The study of DNA damage repair in *Arabidopsis thaliana*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Chun-Hsin Liu

(Taipei; Taiwan)

Köln, 2015

Abstract

DNA is the major storage of genetic information common to all cellular organisms. Endogenous and exogenous factors may damage DNA and affect plant growth and fitness.

To dissect the genetic variation of *Arabidopsis thaliana* to DNA replication stress, 400 Arabidopsis natural accessions were exposed to hydroxyurea (HU), a drug interfering with deoxyribonucleotide synthesis. This revealed a great variation in relative survival of Arabidopsis accessions to HU. However, extremely hypersensitive accessions were rare. To identify the molecular basis of this natural variation, quantitative trait locus (QTL) mapping with the recombinant inbred line population derived from crosses between HU-sensitive and resistant parents was performed. Two significant QTLs with additive effects and one minor QTL with epistatic effect were identified. Using heterogeneous inbred family analysis, one QTL was validated and fine mapped to a 22 kb region containing eight candidate genes which require further analysis.

Genes involved in related biological processes or pathways are often expressed coordinately, thus the uncharacterized candidate genes might be co-expressed with the functionally known genes. This co-expression strategy was applied to search for novel DNA damage repair genes in Arabidopsis and revealed increased sensitivity of mutants in *CHROMATIN REMODELING 31* to DNA inter-strand cross-linker, mitomycin C. Furthermore, the *chr31* mutant showed delayed cell cycle progression and increased frequencies of somatic homologous recombination (HR). This provides a novel connection between chromatin remodeling and DNA damage responses by so far unknown mechanisms.

Nucleoside analogues are frequently used in basic and medical research. However,

their mode of action and spectra of effects are not well understood. Analysis of the molecular effects of non-methylable cytidine analogue zebularine in Arabidopsis revealed induction of DNA damage response, post-replicative cell cycle arrest, and increased endoreduplication, but without obvious changes in DNA methylation. Molecular and genetic data suggested that zebularine-induced damage specifically occur during DNA replication in the course of DNA strand synthesis. The signaling of this damage was mediated by additive activity of ATAXIA TELANGIECTASIA MUTATED and ATAXIA_TELANGIECTASIA MUTATED AND RAD3-RELATED kinases. The repair required a functional STRUCTURAL MAINTENANCE OF CHROMOSOMES (SMC) 5-SMC6 complex and was accomplished predominantly by synthesis-dependent strand annealing type of HR. Hence, commonly used "epigenetic inhibitor" zebularine causes specific type of DNA damage.

Zusammenfassung

DNA ist der wichtigste Träger der Erbinformationen für alle zellulären Organismen. Endogene und exogene Faktoren können Schädigungen der DNA verursachen und damit das Pflanzenwachstum und die Anpassungsfähigkeit beeinträchtigen.

Um die genetische Variation im Zusammenhang von DNA-Replikationsstress in *Arabidopsis thaliana* zu untersuchen, wurden 400 *A. thaliana* Ökotypen der Chemikalie Hydroxyurea (HU), einem Stoff der die DNA-Synthese behindert, ausgesetzt, was deutliche Unterschiede in der Überlebensrate von verschiedenen Arabidopsis Ökotypen aufdeckte. Um die genetische Grundlage dieser Unterschiede zu identifizieren, wurden QTL-Kartierungen mit einer rekombinanten Inzuchtline durchgeführt, die aus der Kreuzung einer HU-sensitiven mit einer HU-resistenten Linie abstammt. Dabei wurden zwei signifikante QTLs ermittelt, die additive Effekte aufwiesen, und ein weiterer QTL der epistatische Effekte aufwies. Mit Hilfe von heterogenen Inzuchtlinien konnte einer dieser QTLs validiert werden und durch eine Feinkartierung die Region auf 22 kb und insgesamt 8 Kandidatengene eingegrenzt werden.

Gene, die in ähnlichen oder zusammenhängenden biologischen Prozessen eingebunden sind, werden oftmals koordiniert exprimiert. Daher können durch Untersuchungen der Expression von nicht charakterisierten Genen und ihrer Koexpression mit bereits bekannten Genen Rückschlüsse auf deren Funktion geschlossen werden. Mit dieser Herangehensweise wurde das bisher in diesem Zusammenhang uncharakterisierte DNA-Reparaturgen *CHROMATIN REMODELING 31* identifiziert, welches eine erhöhte Sensitivität gegenüber der Chemikalie Mitomycin C aufwies, einem Stoff der zu einer erhöhten Anzahl von intermolekularen Bindungen zwischen zwei DNA-Moleküle führt. Des weiteren zeigten *chr31* Mutanten einen

verzögerten Zellzyklus und erhöhte Raten homologer Rekombination in somatischen Zellen. Damit konnte ein neuer Zusammenhang zwischen Chromatin-Remodellierung und DNA-Reparatur hergestellt werden.

Nukleosidanaloga werden häufig in der Grundlagenforschung angewendet, jedoch ohne ihren genauen Wirkungsmechanismus zu kennen. Daher wurden Untersuchungen zum molekularen Wirkungsmechanismus des nicht methylierbaren Cytidinanalogons Zebularin in Arabidopsis durchgeführt. Diese zeigten, dass Zebularin die DNA-Reparaturantwort induziert, zum Stopp des Zellzyklus nach Replikation führt und die Endoreduplikationsrate erhöht, ohne dabei die DNA-Methylierung sichtbar zu beeinflussen. Molekulare und genetische Daten deuten darauf hin, dass durch Zebularin induzierte DNA-Schäden spezifisch während der Synthese der DNA-Stränge im Verlauf der DNA-Replikation auftreten. Die Signalweiterleitung dieser Schädigungen wird additiv durch die Aktivität der Kinasen ATAXIA TELANGIECTASIA MUTATED und TAXIA_TELANGIECTASIA MUTATED AND RAD3-RELATED vermittelt. Außerdem wird für die Reparatur ein funktionsfähiger STRUCTURAL MAINTENANCE OF CHROMOSOMES (SMC) 5-SMC6 Komplex benötigt. Die Reparatur selbst erfolgt überwiegend in Form des *synthesis-dependent strand annealing* der homologen Rekombination. Somit konnte gezeigt werden, dass Zebularin spezifisch Schäden an der DNA verursacht.

Table of Contents

List of Abbreviation

A. thaliana	Arabidopsis thaliana
ADH	Alcohol dehydrogenase
AP	Apurinic/apyrimidinic
ATM	ATAXIA TELANGIECTASIA MUTATED
ATR	ATAXIA_TELANGIECTASIA MUTATED AND RAD3-RELATED
BER	Base excision repair
CAPS	Cleaved amplified polymorphic sequences
CHR31	CHROMATIN REMODELING 31
CLSY 3	CLASSY 3
CMT3	CHROMOMETHYLASE
CPD	cyclobutane pyrimidine dimmer
CYCB1;1	MITOTIC CYCLIN-DEPENDENT KINASE CYCB1;1
ddi	DNA damage-inducing
DDM1	DECREASED DNA METHYLATION 1
DDR	DNA damage response
DEFL	defensin-like
DFC	DHFS-FPGS HOMOLOG C
DMR	DOMAIN-REARRANGED METHYLTRANSFERASE
dNDP	Deoxyribonucleoside
DNMT	DNA methyltransferase
dNTP	Deoxyribonucleoside triphosphate
DSB	Double-strand breaks
EMMAX	Efficient mixed-model association eXpedited
FA	Formaldehyde
GMI1	GAMMA-IRRADIATION AND MITOMYCIN C INDUCED 1
GO	Gene ontology
GUS	β-glucuronidase
GWA studies	Genome-wide association studies
hhu	Hypersensitive to HU
HIF	Heterogeneous inbred family
HR	Homologous recombination
HU	Hydroxyurea
INDEL	Insertions and deletions
LD	Linkage disequilibrium
LOD	Log of odds

MET1	DNA METHYLTRANSFERASE
MMC	Mitomycin C
MMEJ	Micro-homology-mediated end joining
MMR	Mismatch repair
MRN complex	MRE11/RAD50/NBS1 complex
MS	Murashige and Skoog
NASC	Nottingham Arabidopsis Stock Center
NDP	Ribonucleotide diphosphate
NHEJ	Non-homologous end joining
NIL	Near isogenic line
NO	Nitric oxide
NPA	Nucleoprotein adduct
nvh	Natural variation to HU
PAP16	PURPLE ACID PHOSPHATASE 16
PARP	poly(ADP-ribose)polymerase
PARP2	POLY (ADP-RIBOSE) POLYMERASE 2
PI	propidium iodide
P-loop NTPases	P-loop containing nucleoside triphosphate hydrolases
Pol IV	RNA polymerase IV
QTG	Quantitative trait gene
QTL	Quantitative trait loci
RAM	root apical meristem
RECA3	RECA HOMOLOG 3
RFLP	Restriction fragment length polymorphisms
RIL	Recombinant inbred lines
RNR	Ribonucleotide reductase
ROS	Reactive oxygen species
RS	Relative survival rate
RT	reverse transcription
RT-qPCR	Reverse transcription quantitative PCR
SDC	SUPPRESSOR OF DRM1 DRM2 CMT3
SDSA	Synthesis-dependent strand annealing
SMC6B	STRUCTURAL MAINTENANCE OF CHROMOSOMES 6B
SNF2	SUCROSE NON-FERMENTING 2
SNP	Single nucleotide polymorphism
SSA	Single strand annealing
SWI2	SWITCH 2
TE	Transposable element

WT	Wild type
XPF	XERODERMA PIGMENTOSUM GROUP F DNA REPAIR
XRCC4	HOMOLOG OF X-RAY REPAIR CROSS COMPLEMENTING 4

List of Figures

xiii

List of Tables

1. Introduction

1.1 DNA damage and repair in plants

Plants as sessile organisms are routinely exposed to external and internal factors that induce DNA damages, leading to genome instability and genotoxic stress. DNA damages include base deletions, pyrimidine dimers, cross-links, strand breaks, and base modification (Tuteja et al., 2009). They require specific repair mechanisms to protect the cells against the damages and to ensure faithful transmission of genetic information from one generation to the next. The DNA repair mechanisms are well understood in *E. coli*, *Saccharomyces cerevisiae*, rodents and human, but only recently have plant DNA repair systems begun to be studied in detail (Singh et al., 2010; Yoshiyama et al., 2013a).

1.1.1 DNA damage response (DDR) – the signal transducers

The presence of DNA damages and the DDR signaling need to be transmitted to switch on the proper repair mechanisms. ATAXIA TELANGIECTASIA MUTATED (ATM) and ATAXIA TELANGIECTASIA MUTATED AND RAD3-RELATED (ATR) are serine/threonine kinases that are rapidly activated in response to DNA damages. They are highly conserved among eukaryotes and are triggered by different types of DNA damages. ATM has been proposed to respond to the damage of double-strand breaks (DSBs), whereas ATR responds to a wide range of DNA lesions, especially those associated with DNA replication. Arabidopsis *atm* mutants are hypersensitive to DSB-inducing agents including gamma-irradiation and methylmethane sulfonate (Garcia et al., 2003), while *atr* mutants are sensitive to replication-blocking agents, such as UV radiation or HU (Culligan et al., 2004), The phosphorylation of the protein kinases

1

amplifies and transduces DDR signals to downstream effectors which then elicit appropriate responses, such as the arrest of cell cycle progression, DNA repair, or programmed cell death (Yoshiyama et al., 2013a).

1.1.2 DDR - endoreduplication and cell cycle checkpoint

Endoreduplication is a cell cycle variant in multicellular eukaryotes in which mitosis is skipped and cells repeatedly replicate their DNA, resulting in cellular polyploidy. It is essential for developmental processes and prominent response to changing physiological conditions (De Veylder et al., 2011).

Upon the presence of DNA damages, the cell cycle progression is delayed or arrested at critical stage before or during DNA replication and before cell division. Activation of the DNA damage checkpoint results in the simultaneous induction of DNA repair genes and inactivation of genes that are required for mitosis and cytokinesis. This coordinated action ensures that cells repair their damaged genome before they proceed into mitosis (De Veylder et al., 2007). Unlike mammals, the activation of DNA damage checkpoint in plants rarely results in apoptosis or programmed cell death but the onset of endoreduplication.

It remains unknown why endoreduplication is induced upon DNA damages. One explanation is that the initiation of the endocycle program might be a mechanism to prevent the transmission of DNA lesions into the pool of meristematic cells by pushing the damaged cell into a non-dividing state, in such way safeguarding the progeny from DNA mutations (De Veylder et al., 2011).

1.1.3 DDR - DNA repair mechanisms

In plants, several DNA repair mechanisms corresponding to different types of DNA damages were proposed. The specificity and efficiency of each pathways are critical to ensure genome stability, and cross-talks between different repair pathways might occur to withstand the same type of DNA damages (Balestrazzi et al., 2011).

Photoreactivation

Ultraviolet (UV)-B existing in the solar radiation induces dimers production between adjacent pyrimidine residues in a DNA strand. Thus, direct repair or photoreactivation is to repair UV-damaged DNA by photolyases. They use 350−450 nm light as an energy source to power cyclic electron transfer to split the pyrimidine dimer ring structure and restore the bases to their normal undamaged form (Jiang et al., 1997; Sancar, 2003).

Excision repair pathways

Excision repair pathways can be divided into three main categories: (1) base excision repair (BER) involves the removal of a single damaged base through the action of glycosylases which leave the DNA an abasic site, then the apurinic/apyrimidinic (AP) endonuclease nicks the DNA at the AP site, and the DNA is restored to its original sequence through the combined actions of exonucleases, a repair polymerase, and DNA ligase (Britt, 1996); (2) in contrast to BER, nucleotide excision repair (NER) acts on a wider spectrum of DNA damage products, in which it detects modifications indirectly by conformational changes to the DNA duplex rather than relying on the recognition of specific DNA damages. The repair reactions include recognition of DNA damage, unwinding double-strand DNA in the neighborhood of the damage, excision of the

damaged nucleotides, and filling of the single-stranded gap by DNA synthesis (Kimura and Sakaguchi, 2006); (3) mismatch repair (MMR) recognizes the mismatch in replicating DNA by discriminating between the existing correct nucleotide in the template strand and the incorrect nucleotide inserted into the replicating DNA strand, and the nascent DNA to a point beyond the mismatch lesion is digested by exonuclease, and the DNA resynthesis and ligation are accomplished by DNA polymerase and DNA ligase, separately (Bray and West, 2005);

DNA double-strand break (DSB) repair

DSBs are most deleterious type of DNA damages, which can lead to cell death. The repair of DSBs mainly comprises the non-homologous end joining (NHEJ) and the homologous recombination (HR) mechanisms.

NHEJ is the major pathway for DSB repair in most of the higher eukaryote cells. The initial recognition of DSB is mediated by KU70 and KU80 complex by binding to the exposed DNA ends. The DNA ends are processed by the MRE11/RAD50/NBS1 (MRN) complex to make them suitable substrates for DNA ligases. Ligation is then catalyzed by a complex of DNA ligase IV and HOMOLOG OF X-RAY REPAIR CROSS COMPLEMENTING 4 (XRCC4). A second, KU-independent pathway was defined lately in both yeast and Arabidopsis, the micro-homology-mediated end joining (MMEJ) (Bleuyard et al., 2006). The detailed repair mechanism was proposed in yeast, in which the DNA ends of DSBs are processed into single-stranded DNA, and the joining of the two ends of the breaks is promoted via hybridization of these single-stranded regions using imperfect homologies of the order of ten bases (micro-homology) (Ma et al., 2003).

HR creates covalent linkages between DNA in regions of highly similar or identical sequence with a dual function in both meiosis and somatic cells. Upon DSBs in somatic cells, two models were proposed for HR. Synthesis-dependent strand annealing (SDSA) is initiated by a 3' resection forming a long single-stranded DNA tails that invades a homologous duplex and primes DNA synthesis. The newly synthesized DNA then reanneals with the other side of the DSB, repairing the break. A gene conversion without loss of sequence information is the final result of the reaction. Single-stranded annealing (SSA) may occur between tandemly repeated sequences (where homologous sequences are available). These molecules can directly anneal to one another, forming a chimeric DNA molecule. The presence of 3'-overhangs will be trimmed, and the single-stranded regions are filled by DNA synthesis (Bray and West, 2005; Roth et al., 2012).

1.1.4 Chemical inducers of DNA damages

Chemicals that induce DNA damages have frequently been applied to suppress viruses, neoplastic diseases, and the immune system in mammals, including human (Deans and West, 2011; Fu et al., 2012). They also provide an alternative tool to study the mechanism of DNA replication and repair, to compensate for the limitations of mutant selection and the complicated process of genetic analysis.

Replication blocking agent – hydroxyurea (HU)

Deoxynucleotide triphosphates (dNTPs) are the building blocks of DNA. Interference with dNTP biosynthesis may limit the progression of DNA replication (Christopherson

et al., 2002). Hydroxyurea (HU) reduces dNTP levels by inhibiting the activity of the small subunits of RIBONUCLEOTIDE REDUCTASE (RNR), an enzyme which catalyzes the reduction of ribonucleotide diphosphates (NDPs) into deoxyribonucleosides (dNDPs), thus affecting the progression of replication fork and inducing DNA replication stress (Roa et al., 2009). In plants, HU transiently arrests the cell cycle at the transition between G1/S and G2/M, providing a useful tool for cell-cycle synchrony (Pan et al., 1993; Cools et al., 2010), and has been used to identify the components in cell-cycle control and ATR-dependent repair pathways (Culligan et al., 2004; Cools et al., 2011; Spadafora et al., 2011).

Alkylating agent – mitomycin C (MMC)

Alkylating agents transfer alkyl carbon groups onto a broad range of biomolecules. They react with the ring nitrogens (N) and extracyclic oxygens (O) in DNA and generate a variety of covalent adducts. Mitomycin C (MMC) is a bifunctional alkylating agent containing two active moieties that can react with separate bases of DNA to form bi-adducts (Fu et al., 2012). Application of MMC (alone or in combination with other drugs) readily induces expression of DNA damage repair genes (AtGenExpress dataset ME00326). In plants, it has been used as one of the major marker drugs for associating gene functions with DNA damage repair, DNA replication, and cell-cycle control (Mannuss et al., 2010; Nezames et al., 2012; Rosa et al., 2013). The repair of MMC-induced inter-strand cross-links is thought to be mediated mainly by ATR and RAD51-dependent HR pathways (Bleuyard et al., 2005; Nezames et al., 2012).

Radiomimetic drug – bleomycin (bleocin)

Radiomimetic drugs are chemicals that imitate the effect of gamma irradiation. By inducing highly specific attack of free radicals, these drugs abstract hydrogen (H) predominantly at the Cl-, C4-, or C5- positions, leading to oxidation of 2-deoxyribose in DNA and DNA strand breaks (Povirk, 1996). Bleomycin (bleocin) is one of the most commonly used radiomimetic drugs in plant research. It was shown to activate DNA damage repair genes and inhibit cell growth (Smetana et al., 2012; Da Ines et al., 2013) and is frequently employed as markers for studying ATM-mediated DSB repair (Lang et al., 2012).

