0	0	0	0	1	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	1	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	1	1	1	0	1	1	1	1
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	1	1	0	1	1	1	1
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	het	0	0	1	0	0	0	0
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	het	0	0	1	0	0	0	0
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	het	0	0	0	1	1	0	0
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	0	1	0	0	0	0
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	1	1	1	0	1	1	1	1
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	0	0	1	0	0	0	0
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0	0	0	0	0	0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	het	1	0	0	1	1	1	1
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	1	1	0	0	0	0
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	0	0	1	1	0	0	0	0
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0	0	0	0	0	0	0	0	0

SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	0	1	1	0	0	0	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	1	0	het	1	0	0	0	het
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0	1	0	1	0	0	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	1	het	1	1	1	0	0	0	1
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0	0	0	0	0	0	0	0
SNP	Chr3	3666840	A	G	Lfiepi-GTAt3	122	0	1	0	0	0	0	1	0	1
SNP	Chr3	4740011	C	T	Lfiepi-GTAt3	122	0	0	0	1	0	1	1	0	1
TE	Chr3	5595169	.	ATENSPM3	Lfiepi-GTAt3	102	0	0	0	1	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTAA	Lfiepi-GTAt3	122	0	0	0	0	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTA	Lfiepi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfiepi-GTAt3	122	0	0	0	1	0	0	1	0	1
TE	Chr3	8759468	.	ATCOPIA78	Lfiepi-GTAt3	102	0	1	1	1	0	1	0	0	0
indel	Chr3	9589063	TA	T	Lfiepi-GTAt3	122	1	1	1	1	1	1	0	1	0
SNP	Chr3	9858537	T	C	Lfiepi-GTAt3, Rbiepi-GTAt3	122	1	1	1	1	1	1	0	1	0
SNP	Chr3	10202051	A	G	Lfiepi-GTAt3, Rbiepi-GTAt3	122	0	0	0	0	0	0	1	0	1
SNP	Chr3	10743038	G	A	Lfiepi-GTAt3, Rbiepi-GTAt3	122	1	1	1	1	1	1	0	1	0
SNP	Chr3	13092809	C	T	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	14288480	A	T	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	1	1	1	0	1	1
SNP	Chr3	14509867	G	T	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	0	0	1	0	0
SNP	Chr3	14518339	C	G	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	0	0	1	0	0
SNP	Chr3	14530903	G	A	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	0	0	1	0	0
SNP	Chr3	14545591	C	A	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	0	0	0	1	0	0
SNP	Chr3	15861145	C	T	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	0	1	1	0	1	1
SNP	Chr3	16205156	C	T	Lfiepi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	1	1	0	1	1	0	het	1
SNP	Chr3	18112512	G	A	Lfiepi-GTAt3	122	1	het	1	0	1	1	0	0	1
SNP	Chr3	18253358	C	G	Lfiepi-GTAt3	122	1	het	1	0	1	1	0	0	1
SNP	Chr3	20306786	C	T	Lfiepi-GTAt3	122	0	het	0	1	0	1	1	1	0

SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	0	1	0	0	1	1
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	het	0	1	0	1	1	0	0
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	het	0	1	0	1	1	0	0
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	1	0	1	0	0	1	1
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	het	0	1	0	1	1	0	0
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	het	1	0	1	0	1	1	1
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	het	1	0	1	0	1	1	1
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	het	0	1	0	0	0	0	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	het	0	1	0	1	0	0	0
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	het	1	0	1	0	1	1	1
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0	0	0	1	0	1	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	1	het	0	0	1	1	het	het	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	het	1	1	0	0	het	het	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	1	1	1	0	1	0	0	1	1
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0	0	0	0	0	0	1	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	1	0	0	0	0	1	1	het	1
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	0	0	0	0	1	1	het	1
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	1	1	1	1	0	0	het	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	1	0	0	0	0	1	1	het	1
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	0	1	1	1	1	0	1	0	0
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	0	1	1	1	1	0	1	0	0
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	0	1	1	1	1	0	1	0	0
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	1	0	0	0	0	1	0	1	1
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	1	0	0	0	0	1	0	1	1
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	1	0	0	0	0	1	0	1	1

SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	1	1	1	1	0	1	0	0
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	1	1	1	1	0	1	0	0

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 480	epi RIL 488*	epi RIL 492*	epi RIL 493	epi RIL 494	epi RIL 495	epi RIL 497	epi RIL 500	epi RIL 503
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	1			0	0	0	0	0	1
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	1	0	1	0	het	het	0	0	1
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	0	1		0	het	1	1	het	0
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	1	0	1	1	het	0	0	het	1
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0			0	0	0	0	1	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0			0	0	0	0	het	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	0	1	0	0	het	1	0	het	0
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1		0	1	1	0	het	1
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	1	0	0	1	het	0
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	1	0	0	1	het	0
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1		0	1	1	0	het	1
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	1	0	0	1	0	0
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1		0	1	1	0	0	1
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	0	1	0	0	1	0	0
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	102	0			0	1	1	0	0	0

indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	0	1	0	1	1	0	het	1
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	1	0	0	1	het	0
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	1	0	0	1	het	0
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0			0	0	0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	1	0	0	0	0	1	het	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	1	0	0	0	0	1	het	1
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0			0	0	0	0	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	0	0	1	1	1	1	1	het	0
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0			0	het	1	0	0	0
SNP	Chr3	3666840	A	G	Lfi1epi-GTAt3	122	1	1	0	1	0	0	0	0	1
SNP	Chr3	4740011	C	T	Lfi1epi-GTAt3	122	0	1	0	1	het	1	0	0	1
TE	Chr3	5595169	.	ATENSPM3	Lfi1epi-GTAt3	102	0			1	0	0	0	0	0
Footprint	Chr3	5595169	C	CTTAA	Lfi1epi-GTAt3	122	0	0	0	0	0	0	1	0	0
Footprint	Chr3	5595169	C	CTTA	Lfi1epi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfi1epi-GTAt3	122	0	0	0	1	het	1	1	0	1
TE	Chr3	8759468	.	ATCOPIA78	Lfi1epi-GTAt3	102	1			0	1	0	0	0	0
indel	Chr3	9589063	TA	T	Lfi1epi-GTAt3	122	1	1		0	het	0	0	1	0
SNP	Chr3	9858537	T	C	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	1	1	1	0	het	0	0	1	0
SNP	Chr3	10202051	A	G	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	0	0		1	het	1	1	0	1
SNP	Chr3	10743038	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3	122	1	1	1	0	1	1	0	1	0
SNP	Chr3	13092809	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0		0	0	0	0	0	1
SNP	Chr3	14288480	A	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	1	0	1	0
SNP	Chr3	14509867	G	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	0	1	0	1
SNP	Chr3	14518339	C	G	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	0	1	0	1
SNP	Chr3	14530903	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	0	1	0	1
SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	0	1	0	1

SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	1	0	1	0
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	1	1	1	0	1	1	0		0
SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	1	1	1	0	het	0	0	1	0
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	1	1	1	0	het	0	0	1	0
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	0	0	0	1	het	1	1	0	1
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	1	1	1	0	1
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	0	0	0	0	1	0
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	0	0	0	0	1	0
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	1	1	1	1	0	1
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	0	0	0	0	1	0
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	1	1	1	1	0	0
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	1	1	1	1	het	0
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	0	0	0	het	1
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	0	0	0	het	1
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	0	1	het	1	1	het	0
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0			0	0	0	0	0	1
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	0	0	1	0	het	0	0	het	1
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	1	1	0	1	het	1	1	0	0
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	0	1	1	0	het	het	1	het	1
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	1			1	0	1	0	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	1	0	0	het	1	1	0	1	0
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	0	0	1	het	1	1	1	0
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	1	1	0	het	0	0	0	1
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	1	0	0	1	0	0	1	1	0
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	0	1		0	1	1	0	0	1

SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	0	1	1	0	1	1	0	0	1
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	0	1	1	0	1	1	0	0	1
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	1	0	0	1	0	0	1	1	0
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	1			1	0	0	1	1	0
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	1	0	0	1	0	0	1	1	0
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	0	1	1	0	1	1	0	het	1
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	0	1		0	1	1	0	0	1

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 506	epi RIL 508*	epi RIL 523	epi RIL 538*	epi RIL 539*	epi RIL 556*	epi RIL 558	epi RIL 559	epi RIL 561*
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0		0				0	0	
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	het	0	1	1	1	0	0	0	0
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	0	1	0	0	0	1	1	0	1
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	1	0	1	1	1	0	0	1	
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0		0				0	0	
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0		0				0	0	
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0	0	0	0	0
SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	0	1	1	0	1		1	0	0
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	het	0	1	0	1	0	0
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	het	1	0	1	0	1	1
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	1	0	het	1	0	1	0	1	1
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfiepi-GTAt1	122	0	1	0	0	0	0	0	0	0

SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	het	1	0	1	0	1	1
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	het	0	1	0	1	0	0
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	het	het	het	0	1	0	1	1
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	102	0		0			0	0		
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	0	1	0	0	1	0	1	0	0
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	1	1	0	1	0	1	1
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	1	1	0	1	0	1	1
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	102	0		0			0	0		
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi2epi-GTAt1	122	1	0	1	1	0	1	0	1	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	1	0	1	1	0	1	0	1	1
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0		0			1	0		
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	0	1	0	0		het	1	1	1
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0		0			0	0		
SNP	Chr3	3666840	A	G	Lfi2epi-GTAt3	122	0	1	1	0	het	0	1	0	1
SNP	Chr3	4740011	C	T	Lfi2epi-GTAt3	122	0	1	1	0	het	0	1	0	1
TE	Chr3	5595169	.	ATENSPM3	Lfi2epi-GTAt3	102	0		0			0	0		
Footprint	Chr3	5595169	C	CTTAA	Lfi2epi-GTAt3	122	0	0	0	0	0	0	1	0	0
Footprint	Chr3	5595169	C	CTTA	Lfi2epi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfi2epi-GTAt3	122	0	1	1	1	1	1	1	0	1
TE	Chr3	8759468	.	ATCOPIA78	Lfi2epi-GTAt3	102	0		0			0	0		
indel	Chr3	9589063	TA	T	Lfi2epi-GTAt3	122	0	0	0	0	0	0	0	1	0
SNP	Chr3	9858537	T	C	Lfi2epi-GTAt3, Rbiepi-GTAt3	122	0	0	0	1	0	0	0	1	0
SNP	Chr3	10202051	A	G	Lfi2epi-GTAt3, Rbiepi-GTAt3	122	1	1	1	0	1	1	1	0	1
SNP	Chr3	10743038	G	A	Lfi2epi-GTAt3, Rbiepi-GTAt3	122	0	0	0	1	0	0	0	1	0

SNP	Chr3	13092809	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	0	1	1	0	0	1
SNP	Chr3	14288480	A	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	0	0	1	0
SNP	Chr3	14509867	G	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	1	1	0	1
SNP	Chr3	14518339	C	G	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	1	1	0	1
SNP	Chr3	14530903	G	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	1	1	0	1
SNP	Chr3	14545591	C	A	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	1	1	0	1	1	1	0	1
SNP	Chr3	15861145	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0	0	0	1	0
SNP	Chr3	16205156	C	T	Lfi1epi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	0	0	0	1	0	0	0	1	0
SNP	Chr3	18112512	G	A	Lfi1epi-GTAt3	122	0	0	0	0	0	0	0	0	0
SNP	Chr3	18253358	C	G	Lfi1epi-GTAt3	122	0	0	0	0	0	1	0	0	0
SNP	Chr3	20306786	C	T	Lfi1epi-GTAt3	122	0	1	1	1	1	0	1	1	1
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	1	1	0	0	0	1
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	0	1	1	1	0
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	0	1	1	1	0
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	0	0	1	1	0	0	0	1
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	1	1	0	0		1	1	0
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	1	0	0	0	1
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	1	0	0	0	1
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	0	0	1	0	0	0
indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	0	0	1	1	1	0
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	1	1	0	0	0	1
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102			0				1	0	
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	1	1	0	0	0	1	1	1	0

SNP	Chr4	10291133	A	T	Pb2epi-At4	122	0	0	1	1	1	0	0	0	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	0	1	1	0	1	1	1	0	1
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0		0				0	0	
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	1	0	0	1	0	0	0	1	0
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	1	0	0	1	0	0	0	1	0
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	0	1	1	0	1	1	1	0	1
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	0	0	0	1	0	0	0	1	0
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	1	1	1	0	1	1	1	0	1
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	1	1	1	0	1	1	1	0	1
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	1	1	1	0	0	1	1	0	1
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	0	0	0	1	1	0	0	1	0
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	0	0	0	1	1	0	0	0	0
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	0	0	0	1	1	0	0	0	0
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	1	1	1	0	0	1	1	1	1
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	1	1	1	0	0	1	1	1	1

Variant Type	Chromosome	Position	Ref	Alt	QTL interval	Lines nb for variant detection	epi RIL 567	epi RIL 570	epi RIL 572	epi RIL 573	epi RIL 579
TE	Chr1	6617937	.	ATCOPIA78	Rbiepi-At1	102	0	0	1	1	0
SNP	Chr1	7241579	G	A	Rbiepi-At1	122	0	0	1	1	0
SNP	Chr1	9497509	G	A	Pb1epi-At1, Rbiepi-At1	122	1	0	0	het	1
SNP	Chr1	9883998	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	1	1	het	0
TE	Chr1	12177658	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0	0	0
TE	Chr1	12617099	.	ATENSPM3/HELITRON	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	102	0	0	0	0	0
Footprint	Chr1	12620247	C	CGTA	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1	122	0	0	0	0	0

SNP	Chr1	13033405	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1	122	1	0	0	0	het
indel	Chr1	13908828	CA	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	0	0	het
SNP	Chr1	13967107	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	1	1	het
SNP	Chr1	14023610	C	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	1	1	het
SNP	Chr1	14590272	T	C	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	0	0	het
SNP	Chr1	15509104	G	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	1	1	het
indel	Chr1	15935454	TGGTGAAG	T	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	0	0	het
indel	Chr1	15972511	AT	A	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	1	1	het
TE	Chr1	16836944	.	ATENSPM3	Pb1epi-At1, Rbiepi-At1, Lfi1epi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0	0
indel	Chr1	17863273	TC	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	1	0	0	0	het
SNP	Chr1	18002233	T	C	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	1	1	het
SNP	Chr1	18173880	C	T	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	1	1	het
TE	Chr1	20150892	.	ATENSPM3/HELITRON	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	102	0	0	0	0	0
SNP	Chr1	20362809	G	A	Rbiepi-At1, Lfi2epi-At1, Lfi1epi-GTAt1	122	0	1	1	1	1
SNP	Chr1	21675141	G	A	Rbiepi-At1	122	0	1	1	1	1
TE	Chr2	10506856	.	ATENSPM3/ATLANTYS	Rbiepi-At2a	106	0	0	1	0	0
SNP	Chr2	11434437	A	C	Rbiepi-At2a	122	0	0	1	1	0
TE	Chr2	12762111	.	ATENSPM3	Rbiepi-At2b	102	0	0	0	0	0
SNP	Chr3	3666840	A	G	Lfi1epi-GTAt3	122	0	0	1	1	1
SNP	Chr3	4740011	C	T	Lfi1epi-GTAt3	122	0	0	1	1	1
TE	Chr3	5595169	.	ATENSPM3	Lfi1epi-GTAt3	102	0	0	1	0	0
Footprint	Chr3	5595169	C	CTTAA	Lfi1epi-GTAt3	122	0	0	0	0	0

Footprint	Chr3	5595169	C	CTTA	Lfiempi-GTAt3	122	0	0	0	0	0
SNP	Chr3	8547929	G	A	Lfiempi-GTAt3	122	1	0	1	1	1
TE	Chr3	8759468	.	ATCOPIA78	Lfiempi-GTAt3	102	0	0	0	0	0
indel	Chr3	9589063	TA	T	Lfiempi-GTAt3	122	0	1	1	0	0
SNP	Chr3	9858537	T	C	Lfiempi-GTAt3, Rbiepi-GTAt3	122	0	1	1	0	0
SNP	Chr3	10202051	A	G	Lfiempi-GTAt3, Rbiepi-GTAt3	122	1	0	0	1	1
SNP	Chr3	10743038	G	A	Lfiempi-GTAt3, Rbiepi-GTAt3	122	0	1	1	0	0
SNP	Chr3	13092809	C	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	0	0	1	0
SNP	Chr3	14288480	A	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	0	0
SNP	Chr3	14509867	G	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	1
SNP	Chr3	14518339	C	G	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	1
SNP	Chr3	14530903	G	A	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	1
SNP	Chr3	14545591	C	A	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	1	0	0	1	1
SNP	Chr3	15861145	C	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3, Lfi1epi-At3	122	0	1	1	0	0
SNP	Chr3	16205156	C	T	Lfiempi-GTAt3, Rbiepi-GTAt3, Lfi2epi-At3	122	1	1	1	0	0
SNP	Chr3	18112512	G	A	Lfiempi-GTAt3	122	1	1	0	0	1
SNP	Chr3	18253358	C	G	Lfiempi-GTAt3	122	1	1	0	0	1
SNP	Chr3	20306786	C	T	Lfiempi-GTAt3	122	0	0	1	1	0
SNP	Chr4	2834559	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	1	0
SNP	Chr4	2920544	G	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	0	0	0	1
SNP	Chr4	3318778	C	G	Pb1epi-At4, Pbepi-GTAt4a	122	1	0	0	0	1
SNP	Chr4	4474111	ACCAGAT	A	Pb1epi-At4, Pbepi-GTAt4a	122	0	1	1	1	0
SNP	Chr4	5115842	T	A	Pb1epi-At4, Pbepi-GTAt4a	122	1	0	0	0	1
indel	Chr4	7214677	C	CATTATTTGT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	1	0
SNP	Chr4	7344157	C	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	1	0
indel	Chr4	7652616	AC	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	1

indel	Chr4	7750735	G	GT	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	0	0	0	0	1
SNP	Chr4	7944780	G	A	Rbiepi-At4, Pb2epi-At4, Pbepi-GTAt4b, Pbepi-GTAt4a	122	1	1	1	1	0
TE	Chr4	9650119	.	ATCOPIA93	Pb2epi-At4	102	0	0	0	0	0
SNP	Chr4	10012411	A	G	Pb2epi-At4	122	0	0	0	0	0
SNP	Chr4	10291133	A	T	Pb2epi-At4	122	1	0	1	1	1
SNP	Chr5	4896800	G	A	Lfi1epi-At5	122	0	0	1	1	0
TE	Chr5	6357944	.	ATCOPIA78	Lfi1epi-At5	102	0	1	1	0	0
SNP	Chr5	7980674	A	T	Lfi1epi-At5	122	0	1	1	0	1
SNP	Chr5	8654576	G	C	Lfi1epi-At5	122	0	1	1	0	1
indel	Chr5	8872537	TG	T	Lfi1epi-At5	122	1	0	0	1	0
SNP	Chr5	9381623	C	A	Lfi1epi-At5	122	0	1	1	0	1
SNP	Chr5	10075739	G	A	Lfi1epi-At5	122	1	0	0	1	0
SNP	Chr5	10214637	G	A	Lfi1epi-At5	122	1	0	0	1	0
SNP	Chr5	11645737	T	C	Lfi1epi-At5	122	1	0	0	1	0
SNP	Chr5	12506158	G	A	Lfi1epi-At5	122	0	1	1	0	1
SNP	Chr5	14215841	A	T	Lfi1epi-At5	122	0	1	1	0	1
SNP	Chr5	14240551	T	A	Lfi1epi-At5	122	0	1	1	0	1
SNP	Chr5	15070083	T	C	Lfi1epi-At5	122	1	0	0	1	0
indel	Chr5	15431687	G	GT	Lfi1epi-At5	122	1	0	0	1	0

File S2. Co-localization of QTL^{epi} confidence intervals with candidate SNP identified by GWAS after clubroot infection. For each QTL^{epi} and QTL^{epi}GT identified the set of significant SNP found in the physical confidence interval were identified. For each SNP, their PERL identification number, chromosomal position, physical position and the trait used for the detection are shown (Jubault *et al.*, submitted).