1.2 DNA methylation and DNA methyltransferase (DNMT) inhibitors

1.2.1 DNA methylation in plants

The methylation of cytosine at the 5th position is a gene-silencing mechanism that protects the genome from selfish DNA elements and regulates gene expression. DNA methylation can be divided into two steps: *de novo* methylation and maintenance methylation. *De novo* methylation refers to the modification of a previously unmethylated DNA. It occurs at CG, CHG, and CHH (where H is A, C or T) sequences, each of which has different genetic requirements for the maintenance of its methylation. Maintenance methylation acts on hemimethylated DNA replication or repair, guided by the modification of still present in the template DNA strand. It occurs most efficiently at CG and CHG sequences (Chan et al., 2005; Allis et al., 2007).

There are mainly three conserved families of DNA methyltransferases (DNMTs) which are all present in plants: (1) DNA METHYLTRANSFERASE (e.g. MET1) family are

considered to be CG maintenance methyltransferases; (2) DOMAIN-REARRANGED METHYLTRANSFERASES (e.g. DMR2) are considered to be *de novo* methyltransferases in all sequence contexts. As indicated by their name, they have rearrange domains which might give them the ability to methylate asymmetric CHH nucleotide groups; (3) CHROMOMETHYLASE (e.g. CMT3) modifies CHG sequences and has been implicated in both *de novo* and maintenance methylation (Allis et al., 2007).

1.2.2 Nucleoside analogues as inhibitors of DNA methyltransferase (DNMT) and DNA methylation inhibitors

The most frequently-used drugs to interfere with DNA methylation in plants are 5-azacytidine and zebularine. Both of the drugs are cytosine analogues and were proposed to have similar mode of action for DNA demethylation. Before being incorporated into DNA, they are converted into the active triphosphate-deoxy form by RNR and cytidine/deoxycytidine kinase. Inside S-phase cells, the deoxy-nucleoside analogue is then incorporated in place of cytosine into DNA, which forms covalent bonds with DNMTs, resulting in the depletion of active enzymes and the demethylation of DNA (Egger et al., 2004).

Both of the nucleoside analogues were shown to reduce DNA methylation and reactivate transcriptionally silenced genes in Arabidopsis (Baubec et al., 2009; Baubec et al., 2014). However, the rapid degradation of 5-aza-cytidine in aqueous solution (10% loss in 2 hours) and its high cytotoxicity make zebularine (less than 7% loss in 48 hours) a more suitable candidate for DNA methylation inhibition (Ben-Kasus et al., 2005).

1.2.3 The cytotoxicity of nucleoside analogues – the formation of nucleotide protein adducts (NPAs) with DNMTs

Both zebularine and 5-aza-cytidine are toxic and were shown to inhibit plant growth (Baubec et al., 2009; Zhong et al., 2013). The cytotoxicity of the two drugs was proposed to be due to the formation of steric hindrance between the nucleoside analogue and DNMT along DNA structure. Due to their structural similarity to cytosine, when incorporated into DNA, the cytidine analogues are targeted by DNMT as hemi-methylated cytidines. A covalent bond is commonly found at the C6 position of the cytosine ring. Due to the differences between the chemical structures of cytidine analogs and physiologic cytidine, the attached DNMTs are entrapped. DNMT-DNA nucleoprotein adducts (NPAs) are formed, which can trigger DNA damage responses and subsequent G2-phase cell-cycle arrest or cytotoxicity (Christman, 2002; Lim et al., 2011).

In vitro study showed that an increased binding between zebularine and DNA and a strong decreased dissociation rate between DNMT and zebularine were observed, comparing to deoxycytidine, suggesting the formation of a more stable NPA between zebularine and DNMT (Champion et al., 2010). 5-aza-cytidine also formed a covalent complex with DNMT which was proposed to be irreversible and the cause of massive loss of DNA methylation (Santi et al., 1984).

1.3 Natural variation in *Arabidopsis thaliana (A. thaliana)*

1.3.1 *A. thaliana*: the model organism in plant genetic analysis

A. thaliana is a small plant in the mustard family (*Brassicaceae*). It has a small genome (125 Mb) which is organized into five chromosomes and contains an estimated 25,000 protein-coding genes (Page and Grossniklaus, 2002). The simple procedures for chemical and insertional mutagenesis, efficient methods for performing crosses and introducing DNA through plant transformation, extensive collections of mutants with diverse phenotypes, as well as a variety of genetic maps of mutant genes and molecular markers, all make *A. thaliana* an excellent model for plant genetic analysis (Meinke et al., 1998).

1.3.2 Naturally occurring variation in *A. thaliana*

A. thaliana is widely distributed throughout the Northern hemisphere. The broad climatic zone and environmental conditions where *A. thaliana* populations occur make it an important resource to uncover the molecular basis of plant adaptation to diverse environment. Comparing to laboratory-induced mutants which have limited genetic background, e.g. functionally null or weak allele in the wild type (WT) might not be detected or epistatic interactions of functional redundant genes ensure only certain phenotypes appear, natural variation provides another resource to identify gene functions and allelic variants that interact with the genetic background and/or the environment, or allele showing small effects on phenotype, particularly for traits related to plant adaptation (Alonso-Blanco and Koornneef, 2000; Koornneef et al., 2011).

According to Koornneef et al. (2011), the analysis of natural variation in *A. thaliana* is relevant in two aspects: first, as wild species, the analysis will increase the understanding in the genetic variation of how plants adapt to various environmental conditions and the complex population dynamics influenced by human activities. Second, as the model species, the study of *A. thaliana* also facilitates the analysis of natural variation of general traits related to adaptation in other wild and cultivated *Brassicaceae* species.

1.3.3 Genetic tools for the study of natural variation

1.3.3.1 Quantitative trait loci (QTL) mapping

To dissect the genetic basis of natural variation among *A. thaliana* accessions is to identify how many loci account for it and where they are located in the genome (mapping). If the variation is of qualitative nature, i.e. the phenotypes can be classified into unambiguous discrete categories and are responsible by a single locus (monogenic), the genetic analysis is similar to the study of mutants and can be achieved by Mendelian linkage analysis and allelism tests by complementation. However, most of the variation is of quantitative nature due to allelic variation at several loci (multigenic), combined with the environmental effect, determines a continuous (quantitative) phenotypic distribution of the trait in the segregation populations. The responsible genotypes can only be indirectly inferred from linked marker loci, i.e. by linkage (QTL) mapping (Koornneef et al., 2004; Bergelson and Roux, 2010).

To initiate QTL mapping requires three major steps: first, the generation of mapping population, preferentially those are homozygous and the recombinant chromosomes have been fixed through inbreeding, such as recombinant inbred lines (RILs) and near

isogenic lines (NILs). Such populations only need to be genotyped once and can be phenotyped repeatedly for many different traits under different environmental conditions. Second, the genotyping with markers throughout the genome and the phenotyping for the trait of interest. Several molecular marker techniques are available in Arabidopsis to detect single nucleotide polymorphism (SNP), including restriction fragment length polymorphisms (RFLPs), cleaved amplified polymorphic sequences (CAPSs), and microsatellites. For example, CAPS polymorphisms are differences in restriction fragment lengths caused by SNPs or insertion and deletion (INDEL) that create or abolish restriction endonuclease recognition sites in PCR amplicons produced by locus-specific oligonucleotide primers (Konieczny and Ausubel, 1993). Third, the association analysis between phenotypic values of the trait and the genotypic classes of the polymorphic markers, i.e. to calculate the number and the position of loci which control the trait variation in the population, their additive effect, and the contribution of genetic interactions between loci (epistasis) (Koornneef et al., 2004; Weigel, 2012).

1.3.3.2 QTL confirmation and fine mapping

The statistical power to detect QTLs can be influenced by several factors, such as genetic properties of QTLs that control traits, environmental effects, population size, and experimental error from genotyping and phenotyping. Thus, the QTL mapping results should be independently confirmed or verified, which may involve independent populations constructed from the same parental genotypes or closely-related genotypes used in the primary QTL mapping study (Collard et al., 2005).

A common approach to confirm QTLs has been using a specific type of population called NILs, in which a region that contains the QTL is introgressed into the isogenic

background of one of the parental lines, and successive generations of recombination are used to narrow down the QTL to a small genomic interval (Kooke et al., 2012). Nonetheless, the time and effort required to develop NILs have limited their use. Tuinstra et al. (1997) proposed an alternative procedure to identify NILs by heterogeneous inbred family (HIF) analysis. Each HIF is isogenic at most of the loci in the genome but ideally has only a small portion remained as heterozygous which segregates for the genomic region of interest. Progeny of HIF will segregate for those loci not yet fixed and will represent a HIF of near-isogenic individuals. The descendants (family) with homozygous genotype at the region of interest will be identified and phenotyped, and the QTL is confirmed as the genotypes are consistently associated with the phenotypes.

Once the QTL for a trait of interest is detected and confirmed, the QTL region needs further delimitation to isolate the underlying genes. A sequential QTL fine-mapping strategy using recombinant-derived progeny was proposed by Yang et al. (2012). Recombinants identified in NILs or HIFs that underwent crossing-over within the confidence interval can be selected for further backcrossing or self-pollination to produce progeny for fine mapping. Values of the trait will be calculated in the progenies, and statistically significant differences in the trait values between homozygous and heterozygous plants (in the donor segment in the case of NILs) or between homozygous plants having each of the parental genotypes (in the segregating region of interest in the case of HIFs) would indicate the presence of the QTL in the targeted region. The comparison of the phenotypes for all recombinants should enable the QTL to be narrowed down. Moreover, new recombinants within the mapped QTL region will be identified by genotyping either large segregating population or multiple generations across intercross/backcross populations and by additional genotyping markers to

resolve recombination breakpoints. The new recombinants are then backcrossed or self-pollinated again to produce the next progeny for further fine mapping. The fine-mapping procedure will be sequentially carried out until the candidate genes underlying the target QTL are identified.

1.3.3.3 The molecular basis of natural variation: from QTL to quantitative trait gene (QTG)

A variety of strategies have been proposed to identify which gene in the QTL interval is the quantitative trait gene (QTG), such as complementation cloning of the QTL to one of the parental lines in order to complement the phenotype, microarray or sequence prediction for inactivated genes within the QTL intervals, or direct sequencing of the segregating population to identify causative mutations (Bergelson and Roux, 2010). Another quantitative complementation experiment was proposed by Long et al. (1996) to test the genetic interactions between the naturally occurring alleles at the QTL and the mutant alleles at candidate loci by making crosses between the natural strain and the mutant strain. The failure to complement the quantitative phenotype can be interpreted as evidence for allelism between the factors on selected chromosome and the candidate loci.

1.3.3.4 Epistasis and quantitative traits: the study of gene-gene interactions

Epistasis—the interaction between genes at different loci—has two related but distinct definitions depending on the way it is revealed. In 1909, Bateson used the term epistasis to describe the masking effect of an allele at one locus upon an allele at another locus. In quantitative genetics, Fisher defined epistasis as the deviation from the additivity of

the contributions of several loci to a quantitative phenotype (Causse et al., 2007).

Several approaches have been proposed to test for the epistatic effects between two QTLs. Interactions with the genetic background could be revealed by multiple crosses between related inbred lines (Causse et al., 2007). For example, the addition of favorable alleles for fruit weight in tomato provided less progress than expected (Eshed and Zamir, 1996). Transformation or allelic replacement could prove that the variants are causal, or to engineer all possible combinations of causal variants to investigate epistasis at nucleotide resolution (Mackay, 2014). These approaches were used in *D. melanogaster* to show that each of the three domains in the *ALCOHOL DEHYDROGENASE (*ADH*)* gene, as well as an intragenic epistatic interaction, contributed to the difference in ADH protein levels in multiple independent transformants of each of the construct types (Stam and Laurie, 1996).

1.3.3.5 Genome-wide association (GWA) studies

Genome-wide association (GWA) studies is a mapping approach that uses natural linkage disequilibrium (LD) to identify polymorphisms which are associated with the phenotypic variation (Bergelson and Roux, 2010). LD here refers to the fact that in most of the species there has not been enough historic recombination to produce all possible combinations of physically adjacent polymorphisms, but rather the sequence variants are found in haplotype blocks of various lengths. Thus, a causal polymorphism of a specific trait can be identified indirectly through its association with any of the other sequence variants in its haplotype blocks (Weigel, 2012).

Once a sufficient number of the investigated accessions has been densely genotyped or

completely sequenced, unlimited amount of traits in different environments can be analyzed in genetically identical material (Nordborg and Weigel, 2008). The 1001 Genome Project was initiated in 2008 (http://1001genomes.org/), to discover the whole-genome sequence variation in at least 1000 accessions of *A. thaliana* (Cao et al., 2011). Up to September 2014, over 1100 lines have been sequenced.

The first GWA studies in plants were reported in 2010 with mapping of 107 phenotypic traits in *A. thaliana*, and the genetic and phenotypic resources are publicly available. Several functionally known genes related to disease resistance, trichome formation, and ionomics were identified, and *a priori* candidate genes were found for several traits related to floral transition (Atwell et al., 2010).

1.4 Co-expression network analysis

1.4.1 The principle of "guilt-by-association"

Biological processes are carried out by coordinated modules of interacting molecules. Eisen et al. (1998) proposed a system of cluster analysis for genome-wide expression data. They proposed that clustering gene expression data groups together efficiently genes of known and similar function, and co-expression of functionally known genes with poorly characterized or novel genes may provide a simple means of gaining leads to the functions of many genes for which information is not available currently. This has informed the guilt-by-association (GBA) principle which is now widely invoked in functional genomics.

The GBA paradigm has records in human, yeast, flies, worm, and also plants, to identify

new genes which are not yet associated with a given biological process in a given organism (Usadel et al., 2009). In Arabidopsis, the GBA paradigm has been applied to associate 'guide genes', i.e. genes involved in a specific pathway or with a given function, with biological processes to identify functionally unknown genes. For example, collections of gene expression data sets were used to ascribe the potential functions of the members in the large cytochrome P450 superfamily (Ehlting et al., 2006). Large-scale mutant screen for multiple stress regulator genes by combining gene expression ranking and co-expression network analysis identified around 40 candidates showing enhanced or reduced stress response comparing to WT (Ransbotyn et al., 2014).

1.4.2 The co-expression analysis tool: ATTED-II

ATTED-II (http://atted.jp) is a database of gene co-expression that was originally developed to identify functionally related genes in Arabidopsis and rice and now can be used to design a wide variety of experiments in additional agriculturally important plants. It included gene expression data from microarray and RNA sequencing studies to prioritize genes for functional identification or for studies of regulatory relationships (Obayashi et al., 2009; Obayashi et al., 2014).

ATTED-II provides two major ways to examine gene co-expression information, a gene list view and a gene network view using CoexSearch and NetworkDrawer tools, respectively. The CoexSearch tool provides a list of genes that are co-expressed with the guide genes (can be a single gene or multiple genes), whereas the NetworkDrawer tool accepts a set of genes and analyzes the internal relationships among the query genes.

Co-expressed genes in a list were defined by continuous value of Mutual Rank (MR)

which was calculated as the geometric mean of the correlation rank of gene A to gene B and of gene B to gene A. The mean value of MR was calculated based on each gene list, then the co-expressed genes were ranked from the smallest (the strongest association) to the largest MR (the weakest association). The top 300 MR genes are usually provided based on practical reasons. A comparative view between Arabidopsis and rice co-expression is also provided using orthologous genes to increase the reliability of co-expression data (Obayashi and Kinoshita, 2010).

2 Aim of the projects

Plants are constantly exposed to environmental factors and endogenous processes that induce genotoxic stress, damage DNA structure, and thus affect genome stability and growth. Though the signaling pathways and repair mechanisms of DDR are conserved among eukaryotes, plants as sessile organisms with postembryonic development require an efficient and specific DDR system to cope with DNA damages. Comparing to the achievement in yeast and animal, the understanding of DDR mechanisms is relatively less advanced in plants.

Natural variation provides a useful resource to dissect the genetic basis of how plants adapt to changing environment and external stimuli. The aim of the first project in this thesis is to describe the genetic variation of the *A. thaliana* natural accessions to DNA replication stress using HU as a genotoxic stress inducer. Genetic mapping including GWA studies and QTL mapping for HU sensitivity were performed using 400 accessions and RIL population, respectively. HIFs was then applied to validate the identified QTL, and the sequential QTL fine-mapping strategy using recombinant-

derived progeny was employed to narrow down the QTL into a region containing eight genes.

Zebularine has been used as an inhibitor of DNA methylation and was shown to be toxic to both animals and plants. However, the cytotoxicity of zebularine to Arabidopsis was shown to be more severe that the mutation in the core gene of DNA methylation, *DDM1*, indicating other effects of zebularine than DNA methylation inhibition. Thus, the aim of the second project in this thesis is to dissect the potential effects other than DNA methylation inhibition and the cytotoxicity of zebularine on Arabidopsis. First, genome-wide expression analysis was performed to investigate zebularine effects on Arabidopsis transcriptome and revealed up-regulation of several core DDR genes. Reverse genetic screen and analysis of HR and endoreduplication frequencies were then employed. The results suggest the formation of zebularine-DNMT NPAs might partially contribute to zebularine cytotoxicity and SDSA-HR plays a crucial role in detoxifying zebularine-induced damages.

Genes involved in the same biological processes or pathways are commonly expressed at the same time and under the same condition. This allows developing a strategy for co-expression analysis and identification of putative novel DDR genes specific to plant system. Known genes associated with DDR were used as guide genes to identify the potential DDR gene lists. Subsequently, reverse genetic screens were performed for mutants of candidate genes using several chemical inducers of DNA damages. This revealed *CHROMATIN REMODELING 31* as a potential candidate of a novel DDR gene and may provide further insight of the connection between chromatin remodeling and DNA damage repair.

3 Material and methods

3.1 Plant materials

391 *A. thaliana* natural accessions from the HapMap (2010) and the Nordborg collections (Nordborg et al., 2005) (Supplement Table1) were obtained from the Nottingham Arabidopsis Stock Center (NASC) (http://arabidopsis.info/). The RIL population of Bla-1 x Col-0 (Supplement Table 2) was provided by Versailles Arabidopsis Stock Center (http://publiclines.versailles.inra.fr/catalog/index) (Simon et al., 2008). T-DNA insertional mutant seeds (Supplement Table 5) were obtained from NASC and GABI-Kat (http://www.gabi-kat.de/).