QTL ^{epi}					Significant SNP GWAS			
Name	Trait	Chr	Confidence interval (pb)		Name	Trait	Chr	Position
Rbi ^{epi} - At1	Root biomass	1	6010670	21754627	PERL0044518	Pathogen DNA quantity	1	6034610
					PERL0056460	Difference leaf length		7726580
					PERL0056466	Difference leaf length		7727158
					PERL0057771	Difference leaf length		7877982
					PERL0057917	Difference leaf length		7895107
					PERL0058893	Difference leaf length		7976793
					PERL0092952	Galle surface		11515504
					PERL0092952	Disease index		11515504
					PERL0092952	Pathogen DNA quantity		11515504
					PERL0108633	Disease index		12811053
					PERL0160323	Disease index		18678528
					PERL0160353	Disease index		18684492
					PERL0195496	Pathogen DNA quantity		22582105
					PERL0203304	Disease index		23186724
					PERL0206544	Galle surface		23445111
PERL0206559	Galle surface	23445466						
PERL0206562	Galle surface	23445623						
Rbi ^{epi} - At2a	Root biomass	2	10533516	12454436	-	-	-	-
Rbi ^{epi} - At2b	Root biomass	2	12449489	15411897	PERL0376265	Difference leaf length	2	12924267
					PERL0384487	Galle surface		14483203
Rbi ^{epi} - At4	Root biomass	4	6723450	8907282	PERL0750950	Disease index	4	7958405
					PERL0756484	Difference leaf length		8292785
					PERL0756549	Difference leaf length		8296004
					PERL0756772	Difference leaf length		8302947
					PERL0756785	Difference leaf length		8303556
					PERL0760746	Difference leaf length		8562940
Lfi1 ^{epi} - At1	Leaf length	1	8931527	17366043	PERL0092952	Galle surface	1	11515504
					PERL0092952	Disease index		11515504
					PERL0092952	Pathogen DNA quantity		11515504
					PERL0108633	Disease index		12811053
Lfi1 ^{epi} - At3	Leaf length	3	11390669	16180190	PERL0525412	Difference leaf length	3	11502036
					PERL0526946	Difference leaf length		11584549
					PERL0530644	Pathogen DNA quantity		11774998
Lfi2 ^{epi} - At1	Leaf length	1	13003203	21455680	PERL0160323	Disease index	1	18678528
					PERL0160353	Disease index		18684492
Lfi2 ^{epi} - At3	Leaf length	3	11390669	16814166	PERL0427225	Difference leaf length	3	1518618
					PERL0448258	Disease index		4328701
					PERL0484841	Pathogen DNA quantity		8591029
					PERL0496867	Disease index		9503591
					PERL0525412	Difference leaf length		11502036
					PERL0526946	Difference leaf length		11584549
					PERL0530644	Pathogen DNA quantity		11774998
					PERL0595278	Pathogen DNA quantity		16280599
Pb1 ^{epi} - At1	Pathogen quantification	1	8931527	17523612	PERL0092952	Galle surface	1	11515504
					PERL0092952	Disease index		11515504
					PERL0092952	Pathogen DNA quantity		11515504
					PERL0108633	Disease index		12811053
Pb1 ^{epi} - At4	Pathogen quantification	4	788271	5470844	-	-	-	-
Pb2 ^{epi} - At4	Pathogen quantification	4	6723450	10529079	PERL0750950	Disease index	4	7958405
					PERL0756484	Difference leaf length		8292785
					PERL0756549	Difference leaf length		8296004
					PERL0756772	Difference leaf length		8302947

					PERL0756785	Difference leaf length		8303556
					PERL0760746	Difference leaf length		8562940
					PERL0770540	Pathogen DNA quantity		9280030
					PERL0775518	Galle surface		9554698
					PERL0776438	Galle surface		9593401
					PERL0782476	Pathogen DNA quantity		10040366
					PERL0782495	Pathogen DNA quantity		10042102
					PERL0782496	Pathogen DNA quantity		10042299
					PERL0782504	Pathogen DNA quantity		10042711
					PERL0782510	Pathogen DNA quantity		10043069
					PERL0782514	Pathogen DNA quantity		10043142
					PERL0782534	Pathogen DNA quantity		10045289
Rbiepi-GTAt3	Root biomass	3	9692187	16814166	PERL0525412	Difference leaf length		11502036
					PERL0526946	Difference leaf length	3	11584549
					PERL0530644	Pathogen DNA quantity		11774998
					PERL0595278	Pathogen DNA quantity		16280599
Pb ^{epi} -GTAt4a	Pathogen quantification	4	1312610	8319386	PERL0704522	Difference leaf length		4605597
					PERL0750950	Difference leaf length		7958405
					PERL0756484	Difference leaf length	4	8292785
					PERL0756549	Difference leaf length		8296004
					PERL0756772	Difference leaf length		8302947
					PERL0756785	Difference leaf length		8303556
Pb ^{epi} -GTAt4b	Pathogen quantification	4	6723450	9485389	PERL0750950	Disease index		7958405
					PERL0756484	Difference leaf length		8292785
					PERL0756549	Difference leaf length		8296004
					PERL0756772	Difference leaf length	4	8302947
					PERL0756785	Difference leaf length		8303556
					PERL0760746	Difference leaf length		8562940
					PERL0770540	Pathogen DNA quantity		9280030

File S3. Statistic analysis of TE and sequence variant effect compared to methylated markers on QTL^{epi} detection for each trait study. Trait indicates the phenotypic trait used for the analysis. Growth room gives the number of the growth room where the phenotypic data were obtained. QTL type indicates if the effect tested includes (GT) or not (G) an interaction between the epigenotype and the temperature. Model I indicate the marker considered in the model (Marker, SNP, Indel, Footprint or the combination of the Marker and the sequence variant). Model II describes the model used in R regression (lm function) with the peak markers used and/or the sequence variants considered in the model. n: indicates the number of epi-lines with phenotypic data for the trait considered. *p-value* β_x were the p-values for each marker, TE or sequence variant used in the associated model. F-statistic Model II give the F-statistic of each model. *p-value* Model II indicates p-values for each model. Adjusted R² gives the adjusted R² for each model. Cells in red indicate the SNP variants that showed a significant impact greater than that brought by the epigenetic marker at the peak of the QTL.

Trait	Growth room	QTL type	n	Model I	Model II	<i>p-value</i> β_1	<i>p-value</i> β_2	<i>p-value</i> β_3	<i>p-value</i> β_4	<i>p-value</i> β_5	F-statistic Model II	<i>p-value</i> Model II	Adjusted R ²
Lfi	1	G	1 2 3	Marker	$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	8.24 E-08	1.30 E-07	5.43 E-05			30.88	7.78E-15	0.44
				SNP	$y_i = \beta_0 + \beta_1g(\text{SNP_9883998})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	3.21 E-03	1.87 E-06	1.90 E-04			20.59	1.16E-10	0.36
					$y_i = \beta_0 + \beta_1g(\text{SNP_13033405})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	9.30 E-05	9.31 E-07	2.39 E-04			21.51	4.89E-11	0.37
					$y_i = \beta_0 + \beta_1g(\text{SNP_13967107})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	4.19 E-05	3.81 E-07	2.71 E-04			23.66	6.51E-12	0.39
					$y_i = \beta_0 + \beta_1g(\text{SNP_14023610})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	4.15 E-05	3.26 E-07	2.25 E-04			23.88	5.06E-12	0.39
					$y_i = \beta_0 + \beta_1g(\text{SNP_14590272})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.20 E-05	2.46 E-07	1.65 E-04			26.06	7.77E-13	0.41
					$y_i = \beta_0 + \beta_1g(\text{SNP_15509104})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	2.33 E-05	2.21 E-07	6.48 E-04			23.75	5.41E-12	0.39
					$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_13092809})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	3.29 E-07	1.47 E-06	1.49 E-04			27.41	2.03E-13	0.42
					$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_14288480})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.69 E-05	3.73 E-04	6.96 E-05			22.15	2.16E-11	0.37
					$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_14509867})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.69 E-05	3.73 E-04	6.96 E-05			22.15	2.16E-11	0.37
					$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_14518339})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.69 E-05	3.73 E-04	6.96 E-05			22.15	2.16E-11	0.37
					$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_14530903})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.69 E-05	3.73 E-04	6.96 E-05			22.15	2.16E-11	0.37
					$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_14545591})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.64 E-05	4.98 E-04	8.68 E-05			22.00	2.59E-11	0.37

			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_15861145})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	9.60 E-06	8.63 E-05	1.90 E-04			23.63	4.99E-12	0.38
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_4896800})_i + \epsilon_i$	1.35 E-06	9.57 E-08	2.41 E-02			21.65	4.29E-11	0.37
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_7980674})_i + \epsilon_i$	2.24 E-06	3.74 E-06	2.72 E-04			22.80	1.82E-11	0.39
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_8654576})_i + \epsilon_i$	6.59 E-07	6.28 E-07	3.39 E-05			27.93	1.66E-13	0.43
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_9381623})_i + \epsilon_i$	4.38 E-07	4.76 E-07	8.80 E-04			25.98	7.44E-13	0.41
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_10075739})_i + \epsilon_i$	9.50 E-07	1.10 E-07	2.86 E-04			27.28	2.27E-13	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_10214637})_i + \epsilon_i$	6.00 E-07	1.03 E-07	2.41 E-04			27.88	1.27E-13	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_11645737})_i + \epsilon_i$	7.16 E-07	5.71 E-08	4.44 E-04			27.22	2.24E-13	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_12506158})_i + \epsilon_i$	7.16 E-07	5.71 E-08	4.44 E-04			27.22	2.24E-13	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_14215841})_i + \epsilon_i$	7.54 E-07	3.92 E-08	2.17 E-04			27.90	1.25E-13	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_14240551})_i + \epsilon_i$	3.92 E-07	3.02 E-08	1.76 E-04			28.85	4.90E-14	0.43
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{SNP_15070083})_i + \epsilon_i$	7.22 E-07	3.16 E-08	9.89 E-05			30.05	2.02E-14	0.44
1	Indel		$y_i = \beta_0 + \beta_1g(\text{indel_13908828})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	4.42 E-05	3.16 E-07	3.65 E-04			23.48	7.69E-12	0.39
2			$y_i = \beta_0 + \beta_1g(\text{indel_15935454})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.50 E-04	8.37 E-07	5.17 E-04			21.86	3.09E-11	0.37
3			$y_i = \beta_0 + \beta_1g(\text{indel_15972511})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.74 E-04	2.00 E-06	1.69 E-04			20.77	9.41E-11	0.36
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{Indel_8872537})_i + \epsilon_i$	9.97 E-07	1.97 E-06	6.47 E-05			28.09	1.28E-13	0.43
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{Indel_15431687})_i + \epsilon_i$	3.56 E-07	1.58 E-08	4.17 E-04			29.61	3.13E-14	0.44
1	TE		$y_i = \beta_0 + \beta_1g(\text{TE_12177658})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	5.08 E-02	4.14 E-05	1.40 E-04			13.67	1.69E-07	0.30
2			$y_i = \beta_0 + \beta_1g(\text{TE_12617099})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	1.21 E-01	2.42 E-05	6.34 E-05			14.73	6.14E-08	0.32
3			$y_i = \beta_0 + \beta_1g(\text{TE_16836944})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	3.82 E-02	1.26 E-04	1.31 E-04			14.30	8.42E-08	0.30
			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{TE_6357944})_i + \epsilon_i$	3.95 E-06	5.45 E-06	2.21 E-03			20.31	2.57E-10	0.38
1	Footprint		$y_i = \beta_0 + \beta_1g(\text{Footprint_12620247})_i + \beta_2g(\text{MM427})_i + \beta_3g(\text{MM728})_i + \epsilon_i$	7.06 E-02	5.41 E-06	8.36 E-06			17.34	2.16E-09	0.31
1	Marker + SNP		$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_9883998})_i + \beta_3g(\text{MM427})_i + \beta_4g(\text{MM728})_i + \epsilon_i$	9.58 E-05	3.23 E-01	5.44 E-07	1.45 E-04		21.71	3.39E-13	0.44
2			$y_i = \beta_0 + \beta_1g(\text{MM52})_i + \beta_2g(\text{SNP_13033405})_i + \beta_3g(\text{MM427})_i + \beta_4g(\text{MM728})_i + \epsilon_i$	4.49 E-04	1.10 E-01	3.77 E-07	3.01 E-04		21.18	6.06E-13	0.44

			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{SNP_13967107})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	4.70 E-04	1.82 E-01	1.18 E-07	3.06 E-04		22.91	8.68E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{SNP_14023610})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	5.77 E-04	1.82 E-01	1.00 E-07	2.29 E-04		22.90	8.19E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{SNP_14590272})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.40 E-02	9.11 E-01	3.01 E-07	8.35 E-05		21.33	4.82E-13	0.44
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{SNP_15509104})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	5.17 E-04	1.36 E-01	7.19 E-08	4.86 E-04		22.87	7.90E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{SNP_13092809})i + \beta_4g(\text{MM728})i + \epsilon_i$	1.24 E-07	2.50 E-02	3.20 E-01	7.62 E-05		22.59	9.92E-14	0.44
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{SNP_14288480})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.77 E-07	2.97 E-05	9.89 E-02	1.27 E-04		23.93	2.19E-14	0.46
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{SNP_14509867})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.77 E-07	2.97 E-05	9.89 E-02	1.27 E-04		23.93	2.19E-14	0.46
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{SNP_14518339})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.77 E-07	2.97 E-05	9.89 E-02	1.27 E-04		23.93	2.19E-14	0.46
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{SNP_14530903})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.77 E-07	2.97 E-05	9.89 E-02	1.27 E-04		23.93	2.19E-14	0.46
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{SNP_14545591})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.51 E-07	2.97 E-05	1.16 E-01	1.61 E-04		23.84	2.61E-14	0.46
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{SNP_15861145})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.51 E-07	4.71 E-05	2.55 E-02	2.15 E-04		24.75	7.92E-15	0.46
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_4896800})i + \epsilon_i$	1.31 E-06	7.64 E-08	1.84 E-04	3.00 E-01		22.05	2.35E-13	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_7980674})i + \epsilon_i$	3.04 E-06	6.78 E-06	1.49 E-02	4.86 E-02		19.46	5.27E-12	0.43
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_8654576})i + \epsilon_i$	7.05 E-07	8.36 E-07	5.66 E-02	1.49 E-02		22.39	1.63E-13	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_9381623})i + \epsilon_i$	4.35 E-07	7.30 E-07	1.41 E-02	1.68 E-01		21.94	2.18E-13	0.44
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_10075739})i + \epsilon_i$	6.82 E-07	2.62 E-07	2.29 E-02	8.76 E-02		22.56	1.03E-13	0.44
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_10214637})i + \epsilon_i$	4.80 E-07	2.33 E-07	2.08 E-02	8.64 E-02		23.10	5.36E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_11645737})i + \epsilon_i$	5.27 E-07	1.83 E-07	1.35 E-02	1.09 E-01		22.93	6.45E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_12506158})i + \epsilon_i$	5.27 E-07	1.83 E-07	1.35 E-02	1.09 E-01		22.93	6.45E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_14215841})i + \epsilon_i$	5.41 E-07	1.73 E-07	1.71 E-02	5.51 E-02		23.27	4.43E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_14240551})i + \epsilon_i$	3.74 E-07	1.17 E-07	1.66 E-02	5.30 E-02		24.02	1.72E-14	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{SNP_15070083})i + \epsilon_i$	6.66 E-07	1.09 E-07	2.52 E-02	3.29 E-02		24.64	1.04E-14	0.46
1	Marker + Indel		$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{indel_13908828})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	8.90 E-04	1.85 E-01	9.95 E-08	3.40 E-04		22.22	1.83E-13	0.45
2 3			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{indel_15935454})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	1.25 E-04	2.51 E-01	1.16 E-07	4.06 E-04		22.50	1.18E-13	0.45

			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{indel_15972511})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	6.62 E-04	1.67 E-01	3.03 E-07	2.30 E-04		20.23	1.64E-12	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{Indel_8872537})i + \epsilon_i$	1.02 E-06	1.75 E-06	4.97 E-02	2.78 E-02		22.60	1.13E-13	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM427})i + \beta_3g(\text{MM728})i + \beta_4g(\text{Indel_15431687})i + \epsilon_i$	2.30 E-07	8.86 E-08	1.71 E-02	1.05 E-01		24.62	1.14E-14	0.47
1 2 3	Marker + TE		$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{TE_12177658})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	2.81 E-05	8.54 E-01	1.52 E-05	1.77 E-04		17.06	1.42E-10	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{TE_12617099})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	1.06 E-04	8.47 E-01	9.91 E-06	8.23 E-05		16.96	1.78E-10	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{TE_16836944})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	3.18 E-05	6.73 E-01	1.08 E-05	1.52 E-04		17.46	8.01E-11	0.42
			$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{MM52})i + \beta_3g(\text{MM427})i + \beta_4g(\text{TE_6357944})i + \epsilon_i$	7.30 E-06	5.18 E-06	5.69 E-04	1.16 E-02		20.23	3.86E-12	0.45
1 2 3	Marker + Footprint		$y_i = \beta_0 + \beta_1g(\text{MM52})i + \beta_2g(\text{Footprint_12620247})i + \beta_3g(\text{MM427})i + \beta_4g(\text{MM728})i + \epsilon_i$	6.53 E-07	6.87 E-01	1.81 E-07	5.44 E-05		22.87	5.64E-14	0.44
1 2 3	Marker		$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{MM515})i + \epsilon_i$	2.64 E-06	2.12 E-05				20.76	1.81E-08	0.26
2	1 2 3	SNP	$y_i = \beta_0 + \beta_1g(\text{SNP_13033405})i + \beta_2g(\text{MM515})i + \epsilon_i$	1.74 E-04	2.37 E-05				16.02	7.70E-07	0.22
			$y_i = \beta_0 + \beta_1g(\text{SNP_13967107})i + \beta_2g(\text{MM515})i + \epsilon_i$	6.25 E-05	3.45 E-05				17.06	3.41E-07	0.23
			$y_i = \beta_0 + \beta_1g(\text{SNP_14023610})i + \beta_2g(\text{MM515})i + \epsilon_i$	4.43 E-05	3.75 E-05				17.25	2.90E-07	0.23
			$y_i = \beta_0 + \beta_1g(\text{SNP_14590272})i + \beta_2g(\text{MM515})i + \epsilon_i$	2.75 E-05	6.70 E-06				19.80	4.32E-08	0.26
			$y_i = \beta_0 + \beta_1g(\text{SNP_15509104})i + \beta_2g(\text{MM515})i + \epsilon_i$	5.08 E-05	6.48 E-05				16.99	3.47E-07	0.23
			$y_i = \beta_0 + \beta_1g(\text{SNP_18002233})i + \beta_2g(\text{MM515})i + \epsilon_i$	1.86 E-05	2.33 E-04				17.81	1.90E-07	0.24
			$y_i = \beta_0 + \beta_1g(\text{SNP_18173880})i + \beta_2g(\text{MM515})i + \epsilon_i$	6.30 E-05	4.01 E-04				15.81	8.83E-07	0.22
			$y_i = \beta_0 + \beta_1g(\text{SNP_20362809})i + \beta_2g(\text{MM515})i + \epsilon_i$	1.13 E-04	4.02 E-04				15.43	1.19E-06	0.21
			$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{SNP_13092809})i + \epsilon_i$	1.24 E-05	3.76 E-04				16.74	4.15E-07	0.23
			$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{SNP_14288480})i + \epsilon_i$	1.55 E-04	4.10 E-02				11.45	2.90E-05	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{SNP_14509867})i + \epsilon_i$	1.55 E-04	4.10 E-02				11.45	2.90E-05	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{SNP_14518339})i + \epsilon_i$	1.55 E-04	4.10 E-02				11.45	2.90E-05	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{SNP_14530903})i + \epsilon_i$	1.55 E-04	4.10 E-02				11.45	2.90E-05	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{SNP_14545591})i + \epsilon_i$	2.64 E-04	3.70 E-02				11.13	3.82E-05	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM91})i + \beta_2g(\text{SNP_15861145})i + \epsilon_i$	1.61 E-04	2.85 E-03				14.23	2.90E-06	0.19

		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_16205156})_i + \epsilon_i$	1.87 E-04	3.61 E-03				13.49	5.48E-06	0.19
1 2 3	Indel	$y_i = \beta_0 + \beta_1g(\text{indel_13908828})_i + \beta_2g(\text{MM515})_i + \epsilon_i$	2.37 E-05	4.50 E-05				17.80	1.94E-07	0.24
		$y_i = \beta_0 + \beta_1g(\text{indel_15935454})_i + \beta_2g(\text{MM515})_i + \epsilon_i$	9.35 E-05	1.22 E-04				16.35	5.74E-07	0.22
		$y_i = \beta_0 + \beta_1g(\text{indel_15972511})_i + \beta_2g(\text{MM515})_i + \epsilon_i$	1.63 E-04	3.99 E-04				14.19	3.21E-06	0.20
		$y_i = \beta_0 + \beta_1g(\text{indel_17863273})_i + \beta_2g(\text{MM515})_i + \epsilon_i$	7.70 E-05	1.97 E-04				16.60	4.85E-07	0.23
1 2 3	TE	$y_i = \beta_0 + \beta_1g(\text{TE_16836944})_i + \beta_2g(\text{MM515})_i + \epsilon_i$	1.73 E-03	4.98 E-03				9.76	1.35E-04	0.16
		$y_i = \beta_0 + \beta_1g(\text{TE_20150892})_i + \beta_2g(\text{MM515})_i + \epsilon_i$	7.02 E-03	5.18 E-03				8.25	4.85E-04	0.14
1 2 3	Marker + SNP	$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_13033405})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	2.62 E-03	7.58 E-02	1.04 E-05			14.66	4.26E-08	0.29
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_13967107})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	4.40 E-03	8.84 E-02	1.45 E-05			14.95	3.08E-08	0.29
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_14023610})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	3.42 E-03	8.73 E-02	1.43 E-05			15.29	2.09E-08	0.29
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_14590272})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	1.04 E-01	6.66 E-01	5.62 E-06			14.29	6.11E-08	0.28
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_15509104})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	2.24 E-03	1.00 E-01	1.70 E-05			15.46	1.71E-08	0.29
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_18002233})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	2.05 E-03	1.95 E-02	5.61 E-05			16.13	8.77E-09	0.30
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_18173880})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	2.47 E-03	3.68 E-02	1.06 E-04			14.54	4.61E-08	0.28
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{SNP_20362809})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	1.29 E-03	3.84 E-02	1.02 E-04			14.82	3.42E-08	0.28
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_13092809})_i + \epsilon_i$	7.37 E-06	5.19 E-02	7.41 E-01			12.73	3.11E-07	0.25
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_14288480})_i + \epsilon_i$	1.33 E-05	4.29 E-04	7.55 E-01			12.82	2.77E-07	0.25
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_14509867})_i + \epsilon_i$	1.33 E-05	4.29 E-04	7.55 E-01			12.82	2.77E-07	0.25
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_14518339})_i + \epsilon_i$	1.33 E-05	4.29 E-04	7.55 E-01			12.82	2.77E-07	0.25
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_14530903})_i + \epsilon_i$	1.33 E-05	4.29 E-04	7.55 E-01			12.82	2.77E-07	0.25
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_14545591})_i + \epsilon_i$	2.46 E-05	4.62 E-04	7.13 E-01			12.53	3.87E-07	0.25
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_15861145})_i + \epsilon_i$	1.69 E-05	1.85 E-03	2.48 E-01			13.61	1.12E-07	0.26
		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{MM515})_i + \beta_3g(\text{SNP_16205156})_i + \epsilon_i$	2.56 E-05	1.85 E-03	2.40 E-01			13.09	2.08E-07	0.26
1 2 3	Marker + Indel	$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{indel_13908828})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	6.24 E-03	5.87 E-02	1.81 E-05			15.18	2.42E-08	0.29

			$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{indel_15935454})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	1.61 E-03	9.44 E-02	2.32 E-05			15.28	2.06E-08	0.29
			$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{indel_15972511})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	2.38 E-03	9.99 E-02	7.38 E-05			13.42	1.56E-07	0.27
			$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{indel_17863273})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	6.59 E-04	6.46 E-02	3.58 E-05			16.28	7.76E-09	0.31
1	Marker + TE		$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{TE_16836944})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	4.70 E-03	7.27 E-02	1.93 E-03			9.79	1.05E-05	0.23
2			$y_i = \beta_0 + \beta_1g(\text{MM91})_i + \beta_2g(\text{TE_20150892})_i + \beta_3g(\text{MM515})_i + \epsilon_i$	1.53 E-03	9.72 E-02	1.82 E-03			9.58	1.32E-05	0.23
3	Marker		$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	3.28 E-04	6.71 E-01				6.90	1.46E-03	0.10
	SNP		$y_i = \beta_0 + \beta_1g(\text{SNP_13967107})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	5.29 E-04	2.81 E-01				6.42	2.29E-03	0.10
			$y_i = \beta_0 + \beta_1g(\text{SNP_14023610})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	1.47 E-03	2.53 E-01				5.42	5.66E-03	0.09
			$y_i = \beta_0 + \beta_1g(\text{SNP_14590272})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	2.60 E-02	4.80 E-01				2.78	6.66E-02	0.05
			$y_i = \beta_0 + \beta_1g(\text{SNP_15509104})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	9.25 E-04	2.93 E-01				5.87	3.75E-03	0.09
			$y_i = \beta_0 + \beta_1g(\text{SNP_18002233})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	1.78 E-05	3.42 E-01				10.17	8.68E-05	0.15
			$y_i = \beta_0 + \beta_1g(\text{SNP_18173880})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	1.87 E-04	4.99 E-01				7.51	8.69E-04	0.12
			$y_i = \beta_0 + \beta_1g(\text{SNP_20362809})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	4.07 E-04	5.46 E-01				6.72	1.75E-03	0.11
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_20362809})_i + \epsilon_i$	1.94 E-04	3.25 E-01				8.31	4.32E-04	0.13
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_4740011})_i + \epsilon_i$	3.66 E-04	2.32 E-02				9.88	1.13E-04	0.15
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_8547929})_i + \epsilon_i$	2.19 E-03	1.63 E-05				17.23	2.84E-07	0.23
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_9858537})_i + \epsilon_i$	1.80 E-03	8.61 E-04				12.79	9.57E-06	0.18
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_10202051})_i + \epsilon_i$	1.75 E-03	1.49 E-03				12.10	1.70E-05	0.17
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_10743038})_i + \epsilon_i$	9.69 E-04	6.33 E-02				8.43	3.79E-04	0.13
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_13092809})_i + \epsilon_i$	6.37 E-04	2.95 E-03				11.24	3.49E-05	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_14288480})_i + \epsilon_i$	4.45 E-03	6.63 E-02				8.08	5.17E-04	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_14509867})_i + \epsilon_i$	4.45 E-03	6.63 E-02				8.08	5.17E-04	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_14518339})_i + \epsilon_i$	4.45 E-03	6.63 E-02				8.08	5.17E-04	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_14530903})_i + \epsilon_i$	4.45 E-03	6.63 E-02				8.08	5.17E-04	0.12

Gx
T1
2
3

SNP

			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_14545591})_i + \epsilon_i$	4.82 E-03	9.23 E-02				7.51	8.60E-04	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_15861145})_i + \epsilon_i$	5.00 E-03	1.05 E-02				10.27	7.73E-05	0.15
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_16205156})_i + \epsilon_i$	4.33 E-03	3.09 E-02				8.61	3.29E-04	0.13
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_18112512})_i + \epsilon_i$	2.95 E-03	3.99 E-03				10.78	5.06E-05	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_18253358})_i + \epsilon_i$	3.22 E-03	4.94 E-03				10.55	6.15E-05	0.15
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_20306786})_i + \epsilon_i$	4.39 E-03	1.12 E-02				9.68	1.29E-04	0.14
1	Indel		$y_i = \beta_0 + \beta_1g(\text{indel_13908828})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	1.31 E-03	2.64 E-01				5.52	5.15E-03	0.09
2			$y_i = \beta_0 + \beta_1g(\text{indel_15935454})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	7.85 E-04	3.42 E-01				6.05	3.20E-03	0.10
3			$y_i = \beta_0 + \beta_1g(\text{indel_15972511})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	1.45 E-03	4.45 E-01				5.37	5.94E-03	0.09
			$y_i = \beta_0 + \beta_1g(\text{indel_17863273})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	2.36 E-04	4.37 E-01				7.32	1.03E-03	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{indel_9589063})_i + \epsilon_i$	1.71 E-03	1.99 E-03				11.52	2.78E-05	0.17
1	TE		$y_i = \beta_0 + \beta_1g(\text{TE_16836944})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	5.52 E-03	6.14 E-01				4.26	1.68E-02	0.08
2			$y_i = \beta_0 + \beta_1g(\text{TE_20150892})_i + \beta_2g(\text{MM547})_i + \epsilon_i$	8.03 E-02	5.06 E-01				1.79	1.73E-01	0.03
3			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{TE_5595169})_i + \epsilon_i$	2.78 E-03	4.03 E-01				4.90	9.37E-03	0.09
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{TE_8759468})_i + \epsilon_i$	2.55 E-03	5.57 E-01				5.25	6.80E-03	0.10
1	Footprint		$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{Footprint_5595169})_i + \epsilon_i$	4.48 E-04	8.20 E-01				6.54	2.04E-03	0.10
2			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{Footprint_5595169})_i + \epsilon_i$	4.48 E-04	8.20 E-01				6.54	2.04E-03	0.10
1	Marker + SNP		$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_13967107})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	2.83 E-02	2.70 E-02	3.75 E-01			6.08	7.22E-04	0.14
2			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_14023610})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	1.59 E-02	5.99 E-02	3.52 E-01			5.77	1.05E-03	0.13
3			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_14590272})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	1.79 E-02	7.56 E-01	5.14 E-01			3.86	1.14E-02	0.09
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_15509104})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	1.72 E-02	5.86 E-02	3.84 E-01			6.03	7.61E-04	0.14
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_18002233})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	1.05 E-01	5.69 E-03	3.74 E-01			7.77	9.25E-05	0.17
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_18173880})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	1.21 E-01	2.10 E-02	5.39 E-01			5.88	9.12E-04	0.14
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{SNP_20362809})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	5.20 E-02	3.71 E-02	5.56 E-01			5.88	9.20E-04	0.14

			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_3666840})_i + \epsilon_i$	1.81 E-04	4.30 E-01	3.20 E-01			5.73	1.11E-03	0.14
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_4740011})_i + \epsilon_i$	3.47 E-04	4.43 E-01	2.37 E-02			6.76	3.17E-04	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_8547929})_i + \epsilon_i$	2.20 E-03	6.07 E-01	1.79 E-05			11.50	1.20E-06	0.23
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_9858537})_i + \epsilon_i$	1.81 E-03	4.85 E-01	8.42 E-04			8.65	3.17E-05	0.18
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_10202051})_i + \epsilon_i$	1.80 E-03	4.47 E-01	1.37 E-03			8.23	5.24E-05	0.18
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_10743038})_i + \epsilon_i$	9.88 E-04	5.18 E-01	5.72 E-02			5.73	1.08E-03	0.13
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_13092809})_i + \epsilon_i$	6.93 E-04	9.02 E-01	3.42 E-03			7.43	1.37E-04	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_14288480})_i + \epsilon_i$	4.37 E-03	5.82 E-01	6.40 E-02			5.46	1.52E-03	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_14509867})_i + \epsilon_i$	4.37 E-03	5.82 E-01	6.40 E-02			5.46	1.52E-03	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_14518339})_i + \epsilon_i$	4.37 E-03	5.82 E-01	6.40 E-02			5.46	1.52E-03	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_14530903})_i + \epsilon_i$	4.37 E-03	5.82 E-01	6.40 E-02			5.46	1.52E-03	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_14545591})_i + \epsilon_i$	4.79 E-03	6.65 E-01	8.99 E-02			5.03	2.59E-03	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_15861145})_i + \epsilon_i$	5.04 E-03	6.04 E-01	1.07 E-02			6.89	2.58E-04	0.15
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_16205156})_i + \epsilon_i$	4.42 E-03	7.58 E-01	3.17 E-02			5.72	1.10E-03	0.13
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_18112512})_i + \epsilon_i$	3.06 E-03	9.22 E-01	4.79 E-03			7.13	1.95E-04	0.16
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_18253358})_i + \epsilon_i$	3.34 E-03	9.36 E-01	5.94 E-03			6.97	2.35E-04	0.15
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{SNP_20306786})_i + \epsilon_i$	4.81 E-03	8.49 E-01	1.32 E-02			6.41	4.68E-04	0.14
1 2 3	Marker + Indel		$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{indel_13908828})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	1.88 E-02	5.61 E-02	3.61 E-01			5.73	1.10E-03	0.13
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{indel_15935454})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	1.67 E-02	3.60 E-02	4.01 E-01			6.17	6.36E-04	0.14
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{indel_15972511})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	2.04 E-02	4.50 E-02	4.87 E-01			5.57	1.35E-03	0.13
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{indel_17863273})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	7.22 E-02	3.52 E-02	4.79 E-01			6.08	7.19E-04	0.14
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{indel_9589063})_i + \epsilon_i$	1.75 E-03	5.54 E-01	1.94 E-03			7.75	9.38E-05	0.17
1 2 3	Marker + TE		$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{TE_16836944})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	3.19 E-02	8.32 E-02	5.38 E-01			4.53	5.13E-03	0.12
			$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{TE_20150892})_i + \beta_3g(\text{MM547})_i + \epsilon_i$	8.27 E-03	3.77 E-01	4.61 E-01			3.69	1.45E-02	0.10

				$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{TE_5595169})_i + \epsilon_i$	2.75 E-03	5.20 E-01	4.18 E-01			3.39	2.12E-02	0.10
				$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{TE_8759468})_i + \epsilon_i$	2.46 E-03	4.60 E-01	6.13 E-01			3.67	1.49E-02	0.10
		1 2 3	Marker + Footprint	$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{Footprint_5595169})_i + \epsilon_i$	4.56 E-04	6.25 E-01	8.34 E-01			4.41	5.64E-03	0.10
				$y_i = \beta_0 + \beta_1g(\text{MM128})_i + \beta_2g(\text{MM547})_i + \beta_3g(\text{Footprint_5595169})_i + \epsilon_i$	4.56 E-04	6.25 E-01	8.34 E-01			4.41	5.64E-03	0.10
		1 2 0	Marker	$y_i = \beta_0 + \beta_1g(\text{MM123})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	1.19 E-05	3.02 E-04				14.06	3.38E-06	0.19
				$y_i = \beta_0 + \beta_1g(\text{SNP_9497509})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	9.95 E-04	2.72 E-03				10.57	6.34E-05	0.16
				$y_i = \beta_0 + \beta_1g(\text{SNP_9883998})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	6.47 E-04	2.65 E-03				10.84	5.13E-05	0.17
				$y_i = \beta_0 + \beta_1g(\text{SNP_13033405})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	3.99 E-02	4.69 E-03				5.52	5.23E-03	0.09
				$y_i = \beta_0 + \beta_1g(\text{SNP_13967107})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	9.70 E-02	1.40 E-02				4.05	2.01E-02	0.07
				$y_i = \beta_0 + \beta_1g(\text{SNP_14023610})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	8.82 E-02	1.56 E-02				3.98	2.14E-02	0.07
		1 2 0	SNP	$y_i = \beta_0 + \beta_1g(\text{SNP_14590272})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	6.68 E-04	3.57 E-03				8.46	3.86E-04	0.13
				$y_i = \beta_0 + \beta_1g(\text{SNP_15509104})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	2.74 E-02	1.17 E-02				4.94	8.82E-03	0.08
				$y_i = \beta_0 + \beta_1g(\text{MM123})_i + \beta_2g(\text{SNP_2834559})_i + \epsilon_i$	3.29 E-04	1.59 E-01				7.08	1.26E-03	0.11
				$y_i = \beta_0 + \beta_1g(\text{MM123})_i + \beta_2g(\text{SNP_2920544})_i + \epsilon_i$	3.51 E-04	2.95 E-02				7.96	5.91E-04	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM123})_i + \beta_2g(\text{SNP_3318778})_i + \epsilon_i$	4.01 E-04	3.37 E-02				7.77	6.91E-04	0.12
				$y_i = \beta_0 + \beta_1g(\text{MM123})_i + \beta_2g(\text{SNP_4474111})_i + \epsilon_i$	3.09 E-04	1.81 E-01				7.12	1.23E-03	0.11
				$y_i = \beta_0 + \beta_1g(\text{MM123})_i + \beta_2g(\text{SNP_5115842})_i + \epsilon_i$	6.58 E-04	1.24 E-01				6.49	2.17E-03	0.11
		1 2 0	Indel	$y_i = \beta_0 + \beta_1g(\text{indel_13908828})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	2.10 E-02	3.67 E-03				6.27	2.66E-03	0.10
				$y_i = \beta_0 + \beta_1g(\text{indel_15935454})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	4.91 E-02	1.66 E-02				4.34	1.53E-02	0.07
				$y_i = \beta_0 + \beta_1g(\text{indel_15972511})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	9.79 E-02	8.17 E-03				4.45	1.38E-02	0.08
		1 2 0	TE	$y_i = \beta_0 + \beta_1g(\text{TE_12177658})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	4.97 E-02	1.24 E-02				4.98	8.76E-03	0.10
				$y_i = \beta_0 + \beta_1g(\text{TE_12617099})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	1.60 E-01	9.27 E-03				4.53	1.32E-02	0.09
				$y_i = \beta_0 + \beta_1g(\text{TE_16836944})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	4.35 E-01	6.98 E-03				3.90	2.36E-02	0.08
		1 2 0	Footprint	$y_i = \beta_0 + \beta_1g(\text{Footprint_12620247})_i + \beta_2g(\text{MM550})_i + \epsilon_i$	4.32 E-01	1.46 E-02				3.36	3.83E-02	0.05

P
b

1

G

			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{SNP_9497509})i + \beta_3g(\text{MM550})i + \epsilon_i$	7.46 E-04	8.82 E-02	9.94 E-05			11.77	9.76E-07	0.24
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{SNP_9883998})i + \beta_3g(\text{MM550})i + \epsilon_i$	1.13 E-03	4.60 E-02	1.32 E-04			11.64	1.16E-06	0.25
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{SNP_13033405})i + \beta_3g(\text{MM550})i + \epsilon_i$	2.63 E-04	8.34 E-01	2.73 E-04			8.88	2.63E-05	0.20
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{SNP_13967107})i + \beta_3g(\text{MM550})i + \epsilon_i$	1.94 E-04	5.16 E-01	5.26 E-04			8.01	7.19E-05	0.18
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{SNP_14023610})i + \beta_3g(\text{MM550})i + \epsilon_i$	3.50 E-04	6.42 E-01	9.45 E-04			7.50	1.31E-04	0.17
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{SNP_14590272})i + \beta_3g(\text{MM550})i + \epsilon_i$	1.26 E-01	9.54 E-01	1.97 E-03			6.50	4.39E-04	0.15
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{SNP_15509104})i + \beta_3g(\text{MM550})i + \epsilon_i$	2.86 E-04	8.42 E-01	7.42 E-04			8.35	4.68E-05	0.18
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{MM550})i + \beta_3g(\text{SNP_2834559})i + \epsilon_i$	2.85 E-05	7.54 E-04	6.81 E-01			9.18	1.75E-05	0.20
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{MM550})i + \beta_3g(\text{SNP_2920544})i + \epsilon_i$	3.34 E-05	3.00 E-03	5.99 E-01			8.77	2.90E-05	0.19
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{MM550})i + \beta_3g(\text{SNP_3318778})i + \epsilon_i$	4.91 E-05	4.42 E-03	6.15 E-01			8.34	4.75E-05	0.18
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{MM550})i + \beta_3g(\text{SNP_4474111})i + \epsilon_i$	2.67 E-05	7.80 E-04	6.41 E-01			9.19	1.76E-05	0.20
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{MM550})i + \beta_3g(\text{SNP_5115842})i + \epsilon_i$	5.36 E-05	9.17 E-04	7.59 E-01			8.61	3.50E-05	0.19
		1 2 0	Marker + SNP								
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{indel_13908828})i + \beta_3g(\text{MM550})i + \epsilon_i$	4.42 E-04	9.39 E-01	1.77 E-04			9.02	2.20E-05	0.20
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{indel_15935454})i + \beta_3g(\text{MM550})i + \epsilon_i$	1.61 E-04	7.67 E-01	6.73 E-04			8.36	4.69E-05	0.19
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{indel_15972511})i + \beta_3g(\text{MM550})i + \epsilon_i$	2.09 E-04	6.15 E-01	5.49 E-04			8.25	5.39E-05	0.19
		1 2 0	Marker + TE								
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{TE_12177658})i + \beta_3g(\text{MM550})i + \epsilon_i$	4.52 E-04	7.72 E-01	4.62 E-04			8.16	6.99E-05	0.21
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{TE_12617099})i + \beta_3g(\text{MM550})i + \epsilon_i$	3.74 E-04	8.96 E-01	3.86 E-04			7.99	8.72E-05	0.21
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{TE_16836944})i + \beta_3g(\text{MM550})i + \epsilon_i$	1.74 E-05	4.94 E-01	1.52 E-04			9.95	9.11E-06	0.24
		1 2 0	Marker + Footprint								
			$y_i = \beta_0 + \beta_1g(\text{MM123})i + \beta_2g(\text{Footprint_12620247})i + \beta_3g(\text{MM550})i + \epsilon_i$	1.90 E-05	7.94 E-01	3.38 E-04			9.24	1.59E-05	0.19
		1 2 2	Marker								
			$y_i = \beta_0 + \beta_1g(\text{MM693})i + \epsilon_i$	2.10 E-04					14.62	2.10E-04	0.11
	2		SNP								
			$y_i = \beta_0 + \beta_1g(\text{SNP_7344157})i + \epsilon_i$	1.32 E-01					2.31	1.32E-01	0.02
			$y_i = \beta_0 + \beta_1g(\text{SNP_7944780})i + \epsilon_i$	3.81 E-02					4.40	3.81E-02	0.04
			$y_i = \beta_0 + \beta_1g(\text{SNP_10012411})i + \epsilon_i$	1.80 E-02					5.76	1.80E-02	0.05
			$y_i = \beta_0 + \beta_1g(\text{SNP_10291133})i + \epsilon_i$	4.84 E-02					3.98	4.84E-02	0.03

Gx T	1 2 2	Indel	$y_i = \beta_0 + \beta_1 g(\text{indel_7214677})_i + \epsilon_i$	1.14 E- 01					2.54	1.14E- 01	0.02	
			$y_i = \beta_0 + \beta_1 g(\text{indel_7652616})_i + \epsilon_i$	1.06 E- 02					6.75	1.06E- 02	0.06	
			$y_i = \beta_0 + \beta_1 g(\text{indel_7750735})_i + \epsilon_i$	2.63 E- 02					5.07	2.63E- 02	0.04	
	1 2 2	TE	$y_i = \beta_0 + \beta_1 g(\text{TE_9650119})_i + \epsilon_i$	1.09 E- 03					11.32	1.09E- 03	0.10	
	1 2 2	Marker + SNP	$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{SNP_7344157})_i + \epsilon_i$	2.94 E- 03	9.71 E- 01					5.86	3.80E- 03	0.09
			$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{SNP_7944780})_i + \epsilon_i$	1.33 E- 02	5.98 E- 01					5.48	5.41E- 03	0.09
			$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{SNP_10012411})_i + \epsilon_i$	8.94 E- 03	9.34 E- 01					6.58	2.00E- 03	0.11
			$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{SNP_10291133})_i + \epsilon_i$	6.86 E- 03	9.56 E- 01					5.90	3.66E- 03	0.09
	1 2 2	Marker + Indel	$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{indel_7214677})_i + \epsilon_i$	3.84 E- 03	8.78 E- 01					5.71	4.34E- 03	0.09
			$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{indel_7652616})_i + \epsilon_i$	2.05 E- 02	5.25 E- 01					6.27	2.62E- 03	0.10
			$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{indel_7750735})_i + \epsilon_i$	7.45 E- 03	4.67 E- 01					6.39	2.35E- 03	0.10
	1 2 2	Marker + TE	$y_i = \beta_0 + \beta_1 g(\text{MM693})_i + \beta_2 g(\text{TE_9650119})_i + \epsilon_i$	8.16 E- 02	2.71 E- 01					7.33	1.08E- 03	0.13
	1 2 0	Marker	$y_i = \beta_0 + \beta_1 g(\text{MM679})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	6.37 E- 03	3.72 E- 05					9.20	1.95E- 04	0.14
			$y_i = \beta_0 + \beta_1 g(\text{SNP_2834559})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	6.46 E- 01	1.87 E- 03					5.21	6.83E- 03	0.08
			$y_i = \beta_0 + \beta_1 g(\text{SNP_2920544})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	5.17 E- 01	4.24 E- 03					4.28	1.62E- 02	0.07
			$y_i = \beta_0 + \beta_1 g(\text{SNP_3318778})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	5.30 E- 01	2.96 E- 03					4.65	1.15E- 02	0.08
			$y_i = \beta_0 + \beta_1 g(\text{SNP_4474111})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	5.28 E- 01	1.96 E- 03					5.09	7.68E- 03	0.08
			$y_i = \beta_0 + \beta_1 g(\text{SNP_5115842})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	6.32 E- 01	3.96 E- 03					4.38	1.48E- 02	0.07
			$y_i = \beta_0 + \beta_1 g(\text{SNP_7344157})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	2.70 E- 01	4.10 E- 03					4.37	1.49E- 02	0.07
			$y_i = \beta_0 + \beta_1 g(\text{SNP_7944780})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	7.11 E- 01	4.89 E- 02					3.34	3.92E- 02	0.06
$y_i = \beta_0 + \beta_1 g(\text{MM679})_i + \beta_2 g(\text{SNP_7344157})_i + \epsilon_i$			5.04 E- 01	5.04 E- 01					0.29	7.48E- 01	0.01	
$y_i = \beta_0 + \beta_1 g(\text{MM679})_i + \beta_2 g(\text{SNP_7944780})_i + \epsilon_i$			1.35 E- 01	3.29 E- 02					2.47	8.94E- 02	0.04	
1 2 0			Indel	$y_i = \beta_0 + \beta_1 g(\text{indel_7214677})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	3.54 E- 01	4.87 E- 03					4.27	1.63E- 02
	$y_i = \beta_0 + \beta_1 g(\text{indel_7652616})_i + \beta_2 g(\text{MM689})_i + \epsilon_i$	1.73 E- 01		5.07 E- 03					4.84	9.61E- 03	0.08	