3.2 Growth conditions

For the phenotypic screening of *A. thaliana* natural accessions on HU, seeds were surface sterilized with sodium hypochlorite (Sigma), thoroughly washed with water, and sowed on full-strength Murashige and Skoog (MS) solid media (Murashige and Skoog, 1962) with or without the treatment of 4 mM HU. The plates were placed at 4 degree for 4 days for stratification and grown in the Broson Climate Chamber at 21 degree under long day conditions (16-hour light: 8-hour dark) for 11 days and then evaluated. The same method was applied for the RIL population for QTL mapping and HIFs for QTL validation and fine mapping.

For T-DNA insertional mutants (Col-0 background), surface-sterilized seeds were sown on half-strength liquid- or solid- MS media (1/2 MS) with or without drug treatment

and then grown under the same condition as described for 12 to 21 days, depending on the experiments.

For cyclin– β-glucuronidase (GUS) containing the *pCYCB1;1::CYCB1;1:GUS* construct (Colón-Carmona et al., 1999), *pGMI1::GUS* reporter line (Böhmdorfer et al., 2011), and the plant materials for quantitative PCR (qPCR), plants were first grown on 1/2 MS solid media for 7 days and then transfer to 1/2 liquid-MS media with or without drug treatment for the appointed time. Reporter lines for HR analysis (651, 11B, and IC9) (Puchta et al., 1995; Molinier et al., 2004) were sown in 1/2 liquid- or solid- MS media with or without drug treatment for 12 or 14 days, whereas the liquid- media were changed every 4 to 5 days.

3.3 Chemicals

Stock solutions for HU (Sigma Aldrich), zebularine (Sigma Aldrich), 5-aza-cytidine (Sigma Aldrich), MMC (Duchefa Biochemie), and bleocin (CALBIOCHEM) were first prepared and then added directly into MS media after cooling down to make drug-treated agar plates or liquid media. Restriction enzymes (Supplement Table 6) were obtained from New England Biolabs and Thermo Scientific. Taq polymerases for PCR were from Ampliqon.

3.4 Histochemical staining and microscopy

GUS histochemical staining was performed as described (Baubec et al., 2009), and

staining with propidium iodide (PI) (Sigma) was according to Yoshiyama (2013b). Images were acquired using MZ16 FA stereomicroscope equipped with DFC490 CCD camera (both Leica).

3.5 Plant phenotyping

3.5.1 Rosette area and root length assays

For root length assay, plants were grown on drug or drug-free media for 7 days and then carefully taken out from the media without breaking the roots and stretched on a new plate for imaging analysis. For rosette area measurement, plants were grown on drug or drug-free media for 14, 15, or 21 days. Photographs were taken with a D90 digital camera (Nikon). Color photographs for rosette area measurement were converted into binary mode using ImageJ (http://rsbweb.nih.gov/ij/), and the selected area were measured as "the number of pixels", and then the true rosette area was converted by the equation:

Rosette area (mm^2) = pixels number / number of plants / 219.03^2*100

Both traits measurements were done using ImageJ calibrated with an internal-size control.

3.5.2 Homologous recombination assay

As shown in Fig. 1, the number of GUS spots (blue spots indicated by red arrows) in line 651, 11B, and IC9 were examined under stereomicroscope (Leica).
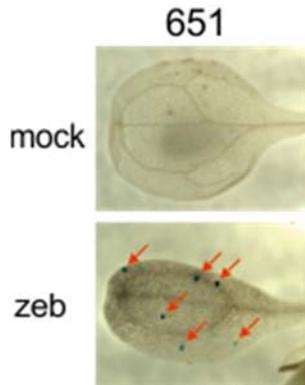
Fig. 1: HR assays. Representative cotyledons of mock and zebularine-treated line 651. HR events visualized as blue dots, are indicated by red arrows.

3.6 Primers for plant genotyping

CAPS markers (Supplement Table 6) were designed based on http://alma.tuebingen.mpg.de/cgi-bin/polymorph-clark20/caps_designer.cgi (Thiel et al., 2004; Clark et al., 2007; Zeller et al., 2008) to identify a restriction site corresponding to a SNP between Col-0 and Ts-1 (Because Bla-1 is not included in this web-based tool, Ts-1 was used to design CAPS markers due to their close relation and hypersensitivity to HU). The Bla-1 haplotype was then examined by 1001 genome browser to confirm whether the SNP exist between Col-1 and Bla-1 (http://alma.tuebingen.mpg.de/cgi-bin/polymorph-clark20/caps_designer.cgi). Primers were designed and the restriction sites were double-checked using SeqBuilder.

PCR primers for T-DNA insertion validation (Supplement Table 7) and qPCR primers (Supplement Table 8) were designed using Ensembl Genome Browser (http://atensembl.arabidopsis.info/index.html) and SeqBuilder.

3.7 Nucleic acids isolation and cDNA synthesis

DNA for normal PCR was extracted using the DNeasy 96 kit (Qiagen) and BioSprint 96 instrument. RNA was extracted using the RNeasy kit (Qiagen) with on column DNase I (Roche) treatment. cDNA was synthesized from 1 μg RNA per sample with Revert Aid H-Minus First Strand cDNA synthesis kit using the oligo-d(T) primer (Thermo Scientific). The purity of cDNA (whether there is genomic DNA contamination) was monitored by PCR with an intron-spanning primer pair.

3.8 Quantitative PCR (qPCR)

The reverse transcription (RT)-qPCR was performed using 1 μl cDNA per 10 μl reaction with SensiMix (PeqLab) kit on CFX384 1000 Touch™ Real-Time PCR Detection System instrument (Bio-Rad). 3 technical replicates were included in every sample, and stand curves were made by serial dilutions of a mixture cDNA samples. *ACT7* was used as internal control and the mock-treated sample as calibrator to calculate fold-changes in gene expression using standard curve method (Livak and Schmittgen, 2001).

3.9 Flow cytometry

For endoploidy analysis, cotyledons of 15-d-old plants growing on drug or drug-free media were dissected. 8 to 10 pieces were chopped with a razor-blade in 300 μL extraction buffer (Partec), filtered through 30 μm nylon mesh (Partec), stained with 900-1800 μL CyStain dye (Partec), and then analyzed with PAS I ploidy analyzer

(Partec). The endopolyploidy cycle value (CV) was calculated using given formula:

CV = ((n 2C*0) + (n 4C*1) + (n 8C+2) + (n 16C*3) + (n 32C*4)) / (n 2C + n 4C + n 8C + n 16C + n 32C), where n = number of counts per given C-value content (Barow and Meister, 2003).

3.10 GWA studies

The GWAs was performed in collaboration with Dr. Ronny Joosen at the Wageningen University (the Netherlands) and Dr. Giang T.H. Vu, Postdoctoral fellow at the Leibniz Institute of Plant Genetics and Crop Plant Research, Department of Cytogenetics and Genome Analysis. The EMMAX (efficient mixed-model association eXpedited) method was applied (Kang et al., 2010) with a kinship matrix for population structure correction. Plotting and $-\log_{10}$ (*P*-value) were accomplished by the R software (http://www.r-project.org/). All accessions used in this study were genotyped with the 250k SNP chip array by Affymetrix with the data publicly available (Kim et al., 2007; Atwell et al., 2010; Li et al., 2010). Every accession contained 214,051 markers, which led to a marker density of about 1 SNP marker / 500 bp, higher than the one commonly used in human studies (1 SNP marker / 6 kb) (Kim et al., 2007).

Gene annotations were according to the Arabidopsis Information Resource (TAIR) 9 (ftp://ftp.arabidopsis.org/home/tair/Genes/TAIR9_genome_release/).

Bonferroni corrected significance threshold (*P* < 0.05) was calculated based on the equation:

$$\text{threshhold} = \log\left(\frac{\text{SNP number}}{\frac{100\% - 95\%}{100\%}}\right),$$ which established the threshold at 6.63.

3.11 QTL mapping

As suggested by R/qtl (http://www.rqtl.org/faq/) for processing RIL population, the heterozygous genotypes were treated as missing and the whole genotype data were converted using the function "*convert2riself*", so the data were treated like a backcross, but the map is expanded before calculating QTL genotype probabilities.

Simple interval mapping and two-dimensional genome scan were performed using R package: qtl ("*scanone*" for simple interval mapping and "*scantwo*" for two-dimensional scans). Log of odds (LOD) threshold values were estimated by permutation test (1000 runs) corresponding to 5% significance.

4. Results

4.1 Natural variation of *Arabidopsis* to DNA replication stress

4.1.1 Phenotypic screening of *Arabidopsis* natural accessions in response to HU

To dissect the natural variation of *A. thaliana* in response to HU, the 391 natural accessions were screened on 4 mM HU. Plant phenotypes were divided into three categories (Fig. 2). The first category consists of seedlings showing a normal and vigorous growth habit with green cotyledons, branched roots, and partial development of true leaves. These seedlings were considered to be viable and given the value of 1. The second category consists of seedlings which looked pale and were smaller than viable plants, having narrowed and back-folded cotyledons, long hypocotyls, and

poorly developed or not elongated roots. These dead plants were given the value of 0. The third category consists of plants showing phenotypes between the first and the second categories. These plants were considered to be intermediate and given the value of 0.5.

Viable	Intermediate	Non-viable
1	0.5	0
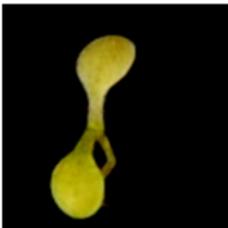
Fig. 2: Classification of Arabidopsis natural accessions of HU-induced phenotype. Seedlings are divided into three groups: viable (left), intermediate (middle), and dead (right).

The relative survival rate (RS) were then calculated as follows:

RS = (Number of viable plants × 1 + number of intermediate plants × 0.5 + number of dead plants × 0) / number of total plants × 100 (%)

To exclude bias in subjective evaluation criteria, the RS were independently re-evaluated by two laboratory members, and the mean values were taken as the final phenotypic data.

Combining the two biological replicates accomplished by Alexander Schott (2012) and the additional one by myself, there were in total three biological replicates in this phenotypic screening. The fourth replicate was performed for the accessions with high standard deviation from the previous three replicates. Accessions with low germination

(germination rate < 33%) and those with ambiguous phenotypes were excluded from further investigations. 312 accessions passed the quality control and revealed a wide range of RS from 0% to 100% (Fig. 3). The majority of the natural accessions showed resistance to HU with RS higher than 80 %, and only 13 accessions showed reproducible hypersensitivity to HU with a RS less than 30%. This includes N6, Wei-0, Jm-1, Go-0, Jl-3, Zu-1, Alc-0, Bla-1, CIBC-17, Ste-3, TOU-C-3, TOU-I-6, and N6, in which TOU-C-3, and Ts-1 were considered to be the most sensitive accessions, showing a RS below 10%.
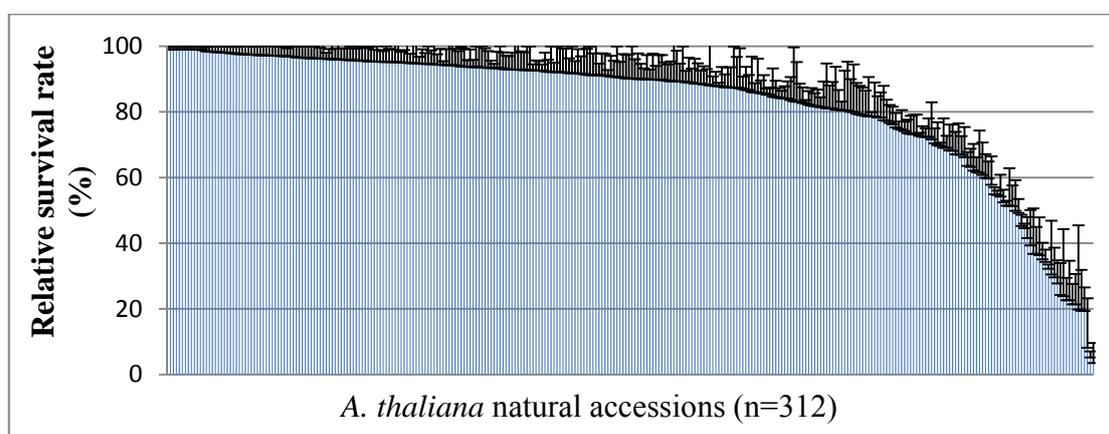
Fig. 3: Distribution of RS of *A. thaliana* natural accessions in response to HU. Error bars represent the standard deviation among three to four biological replicates. The original data is listed in Supplement Table 1.

4.1.2 GWA studies of HU sensitivity

In order to dissect the genetic variation of the natural accessions in response to HU, GWA studies were performed. Three different phenotypic profiles were made, including the total combination of all four replicates (Fig. 4a), the third replicate which was done by the same person (Fig. 4b), and the combination of all the replicates after stringent quality control (standard deviation < 15) (Fig. 4c).

The GWA maps differed among the three profiles. Both Fig. 4a and 4c show several peaks of associations with significant SNP markers, whereas Fig. 4c shows a stronger peak of association at the bottom arm of chromosome 5, indicating higher quality of phenotypic data might increase the mapping power in GWA studies (Zhu et al., 2008). No significant peaks of association were detected in Fig. 4b, the phenotypic profile with only one biological replicate.
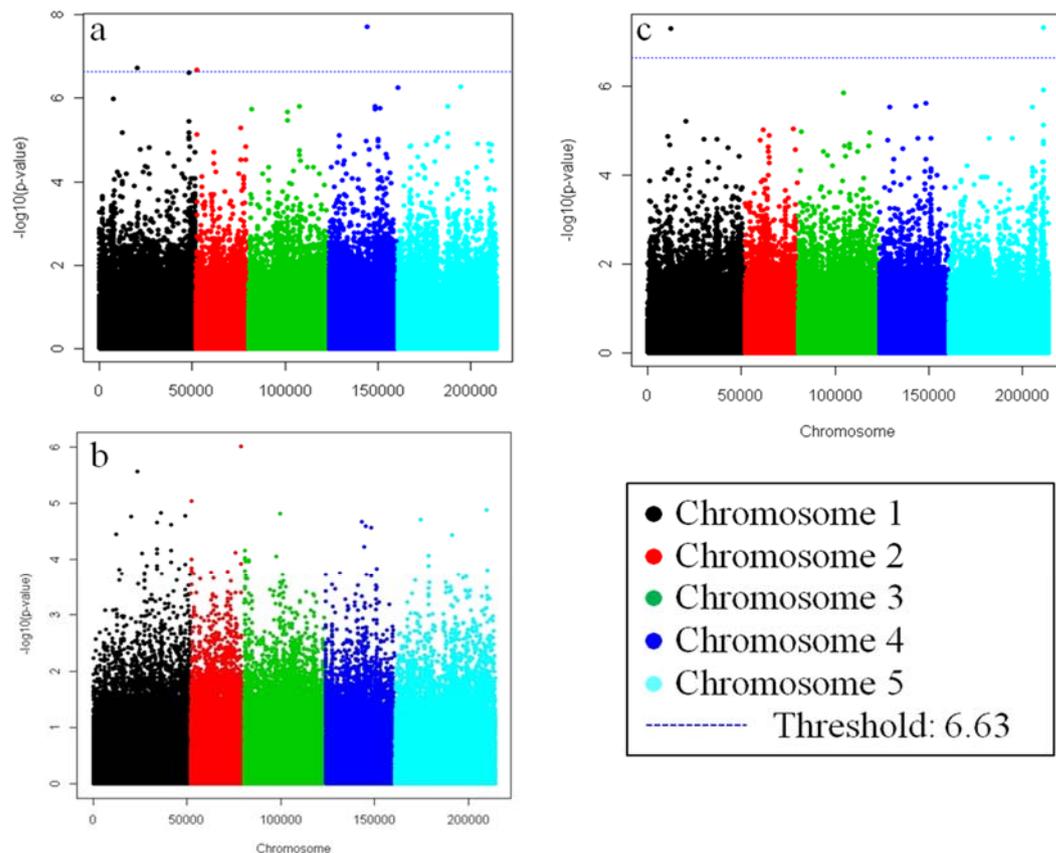
Fig. 4: GWA maps of HU sensitivity. The x-axis shows the position of SNP markers on each chromosome represented by five different colors, and the y-axis shows the -log$_{10}$ (*P*-values). a) The phenotypic profile of total combination of the four biological replicates, b) the profile of third replicate, and c) the profile with stringent quality filtering.

4.1.3 Candidate genes selection suggested by GWA studies

Based on the results of GWA studies, the associated *P*-values corresponded to each SNP marker were manually examined. Candidate genes were selected based on three criteria: first, genes underlie the region of significant peak of association; second, genes closely linked to the SNP marker with significant *P*-value; and third, genes with known or potential function in the response to DNA replication stress. The pseudo-reference sequences of the underlying genes in the candidate region were aligned and compared in HU-resistant and sensitive accessions using 1,001 Genome Browser (http://signal.salk.edu/atg1001/3.0/gebrowser.php).

As depicted in Fig. 5, focusing on the significantly associated region at chromosome 5 (Fig. 4c), an *OXIDOREDUTASE* gene and *CYCLIN DEPENDENT KINASE E* gene underlie in the vicinity. Alignment of the sequences in this region in two HU-sensitive accessions, Bla-1 and Ts-1, shows the haplotypes of the underlying genes are similar, comparing to the haplotypes in HU-resistant accessions, Col-0 and Cvi-0. This implies that the presence or absence of polymorphisms underlying the candidate genes might be responsible for HU sensitivity or resistance.
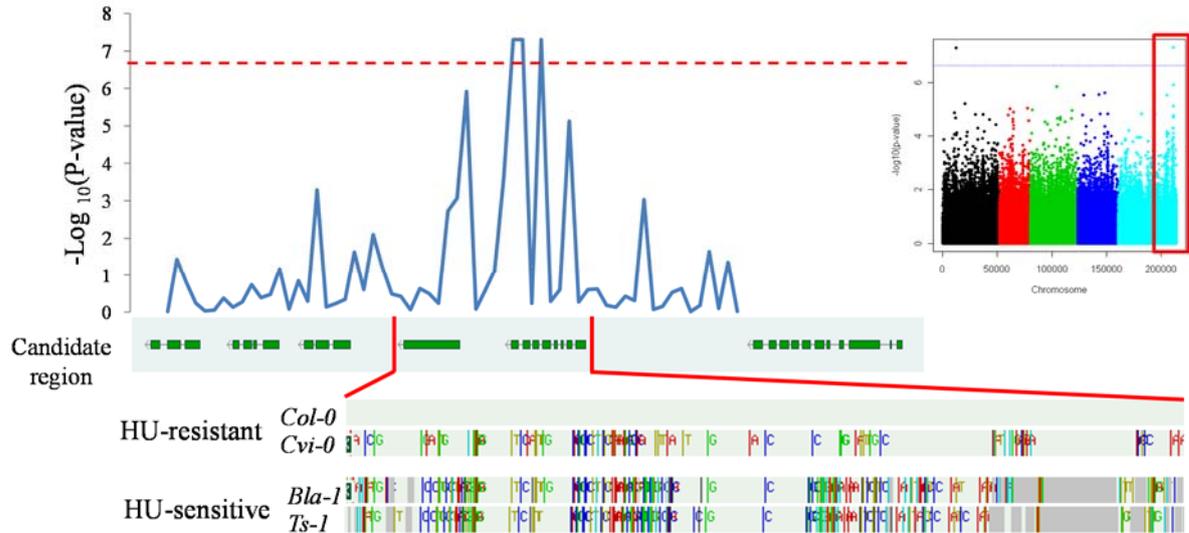
Fig. 5: The depiction of how the candidate genes were selected using the significantly associated region at chromosome 5 (Fig. 4c) as an example. The red rectangular in GWA map (left) marks the selected peak of association; the main figure in the middle is the magnification of the selected associated region, and the red dash line corresponds to the significant threshold in GWA maps; the green boxes below the main figure correspond to the protein-coding genes lying in the associated region; and the sequences alignments of the accessions at the bottom show the polymorphisms between the accessions and Col-0.