				$y_i = \beta_0 + \beta_1g(\text{indel_7750735})_i + \beta_2g(\text{MM689})_i + \epsilon_i$	7.81 E-01	1.68 E-02				3.60	3.05E-02	0.06
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{indel_7214677})_i + \epsilon_i$	5.75 E-01	4.59 E-01				0.29	7.47E-01	0.01
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{indel_7652616})_i + \epsilon_i$	4.96 E-02	2.30 E-02				2.70	7.16E-02	0.05
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{indel_7750735})_i + \epsilon_i$	2.59 E-01	1.24 E-01				1.27	2.85E-01	0.02
		1 2 0	Marker + SNP	$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{SNP_2834559})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	5.46 E-03	4.59 E-01	3.20 E-05			6.37	5.03E-04	0.15
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{SNP_2920544})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	6.28 E-03	6.08 E-01	8.23 E-05			5.61	1.28E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{SNP_3318778})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	6.25 E-03	5.93 E-01	5.69 E-05			5.88	9.24E-04	0.14
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{SNP_4474111})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	4.73 E-03	5.53 E-01	2.92 E-05			6.38	4.96E-04	0.15
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{SNP_5115842})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	5.71 E-03	4.78 E-01	7.21 E-05			5.76	1.08E-03	0.14
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{SNP_7344157})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	8.42 E-03	8.53 E-01	1.29 E-04			5.47	1.52E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{SNP_7944780})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	1.87 E-03	2.72 E-01	7.76 E-04			5.80	1.03E-03	0.14
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{MM689})_i + \beta_3g(\text{SNP_7344157})_i + \epsilon_i$	8.42 E-03	1.29 E-04	8.53 E-01			5.47	1.52E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{MM689})_i + \beta_3g(\text{SNP_7944780})_i + \epsilon_i$	1.87 E-03	7.76 E-04	2.72 E-01			5.80	1.03E-03	0.14
		1 2 0	Marker + Indel	$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{indel_7214677})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	8.60 E-03	9.80 E-01	1.44 E-04			5.39	1.68E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{indel_7652616})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	1.20 E-02	7.46 E-01	1.34 E-03			5.57	1.36E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{indel_7750735})_i + \beta_3g(\text{MM689})_i + \epsilon_i$	4.34 E-03	5.68 E-01	3.98 E-04			5.40	1.69E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{MM689})_i + \beta_3g(\text{indel_7214677})_i + \epsilon_i$	8.60 E-03	1.44 E-04	9.80 E-01			5.39	1.68E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{MM689})_i + \beta_3g(\text{indel_7652616})_i + \epsilon_i$	1.20 E-02	1.34 E-03	7.46 E-01			5.57	1.36E-03	0.13
				$y_i = \beta_0 + \beta_1g(\text{MM679})_i + \beta_2g(\text{MM689})_i + \beta_3g(\text{indel_7750735})_i + \epsilon_i$	4.34 E-03	3.98 E-04	5.68 E-01			5.40	1.69E-03	0.13
R bi	2	G	1 2 3	Marker	$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	5.08 E-04	8.05 E-08	1.05 E-06	1.81 E-05	20.05	1.27E-12	0.40
				SNP	$y_i = \beta_0 + \beta_1g(\text{SNP_7241579})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	4.67 E-02	2.47 E-05	7.99 E-05	2.76 E-05	14.34	1.95E-09	0.34
					$y_i = \beta_0 + \beta_1g(\text{SNP_9497509})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	1.23 E-01	2.25 E-06	1.24 E-05	6.99 E-05	14.35	1.79E-09	0.34
					$y_i = \beta_0 + \beta_1g(\text{SNP_9883998})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	1.69 E-01	2.20 E-06	1.48 E-05	7.71 E-05	14.08	2.71E-09	0.34
					$y_i = \beta_0 + \beta_1g(\text{SNP_13033405})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	1.21 E-01	8.76 E-05	2.02 E-04	9.78 E-05	11.47	8.25E-08	0.30

			$y_i = \beta_0 + \beta_1g(\text{SNP}_{13967107})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	5.26 E-02	4.76 E-06	2.35 E-05	9.91 E-04		12.16	3.18E-08	0.31
			$y_i = \beta_0 + \beta_1g(\text{SNP}_{14023610})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	4.36 E-02	5.78 E-06	2.61 E-05	5.65 E-04		12.53	1.92E-08	0.31
			$y_i = \beta_0 + \beta_1g(\text{SNP}_{14590272})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	3.84 E-02	6.60 E-06	2.43 E-05	1.20 E-04		13.84	3.57E-09	0.33
			$y_i = \beta_0 + \beta_1g(\text{SNP}_{15509104})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	5.33 E-02	7.05 E-06	2.69 E-05	2.29 E-04		12.83	1.25E-08	0.31
			$y_i = \beta_0 + \beta_1g(\text{SNP}_{18002233})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	5.63 E-02	2.37 E-06	6.93 E-06	3.76 E-05		15.88	2.62E-10	0.36
			$y_i = \beta_0 + \beta_1g(\text{SNP}_{18173880})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	6.96 E-02	2.50 E-06	6.58 E-06	4.91 E-05		15.51	4.16E-10	0.36
			$y_i = \beta_0 + \beta_1g(\text{SNP}_{20362809})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	6.54 E-02	3.16 E-06	6.21 E-06	3.15 E-05		15.77	3.01E-10	0.36
			$y_i = \beta_0 + \beta_1g(\text{SNP}_{21675141})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	3.38 E-02	1.60 E-06	2.39 E-06	6.80 E-05		16.58	1.28E-10	0.38
			$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{11434437})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	1.44 E-03	8.55 E-04	1.55 E-01	1.81 E-05		12.74	1.62E-08	0.32
			$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{SNP}_{7344157})_i + \epsilon_i$	3.22 E-03	3.40 E-08	1.32 E-07	1.16 E-01		13.19	7.72E-09	0.32
			$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{SNP}_{7944780})_i + \epsilon_i$	6.07 E-03	7.14 E-06	1.70 E-05	4.19 E-02		9.85	7.71E-07	0.27
1	Indel		$y_i = \beta_0 + \beta_1g(\text{indel}_{13908828})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	3.43 E-02	7.22 E-06	3.83 E-05	3.94 E-04		12.82	1.35E-08	0.32
2			$y_i = \beta_0 + \beta_1g(\text{indel}_{15935454})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	9.58 E-02	6.81 E-06	3.61 E-05	1.18 E-04		12.65	1.59E-08	0.31
3			$y_i = \beta_0 + \beta_1g(\text{indel}_{15972511})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	7.30 E-02	1.46 E-06	9.39 E-06	1.52 E-04		12.78	1.41E-08	0.32
			$y_i = \beta_0 + \beta_1g(\text{indel}_{17863273})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	6.93 E-02	3.24 E-05	6.69 E-05	8.56 E-05		12.48	2.10E-08	0.31
			$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{indel}_{7214677})_i + \epsilon_i$	1.44 E-03	2.39 E-08	1.16 E-07	1.11 E-01		13.99	2.73E-09	0.33
			$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{indel}_{7652616})_i + \epsilon_i$	2.22 E-03	7.11 E-07	4.01 E-06	1.65 E-04		16.32	1.47E-10	0.37
			$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{indel}_{7750735})_i + \epsilon_i$	1.99 E-03	3.00 E-07	1.30 E-06	4.54 E-02		12.75	1.47E-08	0.32
1	TE		$y_i = \beta_0 + \beta_1g(\text{TE}_{6617937})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	8.57 E-01	6.38 E-08	1.48 E-06	2.31 E-06		17.16	1.13E-10	0.41
2			$y_i = \beta_0 + \beta_1g(\text{TE}_{12177658})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	8.05 E-02	7.76 E-07	9.64 E-06	5.93 E-06		14.51	2.78E-09	0.38
3			$y_i = \beta_0 + \beta_1g(\text{TE}_{12617099})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	9.27 E-02	3.78 E-06	3.10 E-05	5.13 E-06		12.19	5.25E-08	0.34
			$y_i = \beta_0 + \beta_1g(\text{TE}_{16836944})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	4.72 E-02	7.04 E-08	2.06 E-06	3.12 E-06		18.87	1.67E-11	0.44
			$y_i = \beta_0 + \beta_1g(\text{TE}_{20150892})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	1.18 E-02	1.27 E-08	4.29 E-07	3.35 E-06		19.96	5.15E-12	0.45
			$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{TE}_{10506856})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	7.27 E-03	7.62 E-01	7.94 E-01	2.86 E-06		8.40	7.55E-06	0.26

		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{TE}_{12762111})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	2.69 E-03	1.29 E-02	4.91 E-01	1.35 E-07		12.15	4.92E-08	0.34
1 2 3	Footprint	$y_i = \beta_0 + \beta_1g(\text{Footprint}_{12620247})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \epsilon_i$	2.67 E-01	2.25 E-07	2.16 E-06	1.64 E-05		15.43	3.65E-10	0.35
1 2 3	Marker + SNP	$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{7241579})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	3.43 E-03	1.86 E-01	1.41 E-05	5.62 E-05	4.06 E-05	14.10	1.32E-10	0.39
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{9497509})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	1.36 E-03	5.33 E-01	2.58 E-06	1.96 E-05	2.54 E-05	14.65	5.35E-11	0.40
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{9883998})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	1.91 E-03	6.32 E-01	2.31 E-06	1.74 E-05	5.05 E-05	14.23	1.09E-10	0.40
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{13033405})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	6.53 E-03	4.76 E-01	4.44 E-05	1.24 E-04	6.87 E-05	11.28	9.13E-09	0.34
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{13967107})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	3.16 E-03	3.81 E-01	1.07 E-06	6.92 E-06	9.20 E-04	12.27	1.91E-09	0.36
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{14023610})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	4.50 E-03	3.12 E-01	1.55 E-06	8.67 E-06	4.66 E-04	12.37	1.58E-09	0.36
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{14590272})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	2.31 E-02	5.36 E-01	2.20 E-06	1.07 E-05	6.15 E-05	12.57	1.22E-09	0.37
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{15509104})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	3.77 E-03	4.51 E-01	1.55 E-06	8.83 E-06	2.16 E-04	12.73	8.87E-10	0.36
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{18002233})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	9.20 E-03	5.82 E-01	6.52 E-07	3.13 E-06	3.63 E-05	14.80	4.32E-11	0.40
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{18173880})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	9.60 E-03	6.39 E-01	6.99 E-07	2.99 E-06	4.80 E-05	14.46	7.06E-11	0.40
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{20362809})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	7.38 E-03	6.80 E-01	7.01 E-07	2.27 E-06	3.09 E-05	14.84	4.09E-11	0.40
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{SNP}_{21675141})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	3.45 E-03	4.90 E-01	3.17 E-07	7.63 E-07	4.68 E-05	16.03	9.10E-12	0.43
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{SNP}_{11434437})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	7.19 E-04	3.57 E-06	2.47 E-02	1.95 E-06	6.13 E-05	17.17	2.06E-12	0.45
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \beta_5g(\text{SNP}_{7344157})_i + \epsilon_i$	1.02 E-03	1.98 E-07	1.41 E-06	4.26 E-04	6.09 E-01	14.34	7.99E-11	0.39
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \beta_5g(\text{SNP}_{7944780})_i + \epsilon_i$	1.65 E-03	3.13 E-05	9.60 E-05	1.63 E-03	6.75 E-01	10.65	2.46E-08	0.33
1 2 3	Maker + Indel	$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{indel}_{13908828})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	3.35 E-03	2.76 E-01	1.88 E-06	1.54 E-05	1.96 E-04	12.80	8.63E-10	0.37
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{indel}_{15935454})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	1.31 E-03	5.60 E-01	1.31 E-06	1.20 E-05	8.16 E-05	13.19	4.43E-10	0.37
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{indel}_{15972511})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	1.51 E-03	5.61 E-01	9.76 E-08	8.91 E-07	8.67 E-05	13.23	4.49E-10	0.38
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{indel}_{1786327})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	1.12 E-02	6.55 E-01	6.20 E-06	1.82 E-05	8.79 E-05	11.83	3.78E-09	0.35
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \beta_5g(\text{indel}_{7214677})_i + \epsilon_i$	6.08 E-04	1.50 E-07	1.24 E-06	3.99 E-04	6.44 E-01	15.10	2.69E-11	0.40
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \beta_5g(\text{indel}_{7652616})_i + \epsilon_i$	1.67 E-03	1.01 E-06	6.13 E-06	1.54 E-01	2.81 E-01	13.59	2.41E-10	0.38
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \beta_5g(\text{indel}_{7750735})_i + \epsilon_i$	1.06 E-03	9.85 E-07	5.77 E-06	2.55 E-03	9.91 E-01	12.90	7.42E-10	0.37

1 2 3	Marker + TE	$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{TE}_{6617937})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	4.16 E-04	4.37 E-01	4.88 E-09	2.38 E-07	1.39 E-06	18.16	1.21E-12	0.49	
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{TE}_{12177658})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	1.65 E-03	2.70 E-01	6.14 E-08	1.23 E-06	3.78 E-06	14.87	9.85E-11	0.44	
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{TE}_{12617099})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	2.22 E-03	2.82 E-01	4.69 E-07	5.24 E-06	3.39 E-06	12.66	2.30E-09	0.41	
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{TE}_{16836944})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	1.71 E-03	1.74 E-01	9.31 E-09	4.51 E-07	1.98 E-06	18.65	6.63E-13	0.49	
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{TE}_{20150892})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	4.97 E-03	1.31 E-01	3.43 E-09	1.74 E-07	2.11 E-06	18.82	5.37E-13	0.49	
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{TE}_{10506856})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	7.79 E-04	8.02 E-09	5.78 E-01	5.88 E-07	2.54 E-06	17.53	3.13E-12	0.48	
		$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{MM383})_i + \beta_3g(\text{MM385})_i + \beta_4g(\text{MM686})_i + \beta_5g(\text{TE}_{12762111})_i + \epsilon_i$	9.42 E-04	1.11 E-07	2.34 E-06	2.87 E-01	1.91 E-06	17.24	4.17E-12	0.48	
	1 2 3	Marker + Footprint	$y_i = \beta_0 + \beta_1g(\text{MM147})_i + \beta_2g(\text{Footprint}_{12620247})_i + \beta_3g(\text{MM383})_i + \beta_4g(\text{MM385})_i + \beta_5g(\text{MM686})_i + \epsilon_i$	8.90 E-04	4.28 E-01	1.12 E-07	1.45 E-06	1.96 E-05	15.80	7.62E-12	0.41
	1 2 2	Marker	$y_i = \beta_0 + \beta_1g(\text{MM432})_i + \epsilon_i$	7.50 E-04					11.97	7.50E-04	0.09
	Gx T	SNP	$y_i = \beta_0 + \beta_1g(\text{SNP}_{9858537})_i + \epsilon_i$	1.55 E-02					6.03	1.55E-02	0.05
$y_i = \beta_0 + \beta_1g(\text{SNP}_{10202051})_i + \epsilon_i$			2.64 E-02					5.06	2.64E-02	0.04	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{10743038})_i + \epsilon_i$			3.63 E-02					4.48	3.63E-02	0.04	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{13092809})_i + \epsilon_i$			1.33 E-03					10.83	1.33E-03	0.09	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{14288480})_i + \epsilon_i$			1.12 E-02					6.64	1.12E-02	0.05	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{14509867})_i + \epsilon_i$			1.12 E-02					6.64	1.12E-02	0.05	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{14518339})_i + \epsilon_i$			1.12 E-02					6.64	1.12E-02	0.05	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{14530903})_i + \epsilon_i$			1.12 E-02					6.64	1.12E-02	0.05	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{14545591})_i + \epsilon_i$			1.56 E-02					6.03	1.56E-02	0.05	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{15861145})_i + \epsilon_i$			1.17 E-02					6.55	1.17E-02	0.05	
$y_i = \beta_0 + \beta_1g(\text{SNP}_{16205156})_i + \epsilon_i$			3.88 E-02					4.37	3.88E-02	0.04	
$y_i = \beta_0 + \beta_1g(\text{MM432})_i + \beta_1g(\text{SNP}_{9858537})_i + \epsilon_i$			8.40 E-03	2.72 E-01				6.77	1.65E-03	0.11	
$y_i = \beta_0 + \beta_1g(\text{MM432})_i + \beta_1g(\text{SNP}_{10202051})_i + \epsilon_i$			7.22 E-03	3.40 E-01				6.41	2.29E-03	0.10	
$y_i = \beta_0 + \beta_1g(\text{MM432})_i + \beta_1g(\text{SNP}_{10743038})_i + \epsilon_i$			7.27 E-03	3.85 E-01				6.10	3.03E-03	0.09	
$y_i = \beta_0 + \beta_1g(\text{MM432})_i + \beta_1g(\text{SNP}_{13092809})_i + \epsilon_i$	1.33 E-03	NA				10.83	1.33E-03	0.09			

ANNEXE II : Nature Communication Supplementary

Methods

Communication title: Natural epiallelic variation associated with quantitative resistance to plant pathogen.

Authors: Benjamin Liégard, Antoine Gravot, Leandro Quadrana, Yoann Aigu, Juliette Benejam, Christine Lariagon, Jocelyne Lemoine, Vincent Colot, Maria Manzanarès-Dauleux and Mélanie Jubault

The following Supplementary Methods are available for this article:

Details of <i>PbAt5.2</i> fine mapping	234
Obtainment and phenotyping of F1 individuals derived from crosses between 10499 and 13499	234
Screening of recombinant individuals in the segregating F2 progeny	234
Phenotyping of recombinant F3 lines.....	235
High-density genotyping of recombinant F2/F3 lines.....	235
F4 lines with fixed alleles in the region of the resistance locus allowed a further reduction of the <i>PbAt5.2</i> resistance locus interval.....	236

Details of *PbAt5.2* fine mapping

Previous identification of *PbAt5.2*. A previous screen among *Arabidopsis thaliana* accessions had identified Bur-0 as partially resistant to the eH isolate of *P. brassicae*, when the canonical accession Col-0 is fully susceptible⁴⁴. Two progenies, were used for the identification of QTL controlling resistance to eH by linkage analysis: the first one was derived from an initial cross Bur-0 (Versailles Id=172AV) x Col-0 (186AV) followed by 6 generation SSD (Recombinant Inbred Line set 20RV, described in⁴⁶). The second one was derived from a progeny from Col-0 (NASC Id N1092) x Bur-0 (N1028) (described in⁴³). Those approaches led to the identification of a series of four additive QTL, including *PbAt5.2* ($R^2=20\%$, Bur-0 allele at his locus conferring partial resistance). The QTL peak of *PbAt5.2* was at the marker C5_19316 (around 19.3 Mb), with a confidence interval of 4.3 cM, between positions 18.7 Mb and 20.3 Mb (157 annotated genes between At5g46260 and At5g47690).

HIF499 for the validation of *PbAt5.2*. Among the RIL set 20RV from⁴⁶, one RIL line (RIL499) displayed a single residual heterozygous region including markers c5_17570, c5_19313 and c5_20318, framed by the homozygous loci c5_14766 and c5_21319. Heterogeneous Inbred Family (HIF) lines 10499 and 13499 were derived from RIL499 to get homozygosity at this locus (Institute Jean Pierre Bourgin, INRA Versailles, France). Those lines are near-isogenic, with identical combinations of Bur-0 and Col-0 homozygous genome sequences at every loci, except in the region between c5_14766 and c5_21319. The use of PCR-based markers allowed us to reduce this interval between markers CL5_15283 and CL5_20983 (excluded). HIF line 10499 displayed a higher level of partial resistance to eH isolate compared to 13499, thus confirming the position of *PbAt5.2* in this interval (Lemarié *et al.* 2015, see also the Figure 1 in the present work).

Obtainment and phenotyping of F1 individuals derived from crosses between 10499 and 13499

Fine mapping of *PbAt5.2* started by the cross between HIF-13499 (allele Col-0) and HIF-10499 (allele Bur-0) lines. Crosses were done in both senses, i.e. using one or the other parent as female. Heterozygosity in the *PbAt5.2* region was checked in several F1 individuals using a series of PCR-based markers. Clubroot index was evaluated in a series of F1 plants using the eH isolate and was found statistically identical to 13499 and higher than for 10499, thus suggesting that the Bur-0 resistant allele at *PbAt5.2* was recessive (Annexe III Figure S1).

Screening of recombinant individuals in the segregating F2 progeny

One validated F1 plant was chosen from each of the two crosses. Those two plants were self-pollinized, and approximately 3200 F2 plants were sown (about 1600 from each cross). Individual F2 from number 1 to 1581 were from a cross 10499 x 13499, and individual F2 plants from 1582 to 3153 were from a cross 13499 x 10499. DNA was extracted from young leaves sampled from those 3152 plants, and then subjected to a first round of genotyping using a series of 10 KAPSPar SNP markers. Analyses

were performed on the GENTYANE platform using a LightCycler 480 device (UMR INRA 1095, Clermont-Ferrand France). Due to low DNA concentration in some samples, about 6 % of the genotyping points were 'negative'. Good or average-quality (i.e. genotyping ambiguity at maximum one marker) data were obtained for 2751 F2 individuals. Among those, 563 plants displayed at least one recombination in the region between At5g37660 and At5g51670, which represented about 20.5 % of the F2 plants, and was consistent with the distance of about 22 cM between those two marker genes. This mean value of 20.5 however masked a clear disparity between genotypes derived from the 10499 x 13499 cross (32 % of individuals with one recombination in the region) from the 13499 x 10499 cross (18 % of individuals with one recombination in the chromosomal region). Nevertheless, most of this disequilibrium was focused in the region between the markers on At5g42520 and At5g44630 (i.e. outside the confidence interval of the clubroot resistance *PbAt5.2* QTL), and the male x female direction of initial hybridization step did not obviously affect the recombination rate in the region corresponding to the peak of the QTL (between At5g46910, At5g47120 and At5g47510).

Phenotyping of recombinant F3 lines

107 recombinant F2 lines were selected based on the presence of a recombination event near the closest markers to the QTL peak (19.3 Mb). Using the seeds derived from the selfing of those individual lines, the clubroot disease index was estimated from 18 inoculated F3 plants (6 plants x 3 biological replicates). Those phenotyping data confirmed the presence of a segregating resistance locus in the region. Lines with Bur-0 or Col-0 homozygous alleles at markers At5g47120/At5g47510 displayed a mean disease index of 37.1 (SE = 7.3) and 87.2 (SE = 14.1), respectively. Lines with heterozygosity at those two markers displayed a mean disease index of 76.7 (SE = 16), thus corroborating the above conclusion that the resistance Bur-0 allele was apparently recessive.

High-density genotyping of recombinant F2/F3 lines

A subset of 69 F2 recombinant lines was selected, based on the presence of a recombination event near the closest markers to the QTL peak (19.3 Mb). For each of them, twelve to eighteen F3 progeny individuals were grown and their leaves were bulk-sampled, for subsequent analysis of 93 SNP. All leaf samples were analysed at the GENTYANE platform. SNP genotyping was performed with the KASPar genotyping chemistry and Dynamic Array™ IFC 96*96 (UMR INRA 1095, Clermont-Ferrand France). Genotyping data obtained from bulked leaves of F3 individuals represents the genotypes of the parental F2 individuals. This genotyping workflow was also applied to a series of non-recombinant F2 lines and parental HIF 10499 and 13499 lines, which were used as controls. The confrontation of the 93 SNP genotyping data and clubroot disease index for all those 69 lines finally allowed to identify a small region between the marker K58 = At5g47230prom (position 19,175,831 bp) and the marker K65 = At5g47360 (position 19,214,446 bp). Due to the recessive status of the resistance allele, a series

of recombinant lines with strategic recombination events (especially in the line 1381) were not exploitable for the reasoning at this stage because of their heterozygosity at the locus.

F4 lines with fixed alleles in the region of the resistance locus allowed a further reduction of the *PbAt5.2* resistance locus interval

DNA was extracted from 12 to 18 individual F3 plants derived from the selfing of a series of F2 recombinant lines. Among them, F3 lines harbouring homozygosity in the region of the resistance locus were screened using the following PCR-based markers: CL5_16921 = At5g42320; CL5_17802 = At5g44200; CL5_18135 = At5g44900; AF-NUD8; CL5_19601 = At5g48375. Homozygous F4 seed stocks were then obtained from the selfing of the selected homozygous F3 lines, and thereafter used for additional clubroot phenotyping assays. Resulting data allowed to reduce the interval to a region between markers K58 = At5g47230prom (position 19,175,831 bp) and K64 (position 19,208,823 bp), as shown in the Figure 1E. Finally, every other SNP and indels in the region were analysed by the sequencing of PCR-amplified fragments from 2313-15, 1381-2, 2509-11, and 1600-5. This allowed to reach a final confidence interval of 26 kb between the marker CLG4 (19,182,401, in the promoter region of At5g47240) and the marker K64 (on SNP at position 19,208,823 bp, in At5g47330). This region contained eight annotated ORF (including three NLR genes), six transposons and one lncRNA.

ANNEXE III : *Nature Communication* Supporting Information

Communication title: Natural epiallelic variation associated with quantitative resistance to plant pathogen.

Authors: Benjamin Liégard, Antoine Gravot, Leandro Quadrana, Yoann Aigu, Juliette Benejam, Christine Lariagon, Jocelyne Lemoine, Vincent Colot, Maria Manzanarès-Dauleux and Mélanie Jubault

The following Supporting Information is available for this article:

Figure S1. Illustration of partial resistance to eH isolate in Bur-0, compared to the full	238
Figure S2. Illustration of partial resistance to eH isolate in 10499, compared to the full susceptibility in 13499.....	239
Figure S3. Comparison of GALA disease index in the F1 progeny and in the parental lines 10499 and 13499.....	240
Figure S4. GALA disease index in RIL499-derived near-isogenic lines 1381-2 and 2313-15 challenged with a series of four European monospore <i>P. brassicae</i> isolates, and with the field isolate P1 ⁺	241
Figure S5. GALA disease index in the mutant lines defective for the expression of At5g47240 or At5g47250.....	242
Figure S6. Methylation of <i>PbAt-5.2</i> region in Col-0 and in mutants in genes involved in RdDM methylation maintenance ⁵¹	243
Table S1. List of natural <i>Arabidopsis</i> accessions showing a genomic like Col-0 structure in the region <i>PbAt-5.2</i>	244
Table S2. Frequency of <i>PbAt-5.2</i> genomic structure and methylation state according to geographic groups.....	250
Table S3. List of <i>Arabidopsis</i> accessions with Col-0 like or Bur-0 like epimutation state.	251
Table S4. List of 127 <i>Arabidopsis</i> accessions used in pathological test.....	255
Table S5. List of primer sets used to CHOP qPCR and qPCR experimentation.	257

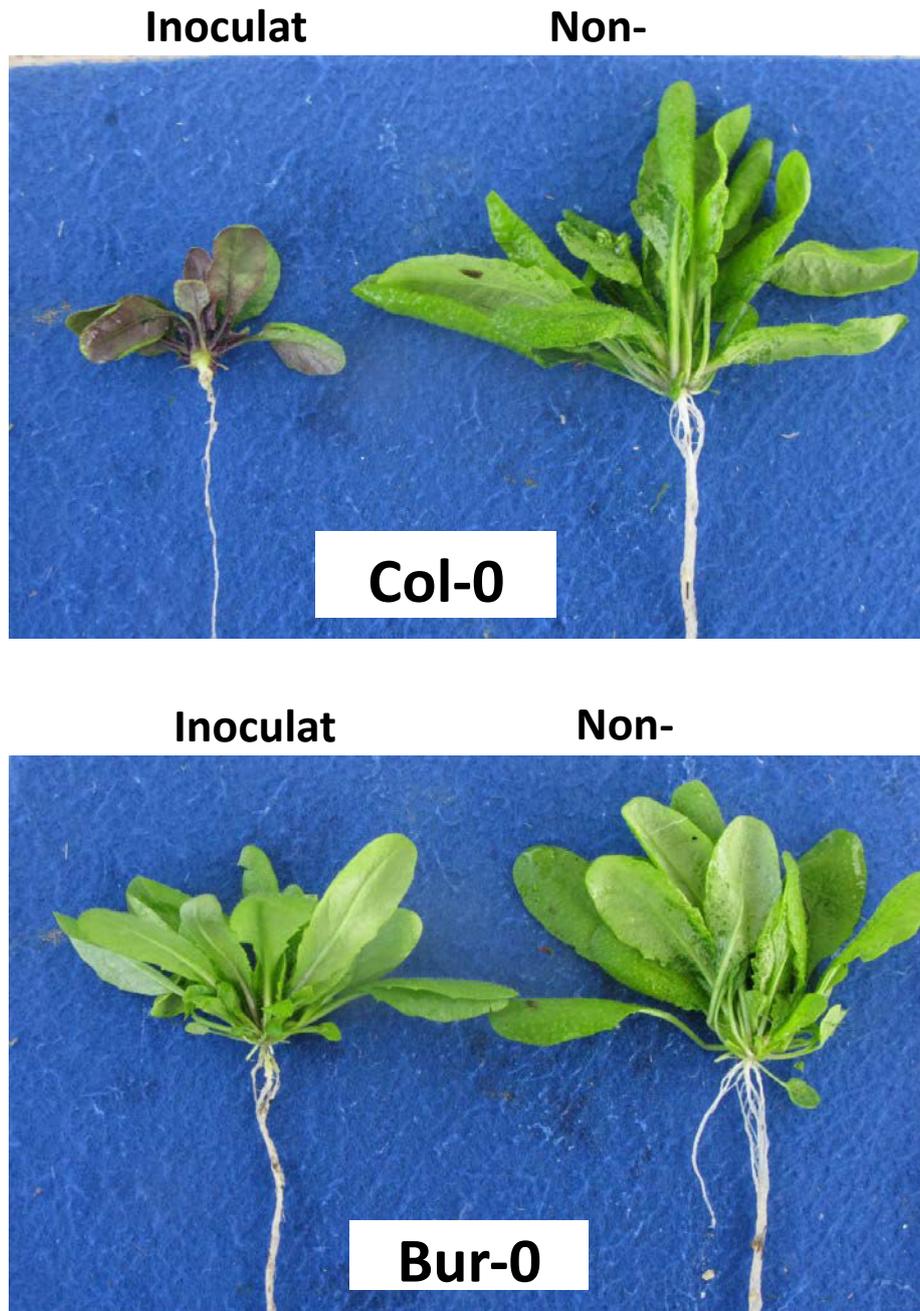


Figure S1. Illustration of partial resistance to eH isolate in Bur-0, compared to the full susceptibility in Col-0. Observations were done at 21 days post-inoculation. The white arrow indicates the presence of limited amount of galls in inoculated Bur-0, reflecting that partial resistance in this accession does not impair pathogen penetration but reduces its subsequent development. Plant individuals are representative of standard observations made in our experimental conditions. Levels of clubroot symptoms can vary depending on pathotests, but levels of clubroot symptoms are robustly found lower in Bur-0.



Figure S2. Illustration of partial resistance to eH isolate in 10499, compared to the full susceptibility in 13499. Observations were done at 21 days post-inoculation. The white arrow indicates the presence of limited amount of galls in inoculated 10499, reflecting that partial resistance in this accession does not impair pathogen penetration but reduces its subsequent development. Plant individuals are representative of standard observations made in our experimental conditions. Levels of clubroot symptoms can vary depending on pathotests, but levels of clubroot symptoms were robustly found lower in 10499 compared to 13499.

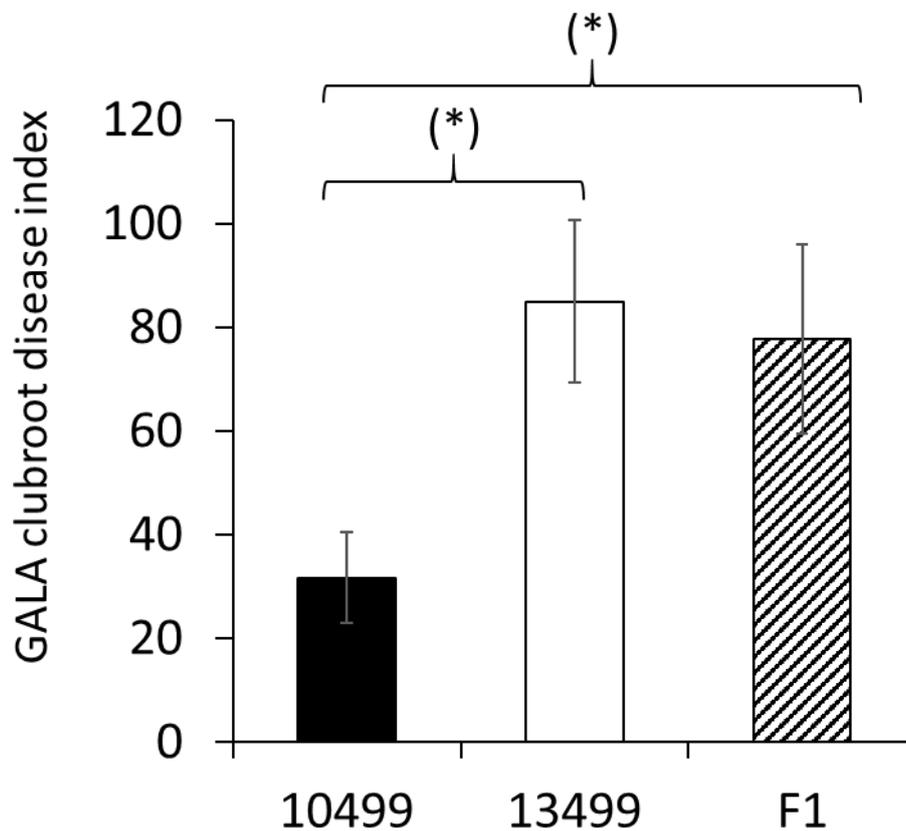


Figure S3. Comparison of GALA disease index in the F1 progeny and in the parental lines 10499 and 13499. GALA disease index is calculated through image analysis (details in the material and method part) from inoculated plants at 21 days post inoculation. Data are from 3 replicates ($n = 3$). For each replicate, GALA disease index was calculated from 6 to 12 inoculated individual plants. Stars indicate statistical differences from the paired Student t-test (p -value = 0.05)

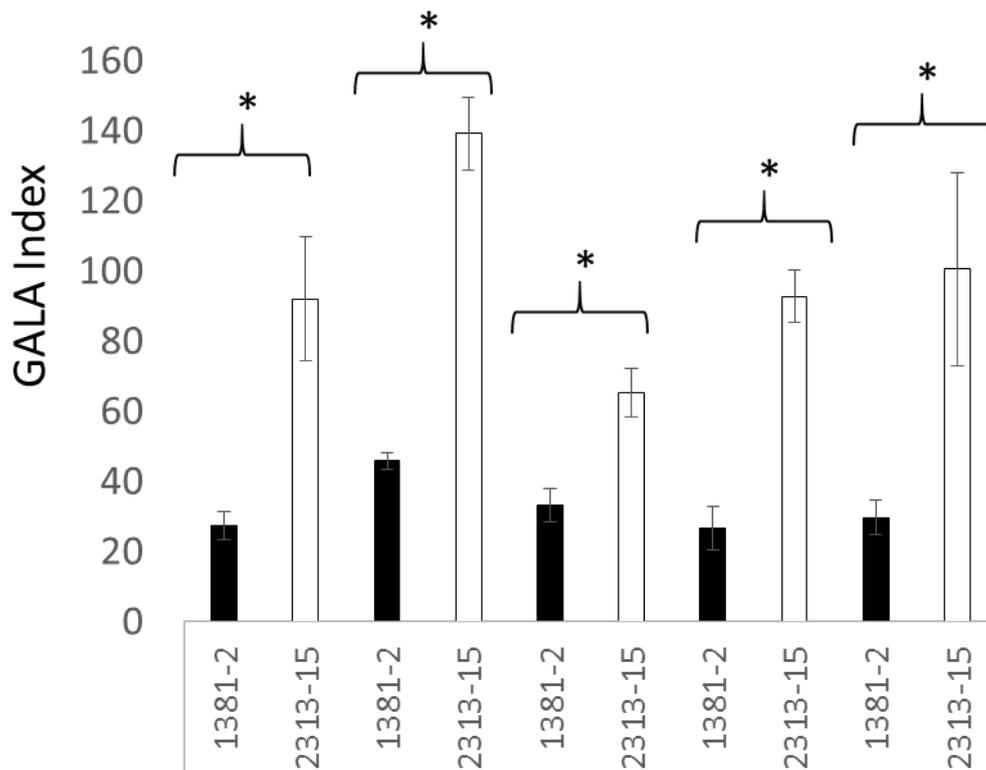


Figure S4. GALA disease index in RIL499-derived near-isogenic lines 1381-2 and 2313-15 challenged with a series of four European monospore *P. brassicae* isolates, and with the field isolate P1⁺. This last isolate is representative of emerging European pathotypes that are virulent on the variety ‘Mendel’ (a clubroot-resistant variety which has been cultivated at large scales in Europe and used as a source of resistance in many breeding programs). Data are means of 4 independent replicates (n = 4). For each replicate, GALA values are means of 6 to 12 individual plants. Error bars indicate standard errors. Statistically different values (from Student t-test) are indicated by stars.

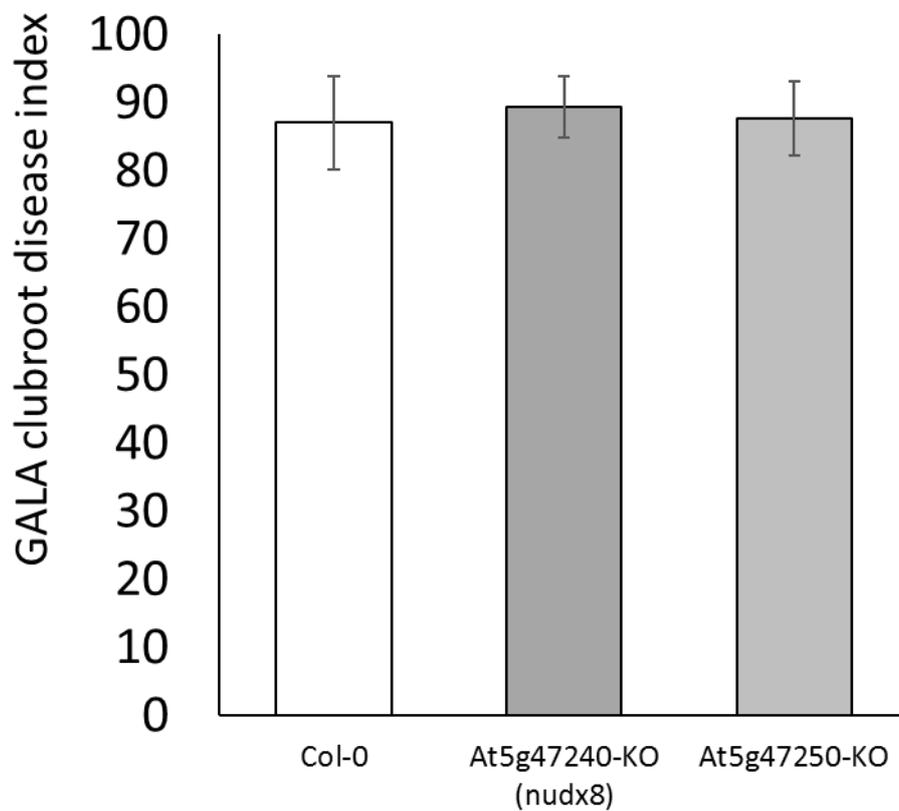


Figure S5. GALA disease index in the mutant lines defective for the expression of At5g47240 or At5g47250. GALA disease index is calculated through image analysis (details in the material and method part) from inoculated plants at 21 days post inoculation. Data are from 4 replicates (n = 4). For each replicate, GALA disease index was calculated from 7 to 11 inoculated individual plants. No statistical differences were found between the genotypes.

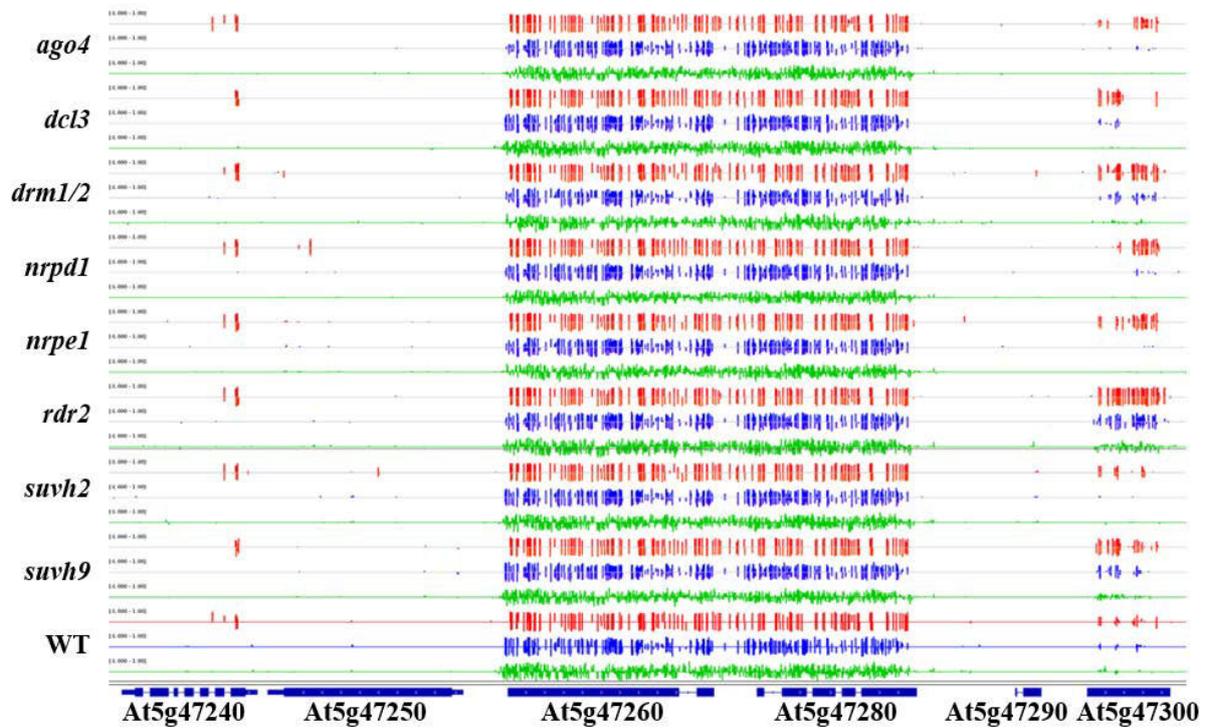


Figure S6. Methylation of *PbAt-5.2* region in Col-0 and in mutants in genes involved in RdDM methylation maintenance⁵¹. In red the methylation in CG context. In blue the methylation in CHG context; in green the methylation in CHH context. WT indicates the methylation profile of Col-0.

Table S1. List of natural *Arabidopsis* accessions showing a genomic like Col-0 structure in the region *PbAt-5.2*. Accession informations were derived of 1001 genome project data.

name	id	country	latitude	longitude	CS_number	group
LDV-46	139	FRA	48.5167	-4.06667	CS77014	western_europe
PYL-6	265	FRA	44.65	-1.16667	CS77198	admixed
TOU-A1-88	350	FRA	46.6667	4.11667	CS77382	western_europe
Doubravnik7	410	CZE	49.4211	16.3497	CS76808	central_europe
Draha2	424	CZE	49.4112	16.2815	CS76812	central_europe
Borky1	428	CZE	49.403	16.232	CS76718	central_europe
Gr-1	430	AUT	47	15.5	CS76496	central_europe
LI-OF-061	628	USA	40.7777	-72.9069	CS78951	western_europe
LIN_S-5	915	USA	41.8972	-71.4378	CS77040	germany
Ale-Stenar-64-24	1002	SWE	55.3833	14.05	CS76654	south_sweden
Brsarp-11-135	1061	SWE	55.7167	14.1333	CS76727	south_sweden
Brsarp-21-140	1063	SWE	55.7167	14.1333	CS76729	south_sweden
Aledal-6-49	1158	SWE	56.7	16.5167	CS76656	south_sweden
Aledal-14-73	1166	SWE	56.7	16.5167	CS77651	south_sweden
MNF-Pot-15	1852	USA	43.595	-86.2657	CS78979	germany
MNF-Pot-75	1872	USA	43.595	-86.2657	CS77100	germany
MNF-Pin-39	2016	USA	43.5356	-86.1788	CS77098	admixed
Pent-23	2202	USA	43.7623	-86.3929	CS77168	admixed
UKSE06-252	5104	UK	51.3	0.5	CS78800	germany
UKNW06-233	5486	UK	54.6	-3.3	CS78794	germany
Cal-2	5720	UK	53.3	-1.6	CS78781	admixed
For-2	5741	UK	56.6	-4.1	CS78783	western_europe
Kil-0	5748	UK	56	-4.4	CS78784	germany
Mc-1	5757	UK	54.6	-2.3	CS78785	admixed
UKID63	5768	UK	54.1	-1.5	CS78786	admixed
Set-1	5772	UK	54.1	-2.3	CS78787	admixed
UKID74	5779	UK	51	-3.1	CS78789	western_europe
UKID96	5800	UK	57.4	-5.5	CS78791	germany
UKID107	5811	UK	52.9	-3.1	CS78778	admixed
Bor-1	5837	CZE	49.4013	16.2326	CS76453	central_europe
DraII-6	5874	CZE	49.4112	16.2815	CS76814	central_europe
DraIV_5-12	5950	CZE	49.4112	16.2815	CS76820	central_europe
DraIV_6-13	5984	CZE	49.4112	16.2815	CS76822	central_europe
Eden-5	6010	SWE	62.877	18.177	CS78000	north_sweden
Eden-6	6011	SWE	62.877	18.177	CS76828	north_sweden
Eden-7	6012	SWE	62.877	18.177	CS76829	north_sweden
Fjñ1-5	6020	SWE	56.06	14.29	CS76861	south_sweden
Gro-3	6025	SWE	62.6437	17.7339	CS76889	north_sweden
Lv-1	6043	SWE	62.801	18.079	CS77049	north_sweden
Lv-5	6046	SWE	62.801	18.079	CS77050	north_sweden
Sr:3	6086	SWE	58.9	11.2	CS77267	admixed
T1020	6092	SWE	55.6514	13.2233	CS77289	south_sweden
T1040	6094	SWE	55.6494	13.2147	CS77290	south_sweden
T1110	6100	SWE	55.6	13.2	CS77294	south_sweden
T1160	6104	SWE	55.7	13.2	CS77296	south_sweden
T450	6105	SWE	55.7967	13.1211	CS77297	south_sweden
T470	6107	SWE	55.7942	13.1222	CS77299	south_sweden
T480	6108	SWE	55.7989	13.1206	CS77300	western_europe
T670	6122	SWE	55.8364	13.3075	CS77308	south_sweden
T680	6123	SWE	55.8369	13.3033	CS78064	south_sweden
T740	6128	SWE	55.8397	13.2881	CS77313	south_sweden
T810	6134	SWE	55.8383	13.2906	CS77318	south_sweden
TAA_03	6153	SWE	62.6425	17.7422	CS77329	north_sweden
TAA_04	6154	SWE	62.6422	17.7406	CS77330	north_sweden
TAA_14	6163	SWE	62.6425	17.7356	CS77331	north_sweden
TAA_17	6166	SWE	62.6425	17.7372	CS77332	north_sweden
TäL_03	6177	SWE	62.6322	17.69	CS77338	north_sweden
TäL_07	6180	SWE	62.6322	17.6906	CS77339	germany
TBû_01	6184	SWE	62.8892	18.4522	CS77343	north_sweden
TDr-2	6189	SWE	55.7686	14.1383	CS77351	south_sweden
TDr-9	6195	SWE	55.7708	14.1342	CS77356	south_sweden
TDr-18	6203	SWE	55.7714	14.1208	CS77350	south_sweden
TEDEN_03	6210	SWE	62.8839	18.1836	CS77359	north_sweden
TFä_04	6214	SWE	63.0175	18.3281	CS78121	north_sweden
TFä_06	6216	SWE	63.0167	18.3283	CS77362	north_sweden
TFä_07	6217	SWE	63.0169	18.3283	CS77363	north_sweden