Genes lying in the associated regions and in the vicinity of 20 kb were considered, since LD in *A. thaliana* decays within 10 kb on average (Kim et al., 2007; Atwell et al., 2010). Especially genes functionally connected to cell cycle control and DNA damage repair, as previously described the HU-induced responses (Wang and Liu, 2006; Cools et al., 2011; Spadafora et al., 2011), were considered as potential candidates.

In total, nine candidate genes (named as *nvh*: natural variation to HU) were selected (Table 1).

Table 1: The list of candidate genes suggested by GWA studies of HU sensitivity. R14, 3R, and R3 represent the three phenotypic profiles for GWA studies, corresponding to the profile combining all four replicates, the profile after stringent quality control, and the replicate done by myself, respectively. *nvh*: natural variation to HU.

Name in this study	SNP position in Col-0	AGI code	*GENE NAME* or description	-LOG10 (*P*-value)		
				R14	3R	R3
nvh2	19807550	AT1G53160	*FLORAL TRANSITION AT THE MERISTEM6*	1.45	4.30	1.27
nvh3	26442421	AT1G70210	*CYCLIN D1*	3.70	2.89	3.83
nvh4	28722132	AT1G76540	*CYCLIN DEPENDENT KINASE B2;1*	0.90	1.40	3.91
nvh6	319734	AT2G01720	Ribophorin I	6.74	2.67	3.54
nvh11	875406	AT3G03620	MATE efflux family protein	4.93	4.97	3.64
nvh13	21929026	AT5G54040	Cysteine/Histidine-rich C1 domain family protein	1.12	3.98	0.61
nvh14	21935221	AT5G54050	Cysteine/Histidine-rich C1 domain family protein	1.46	5.53	1.41
nvh18	25465538	AT5G63610	*CYCLIN DEPENDENT KINASE E (CDK8; HEN3)*	2.38	5.92	2.51
nvh19	25467896	AT5G63620	GroES-like zinc-binding alcohol dehydrogenase family protein	3.85	7.31	3.29

To validate the candidate genes responsible for HU-induced phenotype (HU-sensitivity), the T-DNA insertion mutant lines with Col-0 (the HU-resistant accession) background were obtained (Supplement Table 5), and the homozygote mutants were screened on 4 mM HU. However, all the mutants showed resistant phenotype as the WT (Col-0) (Fig. 6), suggesting a more complex trait of HU sensitivity and the genetic basis might be multigenic rather than monogenic.
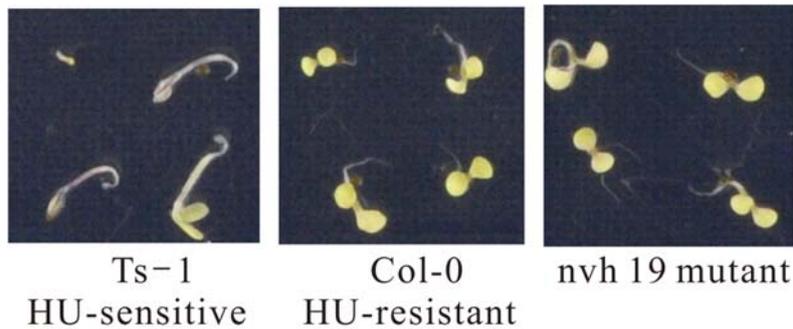
Fig. 6: Representative phenotypes of *nvh 19*, Ts-1, and Col-0 growing on HU for 11 days.

4.1.4 QTL mapping of HU sensitivity using Bla-1×Col-0 RIL population

With widely-distributed collections of mapping population and the publicly available SNP genotype data, GWA studies provide a shortcut to analyze the genetics basis underlying the phenotypic variation with high resolution, however, the statistical power is unpredictable due to confounding effect of population structure and the inflation of false positive association and the limitation of genetic heterogeneity when it comes to complex genetic architecture (Bergelson and Roux, 2010). Thus, to complement the limitation of GWA studies and to dissect the multigenic basis of HU sensitivity, QTL mapping was performed. A publicly available population of 163 RILs derived from a cross between HU-sensitive and HU-resistant accessions, Bla-1 and Col-0, respectively, were screened on 4 mM HU and evaluated for RS of individual lines. The distribution of RS among the RIL population was continuous and showed a well transgressive phenotype beyond the parental values (Fig. 7).
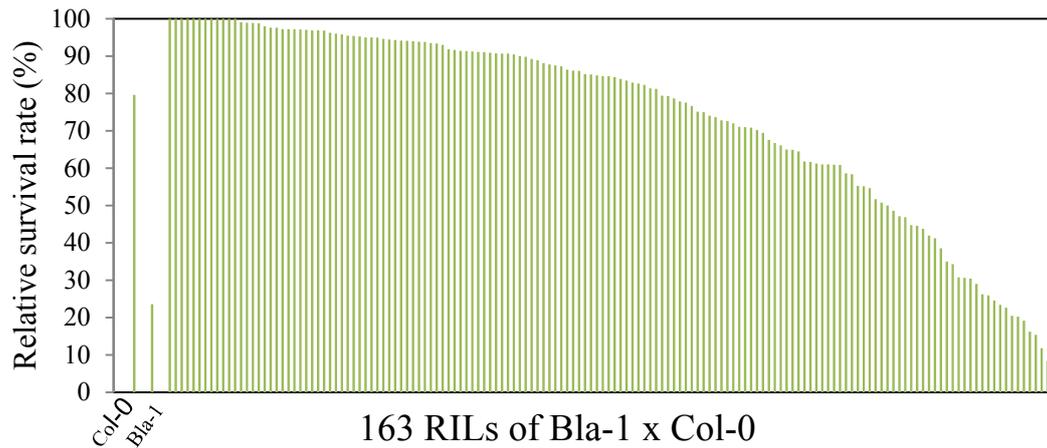
Fig. 7: Phenotypic distribution of Bla-1x Col-0 RIL population to HU. The original data is listed in Supplement Table 2.

With a genetic map of 78 markers established by Simon et al. (2008), QTL analysis was performed. Two major QTLs at the top arm of chromosome 3 and the bottom arm of chromosome 5, respectively, were identified with the LOD scores above the significant threshold (2.48). Minor QTLs appeared at chromosome 1, 2, and 4, respectively (Fig. 8). The intervals of the two major QTLs contain loci of around 8 Mb, with the Bla-1 allele responsible for HU sensitivity in both loci.
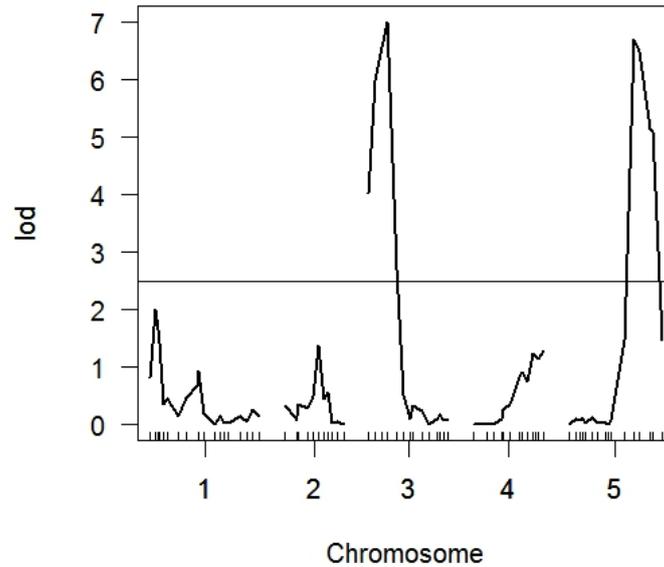
Fig. 8: QTL map of HU sensitivity in Bla-1 x Col-0 RIL population. Ticks along the x-axis represent the positions of genetic markers on each chromosome, and the numbers on y-axis represent the LOD scores. The LOD scores were calculated by standard interval mapping with a single-QTL model. The horizontal line represents the significant threshold for the LOD score: 2.48, estimated after 1000 permutations.

Two-dimensional genome scan was performed with a two-QTL model (Fig. 9). An additive effect between the two major loci was identified, for which the Bla-1 allele on both loci interacts additively to promote HU sensitivity.
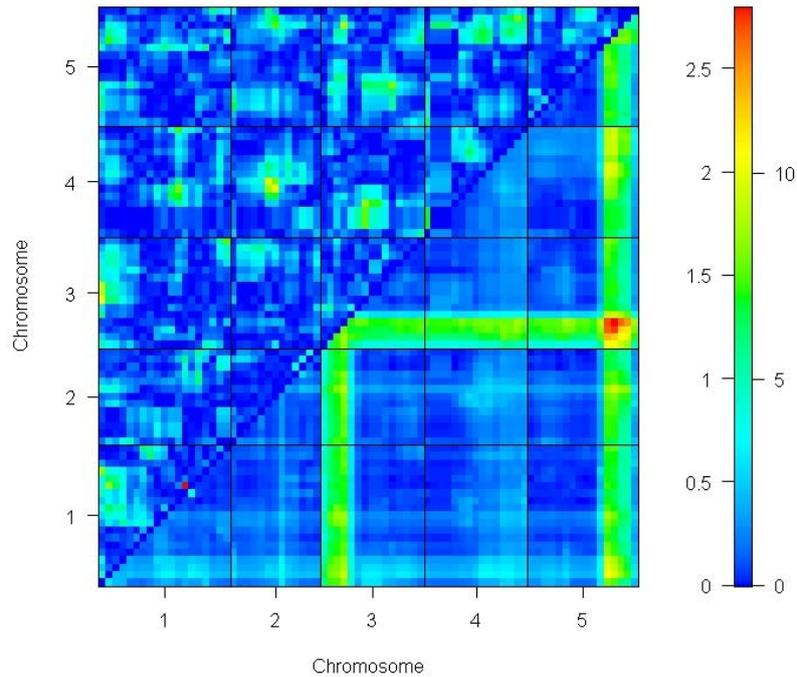
Fig. 9: Heat map for two-dimensional genome scan of HU sensitivity in Bla-1 x Col-0 RILs. Values in the lower right-triangle are for joint two-locus LOD scores, showing additive interactions between markers; values in the upper-left triangle are LOD scores for the test of epistasis, showing epistatic interactions between markers. The color scale at the right indicates scales for the joint LOD scores (numbers at the right side) and the epistasis LOD scores (numbers at the left side).

To further dissect the additive or epistasis interactions between specific markers, markers located in the QTLs region were selected based on the results of two-dimensional genome scan, and the function of "*effecplot*" and "*plot.pxg*" in R were employed.

Fig. 10a shows an additive effect between two selected markers at QTL3 and 5. This analysis confirmed that Bla-1 homozygote genotype (BB) at QTL3 and 5 alone would cause an increase in HU sensitivity (lower RS). This also validated the additive effect contributing by QTL3 and 5, in which both of the loci having BB reduced the RS to 55%.

The range of RS observed from the RILs was from 0% to 100%, however, the lowest RS contributed by the additive effect of QTL3 and 5 was only 55%, suggesting other QTLs contributing to HU sensitivity. The minor QTL on chromosome 1 showed epistasis effects on both QTL3 and 5. The contribution of QTL3 to HU sensitivity slightly increased when combining Col-0 homozygote genotype (AA) at QTL1 (Fig. 10b). A stronger epistatic effect was observed between QTL1 and 5 (Fig. 10c), in which RS dropped 20% when AA at QTL1 was considered.
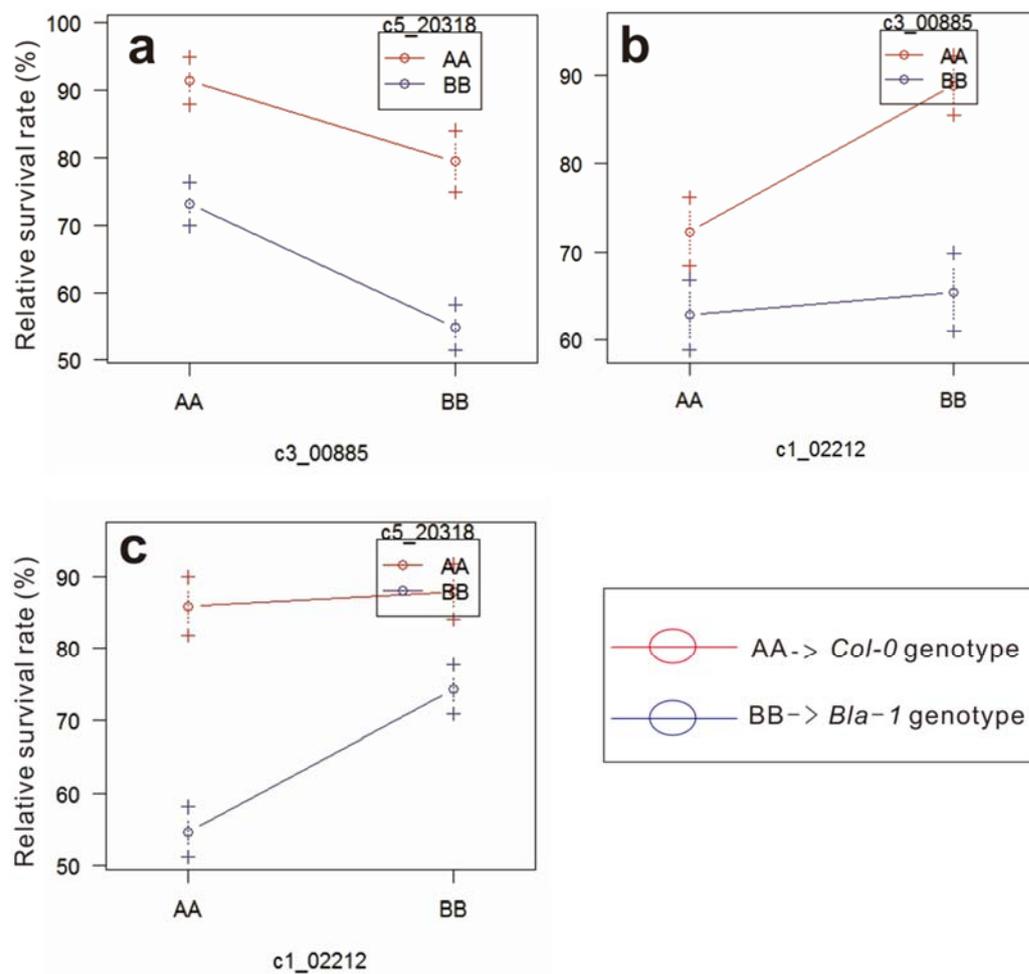
Fig. 10: Two-locus effect plots of QTLs for HU sensitivity in Bla-1 x Col-0 RILs. The x-axis represents the genotype of the selected marker, and y-axis represents RS values of the plants carrying specific genotype at given marker position. Blue and red colors represent Bla-1 (BB) and Col-0 (AA) genotypes, respectively. c1_02212 corresponds to the marker associated with QTL 1, c3_00885 with QTL3, and c5_20318 with QTL5. The circles and crosses indicate the estimated RS mean values and the standard deviation, respectively.

To further analyze the effects of QTL1, 3, and 5 on HU sensitivity, the RS of 163 RILs were divided into eight recombinant classes based on the three QTLs (Fig. 11). Individuals having the recombinant of BB at QTL3 and 5 and AA at QTL1 showed lowest RS mean value, whereas the highest RS mean value was observed in the recombinant of AA at QTL3 and 5 and BB at QTL1. Other recombinant classes showed mean values in between. This further confirms that the genetic basis of HU sensitivity is contributed by an additive effect of QTL3 and 5 with epistatic effect of QTL1.
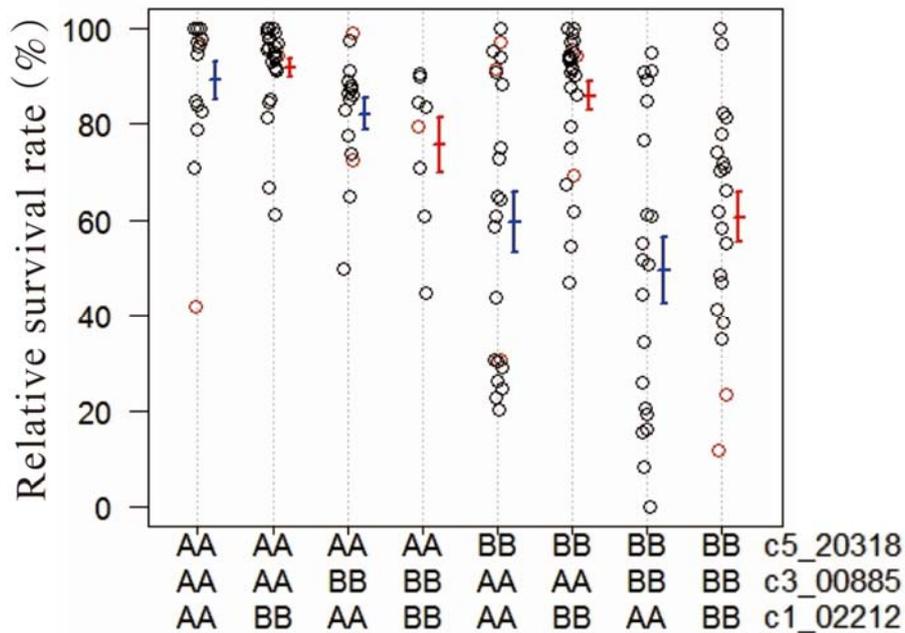
Fig. 11: Dot plot of RS to HU in Bla-1 x Col-0 RILs as a function of markers genotypes at QTL1, 3, and 5. Black dots correspond to the observed genotypes, and red dots correspond to the missing (and so imputed) genotypes. AA represents homozygote Col-0 genotype, and BB represents homozygote Bla-1 genotype. c1_02212 corresponds to the marker associated with QTL 1, c3_00885 with QTL3, and c5_20318 with QTL5.