TFä_08	6218	SWE	63.0172	18.3283	CS77364	north_sweden
Tomegap-2	6242	SWE	55.7	13.2	CS77377	south_sweden
TV-4	6252	SWE	55.5796	14.3336	CS78771	germany
TV-7	6255	SWE	55.5796	14.334	CS78772	germany
TV-10	6258	SWE	55.5796	14.3336	CS78767	admixed
TV-22	6268	SWE	55.5796	14.3336	CS78768	germany
TV-30	6276	SWE	55.5796	14.3336	CS78769	germany
UduI_1-11	6296	CZE	49.2771	16.6314	CS78774	central_europe
UduI_3-36	6390	CZE	49.2771	16.6314	CS78775	central_europe
Zdri_1-23	6424	CZE	49.3853	16.2544	CS78875	central_europe
Zdri_2-21	6445	CZE	49.3853	16.2544	CS78876	central_europe
ANH-1	6680	GER	51.85	6.4333	CS76437	western_europe
An-1	6898	BEL	51.2167	4.4	CS76435	admixed
Bil-5	6900	SWE	63.324	18.484	CS76709	north_sweden
Bil-7	6901	SWE	63.324	18.484	CS76710	north_sweden
Bor-4	6903	CZE	49.4013	16.2326	CS76454	central_europe
Br-0	6904	CZE	49.2	16.6166	CS76455	western_europe
Col-0	6909	USA	38.3	-92.3	CS76778	germany
Eden-2	6913	SWE	62.877	18.177	CS76827	north_sweden
Fib-2	6917	SWE	63.0165	18.3174	CS76850	north_sweden
Fib-4	6918	SWE	63.0165	18.3174	CS76851	north_sweden
Ga-0	6919	GER	50.3	8	CS76490	central_europe
LL-0	6933	ESP	41.59	2.49	CS77047	spain
NFA-10	6943	UK	51.4083	-0.6383	CS77126	western_europe
Pu2-7	6956	CZE	49.42	16.36	CS76580	central_europe
Pu2-8	6957	CZE	49.42	16.36	CS77192	central_europe
Rennes-1	6959	FRA	48.5	-1.41	CS77210	western_europe
Rennes-11	6960	FRA	48.5	-1.41	CS77211	western_europe
Se-0	6961	ESP	38.3333	-3.53333	CS76597	spain
Ull2-3	6973	SWE	56.0648	13.9707	CS78817	germany
Ak-1	6987	GER	48.0683	7.62551	CS76431	admixed
Alst-1	6989	UK	54.8	-2.4333	CS76432	western_europe
Aa-0	7000	GER	50.9167	9.57073	CS76428	germany
Bd-0	7013	GER	52.4584	13.287	CS76445	south_sweden
Ba-1	7014	UK	56.5459	-4.79821	CS76441	germany
Bl-1	7025	ITA	44.5041	11.3396	CS76450	central_europe
Boot-1	7026	UK	54.4	-3.2667	CS76452	western_europe
Bch-1	7028	GER	49.5166	9.3166	CS76444	admixed
Bsch-0	7031	GER	50.0167	8.6667	CS76457	germany
Bu-0	7036	GER	50.5	9.5	CS78889	germany
Bur-0	7058	IRL	54.1	-6.2	CS76734	admixed
Cal-0	7061	UK	53.2699	-1.64293	CS76460	admixed
Ca-0	7062	GER	50.2981	8.26607	CS76459	germany
Can-0	7063	ESP	29.2144	-13.4811	CS76740	relict
Chi-0	7072	RUS	53.7502	34.7361	CS76464	admixed
Co	7081	POR	40.2077	-8.42639	CS78895	spain
Da-0	7094	GER	49.8724	8.65081	CS76791	germany
Do-0	7102	GER	50.7224	8.2372	CS76474	germany
Dra-0	7103	CZE	49.4167	16.2667	CS76476	central_europe
Dr-0	7106	GER	51.051	13.7336	CS78897	central_europe
El-0	7117	GER	51.5105	9.68253	CS76479	germany
Er-0	7125	GER	49.5955	11.0087	CS78898	germany
Es-0	7126	FIN	60.1997	24.5682	CS76484	admixed
Gel-1	7143	NED	51.0167	5.86667	CS76492	germany
Gr-5	7158	AUT	47	15.5	CS76885	central_europe
Hn-0	7165	GER	51.3472	8.28844	CS76513	germany
Kil-0	7192	UK	55.6395	-5.66364	CS76526	germany
Krot-0	7203	GER	49.631	11.5722	CS76534	central_europe
Kyoto	7207	JPN	35.0085	135.752	CS76535	central_europe
Lan-0	7208	UK	55.6739	-3.78181	CS76539	germany
Li-2:1	7223	GER	50.3833	8.0666	CS76541	central_europe
Li-7	7231	GER	50.3833	8.0666	CS77035	germany
Mnz-0	7244	GER	50.001	8.26664	CS76552	germany
Np-0	7268	GER	52.6969	10.981	CS76563	admixed
Ob-0	7276	GER	50.2	8.5833	CS76566	germany
Old-1	7280	GER	53.1667	8.2	CS76567	germany
Oy-0	7288	NOR	60.38554	6.193019	CS77156	admixed
Petergof	7296	RUS	59	29	CS77170	central_europe
Pt-0	7305	GER	53.476	10.6065	CS78915	admixed
Pog-0	7306	CAN	49.2655	-123.206	CS76576	western_europe
Pn-0	7307	FRA	48.0653	-2.96591	CS77182	western_europe
Sten-0	7346	GER	52.6058	11.8558	CS77277	south_sweden
Tha-1	7353	NED	52.08	4.3	CS76611	admixed
Ting-1	7354	SWE	56.5	14.9	CS76612	south_sweden
Uk-1	7378	GER	48.0333	7.7667	CS76620	admixed
Utrecht	7382	NED	52.0918	5.1145	CS76622	admixed

Wa-1	7394	POL	52.3	21	CS76626	admixed
Jl-3	7424	CZE	49.2	16.6166	CS76519	central_europe
Nc-1	7430	FRA	48.6167	6.25	CS76559	admixed
Da(1)-12	7460	CZE			CS76470	central_europe
WAR	7477	USA	41.7302	-71.2825	CS78853	admixed
VNr2-6	7517	SWE	55.58	14.334	CS78831	germany
Lp2-2	7520	CZE	49.38	16.81	CS76546	central_europe
Lp2-6	7521	CZE	49.38	16.81	CS77052	central_europe
Lis-2	8222	SWE	56.0328	14.775	CS77043	south_sweden
Br1-6	8231	SWE	56.3	16	CS76726	south_sweden
Gul1-2	8234	SWE	56.4606	15.8127	CS76896	south_sweden
HSm	8236	CZE	49.33	15.76	CS76941	central_europe
PHW-3	8239	GER	51	7	CS76976	germany
PHW-2	8243	ITA	43.7703	11.2547	CS77173	western_europe
BÑ1-2	8256	SWE	56.4	12.9	CS76676	admixed
Bla-1	8264	ESP	41.6833	2.8	CS76451	spain
Dra3-1	8283	SWE	55.76	14.12	CS76811	south_sweden
DraII-1	8284	CZE	49.4112	16.2815	CS76813	central_europe
DraIII-1	8285	CZE	49.4112	16.2815	CS76815	central_europe
Hovdala-6	8307	SWE	56.1	13.74	CS76938	south_sweden
Lis-1	8326	SWE	56.0328	14.775	CS77042	south_sweden
Mir-0	8337	ITA	44	12.37	CS76551	western_europe
Na-1	8343	FRA	47.5	1.5	CS76558	admixed
Pla-0	8357	ESP	41.5	2.25	CS76573	spain
Rak-2	8365	CZE	49	16	CS77201	central_europe
Sanna-2	8376	SWE	62.69	18	CS77234	north_sweden
Sr:5	8386	SWE	58.9	11.2	CS77268	central_europe
Vinslv	9057	SWE	56.1	13.9167	CS78847	south_sweden
Xan-3	9067	AZE	38.6536	48.7992	CS78860	italy_balkan_caucasus
Xan-5	9069	AZE	38.6536	48.7992	CS78861	italy_balkan_caucasus
Xan-6	9070	AZE	38.6536	48.7992	CS78862	italy_balkan_caucasus
Lerik1-4	9075	AZE	38.7406	48.6131	CS77023	italy_balkan_caucasus
Lerik1-7	9078	AZE	38.7406	48.6131	CS77024	italy_balkan_caucasus
Nar-3	9089	AZE	38.9522	48.925	CS77119	italy_balkan_caucasus
Istisu-5	9095	AZE	38.9786	48.5594	CS76950	italy_balkan_caucasus
Lag1-2	9100	GEO	41.8296	46.2831	CS76998	italy_balkan_caucasus
Lag1-4	9102	GEO	41.8296	46.2831	CS76999	italy_balkan_caucasus
Lag1-5	9103	GEO	41.8296	46.2831	CS77000	italy_balkan_caucasus
Lag1-6	9104	GEO	41.8296	46.2831	CS77001	italy_balkan_caucasus
Lag1-7	9105	GEO	41.8296	46.2831	CS77002	italy_balkan_caucasus
Lag1-8	9106	GEO	41.8296	46.2831	CS77003	italy_balkan_caucasus
Lag2-6	9113	GEO	41.8296	46.2831	CS77006	italy_balkan_caucasus
Lag2-7	9114	GEO	41.8296	46.2831	CS77007	italy_balkan_caucasus
Lag2-10	9115	GEO	41.8296	46.2831	CS77004	italy_balkan_caucasus
Bak-5	9121	GEO	41.7942	43.4767	CS76685	italy_balkan_caucasus
Geg-14	9125	ARM	40.1408	44.8203	CS76876	asia
Gol-2	9314	UK	57.9672	-3.96722	CS76883	western_europe
Bt_1	9339	SWE	57.7133	15.0689	CS76720	south_sweden
Dju-1	9343	SWE	57.3089	18.1512	CS78896	admixed
Dd_2	9352	SWE	57.2608	16.3675	CS76797	south_sweden
Dd_3	9353	SWE	57.2608	16.3675	CS76798	south_sweden
FriL_1	9371	SWE	63.016	18.3175	CS76852	north_sweden
Fri_1	9381	SWE	55.8106	14.2091	CS76868	admixed
Fri_2	9382	SWE	55.8106	14.2091	CS76869	admixed
Fri_3	9383	SWE	55.8106	14.2091	CS76870	admixed
Grn_12	9386	SWE	62.806	18.1896	CS76891	north_sweden
Hadd-2	9391	SWE	57.3263	15.8979	CS76905	south_sweden
Hag-2	9394	SWE	56.5804	16.4063	CS76907	admixed
Hal-1	9395	SWE	57.5089	15.0105	CS76908	south_sweden
Kia_1	9409	SWE	56.0573	14.302	CS76968	south_sweden
Kru-3	9416	SWE	57.7215	18.3837	CS76986	south_sweden
Lan_1	9421	SWE	55.9745	14.3997	CS77009	south_sweden
Puk-1	9436	SWE	56.1633	14.6806	CS77194	south_sweden
Puk-2	9437	SWE	56.1633	14.6806	CS77195	south_sweden
Sim-1	9442	SWE	55.5678	14.3398	CS77250	germany
Spro_1	9450	SWE	57.2545	18.2109	CS77263	south_sweden
Stenk-3	9454	SWE	57.8009	18.5162	CS77275	south_sweden
Tur-4	9470	SWE	57.6511	14.8043	CS77399	south_sweden
VNrA_1	9476	SWE	55.5796	14.3336	CS78832	germany
Yst-1	9481	SWE	55.4242	13.8484	CS78869	south_sweden
11C1	9503	UK	55.8877	-3.21072	CS76640	western_europe
IP-Alo-0	9506	POR	40.11	-7.47	CS76662	admixed
IP-Coa-0	9507	POR	38.45	-7.5	CS76775	spain
IP-Mos-1	9508	POR	40.04	-7.11	CS77108	admixed
IP-Reg-0	9509	POR	39.29	-7.4	CS77207	spain
IP-Rei-0	9510	POR	38.75	-7.59	CS77208	spain

IP-Vav-0	9511	POR	38.53	-8.02	CS78835	spain
IP-Adm-0	9514	ESP	39.15	-4.54	CS76647	spain
IP-Ala-0	9515	ESP	39.72	-6.89	CS76650	spain
IP-Alm-0	9518	ESP	39.88	-0.36	CS76660	spain
IP-Ben-0	9523	ESP	38.37	-2.66	CS76700	admixed
IP-Cab-3	9526	ESP	41.54	2.39	CS76738	spain
IP-Cad-0	9527	ESP	40.37	-5.74	CS76739	admixed
IP-Car-1	9530	ESP	38.25	-4.32	CS76742	admixed
IP-Cem-0	9533	ESP	41.15	-4.32	CS76763	relict
IP-Cmo-3	9534	ESP	40.05	-4.65	CS76774	spain
IP-Coc-1	9535	ESP	42.31	3.19	CS76776	spain
IP-Cum-1	9537	ESP	38.07	-6.66	CS76787	spain
IP-Deh-1	9539	ESP	40.29	-6.67	CS76793	spain
IP-Fue-2	9541	ESP	38.26	-5.42	CS76871	spain
IP-Gra-0	9543	ESP	36.77	-5.39	CS76886	relict
IP-Her-12	9545	ESP	39.4	-5.78	CS76920	relict
IP-Hum-2	9549	ESP	42.23	-3.69	CS76943	relict
IP-Iso-4	9550	ESP	43.05	-5.37	CS76946	relict
IP-Lab-7	9552	ESP	40.87	-4.5	CS76995	admixed
IP-Mar-1	9555	ESP	39.58	-3.93	CS77068	relict
IP-Men-2	9556	ESP	39.66	-4.34	CS77081	spain
IP-Mon-5	9559	ESP	38.06	-4.38	CS77107	admixed
IP-Orb-10	9565	ESP	42.97	-1.23	CS77152	admixed
IP-Pal-0	9567	ESP	42.34	1.3	CS77159	spain
IP-Sdv-3	9581	ESP	42.84	-5.12	CS77242	admixed
IP-Ses-0	9582	ESP	41.48	-1.63	CS77244	spain
IP-Stp-0	9584	ESP	41.19	-3.58	CS77283	spain
IP-Svi-0	9585	ESP	43.4	-7.39	CS77287	western_europe
IP-Tam-0	9586	ESP	41.03	-3.27	CS77340	spain
IP-Tdc-0	9587	ESP	41.5	-1.88	CS77344	spain
IP-Tol-7	9588	ESP	42.11	0.6	CS77371	spain
IP-Trs-0	9590	ESP	43.37	-5.49	CS77387	western_europe
IP-Vae-2	9592	ESP	42.1	-5.44	CS78827	admixed
IP-Vin-0	9599	ESP	42.8	-5.77	CS78846	western_europe
IP-Vis-0	9600	ESP	39.85	-6.04	CS78848	relict
IP-Vpa-1	9602	ESP	40.5	-3.96	CS78850	spain
Karag-2	9608	RUS	51.37	59.44	CS76961	asia
Parti-1	9615	RUS	52.99	52.16	CS77163	asia
Krazo-1	9616	RUS	53.06	51.96	CS76984	asia
Chaba-2	9624	RUS	53.6	79.37	CS76767	asia
Lebja-2	9632	RUS	51.67	80.82	CS77016	asia
Masl-1	9634	RUS	54.13	81.31	CS77073	asia
Nosov-1	9635	RUS	51.87	80.6	CS77130	asia
Noveg-3	9638	RUS	51.73	80.86	CS77133	asia
Rakit-1	9640	RUS	51.87	80.06	CS77202	asia
Rakit-2	9641	RUS	51.9	80.06	CS77203	asia
Rakit-3	9642	RUS	51.84	80.06	CS77204	asia
Zupan-1	9644	CRO	45.07	18.72	CS78882	central_europe
Aiell-1	9646	ITA	39.1269	16.17019	CS76648	italy_balkan_caucasus
Basen-1	9647	ITA	40.37	16.77	CS76690	italy_balkan_caucasus
Bisig-1	9648	ITA	39.48	16.28	CS76712	italy_balkan_caucasus
Giffo-1	9653	ITA	38.44	16.13	CS76878	italy_balkan_caucasus
Marce-1	9655	ITA	38.92	16.47	CS77071	italy_balkan_caucasus
Mitterberg-3-187	9671	ITA	46.3668	11.2837	CS77088	central_europe
Mitterberg-3-188	9672	ITA	46.3668	11.2837	CS77089	central_europe
Mitterberg-3-189	9673	ITA	46.3668	11.2837	CS77090	central_europe
Mitterberg-4-193	9677	ITA	46.3718	11.2866	CS77094	central_europe
Castelfed-2-202	9686	ITA	46.3378	11.2928	CS76751	central_europe
Castelfed-3-205	9689	ITA	46.3378	11.2928	CS76754	central_europe
Castelfed-3-206	9690	ITA	46.3378	11.2928	CS76755	central_europe
Castelfed-3-208	9692	ITA	46.3378	11.2928	CS76757	central_europe
Castelfed-3-209	9693	ITA	46.3378	11.2928	CS76758	central_europe
Dolen-1	9697	BUL	41.62	23.94	CS76802	italy_balkan_caucasus
Goced-1	9698	BUL	41.57	23.85	CS76882	italy_balkan_caucasus
Dolna-1-10	9700	BUL	42.32	23.1	CS76803	italy_balkan_caucasus
Ivano-1	9701	BUL	43.7	25.91	CS76954	italy_balkan_caucasus
Melni-1	9703	BUL	41.53	23.39	CS77079	italy_balkan_caucasus
Melni-2	9704	BUL	41.53	23.39	CS77080	italy_balkan_caucasus
Podvi-1	9707	BUL	41.57	24.84	CS77187	italy_balkan_caucasus
Kardz-1	9708	BUL	41.62	25.35	CS76962	italy_balkan_caucasus
Zerev-1-34	9709	BUL	41.85	23.13	CS78878	italy_balkan_caucasus
Dolna-1-39	9711	BUL	42.32	23.1	CS76804	italy_balkan_caucasus
Dolna-1-40	9712	BUL	42.32	23.1	CS76805	italy_balkan_caucasus
Leska-1-44	9716	BUL	41.54	24.98	CS77030	italy_balkan_caucasus
Kardz-2	9717	BUL	41.66	25.47	CS76963	italy_balkan_caucasus
Groch-1	9722	BUL	41.71	24.41	CS76890	italy_balkan_caucasus

Epidaurus-1	9725	GRC	37.6	23.08	CS76844	italy_balkan_caucasus
Faneronemi-3	9726	GRC	37.07	22.04	CS76853	italy_balkan_caucasus
Stiav-2	9729	SVK	48.46	18.9	CS77280	central_europe
Stiav-3	9731	SVK	48.46	18.9	CS77281	central_europe
Bela-2	9733	SVK	48.47	18.94	CS76697	central_europe
Bran-1	9738	ROU	45.57	25.42	CS76722	admixed
Orast-1	9741	ROU	45.84	23.16	CS77151	central_europe
Iasi-1	9744	ROU	47.16	27.59	CS76944	admixed
Zagub-1	9748	SRB	44.23	21.71	CS78871	admixed
Knjas-1	9749	SRB	43.54	22.29	CS76971	italy_balkan_caucasus
Staro-1	9757	SRB	44.3	21.08	CS77272	italy_balkan_caucasus
Anz-0	9759	IRN	37.47	49.47	CS76439	italy_balkan_caucasus
Etna-2	9762	ITA	37.69	14.98	CS76487	admixed
Qar-8a	9764	LBN	34.1	35.84	CS76581	admixed
HE-1	9769	GER	48.55	8.99	CS76916	central_europe
KBG2-13	9770	GER	48.53	9.01	CS76966	central_europe
Pfn-N2.2-6	9771	GER	48.56	9.11	CS77172	central_europe
Gn-1	9777	GER	48.57	9.17	CS76880	central_europe
Bai-10	9779	GER	48.5	8.78	CS76682	central_europe
Fell2-4	9780	GER	48.43	8.79	CS76856	central_europe
Kus2-2	9781	GER	48.52	9.11	CS76990	central_europe
Erg2-6	9784	GER	48.5	8.8	CS76845	central_europe
Ha-HBT1-2	9785	GER	48.54	9.02	CS76898	central_europe
Haes-1	9791	GER	48.6	9.2	CS76914	central_europe
Ha-HBT2-10	9797	GER	48.54	9.02	CS76899	central_europe
Ha-P2-1	9798	GER	48.54	9.01	CS76902	central_europe
Pfn-10	9805	GER	48.54	9.09	CS77171	central_europe
Schl-7	9807	GER	48.6	9.22	CS77240	central_europe
Tu-B2-3	9808	GER	48.52	9.08	CS77392	central_europe
Tu-KB-6	9809	GER	48.52	9.05	CS77393	western_europe
Tu-W1	9812	GER	48.52	9.03	CS77397	central_europe
BI-4	9813	GER	48.4	8.77	CS76706	central_europe
Ha-HBT3-11	9815	GER	48.54	9.02	CS76900	central_europe
IP-Ace-0	9817	ESP	39.84	-6.6	CS76642	spain
IP-Anu-0	9819	ESP	42.35	-3.03	CS76664	spain
IP-Are-0	9820	ESP	41	-4.71	CS76671	spain
IP-Aru-0	9821	ESP	41.81	2.49	CS76674	spain
IP-Aul-0	9822	ESP	40.52	-4.02	CS76675	spain
IP-Bae-0	9823	ESP	43.34	-5.84	CS76681	western_europe
IP-Boa-0	9825	ESP	40.4	-3.88	CS76714	spain
IP-Bos-0	9827	ESP	42.78	0.69	CS76719	western_europe
IP-Cas-0	9831	ESP	38.54	-3.39	CS76743	admixed
IP-Cat-0	9832	ESP	40.54	-3.69	CS76759	relict
IP-Cha-0	9833	ESP	40.38	-4.21	CS76764	spain
IP-Cho-0	9834	ESP	40.51	-3.9	CS76768	spain
IP-Dar-0	9840	ESP	41.13	-1.43	CS76792	spain
IP-Ees-0	9841	ESP	40.59	-4.15	CS76836	spain
IP-Elp-0	9843	ESP	40.53	-3.92	CS76840	spain
IP-Evs-0	9845	ESP	40.48	-3.96	CS76848	spain
IP-Hue-3	9851	ESP	42.96	-6.1	CS76942	western_europe
IP-Ini-0	9852	ESP	40.46	-3.75	CS76945	spain
IP-Lam-0	9855	ESP	40.57	-3.89	CS77008	spain
IP-Lch-0	9856	ESP	40.51	-4	CS77010	spain
IP-Leg-0	9857	ESP	40.33	-3.8	CS77019	spain
IP-Lro-0	9859	ESP	40.5	-3.88	CS77054	spain
IP-Lum-0	9860	ESP	42.24	-2.62	CS77059	spain
IP-Mac-0	9861	ESP	40.72	-3.21	CS77061	spain
IP-Mad-0	9862	ESP	40.45	-3.67	CS77062	western_europe
IP-Mie-1	9867	ESP	40.94	-3.22	CS77083	spain
IP-Moj-0	9869	ESP	36.76	-5.28	CS77105	relict
IP-Pdl-0	9877	ESP	43.02	-5.6	CS77165	admixed
IP-Pee-0	9878	ESP	40.78	-3.62	CS77167	spain
IP-Per-0	9879	ESP	37.6	-1.12	CS77169	relict
IP-Pib-1	9880	ESP	42.72	-3.44	CS77175	spain
IP-Pil-0	9882	ESP	40.46	-4.26	CS77178	spain
IP-Pva-1	9888	ESP	40.93	-3.31	CS77197	spain
IP-Sen-0	9894	ESP	42.59	0.76	CS77243	admixed
IP-Sfb-6	9895	ESP	41.78	2.57	CS77248	spain
IP-Smt-1	9897	ESP	40.95	-5.63	CS77257	admixed
IP-Som-0	9898	ESP	41.14	-3.58	CS77259	spain
IP-Tau-0	9899	ESP	42.54	0.84	CS77342	spain
IP-Tri-0	9900	ESP	37.38	-6.01	CS77386	spain
IP-Urd-1	9901	ESP	42.27	-2.98	CS78824	spain
IP-Usa-0	9902	ESP	40.71	-3.24	CS78825	spain
IP-Vas-0	9904	ESP	40.95	-3.31	CS78833	spain
IP-Mah-6	9906	ESP	40	4.25	CS77063	admixed

ESP-1-11	9908	FRA	50.72	3.47	CS76847	western_europe
BRI-2	9910	FRA	50.68	3.52	CS76725	western_europe
ARGE-1-15	9911	FRA	47.16	4.28	CS76672	western_europe
SAUL-24	9918	FRA	47.43	5.21	CS77237	western_europe
RUM-20	9925	FRA	48.91	4.52	CS77226	western_europe
LEC-25	9930	FRA	43.91	4.14	CS77018	central_europe
Leo-1	9945	ESP	41.8	-3.11	CS76413	admixed
Mer-6	9946	ESP	38.92	-6.34	CS76414	spain
Ped-0	9947	ESP	40.74	-3.9	CS76415	relict
Stepn-2	9955	RUS	54.09	60.46	CS76377	asia
Borsk-2	9957	RUS	53.04	51.75	CS76421	asia
Valsi-1	9969	ITA	40.18	16.45	CS76425	italy_balkan_caucasus
Bozen-1.2	9972	ITA	46.513	11.331	CS76358	central_europe
Rovero-1	9976	ITA	46.2543	11.167	CS76351	central_europe
Angit-1	9981	ITA	38.76	16.24	CS76366	italy_balkan_caucasus
Ciste-1	9983	ITA	41.62	12.87	CS76359	admixed
Ciste-2	9984	ITA	41.62	12.87	CS76360	admixed
Slavi-1	9985	BUL	41.43	23.65	CS76419	italy_balkan_caucasus
Jablo-1	9986	BUL	41.59	25.2	CS76372	italy_balkan_caucasus
Lag2-2	9990	GEO	41.8296	46.2831	CS76390	italy_balkan_caucasus
TueV-13	10001	GER	48.52	9.05	CS76407	western_europe
Istisu-1	10012	AZE	38.9786	48.5594	CS76389	italy_balkan_caucasus
Xan-1	10014	AZE	38.6536	48.7992	CS76387	italy_balkan_caucasus
Dobra-1	10018	SRB	44.84	20.16	CS76369	central_europe
Jl-2	10020	CZE	49.17	16.5	CS76956	central_europe
Strand-1	10023	NOR	68.8	15.45	CS77284	admixed
OOE3-1	15592	AUT	48.33147	14.71587	CS78940	central_europe
OOE3-2	15593	AUT	48.33147	14.71587	CS78941	central_europe
OOE2-1	19949	AUT	48.32683	14.71817	CS79036	central_europe
OOE2-2	19950	AUT	48.32683	14.71817	CS79037	central_europe
OOE23	19951	AUT	48.32683	14.71817	CS79038	central_europe

Table S2. Frequency of *PbAt-5.2* genomic structure and methylation state according to geographic groups. (A) Proportion of Col-0/Bur-0 structure like at At5g47260/At5g47280 region among the 1135 accessions of the 1001 genome project. **(B)** Proportion of Col-0-methylation like at At5g47260/At5g47280 region among the 287 Col-0/Bur-0 structure-like accessions of the 1001 genome project.

A

Arabidopsis group	Nb accessions with Col/Bur-structure like at At5g47260/280	Total number of accession per category	% Col/Bur-structure like accessions per Arabidopsis group
admixed	57	137	41.6
asia	14	79	17.7
central_europe	75	184	40.8
germany	35	171	20.5
italy_balkan_caucasus	48	92	52.2
north_sweden	25	64	39.1
relict	12	25	48.0
south_sweden	44	156	28.2
spain	58	110	52.7
western_europe	33	117	28.2
TOTAL	401	1135	35.3

B

Group	Bur-like methylation	Col-like methylation	others	total	% Col-like methylation among Col/Bur-structure like accessions	% Col-like structure AND methylation among 1135 accessions	% Bur-like structure AND methylation among 1135 accessions
admixed	36	3	1	40	7.9	3.28	38.3
asia	13	1	0	14	7.1	1.27	16.5
central_europe	23	30	2	55	54.5	22.23	18.5
germany	8	9	0	17	52.9	10.84	9.6
italy_balkan_caucasus	40	3	0	43	7.0	3.64	48.5
north_sweden	10	0	0	10	0	0.00	39.1
relict	11	0	0	11	0	0.00	48.0
south_sweden	12	2	0	14	14.3	4.03	24.2
spain	51	2	1	54	3.7	1.95	50.8
western_europe	22	7	0	29	24.1	6.81	21.4
TOTAL	141	46	38	190	24.2	8.55	26.8

Table S3. List of *Arabidopsis* accessions with Col-0 like or Bur-0 like epimutation state. Methylation level was calculated on the interval of -1 kb before At5g47260 to +1 kb after At5g47280. Epimutation type was indicated according to the clustering obtained (Figure IV.6C). ID indicates the login number of the *Arabidopsis*. CG indicate the CG methylation level. CHG indicate the CHG methylation level. CHH indicate the CHH methylation level

ID	CG	CHG	CHH	Epimutation type
410	0.7753108	0.54875752	0.20438119	Col-0 like
424	0.81174136	0.50244365	0.18517055	Col-0 like
428	0.8457432	0.64093603	0.24412336	Col-0 like
430	0.82450898	0.62159857	0.22852184	Col-0 like
915	0.85990306	0.57014052	0.16520303	Col-0 like
5104	0.77983175	0.54049835	0.19182033	Col-0 like
5486	0.8135243	0.60908674	0.24749872	Col-0 like
5748	0.74809756	0.57148939	0.23411721	Col-0 like
5779	0.79304864	0.50584818	0.19946865	Col-0 like
5874	0.84784964	0.59976596	0.2019615	Col-0 like
5984	0.80185488	0.51471763	0.17373452	Col-0 like
6107	0.76180405	0.53220115	0.17002244	Col-0 like
6296	0.79102351	0.56236858	0.16022947	Col-0 like
6390	0.74428206	0.54093333	0.18843957	Col-0 like
6445	0.80945838	0.49174961	0.18858085	Col-0 like
6680	0.72336316	0.47302431	0.14664895	Col-0 like
6909	0.84508883	0.6453762	0.30078247	Col-0 like
6956	0.75839201	0.52695061	0.19268008	Col-0 like
6957	0.86628246	0.63416019	0.26328867	Col-0 like
6973	0.83517098	0.53994682	0.22643396	Col-0 like
7013	0.8195499	0.56604296	0.19824651	Col-0 like
7026	0.55805673	0.27072502	0.11537027	Col-0 like
7094	0.79958138	0.54257346	0.19045134	Col-0 like
7103	0.79891327	0.49652295	0.19953313	Col-0 like
7106	0.81985071	0.5861668	0.19936823	Col-0 like
7158	0.80073344	0.53301403	0.16836396	Col-0 like
7244	0.80189132	0.50857985	0.18975096	Col-0 like
7424	0.7920344	0.50609624	0.17314575	Col-0 like
7430	0.84908473	0.59133353	0.26047561	Col-0 like
7477	0.79158286	0.52786959	0.18632183	Col-0 like
7521	0.77225135	0.5599462	0.20837347	Col-0 like
8236	0.81944086	0.56181146	0.20917081	Col-0 like
8239	0.63920143	0.39584212	0.16122027	Col-0 like
8285	0.81150197	0.54823322	0.23173841	Col-0 like
8343	0.79816476	0.51141069	0.22463718	Col-0 like
8365	0.76961479	0.53672498	0.17200957	Col-0 like
9121	0.82043976	0.56610216	0.19724543	Col-0 like
9584	0.75560627	0.51973745	0.16268903	Col-0 like
9616	0.77342998	0.55450726	0.17374667	Col-0 like
9716	0.52764212	0.37661623	0.12610664	Col-0 like
9722	0.79550705	0.54672489	0.180345	Col-0 like
9731	0.83229989	0.5623722	0.17181175	Col-0 like
9769	0.79990418	0.52998092	0.2027516	Col-0 like
9779	0.80138463	0.55823811	0.21070446	Col-0 like
9780	0.78385846	0.55234601	0.22601751	Col-0 like
9781	0.78519618	0.53560517	0.18511313	Col-0 like
9785	0.7446359	0.51232772	0.14980665	Col-0 like
9798	0.75927611	0.51542841	0.16850392	Col-0 like
9808	0.82086445	0.52305399	0.07626947	Col-0 like
9809	0.70949427	0.46708279	0.15288726	Col-0 like
9812	0.73418941	0.41969699	0.13620419	Col-0 like
9813	0.82482738	0.53615678	0.23285039	Col-0 like
9815	0.79224951	0.52723315	0.2271669	Col-0 like
9898	0.81293944	0.50571788	0.21082465	Col-0 like
9908	0.79017833	0.55008835	0.20790653	Col-0 like
9925	0.83627485	0.53757729	0.22291216	Col-0 like
10001	0.77983236	0.53355192	0.18313117	Col-0 like
10020	0.79497872	0.4970872	0.21110546	Col-0 like
265	0.18960142	0.00402903	0.00464801	Bur-0 like
350	0.27945621	0.00521332	0.004075	Bur-0 like

1002	0.26422709	0.00217506	0.00342726	Bur-0 like
1158	0.16901505	0.0037716	0.00392648	Bur-0 like
1872	0.2385323	0.00586803	0.01259992	Bur-0 like
2016	0.2074762	0.00311486	0.00424868	Bur-0 like
2202	0.19350011	0.00692096	0.00808374	Bur-0 like
5720	0.21398046	0.00625978	0.0029624	Bur-0 like
5741	0.18209186	0.01028391	0.003855	Bur-0 like
5757	0.21084042	0.0034783	0.00499672	Bur-0 like
5772	0.20646637	0.00424923	0.00317777	Bur-0 like
5800	0.26107304	0.00931972	0.00356397	Bur-0 like
5811	0.00367647	0.00393682	0.00906465	Bur-0 like
5837	0.2663838	0.01031805	0.00582426	Bur-0 like
5950	0.28390507	0.00376999	0.00394349	Bur-0 like
6025	0.23116081	0.00288222	0.00453831	Bur-0 like
6086	0.18468645	0.00351112	0.0040535	Bur-0 like
6108	0.21926987	0.0046386	0.00408907	Bur-0 like
6154	0.24252202	0.00438619	0.00414038	Bur-0 like
6163	0.22377166	0.00928597	0.0086525	Bur-0 like
6166	0.2354933	0.00222081	0.00432915	Bur-0 like
6177	0.32945579	0.00470813	0.00285794	Bur-0 like
6195	0.14032794	0.00677432	0.00400657	Bur-0 like
6203	0.05364425	0.0017777	0.00388618	Bur-0 like
6217	0.2975652	0.00747762	0.00267472	Bur-0 like
6218	0.25307173	0.00301994	0.00271496	Bur-0 like
6255	0.23753712	0.0038823	0.00398556	Bur-0 like
6268	0.2581363	0.00403883	0.00308881	Bur-0 like
6900	0.24851817	0.00586511	0.00420033	Bur-0 like
6917	0.26116357	0.00387745	0.0043821	Bur-0 like
6920	0	0.00280112	0.00113636	Bur-0 like
6933	0.23084656	0.0021164	0.00322088	Bur-0 like
6943	0.02045884	0.00265613	0.00337215	Bur-0 like
6959	0.26621053	0.00438723	0.0034639	Bur-0 like
6960	0.24676579	0.004013	0.00483791	Bur-0 like
6987	0.29064627	0.00589554	0.00231058	Bur-0 like
6989	0.2575069	0.00251418	0.00310868	Bur-0 like
7028	0.23748559	0.00561762	0.00433335	Bur-0 like
7058	0.21924898	0.01093594	0.00249537	Bur-0 like
7063	0.16724567	0.00804935	0.00646453	Bur-0 like
7081	0.24694283	0.00389188	0.00715494	Bur-0 like
7102	0.25886712	0.00623491	0.00476594	Bur-0 like
7125	0.06198616	0.00295545	0.00853184	Bur-0 like
7165	0.19497079	0.00235163	0.0031337	Bur-0 like
7231	0.26003908	0.00892677	0.0165103	Bur-0 like
7288	0.23693419	0.00554432	0.00371054	Bur-0 like
7296	0.25160796	0.00732673	0.00973279	Bur-0 like
7305	0.32404853	0.12940803	0.05308662	Bur-0 like
7307	0.24007937	0.01688713	0.00158349	Bur-0 like
7346	0.19353865	0.00202492	0.00381277	Bur-0 like
7353	0.27852385	0.006991	0.00441488	Bur-0 like
7354	0.04637188	0.00499621	0.00281973	Bur-0 like
8231	0.17733905	0.00380538	0.00409151	Bur-0 like
8234	0.23868291	0.00680989	0.00543864	Bur-0 like
8243	0.23808785	0.0020645	0.00295529	Bur-0 like
8284	0.31847298	0.00841935	0.00610003	Bur-0 like
8326	0.01072727	0.00345719	0.00386201	Bur-0 like
8337	0.2552037	0.00344955	0.00303947	Bur-0 like
8357	0.21599879	0.01136761	0.01227515	Bur-0 like
8386	0.54407783	0.08496504	0.00389211	Bur-0 like
9067	0.29389522	0.01109547	0.00779884	Bur-0 like
9069	0.06761677	0.00271308	0.00358178	Bur-0 like
9070	0.09794235	0.00265392	0.00538272	Bur-0 like
9075	0.08594328	0.00365722	0.00340793	Bur-0 like
9078	0.06910823	0.00289543	0.00483589	Bur-0 like
9089	0.22752641	0.00460168	0.00318634	Bur-0 like
9095	0.27063932	0.00347645	0.002678	Bur-0 like
9100	0.24459064	0.00549233	0.00371063	Bur-0 like
9102	0.30525888	0.00677824	0.00654068	Bur-0 like
9103	0.20379179	0.00116094	0.00360017	Bur-0 like
9104	0.26687146	0.00375907	0.00418764	Bur-0 like
9105	0.20885482	0.00392487	0.00370457	Bur-0 like
9106	0.20382085	0.00594523	0.00306846	Bur-0 like
9113	0.28656064	0.00097902	0.00496253	Bur-0 like
9114	0.19647598	0.00251223	0.00325124	Bur-0 like
9115	0.18761557	0.00288369	0.00350075	Bur-0 like
9125	0.30239347	0.00590445	0.00448331	Bur-0 like
9314	0.05020777	0.05509984	0.00730531	Bur-0 like

9386	0.32500958	0.00324869	0.00260546	Bur-0 like
9409	0.22326081	0.00300772	0.00445355	Bur-0 like
9421	0.04330814	0.00628194	0.00386498	Bur-0 like
9450	0.27595702	0.00586476	0.00365756	Bur-0 like
9503	0.31298068	0.00327627	0.00490439	Bur-0 like
9506	0.22165731	0.00266417	0.00498302	Bur-0 like
9507	0	0.00269993	0.00527809	Bur-0 like
9508	0.01319444	0.00135802	0.00385263	Bur-0 like
9509	0.26866063	0.00348048	0.00356244	Bur-0 like
9510	0.28663861	0.00456423	0.00445276	Bur-0 like
9511	0.18886981	0.00461972	0.00264884	Bur-0 like
9514	0.26771815	0.00588468	0.00395246	Bur-0 like
9515	0.2221067	0.0058716	0.00387114	Bur-0 like
9518	0.16049942	0.00380204	0.0037905	Bur-0 like
9523	0.28180861	0.00259282	0.00486977	Bur-0 like
9526	0.03329604	0.00577036	0.00345476	Bur-0 like
9527	0.04460709	0.00306897	0.00438995	Bur-0 like
9530	0.2024623	0.00363426	0.00536398	Bur-0 like
9533	0.23689767	0.00827581	0.0026446	Bur-0 like
9534	0.22499448	0.00647803	0.00354458	Bur-0 like
9535	0.22961756	0.00618578	0.00309539	Bur-0 like
9537	0.18402208	0.00314743	0.0037633	Bur-0 like
9541	0.15333803	0.0066385	0.00424489	Bur-0 like
9543	0.22582785	0.00171893	0.00392408	Bur-0 like
9545	0.19501495	0.00304061	0.00388967	Bur-0 like
9549	0.21292509	0.00418635	0.00356363	Bur-0 like
9550	0.21308374	0.00517206	0.00425985	Bur-0 like
9552	0.14210712	0.00204531	0.0031216	Bur-0 like
9556	0.18406504	0.00477923	0.00419848	Bur-0 like
9559	0.21856648	0.00310507	0.00426284	Bur-0 like
9565	0.20178971	0.00400121	0.00470431	Bur-0 like
9567	0.18150243	0.00426038	0.0042607	Bur-0 like
9581	0.1889676	0.00739573	0.00462141	Bur-0 like
9582	0.31244663	0.00442386	0.00370545	Bur-0 like
9585	0.22527543	0.00823437	0.00430403	Bur-0 like
9586	0.2302093	0.0058358	0.00336108	Bur-0 like
9587	0.16380871	0.00262978	0.00341238	Bur-0 like
9588	0.21970956	0.00773843	0.00400371	Bur-0 like
9590	0.21791899	0.002167	0.00291042	Bur-0 like
9592	0.0509105	0.00934986	0.00254881	Bur-0 like
9599	0.16965952	0.00484818	0.00562846	Bur-0 like
9600	0.23972561	0.00862018	0.00415864	Bur-0 like
9608	0.19238442	0.00567812	0.00280147	Bur-0 like
9615	0.02702703	0.01210819	0.0024992	Bur-0 like
9624	0	0.00461057	0.00370276	Bur-0 like
9632	0.20562695	0.00358135	0.00252736	Bur-0 like
9634	0.1838027	0.00499688	0.00466308	Bur-0 like
9635	0.21526628	0.00146042	0.00299661	Bur-0 like
9638	0.22086931	0.00274195	0.00184279	Bur-0 like
9640	0.16041235	0.00192075	0.00132529	Bur-0 like
9641	0.19178118	0.00436486	0.0032514	Bur-0 like
9642	0.19329574	0.00219269	0.00513765	Bur-0 like
9644	0.03244145	0.00128791	0.00400423	Bur-0 like
9646	0.303462	0.00420975	0.00269696	Bur-0 like
9647	0.21380195	0.00362528	0.00253616	Bur-0 like
9648	0.03666667	0.00211864	0.00256714	Bur-0 like
9653	0.26667457	0.00663061	0.00367827	Bur-0 like
9655	0.12862166	0.00564039	0.00510043	Bur-0 like
9671	0.23781872	0.00451061	0.00240042	Bur-0 like
9689	0.22458937	0.0033106	0.00382506	Bur-0 like
9690	0.22138876	0.00471883	0.00450071	Bur-0 like
9693	0.20075157	0.00392903	0.00439678	Bur-0 like
9697	0.03407224	0.00535948	0.00503775	Bur-0 like
9698	0.24410237	0.00582455	0.00410486	Bur-0 like
9700	0.18780645	0.00215046	0.00406439	Bur-0 like
9701	0.2893456	0.00231236	0.00365713	Bur-0 like
9703	0.24536796	0.00289524	0.00409647	Bur-0 like
9704	0.22434932	0.00139164	0.0029908	Bur-0 like
9707	0.18039278	0.00369346	0.00233012	Bur-0 like
9708	0.17810966	0.00669348	0.00434147	Bur-0 like
9709	0.29612765	0.00339076	0.00403045	Bur-0 like
9717	0.31708739	0.00342193	0.0033808	Bur-0 like
9725	0.04977169	0.0103842	0.00445888	Bur-0 like
9726	0.24731256	0.00576935	0.00305042	Bur-0 like
9729	0.04396648	0.00121456	0.00328033	Bur-0 like
9733	0.22097508	0.00521988	0.00404021	Bur-0 like