4.1.5 Confirmation and fine mapping of QTL3 by HIF analysis

To validate the identified QTLs, HIF analysis was employed. Due to the limited number of available RILs, only HIFs segregating for QTL3 were selected. Genotypes of the

segregating progenies were characterized by CAPS markers, and the seeds from each individual line were sown on 4 mM HU for their RS. HIF383 showed the expected segregation phenotype (Fig. 12). The progeny with Col-0 allele showed resistant phenotype to HU, whereas the progeny with Bla-1 allele was sensitive to HU. The genotypes of the segregating locus on the long arm of chromosome 2 were characterized by CAPS markers but did not show any obvious link to HU sensitivity.
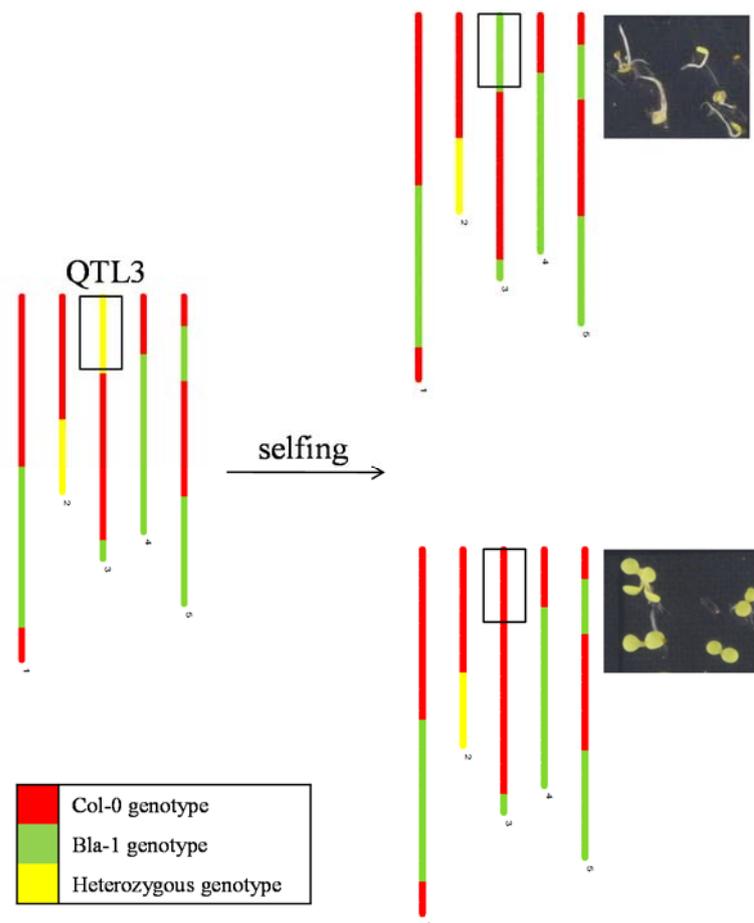
Fig. 12: The scheme showing QTL3 confirmation by HIF analysis. HIF383 with segregating region at QTL 3 confirmed QTL3 is responsible for HU sensitivity. The five bars represent the five chromosomes, indicated by the numbers below. The photos at right show the segregating phenotypes of the progeny after 4 mH HU treatment for 11 days.

About forty lines of the progeny of HIF383 were genotyped and phenotyped on HU. The difference in RS for the homozygous plants (homozygous for Col-0 and Bla-1) and

between the Bla-1 homozygous plants and heterozygous plants were both significant, while the differences between heterozygous plants and plants homozygous for Col-0 allele were not (Fig. 13), indicating that the Bla-1 allele responsible for HU sensitivity is recessive.
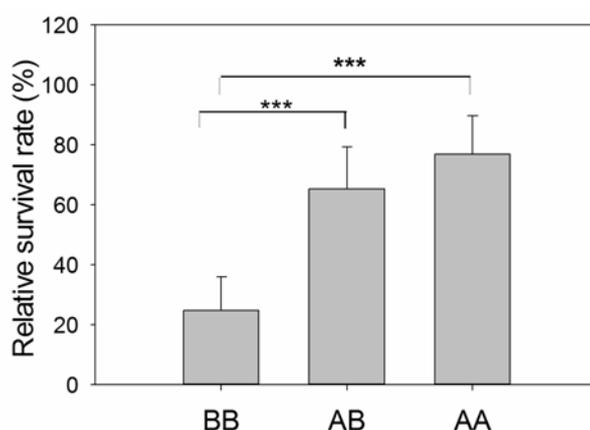
Fig. 13: RS of HIF 383 of Bla-1 homozygous plants (BB), heterozygous plants (AB), and Col-0 homozygous plants (AA). Error bars represent the standard deviation of eight to nine biological replicates. Asterisks indicate significantly difference in RS mean values (***Turkey Test, P≤0.001).

To fine map the causal genes underlying QTL3, a sequential QTL fine-mapping strategy using recombinant-derived progeny was employed (Yang et al., 2012). New markers were added to delimit the QTL region (Supplement Table 6), and large amount of descendants of heterozygous plants at the QTL interval were genotyped and phenotyped to accumulate more recombinants. New recombinants were self-pollinated again to produce the next progeny for further fine mapping. This fine-mapping process was sequentially carried out from F7 to F13 generation, and approximately 1500 individuals were screened until the eight candidate genes (named as *hhu*: hypersensitive to HU) underlying QTL3 were identified (Table 2).

Table 2: The list of candidate genes underlying QTL3 for HU sensitivity. *hhu*: hypersensitive to HU.

Name	AGI code	*GENE NAME* or description
hhu1	AT3G10130	SOUL heme-binding family protein
hhu2	AT3G10140	*RECA HOMOLOG 3* (*RECA3*)
hhu3	AT3G10150	*PURPLE ACID PHOSPHATASE 16* (*PAP16*)
hhu4	AT3G10160	*DHFS-FPGS HOMOLOG C* (*DFC*)
hhu5	AT3G10180	P-loop containing nucleoside triphosphate hydrolases (P-loop NTPases) superfamily protein
hhu6	AT3G10185	Gibberellin-regulated GASA/GAST/Snakin family protein
hhu7	AT3G10190	Calcium-binding EF-hand family protein
hhu8	AT3G10195	defensin-like (DEFL) family protein

4.2 The effect of zebularine treatment on Arabidopsis genome stability

4.2.1 Genome-wide expression analysis of Arabidopsis to zebularine (by Dr. Pecinka)

To understand the effect of zebularine on Arabidopsis, a genome-wide transcription analysis was performed. RNA was extracted from dissected shoot apices of 12-day-old WT after 24 hours (short) and 5 days (long) treatment of zebularine. 31 and 678 genes were up-regulated and 12 and 392 genes were down-regulated after short- and long-zebularine treatment, respectively (Fig. 14a). 38.7% of up- and 50% of down-regulated genes after short zebularine treatment overlapped with the set of genes differentially transcribed after long exposure, indicating a duration-dependent and contrasting effects

of zebularine treatment on the Arabidopsis transcriptome.

Zebularine was shown to reactivate several transcriptionally silenced genes (Baubec et al., 2014). To investigate whether the transcriptional changes by zebularine treatment were due to reactivation of transcriptional gene silencing (TGS), the RNA-sequencing dataset in this study were compared with the dataset in *ddm1* (the mutant of *DECREASED DNA METHYLATION 1*, *DDM1*). *DDM1* is a nucleosome remodeler required for DNA methylation and stable silencing of transposable elements (Zemach et al., 2013). No overlap was found for short zebularine treatment and only 4 (TE_gene AT1G42050; *MuDr* AT2G15810, *LINE1-6* AT3G28915, and *Gypsy-like* AT5G35057) out of 908 genetic elements up-regulated in *ddm1* were also significantly up-regulated after the long zebularine treatment. Hence, less than 1% of the zebularine-up regulated genes were TGS targets.
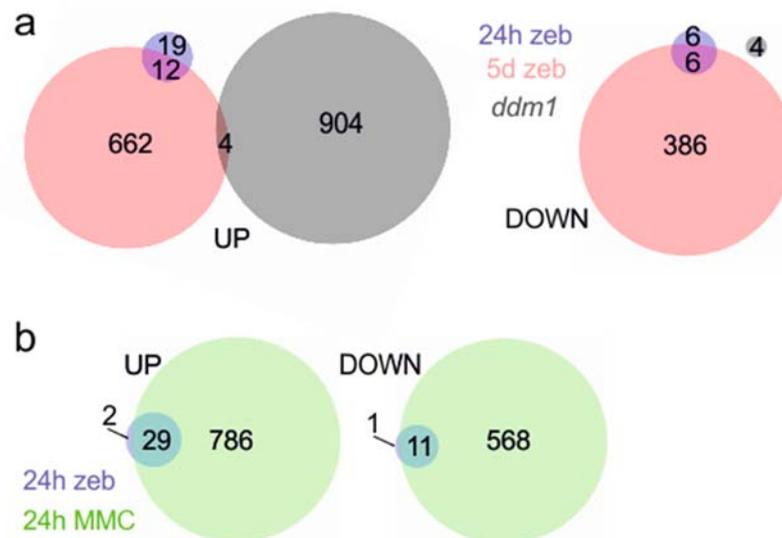
Fig. 14: Zebularine-or MMC-effect on Arabidopsis transcriptome. a) Genes significantly up- or down-regulated in response to 24 hours (blue) and 5 days (pink) 20 µM zebularine treatment of WT. The number in the gray circle indicates the genes up-regulated in *ddm1*. b) Significantly up- and down-regulated genes in response to 24 hours zebularine (blue) and 24 hours 10 µM MMC treatment (green).

Functional analysis (TAIR10) of the 31 genes induced by short- zebularine treatment revealed that 25.8% are linked with DNA metabolism and DNA damage repair. Furthermore, 6 out of 15 genes commonly up-regulated after both short- and long-zebularine treatment were functionally connected to DNA damage repair (Table 3).

Table 3: List of common genes up-regulated after 24 hours and 5 days zebularine treatment. Genes functionally connected to DDR are in bold.

AGI Locus	Gene annotation
AT1G20750	***RAD3-LIKE***
AT2G21790	***RIBONUCLEOTIDE REDUCTASE LARGE SUBUNIT 1***
AT3G03780	*METHIONINE SYNTHASE 2*
AT3G07800	***THYMIDINE KINASE 1A (TK1A)***
AT3G13470	*CHAPERONIN-60BETA2*
AT3G15950	*NAI2*
AT3G19680	Protein of unknown function
AT3G27060	***TSO MEANING 'UGLY' IN CHINESE 2 (TSO2)***
AT3G27630	***SIAMESE-RELATED 7 (SMR7)***
AT3G54810	*BLUE MICROPYLAR END 3*
AT4G21070	***BREAST CANCER SUSCEPTIBILITY1 (BRCA1)***
AT4G22410	*UBIQUITIN C-TERMINAL HYDROLASE PROTEIN*
AT4G22880	*LEUCOANTHOCYANIDIN DIOXYGENASE*
AT5G20850	***RAS ASSOCIATED WITH DIABETES 51 PROTEIN (RAD51)***
AT5G42800	*DIHYDROFLAVONOL 4-REDUCTASE*

The transcriptional changes of these common DDR genes were validated by RT-qPCR using the same RNA samples for RNA sequencing (Fig. 15). DDR genes showed 2- to 3-fold increase in transcription after short-zebularine treatment and 5- up to 30-fold increase after long-zebularine treatment.
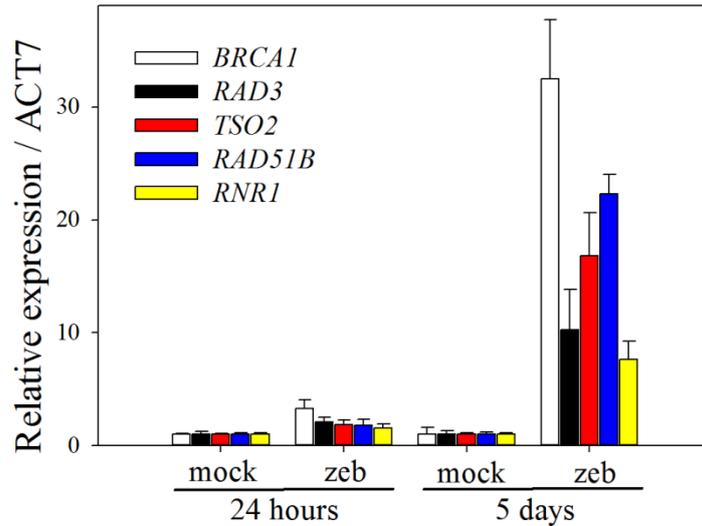
Fig. 15: RT-qPCR measurement to validate the transcriptional changes in common DDR genes up-regulated after both short- and long-zebularine treatment in RNA sequencing. Error bars represent standard deviation of two biological replicates.

To confirm whether the transcriptional changes by zebularine represent a bona fide response to DNA damage stimuli, plants were exposed to 10 μM MMC for 24 hours. 815 and 579 genes were significantly up- and down-regulated, respectively (Fig. 14b), including numerous DNA damage repair genes. The sets of genes up- and down-regulated in response to 24 hours of zebularine exposure were compared with the sets after MMC treatment. 93.1% (29 out of 31) and 91.7% (11 out of 12) were overlapped in up- and down- regulated genes, respectively.

4.2.2 Histochemical staining and expression analysis of DNA damage repair genes after zebularine treatment

To dissect the dynamics of zebularine-induced DDR in Arabidopsis and its tissue specificity, a *pGMI1::GUS* reporter line was used, which allows the visualization of the spatial expression of drug-induced DDR. *GMI1* was characterized as *GAMMA-IRRADIATION AND MITOMYCIN C INDUCED 1*. It was shown to be activated after

treatment with several genotoxic agents, including long-zebularine treatment in this study, and was shown to be involved in DNA repair (Böhmdorfer et al., 2011). The reporter lines were exposed to zebularine, MMC, and bleocin for the appointed time points. There was no GUS detected after one hour drug treatment. Three hours of bleocin and six hours of MMC or zebularine treatment were sufficient to induce GUS accumulation, mainly in the shoot apices, petioles of the youngest leaves, and in cotyledons (Fig. 16). GUS accumulation became more prominent over time in true leaves and cotyledons, also in root apical meristems (RAMs) in MMC- and bleocin-treated plants, but not in zebularine-treated plants. These results suggest a rapid induction of *GMI1* by zebularine and its different processing or stability in root and shoot apical meristems.
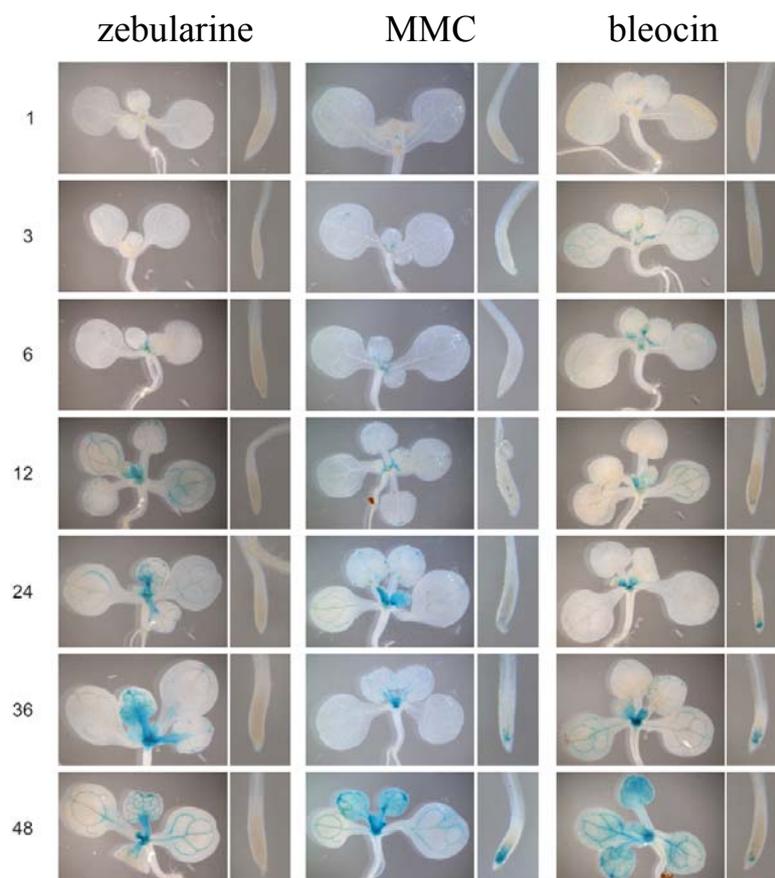
Fig. 16: Histochemical staining of *pGMI1::GUS* reporter line after given hours of treatment (as indicated by the numbers at the left side) with 20 µM zebularine, 10 µM MMC and 100 nM bleocin.

To further assess the transcriptional dynamics of *GMI1* and the DDR genes activated by zebularine- and other drug-treatments in more details, RT-qPCR was performed in dissected shoot apices of mock- and drug-treated plants over 24-hour time series. *RAD51* and *BRCA1* were up-regulated after both short- and long-zebularine treatment, whereas *GMI1* and *POLY (ADP-RIBOSE) POLYMERASE 2* (*PARP2*) were activated after long-zebularine treatment. The amount of transcripts did not follow a simple accumulation over time as observed in the case of GUS staining, probably reflecting the higher stability of the GUS protein compared to *GMI1* mRNA. Other tested DNA damage repair genes were also activated in response to the drug- treatments, and zebularine treatment showed similar kinetics and amplitudes to MMC- and bleocin-treatments (Fig. 17).
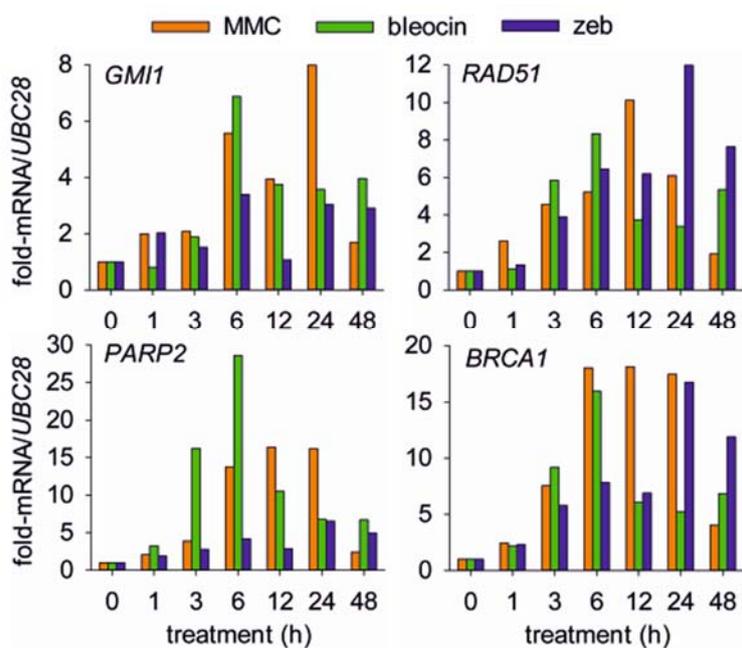
Fig. 17: RT-qPCR analysis of DNA damage repair marker genes in dissected shoot apices after given hours of treatment with 20 µM zebularine, 10 µM MMC and 100 nM bleocin. The bars represent the mean values of transcription from a pool of five to ten seedlings.

4.2.3 Analysis of the effect of zebularine on DNA methylation (by Dr. Finke) and the formation of zebularine-DNMT NPAs

Zebularine was shown to inhibit DNMT activity in human cancer cells (Billam et al., 2010) and to reduce the level of DNA methylation in Arabidopsis (Baubec et al., 2009). To understand whether the activation of DDR genes by 20 µM zebularine treatment was due to DNA de-methylation, methylated DNA regions < 1 kb upstream of *TSO2* and *RAD51*, two DNA damage repair genes activated by zebularine treatment, were identified, and methylation analysis of these regions by bisulfite sequencing was performed. With the same treatment and plant material used in RNA sequencing, dissected shoot apices of mock, short, and long zebularine-treated plants revealed less than 5% reduction in DNA methylation. Similarly, *LINE1-6* retrotransposon (AT3G28915/AT3TE45385) identified as a common target of zebularine and *ddm1* activation, and the repetitive region upstream of the *SUPPRESSOR OF DRM1 DRM2 CMT3* (*SDC*) gene (Henderson and Jacobsen, 2008) up-regulated by long zebularine treatment, the DNA methylation in both cases were maintained and didn't show significant changes (Fig. 18). This suggests zebularine-induced transcriptional activation of DDR genes may occur without changes in DNA methylation.
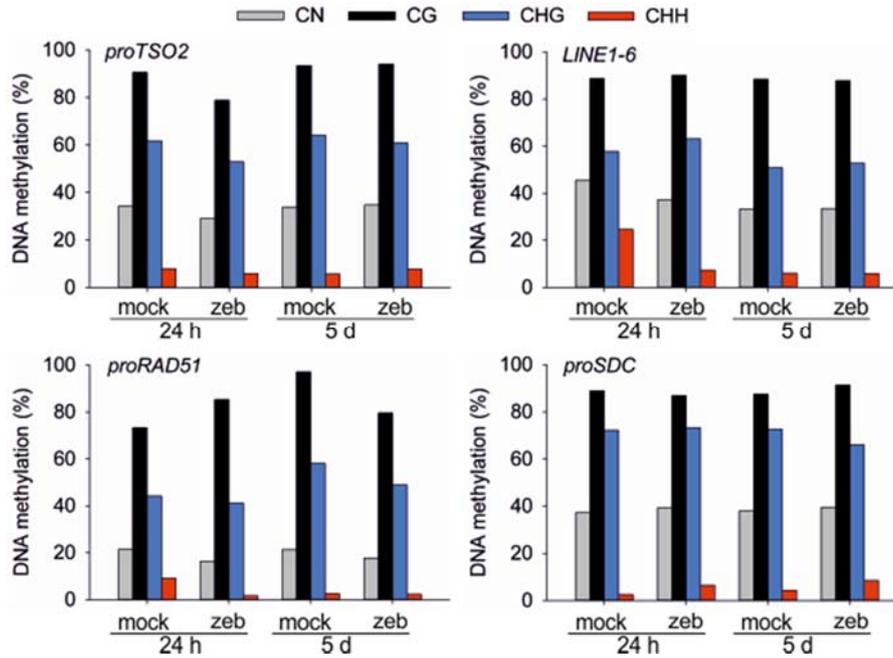
Fig. 18: Percentage of DNA methylation in dissected shoot apices based on bisulfite sequencing of 24 hours and 5 days mock- and 20 µM zebularine-treated samples. A minimum number of 15 reads per experimental point has been analyzed.

To investigate whether the formation of NPAs is the cause of zebularine-induced DDR, RT-qPCR analysis was performed in the *CMT3*, *DRM1*, and *DRM2* triple homozygous mutant (*ddc*) (Henderson and Jacobsen, 2008). If the reduction of available DNMTs , i.e. limited formation of NPAs between zebularine and DNMT in *ddc*, would restrict the transcriptional activation of DDR genes by zebularine, the formation of NPAs could be a cause of zebularine-induced DDR. WT and *ddc* were exposed to mock and 20 µM zebularine for 24 hours, and the shoot apices were dissected for RT-qPCR analysis. *TSO2*, *BRCA1*, *PARP2*, and *RAD51B* were 3.5 to 5.5-fold up-regulated in response to zebularine in WT, whereas only a less than 2-fold up-regulation was found in zebularine-treated *ddc* (Fig. 19a), suggesting the formation of NPAs between zebularine and DNMTs could contribute to the DDR stimuli. To further test the effects of NPAs formation on Arabidopsis, root length assay was performed in WT and *ddc* by exposing the plants to zebularine, MMC, and bleocin for 7 days. Comparing to the other two drugs, zebularine-induced root growth inhibition was significantly reduced in *ddc*

48

compared to WT (Fig. 19b). Therefore, the DNMT-zebularine NPAs seem to be at least partially responsible for the activation of DDR and toxicity by zebularine.
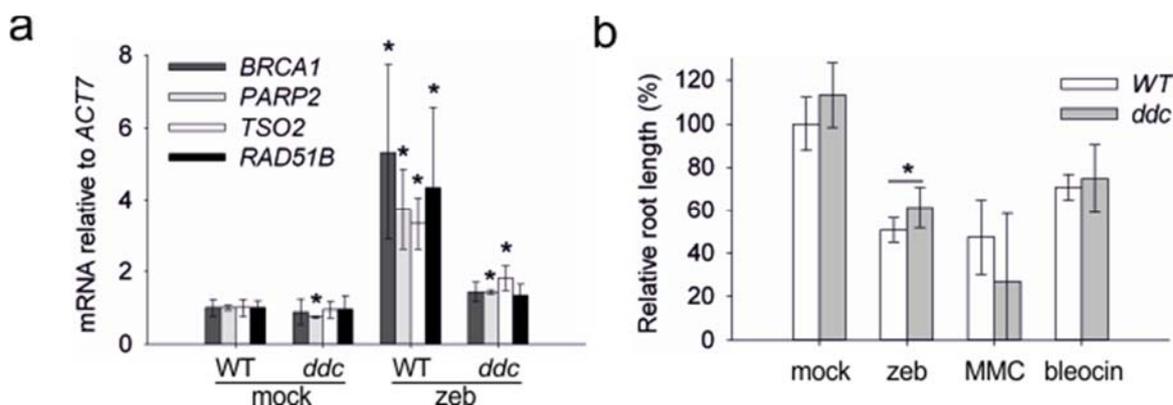
Fig. 19: The effect of *ddc* on zebularine-induced DDR and cytotoxicity. a) RT-qPCR measurement of DNA damage marker genes in dissected shoot apices of WT and *ddc* after 24 hours treatment with mock and 20 µM zebularine normalized to *ACTIN7*. b) Relative root length of WT and *ddc* in response to 20 µM zebularine, 15 µM MMC or 50 nM bleocin treatment. Relative changes were calculated as mean values of drug-treated plants / mean values of mock-treated plants. Error bars represent standard deviation of three biological replicates. Asterisks indicate statistically significant differences in *t*-test ($P < 0.05$).

4.2.4 The activation of DDR signaling transducers by zebularine treatment

The high percentage of overlap (> 90%) between MMC and zebularine-induced transcriptional changes indicated that the repair of DDR induced by the two drugs share partially a common pathway. The repair of MMC-induced inter-strand crosslinks is thought to be mediated mainly by ATR (Culligan et al., 2004; Nezames et al., 2012), thus a genome-wide transcriptional analysis was performed in the shoot apices of *atr* mutant treated with 24 hours of MMC or zebularine, and the transcriptional changes in *atr* were compared with WT. 227 and 119 genes were significantly up- and down-

regulated in mock-treated *atr*, and gene ontology (GO) analysis (by Dr. Pecinka) showed 70 and 20 significantly enriched GO term categories corresponding to stress and immune responses, respectively, suggesting the role of ATR in controlling responses toward environmental stimuli. 62 and 78 genes (29 in common) were significantly up-regulated, whereas 363 and 421 genes (225 in common) were significantly down-regulated in *atr* after 24-hour zebularine and MMC treatment, respectively (Fig. 20a), confirming ATR as a positive regulator of transcriptional responses to stresses (Yoshiyama et al., 2013a). Notably, among the common genes up- and down-regulated in zebularine-treated WT and *atr*, very low numbers were found (4 and 2 in respect to up- and down-regulated), indicating the transcriptional responses to zebularine is ATR-dependent (Fig. 20b). The same phenomenon was less pronounced in MMC treatment, where 50% of up-regulation (408 out of 815) and 61% of down-regulation (353 out of 579) occurred in an ATR-independent manner (Fig. 20c).
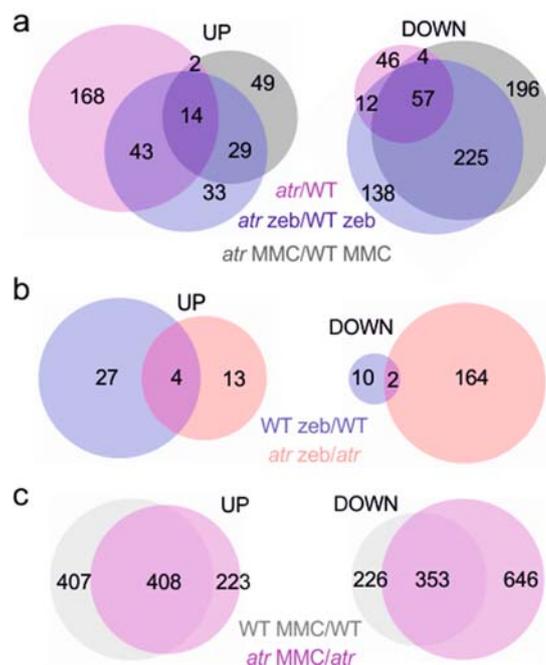
Fig. 20: *atr* effects on gene transcription after zebularine or MMC treatment. a) Numbers of genes significantly up- or down- (UP and DOWN, respectively) regulated in atr in response to mock, 24-hour 20 µM zebularine or 10 µM MMC treatment. The *atr* samples were compared to WT samples in the same treatment. b) Zebularine-*atr* effects on gene transcription. Blue circles show genes significantly up- or down-regulated in response to 24-hour zebularine exposure. Pink circles depict genes significantly up- or down-regulated in zebularine-treated relative to mock-treated *atr*. The genes in overlap are up- or down-regulated in response to zebularine and independent of ATR. c) MMC-*atr* effects on gene transcription analyzed as described in b) (by Dr. Pecinka).

Several up-regulated genes induced by zebularine were also identified as ATM targets (Culligan et al., 2006). Thus, a reverse genetic screen was initiated in *atr* and *atm* single mutant lines and segregating lines for *atr atm* double mutants. Since the double mutant is completely sterile (Culligan et al., 2006), segregating lines for *atr atm* is a population of plants homozygous for *atr* (ATR$^{-/-}$) and with segregating *atm* alleles (ATM$^{-/+}$). To test whether both of the DNA damage transducers were involved in detoxifying zebularine-induced damages, plants were exposed to zebularine, MMC, and bleocin with various concentrations for the appointed time and evaluated as relative changes in rosette area (Fig. 21) and root length (Fig. 22). Upon 20 µM or higher doses of

zebularine treatment, both of the single mutants showed slightly reduced rosette area and root length, and such growth inhibition was stronger in the segregating double mutant lines (1/4 of plants were expected to be double homozygous), suggesting that both of the protein kinases were involved in signaling zebularine-induced DDR. Upon MMC treatment, *atr* showed extreme hypersensitivity whereas *atm* only showed slightly reduced growth, and the segregating double mutant lines showed similar phenotype as *atr*, confirming ATR as the main damage transducer in MMC-induced genotoxic stress. No statistically significant growth inhibition was observed in the mutants after bleocin treatment.
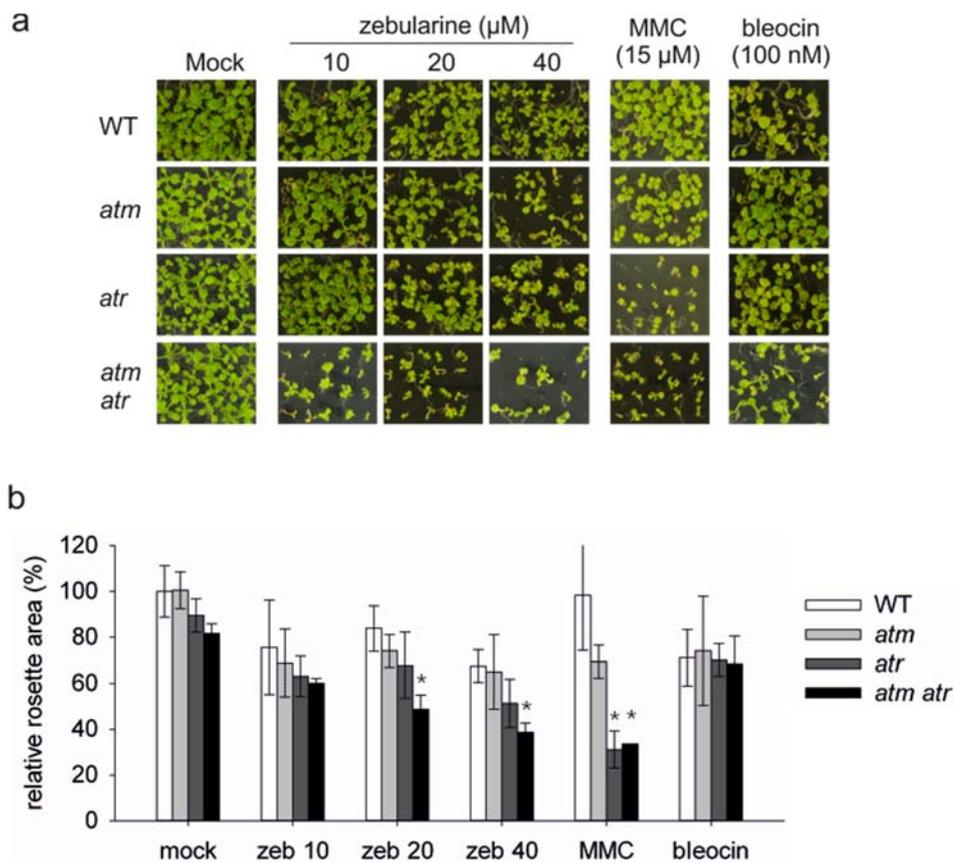
Fig. 21: a) Representative phenotypes of rosette area of *atm*, *atr*, and *atm*(+/-) *atr*(-/-) segregating lines growing on mock, 10, 20 and 40 µM zebularine (zeb), 15 µM MMC ,and 100 nM bleocin for 15 days. b) Quantitative data of relative rosette area displayed in a). Relative changes were calculated in comparison to mock-treated WT (100%). Error bars show standard deviation of three to four biological replicates. Asterisks indicate statistically significant differences in t-test (P < 0.05).
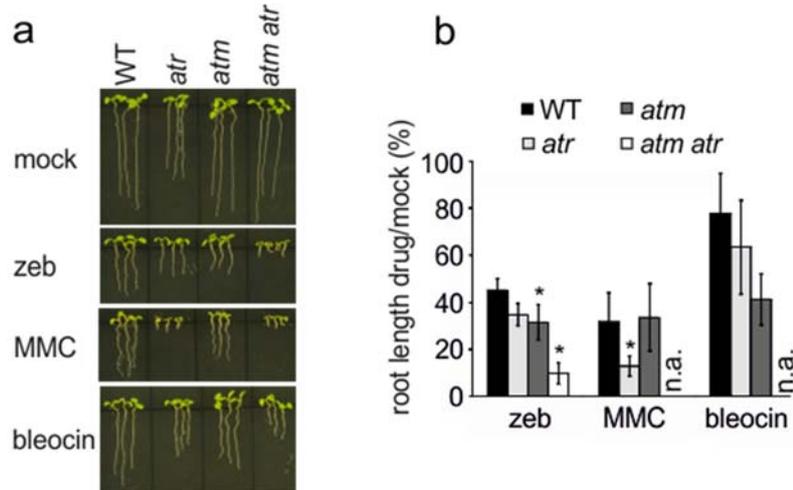
Fig. 22: a) Representative phenotypes of root length of WT, *atr*, *atm* and *atm atr* growing on 20 µM zebularine (zeb), 15 µM MMC and 50 nM bleocin. b) Quantitative data of relative root length for individual genotypes. Asterisks indicate statistically significant difference in *t*-test ($P < 0.05$) and error bars denote standard deviation of three biological replicates. n.a., not determined.

4.2.5 Reverse genetic screen of DDR mutants in response to zebularine, MMC, and bleocin treatment

To further test the induced damages and repair mechanism by zebularine, several mutants of genes related to DNA damage repair were screened on various doses of zebularine, MMC, and bleocin and evaluated as relative changes in rosette area and root length (Fig. 23 and Supplement Table 3). Mutation in the gene encoding *FASCIATA 1* (*FAS1*), the p150 subunit of chromatin assembly factor1, showed strongly reduced growth under mock condition. This was in agreement with previously described, a compensation between decreased cell number and increased cell size in the *fas1* mutant (Hisanaga et al., 2013). The *fas1* growth inhibition was more severe after zebularine treatment. Mutant plants also showed reduced growth upon bleocin treatment (50 nM or higher), but not as severe as those observed in zebularine treatment. No significantly reduced growth was observed in MMC-treated plants.

Plants defective in a homolog of the human *XERODERMA PIGMENTOSUM GROUP F DNA REPAIR* (*XPF*) (Yoshiyama et al., 2009) were screened on various doses of the drugs. *xpf* were extremely hypersensitive to MMC, even on lower dose (10µM). Minor reduced growth was observed in zebularine-treated plants, and bleocin-treated plants showed significant changes in relative roots length and rosette area.

A very sensitive phenotype upon zebularine treatment was observed in the plants mutated in *STRUCTURAL MAINTENANCE OF CHROMOSOMES 6B* (*SMC6B*) gene, whereas the mutant plants only showed weaker sensitivity upon treatments of MMC and bleocin, indicating an important role of *SMC6B* in detoxifying zebularine-specific toxicity.