9738	0.25644339	0.00521561	0.00332107	Bur-0 like
9741	0.25521691	0.00678893	0.00264841	Bur-0 like
9744	0.25280865	0.00448202	0.00330346	Bur-0 like
9748	0.0333505	0.00261814	0.00163981	Bur-0 like
9749	0.09336909	0.00349147	0.00347249	Bur-0 like
9762	0.07416154	0.00357123	0.00508378	Bur-0 like
9770	0.0018315	0.00284485	0.0031647	Bur-0 like
9771	0.33242289	0.0074337	0.00318589	Bur-0 like
9777	0.22546534	0.00243591	0.00474765	Bur-0 like
9784	0.31958466	0.00302226	0.00344611	Bur-0 like
9791	0.24189755	0.00367482	0.00302178	Bur-0 like
9797	0.35719916	0.00472322	0.00435188	Bur-0 like
9805	0.32845743	0.00430097	0.00447871	Bur-0 like
9807	0.22916488	0.00534525	0.00345576	Bur-0 like
9817	0.24200461	0.00279245	0.00388351	Bur-0 like
9819	0.02376347	0.0024402	0.00456764	Bur-0 like
9820	0.09442546	0.00560767	0.00450356	Bur-0 like
9821	0.24742445	0.0040906	0.00305195	Bur-0 like
9822	0.23160031	0.00407594	0.00339769	Bur-0 like
9823	0.21489772	0.00783969	0.00460904	Bur-0 like
9825	0.28388385	0.00680932	0.00316586	Bur-0 like
9827	0.31434059	0.0043701	0.00401182	Bur-0 like
9831	0.19412632	0.00398856	0.00245224	Bur-0 like
9832	0.20821761	0.00431672	0.00429535	Bur-0 like
9833	0.24737518	0.00476461	0.00363853	Bur-0 like
9834	0.16050129	0.00361951	0.00352791	Bur-0 like
9840	0.21578349	0.00441709	0.00436313	Bur-0 like
9841	0.2297241	0.00435087	0.00255488	Bur-0 like
9843	0.2313566	0.0058824	0.00478246	Bur-0 like
9845	0.23487335	0.00439568	0.00337459	Bur-0 like
9851	0.26774732	0.00440711	0.00366609	Bur-0 like
9852	0.35456953	0.13686941	0.05483145	Bur-0 like
9855	0.2200947	0.0062232	0.00376551	Bur-0 like
9856	0.23649632	0.00526779	0.00346334	Bur-0 like
9857	0.2403788	0.00386348	0.00303569	Bur-0 like
9859	0.1917782	0.00459935	0.00370005	Bur-0 like
9860	0.21404115	0.00276777	0.00420435	Bur-0 like
9861	0.18536598	0.00578822	0.0036462	Bur-0 like
9862	0.28047489	0.00233005	0.00374311	Bur-0 like
9867	0.19126378	0.0052431	0.00466808	Bur-0 like
9869	0.2682507	0.00181957	0.00354329	Bur-0 like
9877	0.1876643	0.00600446	0.00424253	Bur-0 like
9878	0.15715788	0.00333391	0.0026732	Bur-0 like
9879	0.25491823	0.00531344	0.00487704	Bur-0 like
9880	0.25439225	0.00394917	0.00248865	Bur-0 like
9882	0.298387	0.00478357	0.00453888	Bur-0 like
9888	0.02826786	0.00565307	0.00481332	Bur-0 like
9894	0.19217209	0.00243595	0.00409667	Bur-0 like
9895	0.2544371	0.00456044	0.00478748	Bur-0 like
9897	0.16982977	0.00472975	0.00521814	Bur-0 like
9899	0.18463638	0.00779345	0.0039291	Bur-0 like
9900	0.22398786	0.00321081	0.00792686	Bur-0 like
9901	0.17027007	0.0054977	0.00296149	Bur-0 like
9902	0.05796041	0.00490745	0.00531321	Bur-0 like
9904	0.19305046	0.00637729	0.00721795	Bur-0 like
9906	0.04835806	0.00576069	0.00299467	Bur-0 like
9910	0.22077333	0.01021927	0.00452103	Bur-0 like
9911	0.05769231	0.00328341	0.00376306	Bur-0 like
9918	0.36578058	0.00985535	0.01049725	Bur-0 like
9945	0.23738007	0.00675354	0.00492017	Bur-0 like
9946	0.20348062	0.00485648	0.00387309	Bur-0 like
9947	0.24239003	0.01343998	0.00566127	Bur-0 like
9955	0.23169068	0.00207203	0.00449102	Bur-0 like
9957	0.09484025	0.01623453	0.00519808	Bur-0 like
9969	0.28885354	0.00661338	0.00336482	Bur-0 like
9972	0.02371114	0.00279748	0.00667992	Bur-0 like
9976	0.21470129	0.00455081	0.0034687	Bur-0 like
9981	0.26801266	0.00187768	0.00297178	Bur-0 like
9983	0.22783552	0.0034247	0.00372013	Bur-0 like
9985	0.06457316	0.01428571	0.00270962	Bur-0 like
9986	0.17054747	0.00395532	0.00431271	Bur-0 like
9990	0.22175284	0.00833726	0.00617139	Bur-0 like
10014	0.28613743	0.00410999	0.00381354	Bur-0 like
10018	0.11411509	0.01028957	0.0057741	Bur-0 like
10023	0.2338396	0.00301604	0.00293356	Bur-0 like

Table S4. List of 127 *Arabidopsis* accessions used in pathological test

name	id	country	latitude	longitude	CS_number	group	Epimutation type
PYL-6	265	FRA	44.65	-1.16667	CS77198	admixed	Bur-0 like
TOU-A1-88	350	FRA	46.6667	4.11667	CS77382	western_europe	Bur-0 like
Doubravnik7	410	CZE	49.4211	16.3497	CS76808	central_europe	Col-0 like
Draha2	424	CZE	49.4112	16.2815	CS76812	central_europe	Col-0 like
Borky1	428	CZE	49.403	16.232	CS76718	central_europe	Col-0 like
LIN_S-5	915	USA	41.8972	-71.4378	CS77040	germany	Col-0 like
MNF-Pot-75	1872	USA	43.595	-86.2657	CS77100	germany	Bur-0 like
MNF-Pin-39	2016	USA	43.5356	-86.1788	CS77098	admixed	Bur-0 like
UKSE06-252	5104	UK	51.3	0.5	CS78800	germany	Col-0 like
Cal-2	5720	UK	53.3	-1.6	CS78781	admixed	Bur-0 like
Set-1	5772	UK	54.1	-2.3	CS78787	admixed	Bur-0 like
UKID74	5779	UK	51	-3.1	CS78789	western_europe	Col-0 like
DraII-6	5874	CZE	49.4112	16.2815	CS76814	central_europe	Col-0 like
DraIV_6-13	5984	CZE	49.4112	16.2815	CS76822	central_europe	Col-0 like
Fjñ1-5	6020	SWE	56.06	14.29	CS76861	south_sweden	Bur-0 like
TAA_14	6163	SWE	62.6425	17.7356	CS77331	north_sweden	Bur-0 like
UduI_1-11	6296	CZE	49.2771	16.6314	CS78774	central_europe	Col-0 like
ZdrI_2-21	6445	CZE	49.3853	16.2544	CS78876	central_europe	Col-0 like
Col-0	6909	USA	38.3	-92.3	CS76778	germany	Col-0 like
LL-0	6933	ESP	41.59	2.49	CS77047	spain	Bur-0 like
Pu2-8	6957	CZE	49.42	16.36	CS77192	central_europe	Col-0 like
Rennes-11	6960	FRA	48.5	-1.41	CS77211	western_europe	Bur-0 like
UII2-3	6973	SWE	56.0648	13.9707	CS78817	germany	Col-0 like
Alst-1	6989	UK	54.8	-2.4333	CS76432	western_europe	Bur-0 like
Bd-0	7013	GER	52.4584	13.287	CS76445	south_sweden	Col-0 like
Bur-0	7058	IRL	54.1	-6.2	CS76734	admixed	Bur-0 like
Da-0	7094	GER	49.8724	8.65081	CS76791	germany	Col-0 like
Dra-0	7103	CZE	49.4167	16.2667	CS76476	central_europe	Col-0 like
Dr-0	7106	GER	51.051	13.7336	CS78897	central_europe	Col-0 like
Mnz-0	7244	GER	50.001	8.26664	CS76552	germany	Col-0 like
Oy-0	7288	NOR	60.38554	6.193019	CS77156	admixed	Bur-0 like
Sten-0	7346	GER	52.6058	11.8558	CS77277	south_sweden	Bur-0 like
Tha-1	7353	NED	52.08	4.3	CS76611	admixed	Bur-0 like
Ting-1	7354	SWE	56.5	14.9	CS76612	south_sweden	Bur-0 like
JI-3	7424	CZE	49.2	16.6166	CS76519	central_europe	Col-0 like
WAR	7477	USA	41.7302	-71.2825	CS78853	admixed	Col-0 like
Lp2-6	7521	CZE	49.38	16.81	CS77052	central_europe	Col-0 like
HSm	8236	CZE	49.33	15.76	CS76941	central_europe	Col-0 like
PHW-3	8239	GER	51	7	CS76976	germany	Col-0 like
Na-1	8343	FRA	47.5	1.5	CS76558	admixed	Col-0 like
Pla-0	8357	ESP	41.5	2.25	CS76573	spain	Bur-0 like
Rak-2	8365	CZE	49	16	CS77201	central_europe	Col-0 like
Sr:5	8386	SWE	58.9	11.2	CS77268	central_europe	Bur-0 like
Lerik1-4	9075	AZE	38.7406	48.6131	CS77023	italy_balkan_caucasus	Bur-0 like
Lerik1-7	9078	AZE	38.7406	48.6131	CS77024	italy_balkan_caucasus	Bur-0 like
Nar-3	9089	AZE	38.9522	48.925	CS77119	italy_balkan_caucasus	Bur-0 like
Istisu-5	9095	AZE	38.9786	48.5594	CS76950	italy_balkan_caucasus	Bur-0 like
Lag1-4	9102	GEO	41.8296	46.2831	CS76999	italy_balkan_caucasus	Bur-0 like
Lag1-6	9104	GEO	41.8296	46.2831	CS77001	italy_balkan_caucasus	Bur-0 like
Lag1-7	9105	GEO	41.8296	46.2831	CS77002	italy_balkan_caucasus	Bur-0 like
Lag2-6	9113	GEO	41.8296	46.2831	CS77006	italy_balkan_caucasus	Bur-0 like
Bak-5	9121	GEO	41.7942	43.4767	CS76685	italy_balkan_caucasus	Col-0 like
Geg-14	9125	ARM	40.1408	44.8203	CS76876	asia	Bur-0 like
IP-Alo-0	9506	POR	40.11	-7.47	CS76662	admixed	Bur-0 like
IP-Reg-0	9509	POR	39.29	-7.4	CS77207	spain	Bur-0 like
IP-Rei-0	9510	POR	38.75	-7.59	CS77208	spain	Bur-0 like
IP-Vav-0	9511	POR	38.53	-8.02	CS78835	spain	Bur-0 like
IP-Adm-0	9514	ESP	39.15	-4.54	CS76647	spain	Bur-0 like
IP-Ala-0	9515	ESP	39.72	-6.89	CS76650	spain	Bur-0 like
IP-Cab-3	9526	ESP	41.54	2.39	CS76738	spain	Bur-0 like
IP-Car-1	9530	ESP	38.25	-4.32	CS76742	admixed	Bur-0 like
IP-Cem-0	9533	ESP	41.15	-4.32	CS76763	relict	Bur-0 like
IP-Cmo-3	9534	ESP	40.05	-4.65	CS76774	spain	Bur-0 like
IP-Fue-2	9541	ESP	38.26	-5.42	CS76871	spain	Bur-0 like
IP-Gra-0	9543	ESP	36.77	-5.39	CS76886	relict	Bur-0 like
IP-Her-12	9545	ESP	39.4	-5.78	CS76920	relict	Bur-0 like
IP-Men-2	9556	ESP	39.66	-4.34	CS77081	spain	Bur-0 like
IP-Mon-5	9559	ESP	38.06	-4.38	CS77107	admixed	Bur-0 like

IP-Stp-0	9584	ESP	41.19	-3.58	CS77283	spain	Col-0 like
IP-Tam-0	9586	ESP	41.03	-3.27	CS77340	spain	Bur-0 like
IP-Tdc-0	9587	ESP	41.5	-1.88	CS77344	spain	Bur-0 like
IP-Trs-0	9590	ESP	43.37	-5.49	CS77387	western_europe	Bur-0 like
IP-Vin-0	9599	ESP	42.8	-5.77	CS78846	western_europe	Bur-0 like
IP-Vis-0	9600	ESP	39.85	-6.04	CS78848	relict	Bur-0 like
Krazo-1	9616	RUS	53.06	51.96	CS76984	asia	Col-0 like
Giffo-1	9653	ITA	38.44	16.13	CS76878	italy_balkan_caucasus	Bur-0 like
Castelfed-3-206	9690	ITA	46.3378	11.2928	CS76755	central_europe	Bur-0 like
Castelfed-3-209	9693	ITA	46.3378	11.2928	CS76758	central_europe	Bur-0 like
Groch-1	9722	BUL	41.71	24.41	CS76890	italy_balkan_caucasus	Col-0 like
Faneronemi-3	9726	GRC	37.07	22.04	CS76853	italy_balkan_caucasus	Bur-0 like
Stiav-3	9731	SVK	48.46	18.9	CS77281	central_europe	Col-0 like
Bela-2	9733	SVK	48.47	18.94	CS76697	central_europe	Bur-0 like
Bran-1	9738	ROU	45.57	25.42	CS76722	admixed	Bur-0 like
Etna-2	9762	ITA	37.69	14.98	CS76487	admixed	Bur-0 like
HE-1	9769	GER	48.55	8.99	CS76916	central_europe	Col-0 like
Pfn-N2-2-6	9771	GER	48.56	9.11	CS77172	central_europe	Bur-0 like
Gn-1	9777	GER	48.57	9.17	CS76880	central_europe	Bur-0 like
Bai-10	9779	GER	48.5	8.78	CS76682	central_europe	Col-0 like
Fell2-4	9780	GER	48.43	8.79	CS76856	central_europe	Col-0 like
Kus2-2	9781	GER	48.52	9.11	CS76990	central_europe	Col-0 like
Erg2-6	9784	GER	48.5	8.8	CS76845	central_europe	Bur-0 like
Ha-HBT1-2	9785	GER	48.54	9.02	CS76898	central_europe	Col-0 like
Haes-1	9791	GER	48.6	9.2	CS76914	central_europe	Bur-0 like
Ha-P2-1	9798	GER	48.54	9.01	CS76902	central_europe	Col-0 like
Schl-7	9807	GER	48.6	9.22	CS77240	central_europe	Bur-0 like
Tu-B2-3	9808	GER	48.52	9.08	CS77392	central_europe	Col-0 like
Tu-W1	9812	GER	48.52	9.03	CS77397	central_europe	Col-0 like
BI-4	9813	GER	48.4	8.77	CS76706	central_europe	Col-0 like
Ha-HBT3-11	9815	GER	48.54	9.02	CS76900	central_europe	Col-0 like
IP-Aru-0	9821	ESP	41.81	2.49	CS76674	spain	Bur-0 like
IP-Bac-0	9823	ESP	43.34	-5.84	CS76681	western_europe	Bur-0 like
IP-Boa-0	9825	ESP	40.4	-3.88	CS76714	spain	Bur-0 like
IP-Cas-0	9831	ESP	38.54	-3.39	CS76743	admixed	Bur-0 like
IP-Cat-0	9832	ESP	40.54	-3.69	CS76759	relict	Bur-0 like
IP-Ees-0	9841	ESP	40.59	-4.15	CS76836	spain	Bur-0 like
IP-Elp-0	9843	ESP	40.53	-3.92	CS76840	spain	Bur-0 like
IP-Hue-3	9851	ESP	42.96	-6.1	CS76942	western_europe	Bur-0 like
IP-Ini-0	9852	ESP	40.46	-3.75	CS76945	spain	Bur-0 like
IP-Lam-0	9855	ESP	40.57	-3.89	CS77008	spain	Bur-0 like
IP-Lro-0	9859	ESP	40.5	-3.88	CS77054	spain	Bur-0 like
IP-Lum-0	9860	ESP	42.24	-2.62	CS77059	spain	Bur-0 like
IP-Mad-0	9862	ESP	40.45	-3.67	CS77062	western_europe	Bur-0 like
IP-Pdl-0	9877	ESP	43.02	-5.6	CS77165	admixed	Bur-0 like
IP-Pil-0	9882	ESP	40.46	-4.26	CS77178	spain	Bur-0 like
IP-Sen-0	9894	ESP	42.59	0.76	CS77243	admixed	Bur-0 like
IP-Sfb-6	9895	ESP	41.78	2.57	CS77248	spain	Bur-0 like
IP-Smt-1	9897	ESP	40.95	-5.63	CS77257	admixed	Bur-0 like
IP-Som-0	9898	ESP	41.14	-3.58	CS77259	spain	Col-0 like
IP-Tri-0	9900	ESP	37.38	-6.01	CS77386	spain	Bur-0 like
IP-Urd-1	9901	ESP	42.27	-2.98	CS78824	spain	Bur-0 like
IP-Mah-6	9906	ESP	40	4.25	CS77063	admixed	Bur-0 like
Leo-1	9945	ESP	41.8	-3.11	CS76413	admixed	Bur-0 like
Ped-0	9947	ESP	40.74	-3.9	CS76415	relict	Bur-0 like
Valsi-1	9969	ITA	40.18	16.45	CS76425	italy_balkan_caucasus	Bur-0 like
Lag2-2	9990	GEO	41.8296	46.2831	CS76390	italy_balkan_caucasus	Bur-0 like
TueV-13	10001	GER	48.52	9.05	CS76407	western_europe	Col-0 like
Dobra-1	10018	SRB	44.84	20.16	CS76369	central_europe	Bur-0 like

Table S5. List of primer sets used to CHOP qPCR and qPCR experimentation.

Gene	Experimentation	LP primer (5' > 3')	RP primer (5' > 3')
At1g47550	qPCR	CTCGCTCTTCCGTCAAATC	CCCCAGTGTGAAAGTGCATC
At1g54610	qPCR	GGTCGGACAGAGGTAGAGCAG	GTATGGTTCACGGGGTTTGT
At5g38470	qPCR	TGACTCGGGTATCCCTGCT	CTGGAGCTGCTGCTTGTG
At5g47260	qPCR	AAGGTGGTCCAATCGGGAAC	GATGGGGCAATCTGGTGTGA
At5g47280	qPCR	AGTCTCTGGCTTGAGAGGGT	ATGGTCGAAGGTAGTTCCGC
At5g47260	CHOP qPCR	TGCGTCGACCTATCGTTACA	CCATGCCGTATCAAGCAAC
At5g13440	CHOP qPCR	ACAAGCCAAATTTTGCTGAGC	ACAACAGTCCGAGTGCATGGT
At5g47400	CHOP qPCR	GAAGCCGAACTGCAAACCTGT	ATGGTCCGGCTCTAGGAAAA

Titre : Rôles des variations épigénétiques transgénérationnelles dans la résistance quantitative à la hernie chez *Arabidopsis thaliana*

Mots clés : méthylation de l'ADN, épimutation, *Plasmodiophora brassicae*, QTL^{épi}

Résumé : Des études récentes ont montré que la variabilité de l'épigénome des plantes est un acteur important dans la réponse des plantes aux stress abiotiques et biotiques. La hernie, causée par le protiste *Plasmodiophora brassicae*, est une maladie racinaire majeure des *Brassicaceae* cultivées dont la résistance quantitative est considérée comme résultant principalement de la ségrégation de multiples allèles. L'objectif de ma thèse est d'établir s'il existe, chez *Arabidopsis thaliana*, une variabilité épigénétique héritable à l'origine de variations de la réponse à l'infection par la hernie. Pour répondre à cet objectif, une approche non ciblée d'épigénétique quantitative a été réalisée en utilisant la population épiRIL *ddm1-2* x Col-0. Dix-sept QTL sous contrôle épigénétique (QTL^{épi}), regroupés en 6 régions génomiques, ont ainsi été détectés, 5 d'entre eux étant sous

la dépendance de la température. Finalement, deux régions identifiées comme impliquées dans la réponse à la hernie ont été caractérisées plus finement. La région du gène majeur de résistance à la hernie *RPB1*, qui colocalise avec 3 QTL^{épi}, présente une variation génomique prépondérante dans les écotypes d'*Arabidopsis* potentiellement due à des mouvements d'éléments transposables. Le QTL *Pb-At5.2* est sous le contrôle d'une épimutation régulant l'état de méthylation et l'expression de deux gènes NLR. Les résultats obtenus montrent que la résistance quantitative à la hernie est associée à des variations de la méthylation de l'ADN stables et héréditaires suggérant un modèle complexe de régulation de la résistance où la combinaison d'allèles et d'épi-allèles favorables est nécessaire pour obtenir une résistance optimale.

Title : Role of the transgenerational epigenetic variation in quantitative resistance to clubroot in *Arabidopsis thaliana*.

Keywords : DNA methylation, epimutation, *Plasmodiophora brassicae*, QTL^{épi}

Abstract : Recent studies have shown that plant epigenome variability is an important factor in plant response to abiotic and biotic stress. Clubroot caused by the protist *Plasmodiophora brassicae* is a major disease of *Brassicaceae* whose quantitative resistance is supposed to result from many allele segregation. The aim of my work is to understand if, in *Arabidopsis thaliana*, an inherited epigenetic variability can lead to variations in clubroot resistance. For that, an untargeted approach of quantitative epigenetics was carried out using the epiRIL population *ddm1-2* x Col-0. Seventeen QTL under epigenetic control (QTL^{épi}), clustered in 6 genomic regions, were detected, 5 of them being temperature-dependant.

Finally, two regions previously identified as involved in clubroot response were finely studied. The region of the major clubroot resistance gene *RPB1*, which colocalizes with three QTL^{épi}, shows major genomic variations in *Arabidopsis* ecotypes potentially due to movements of transposable elements. The QTL *Pb-At5.2* is depending on one epimutation controlling the methylation state and the expression of two NLR genes. The results obtained demonstrate that the clubroot quantitative resistance is associated with inherited stable DNA methylation variations suggesting a complex model of resistance regulation where favourable alleles and epialleles association is necessary to obtain an optimal resistance.