Mutants related to NHEJ, *ku70* and *lig4*, were tested on the drugs. Only upon bleocin treatments, the plants showed significantly reduced growth. On higher dose of zebularine (20 µM in root length and 40 µM in rosette area) treatment, the mutants showed sensitive phenotypes.
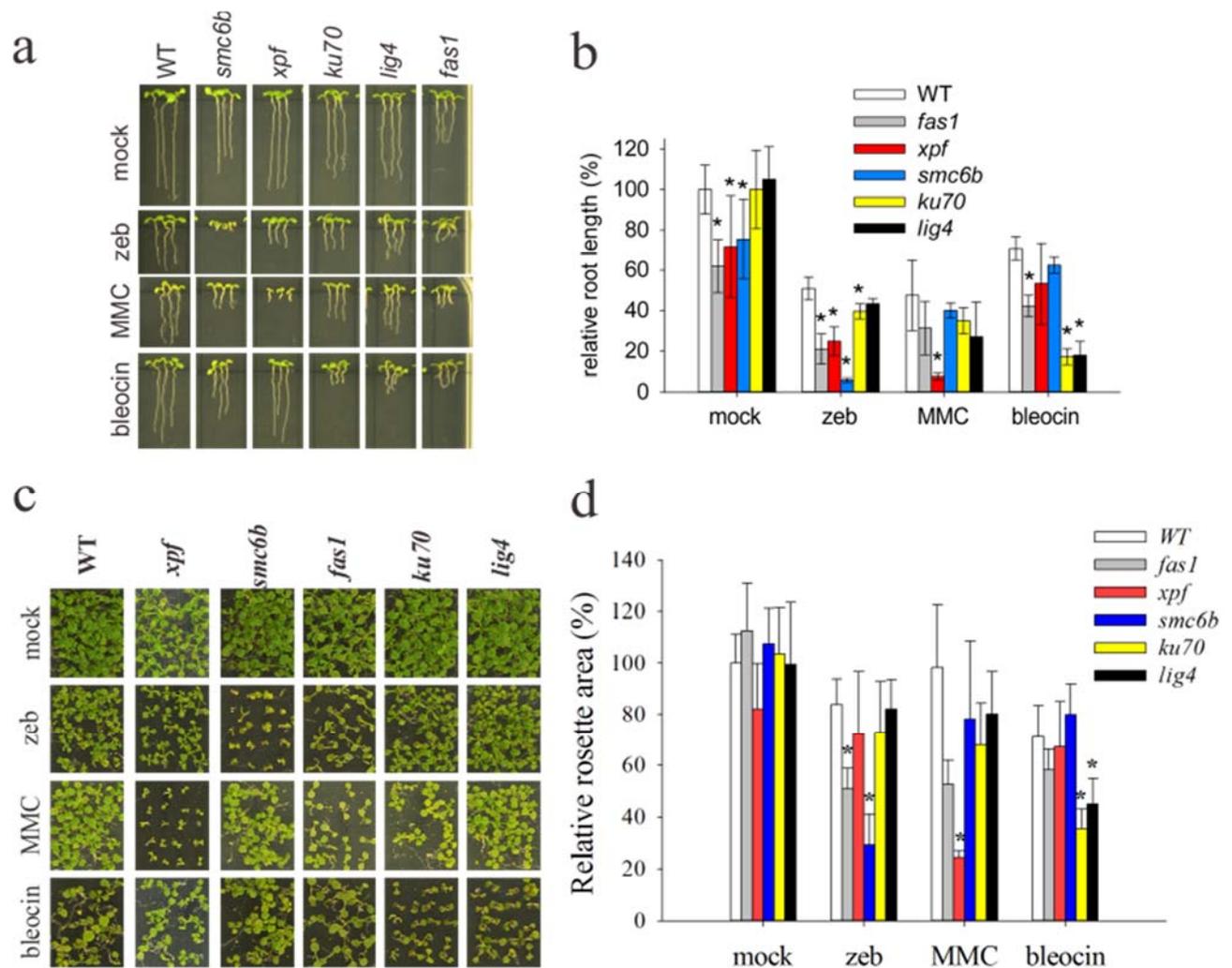
Fig. 23: Representative phenotypes of root length a) and rosette area c) of WT and DDR mutants growing on mock, 20 µM zebularine (zeb), 15 µM MMC, and 100 nM bleocin for 7 and 15 days, respectively. Quantitative data of b) relative root length and d) rosette area presented in graphs are based on three to five biological replicates with the standard deviation indicated by error bars. Statistically significant differences in *t*-test ($P < 0.05$) are labeled by asterisk.

4.2.6 Analysis of homologous recombination and endoreduplication frequencies of DDR mutants upon zebularine, MMC, and bleocin treatments

The *MITOTIC CYCLIN-DEPENDENT KINASE CYCB1;1* (*CYCB1;1*) is one of the regulators controlling cell cycle progression. It is expressed at the G2/M transition and

its levels greatly increase upon DNA damages (ATM-dependent) or DNA replication stress (ATR-dependent) (De Veylder et al., 2007). To test the drugs' effects on cell cycle progression, a cyclin–GUS (*pCYCB1;1::CYCB1;1:GUS*) reporter line was used. The reporter line was exposed to zebularine, MMC, and bleocin for several time intervals. GUS accumulation was observed in root apical meristem (RAM) cells in a time-dependent manner upon all three drug-treatments (Fig. 24). The strongest interference with the cell cycle occurred after MMC treatment followed by zebularine and then bleocin treatments.
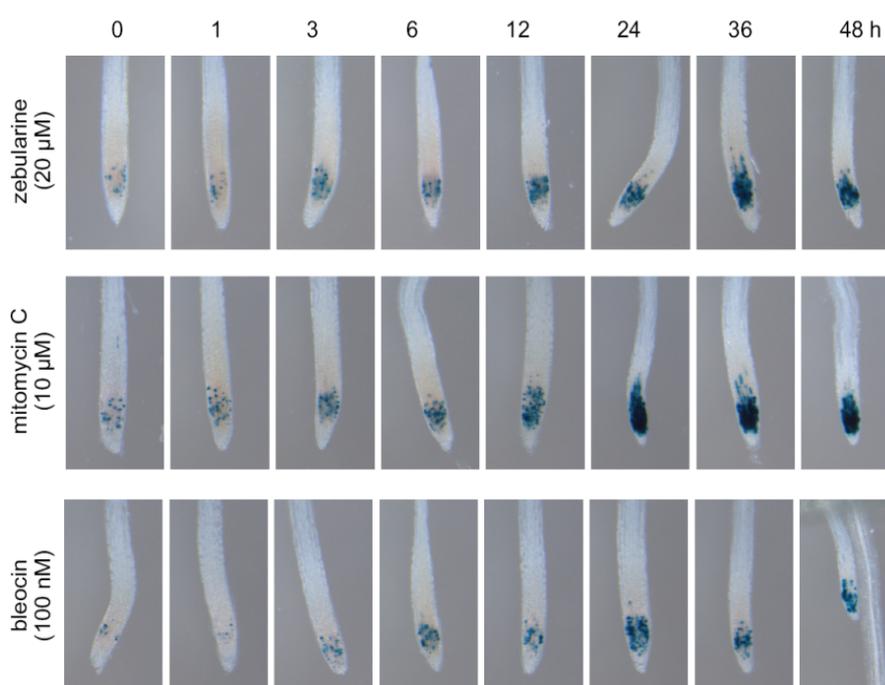
Fig. 24: Representative GUS-stained root tips of the cyclin–GUS (*pCYCB1;1::CYCB1;1:GUS*) reporter line after treatment with 20 µM zebularine, 10 µM MMC, or 100 nM bleocin for given number of hours (as shown at the top in the figure).

In plants, the activation of cycle arrest is usually accompanied by the induction of endoreduplication. To test whether the drug-treatments would induce endoreduplication, WT and several DDR mutants were exposed to 10 µM of zebularine or MMC, or 50 nM bleocin for 15 days, and samples were freshly prepared to estimate the DNA ploidy

levels by flow cytometry (Fig. 25). Both zebularine and MMC treatment significantly increased CV in WT comparing to mock treatment, while no significant changes were observed upon bleocin treatment.

For the DDR mutants growing under mock condition, only *atr* showed a slightly decreased and *fas1* increased CV, and all the other mock-treated mutants were similar as WT.

For the mutants of ATM and ATR protein kinases, MMC treatment greatly increased the CV in *atm*(+/-) *atr*(-/-), and slight increase was observed in *atr*, whereas in *atm* the CV was significantly reduced comparing to mock-treated *atm*. Upon bleocin treatment, both *atm* and *atm*(+/-) *atr*(-/-) showed significant increase in CV. For zebularine-treated plants, only CV in the segregating double mutant lines significantly increased, no significant changes were observed in the single mutants.

For other DDR mutants upon drug-treatments, the increased CV in *fas1* was enhanced by zebularine and MMC treatment, no significant changes were observed upon bleocin treatment. The zebularine-hypersensitive mutant, *smc6b*, didn't show significant changes in CV upon all drug-treatments, whereas the MMC-hypersensitive mutant, *xpf*, a strong increase was observed after MMC treatment. No significant changes were observed in bleocin-treated *smc6b* and *xpf*.
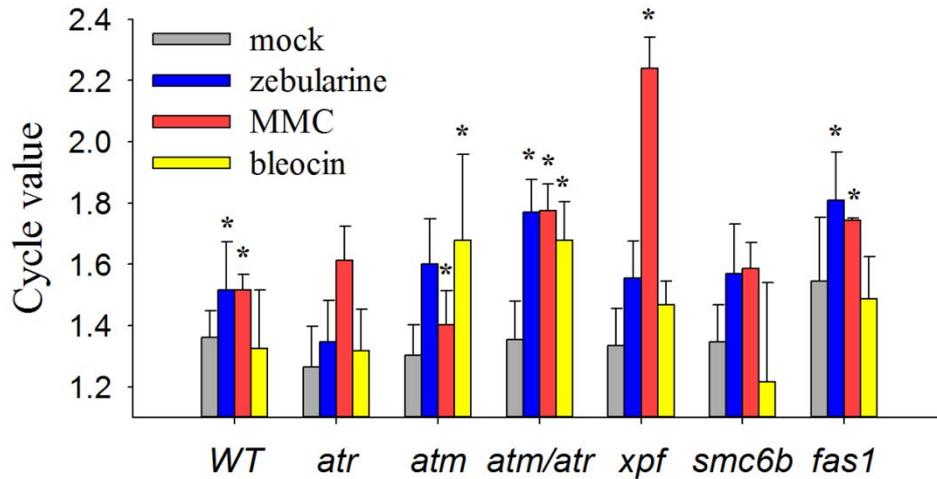
Fig. 25: Mean CV of nuclei isolated from cotyledons of WT and DDR mutants after 15 days of treatment with 10 μM zebularine, 10 μM MMC and 50 nM bleocin. Error bars indicate standard deviation of three to five biological replicates and asterisks denote statistically significant differences in *t*-test ($P < 0.05$) comparing to WT.

Previous study has shown that zebularine treatment induced very high frequencies of somatic HR in Arabidopsis, higher than the treatment of other genotoxic substances and oxidizing agents (Pecinka et al., 2009). However, a detailed analysis and comparison to different types of drug-induced DNA damages were missing. Two HR reporter lines, 651 and IC9, were selected to analyze the HR frequencies induced by zebularine, MMC, and bleocin. Line 651 contains a direct repeat of recombination substrate and allows detection of intra-molecular HR by SSA (Puchta et al., 1995). In contrast, an inverted repeat reporter region in the IC9 line is repaired by inter-molecular HR of SDSA (Molinier et al., 2004; Mannuss et al., 2010).

All drug-treatments significantly increased HR frequencies in both reporter lines (Fig. 26). MMC and bleocin treatment increased higher HR frequencies in line 651, whereas zebularine-treated plants showed higher HR frequencies in line IC9. The comparison of HR frequencies between line 651 and IC9 upon zebularine and MMC treatment was significantly different, the same result was obtained between zebularine and bleocin

(Fisher's exact test, $P < 0.05$). In contrast, the comparison between MMC and bleocin treatment of HR frequencies in the two reporter lines was not significantly different. This suggests that zebularine-induced HR was dependent of DNA synthesis whereas MMC- and bleocin-induced HR were independent of DNA synthesis.
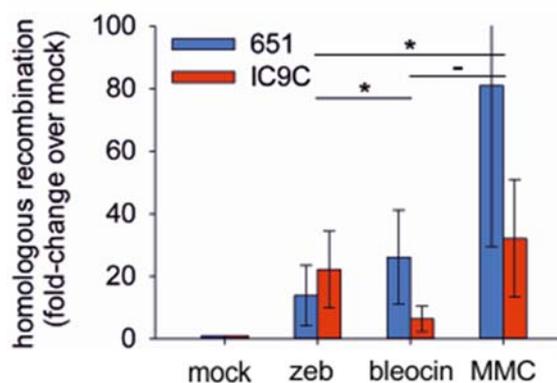
Fig. 26: HR frequencies of SSA reporter line 651 and SDSA reporter line IC9 after 20 µM zebularine, 15 µM MMC, or 100 nM bleocin treatment. Asterisk indicates significant differences in Fisher's exact test ($P < 0.0001$) in the association between SSA and SDSA and between two drug-treatments. Error bars denote standard deviation of three biological replicates.

4.2.7 The effect of 5-azacytidine on Arabidopsis

To compare the effect of other cytosine analogue in Arabidopsis with zebularine, 5-aza-cytidine was applied. First, to test whether 5-aza-cytidine would trigger DDR as zebularine treatment, 7-day-old WT were treated with zebularine or 5-aza-cytidine for 24 hours, and shoot apices were dissected for RT-qPCR measurement. As shown before (Fig. 15), zebularine treatment significantly up-regulated the DDR genes from 2-to 4-fold, however, these up-regulations were not observed in 5-aza-cytidine-treated plants (Fig. 27).
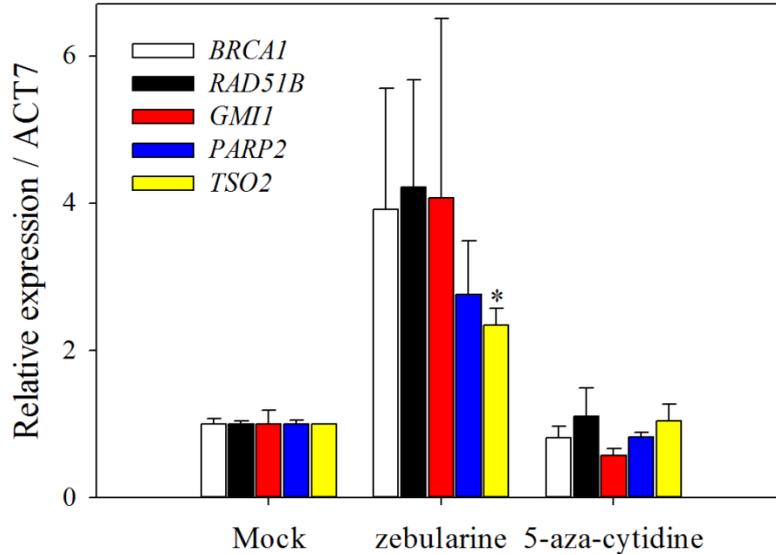
Fig. 27: RT-qPCR measurement of DNA damage marker genes in dissected shoot apices of WT after 24 hours treatment with mock, 20 μM zebularine, or 2 μM 5-aza-cytidine normalized to *ACTIN7*. Error bars represent standard deviation of two biological replicates, and asterisk denotes statistical difference in *t*-test ($P < 0.05$) comparing to mock treatment.

Reverse genetic screen was performed in DDR mutants with zebularine and 5-aza-cytidine treatment. Due to the short half-life of 5-aza-cytidine in aqueous solutions (a 10% loss in 2 to 3 hours at room temperature), seeds were sown directly in liquild-1/2 MS with or without drug treatment, and the solution were changed every four days (Fig. 28). Both of the drug-treatments significantly reduced root length in WT, and 5-aza-cytidine (61% reduction in root length) seems to be more toxic comparing to zebularine (55% reduction in root length). *smc6b* and *atm*(+/-) *atr*(-/-) were sensitive to zebularine treatment, and *ddc* was similar to WT. Upon 5-aza-cytidine treatment, plant growth was strongly inhibited in *ddc*, a weaker growth inhibition was observed in *atm*(+/-) *atr*(-/-), and *smc6b* was not sensitive to 5-aza-cytidine treatment as to zebularine.
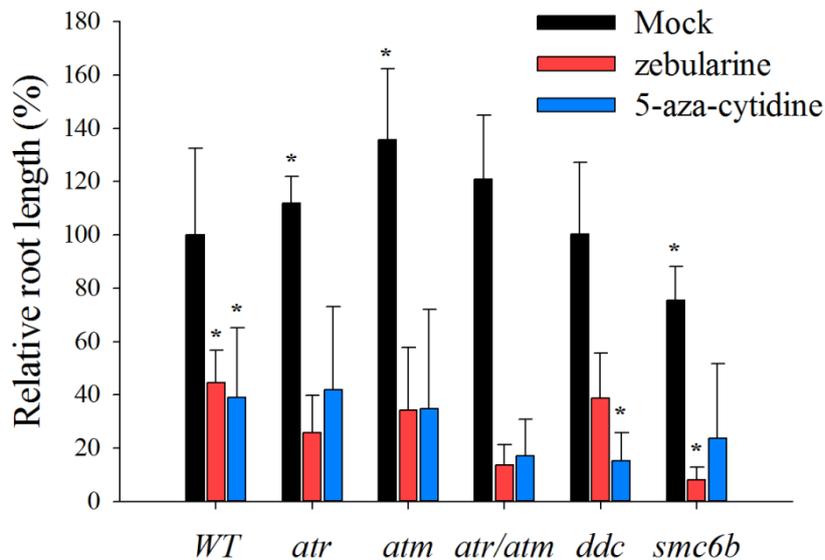
Fig. 28: Relative root length of WT and DDR mutants growing in 1/2 liquid-MS of mock, 20 µM zebularine, and 10 µM 5-aza-cytidine for 7 days. Relative changes were calculated in comparison to mock-treated WT (100%). The standard deviation indicated by error bars were based on five biological replicates. Statistically significant differences in *t*-test ($P < 0.05$) were labeled by asterisk, for the comparison between drug-treated with mock-treatment WT, or mock-treated mutants with mock-treated WT, or drug-treated mutants with drug-treated WT.

Zebularine was shown to induce relatively high frequencies of HR (Pecinka et al., 2009) and preferentially SDSA-dependent (Fig. 26). To test whether 5-aza-cytidine would have similar effects, reporter lines to detect HR frequencies of SSA and SDSA were grown under drug or drug-free 1/2 liquid-MS for 14 days and then evaluated. As shown in Fig. 29, zebularine treatment induced high HR freqencies of SSA and SDSA. Under 5-aza-cytidine treatment, SSA reporter lines showed higher HR freqeuncies but very basal HR frequencies were observed in SDSA reporter lines. These results suggest that the two nucleoside analogues, though were proposed to share the same mechanism in DNA demethylation, their effects on Arabidopsis seem to be very distinct.
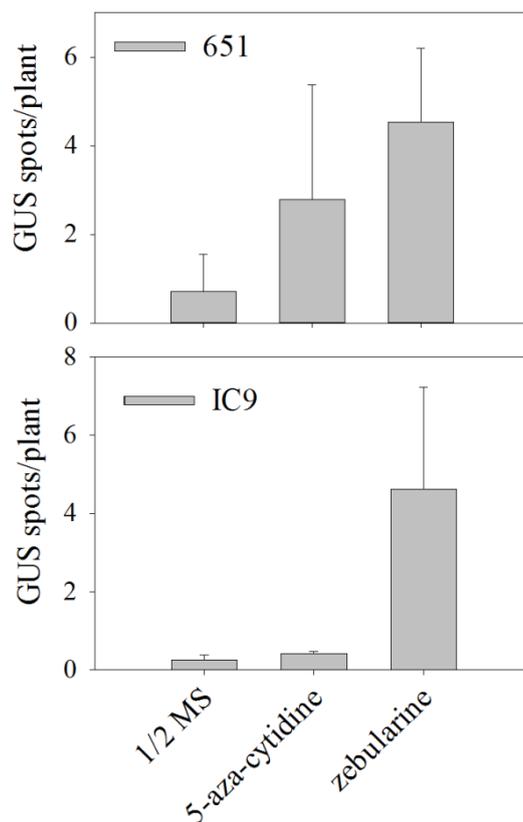
Fig. 29: HR frequencies of SSA reporter line 651 and SDSA reporter line IC9 under mock, 20 µM zebularine, and 10 µM 5-aza-cytidine treatment for 14 days. Error bars denote the standard deviation of two biological replicates.

4.3 Identifying a novel DNA damage repair gene based on the co-expression with known DNA damage repair genes

4.3.1 Candidate genes selection based on co-expression analysis

To choose guide genes involved in DDR and potentially with strong co-expression with each other, a publicly available data set of global stress expression was used (Kilian et al., 2007). Five markers genes were selected (Table 4). Genes encoding *BRCA1*, *RAD51*, and *PARP2* were proposed to be transcriptional hallmarks of DDR, up-regulated by the treatment of HU, bleomycin, and gamma-irradiation (Yi et al., 2014); *TSO2* is the small subunit of *RNR* and its coding gene is activated upon the presence of DSBs (Roa et al.,

2009); *RECQl3* is one of the RecQ helicases and proposed to have essential roles in rewinding DNA structure during DNA replication and DNA repair (Klaue et al., 2013), and its gene expression is activated under gamma-irradiation (Culligan et al., 2006).

Table 4: The list of five guide genes for co-expression analysis for novel DNA damage repair genes.

Gene	AGI code	GENE NAME
BRCA1	AT4G21070	*ARABIDOPSIS THALIANA BREAST CANCER SUSCEPTIBILITY1*
RAD51	AT5G20850	*RAS ASSOCIATED WITH DIABETES PROTEIN 51*
PARP2	AT4G02390	*POLY(ADP-RIBOSE) POLYMERASE 2*
RECQl3	AT4G35740	*RECQ HELICASE 3*
TSO2	AT3G27060	*TSO MEANING 'UGLY' IN CHINESE 2*

The five markers genes were then used as the guide genes in CoexSearch tool (http://atted.jp/top_search.shtml#CoExSearch) in ATTEDII. The top 100 candidates were first picked out. To narrow down the candidate numbers and to find the most potential candidates, genes that are associated with DNA-dependent biological process but not functionally identified as DNA damage repair genes were selected. The Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) was then used to search for conserved domain in other species, in which proteins with similar domain architecture in other species related to DDR might be identified in Arabidopsis. Eventually nine candidates were selected, named as *DNA DAMAGE INDUCED* (*DDI*) and used for more detailed analysis (Table 5).

Table 5: The list of 9 candidate genes co-expressed with the guide genes related to DNA damage repair. ***ddi***: DNA damage induced.

Name	*GENE NAME* or description	Mutual Rank					
		BRCA1	RAD51	PARP2	RECQl3	TSO2	Mean
ddi 1	SMAD/FHA domain-containing protein	9.4	11.2	28.7	4	29	16.5
ddi 2	*POLYMERASE DELTA 4*	18.6	20.8	15.7	21.9	14.9	18.4
ddi 3	*NAC DOMAIN CONTAINING PROTEIN 103*	15.6	19.1	19.8	22.4	38.3	23.0
ddi 4	*GOLGIN CANDIDATE 4*	30	28.3	34.1	12.7	45.9	30.2
ddi 5	zinc knuckle (CCHC-type) family protein	16.9	19.1	29.8	21.6	64	30.3
ddi 6	*CALLOSE DEFECTIVE MICROSPORE1*	27.3	20.6	24.6	35	102.2	41.9
ddi 7	*CHROMATIN REMODELING 31*	40.6	57.3	51.9	48.7	77.4	55.2
ddi 8	DNA binding; ATP binding	46.2	69.1	49.1	42.1	124.4	66.2
ddi 9	RNA-binding (RRM/RBD/RNP motifs) family protein	54.3	44.1	136.2	111.2	35.1	76.2

* The MR values were downloaded from ATTEDII.

The result of co-expression analysis was tested by a reverse genetic screen. Two independent mutant alleles of each candidate gene were obtained (Supplement Table 5) and tested for hypersensitivity on various doses of DNA damaging agents, including MMC, zebularine, and bleocin. Among the nine candidates, only *ddi 7* mutant lines (the mutant of *CHROMATIN REMODELING 31*: *CHR31*) showed hypersensitivity to high doses of MMC. All the other mutant lines were similar to Col-0 phenotype upon the drug-treatments.

4.3.2 Identifying candidate genes by reverse genetic screen

The reverse genetic screen of the candidate mutants revealed *CHR31* as a potential candidate of an unidentified DNA damage repair co-expressed gene. Two independent mutant alleles were obtained and tested. *ddi 7-1* mutant allele harbors a T-DNA insertion in the second exon of *CHR31* gene, whereas *ddi 7-2* insert is located in the fourth exon of the gene (Fig. 30).
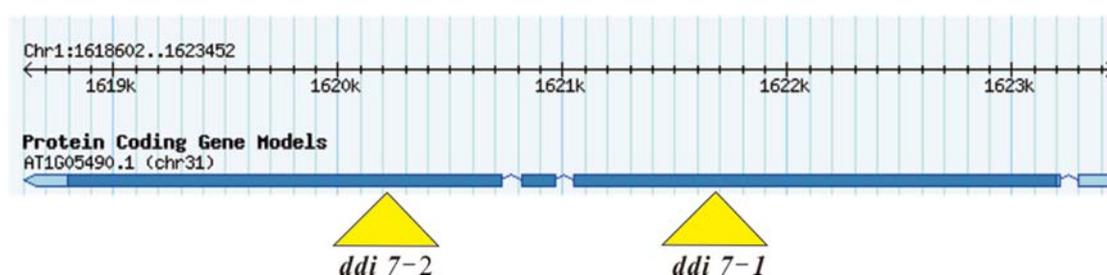
Fig. 30: Intron–exon organization of the *CHR31* gene. Blue boxes represent exons, and the blue lines indicate introns. The yellow triangles correspond to the T-DNA insertion sites of the two different mutant alleles. The arrow lines and the numbers at the top indicate the positions across chromosome 1. (The figure was modified from https://www.arabidopsis.org/index.jsp)

Upon treatment of DNA damaging agents, both *ddi 7-1* and *ddi 7-2* were hypersensitive to high dose of MMC (Fig. 31). Under mock condition, the mutant plants displayed slightly reduced leaf size comparing to WT (Fig. 32). Upon MMC treatment, leaf development in the mutants was significantly decreased, indicating the potential role of *CHR31* in detoxifying MMC-induced damages.
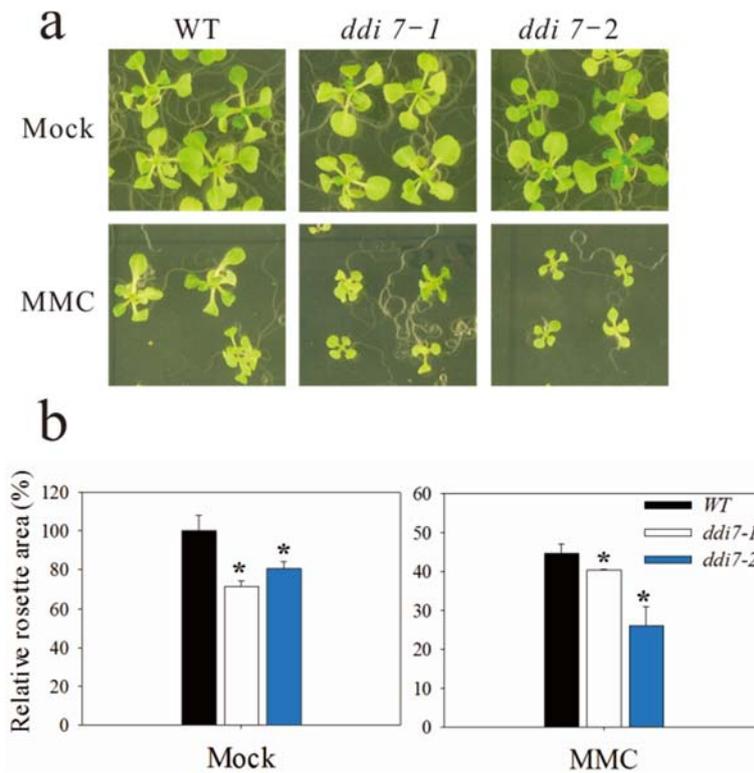
Fig. 31: a) Representative phenotypes of WT and *ddi 7-1* and *ddi 7-2* growing on mock and 35 μM MMC for 21 days. b) Relative rosette area in a). For mock-treated plants, relative changes were calculated in comparison to WT (100%); for MMC-treated plants, relative changes were calculated as mean values of MMC-treated plants / mean values of mock-treated plants. Error bars show standard deviation of four biological replicates. Asterisks indicate statistically significant differences (*t*-test, *P < 0.01).

To validate transcriptional activation of *CHR31* by DNA damaging treatment, expression analysis was performed in dissected shoot apices of 7-day-old WT with or without 24 hours MMC treatment. *CHR31* was up-regulated under the treatment of MMC, suggesting its role in the response to MMC-induced DNA damages (Fig. 33a). The transcriptional changes of core DDR genes, *BRCA1*, *RAD51B*, and *GMI1*, were measured in WT and *ddi 7-1* plants after 24 hours MMC treatment. They were 1.5- to 2-fold up-regulated in *ddi 7-1* comparing to WT (Fig. 33b). This suggests *CHR31* as a negative regulator of MMC-induced response or increased amount of DNA damage in *ddi7* leading to stronger transcriptional responses.
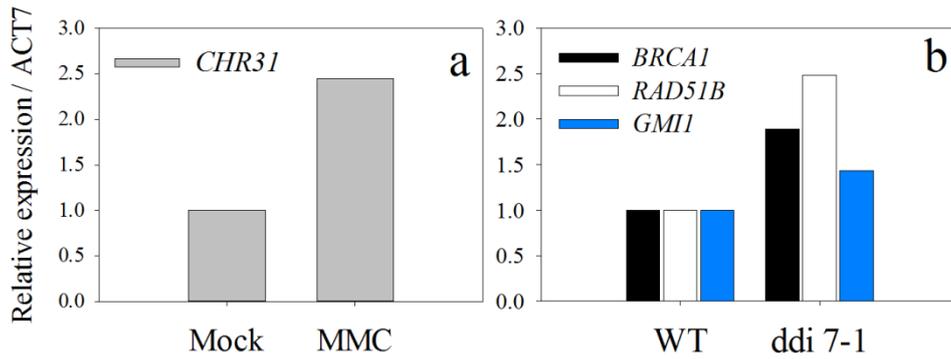
Fig. 32: RT-qPCR measurement of a) *CHR31* in dissected shoot apices of WT after 24 hours treatment with mock or 10 µM MMC; b) DNA damage marker genes in WT and *ddi 7-1* after 24 hours treatment with 10 µM MMC normalized to *ACTIN7*. The bars represent the mean values of transcription from a pool of 20 seedlings.

4.3.3 Functional characterization of the *CHR31* gene

To functionally dissect the possible role of *CHR31* in DNA damage repair, *ddi 7-1* were crossed with the reporter lines for HR analysis (11B for SSA analysis and IC9 for SDSA analysis), and plants homozygous for both the mutation and the reporter substrate were grown on mock and MMC treatment for 12 days (Fig. 34). Under mock and MMC conditions, 11B *ddi7-1* showed significant increase in SSA-HR, comparing to 11B WT, indicating the involvement of *CHR31* in SSA-HR. For IC9 reporter lines, *ddi 7-1* showed no statistically significant difference from the WT under mock and MMC condition.
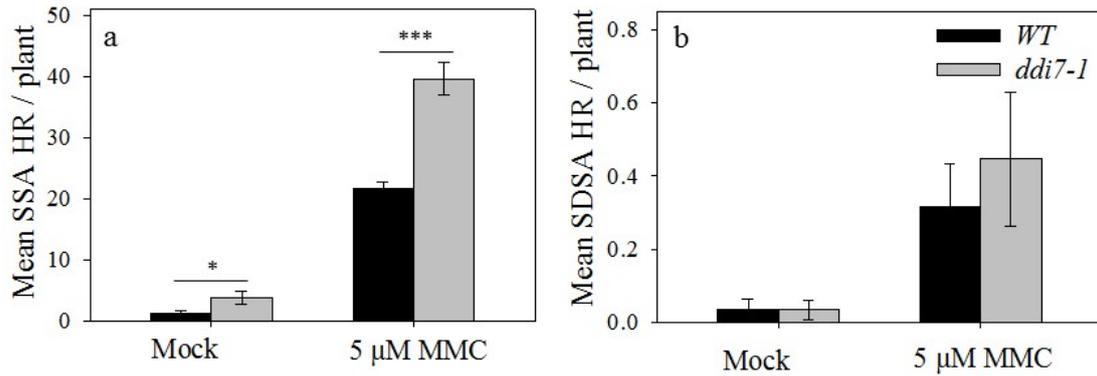
Fig. 33: HR frequencies of SSA reporter line 11B a) and SDSA reporter line IC9 b) growing under mock or 5 μM MMC treatment for 12 days. Asterisk indicates significant differences (*t*-test, *P* < 0.05; *** *P* < 0.001) between WT and *ddi7-1* background. Error bars denote standard deviation of three biological replicates.

Previous results suggested a more severe DNA damages were observed in *ddi 7-1* upon MMC treatment comparing to WT. The presence of DNA damages in plants often triggers cell cycle checkpoint and cell cycle arrest, to investigate the role of *CHR31* in cell cylce control, the *pCYCB1;1::CYCB1;1:GUS* construct was introgressed into the *ddi 7-1* background, and the plants homozygous for both the mutation and the reporter construct were obtained. 7-day-old plants were treated with 10 μM MMC for the appointed time points. The GUS signals accumulated in WT upon MMC treatment in a time-dependent manner. However, in the *ddi 7-1*, GUS accumulation was less pronounced, especially after 12 hours or longer treatment, indicating the involvement of *CHR31* in cell cycle control upon DNA damages.
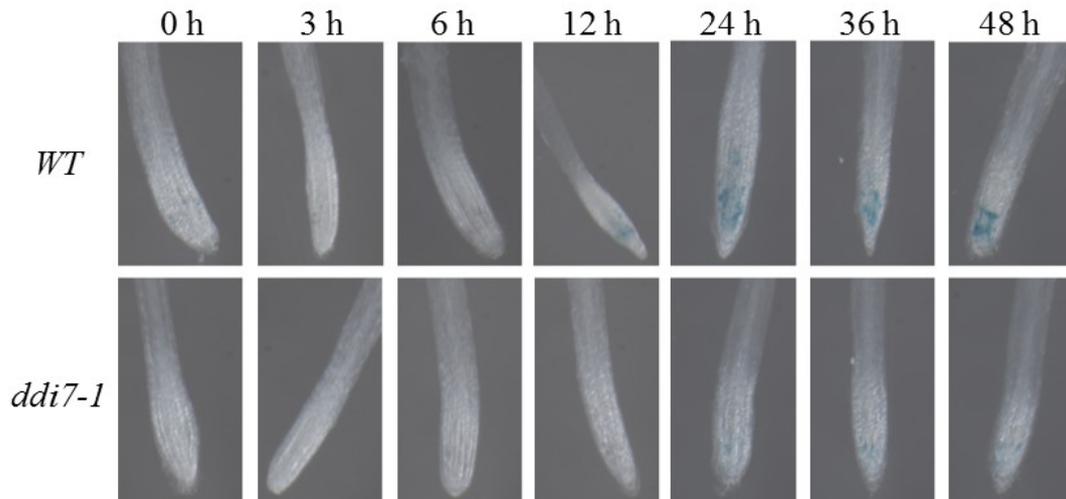
Fig. 34: Representative GUS-stained root tips of the cyclin–GUS reporter line (WT and *ddi 7-1* background) after treatment with 10 µM MMC for given number of hours (as shown at the top of the figure).

To further investigate the involvement of *CHR31* in cell-cycle control, endoploidy analysis was performed in WT and *ddi 7-1* with or without 14 days of MMC treatment. MMC treatment induced endoreduplication in WT (105% in relative CV), while *ddi 7-1* significantly reduced this induction (96% in relative CV), further suggesting the involvement of *CHR31* in cell-cycle regulation (Fig. 36).
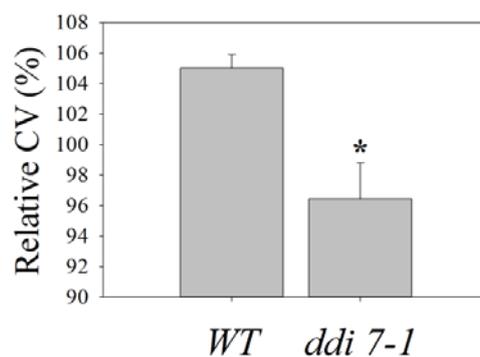
Fig. 35: Relative CV of nuclei isolated from cotyledons of WT and *ddi 7-1* after 14 days of treatment with 10 µM MMC. Relative changes were calculated in comparison to mock-treated plants (100%). Error bars indicate standard deviation of two biological replicates and asterisks denote statistically significant differences in *t*-test ($P < 0.05$).

The presence of DNA damages could result in cell death. To compare the presence of cell death in the two genetic background and upon different treatments, PI was used to identify dead cells in root tips. PI is a fluorescent dye that is generally excluded from viable cells. It outlines the cells by staining the cell wall and stains nuclei of non-viable cells. Seven-day-old WT and *ddi 7-1* were treated with or without 10 µM MMC for 24 hours and then stained with PI for two minutes. Small trace amount of red fluorescence (indicated by the arrowheads) was detected under mock condition at RAM in both genotypes, and such signals were stronger and the root cells were enlarged upon MMC treatment (Fig. 37).
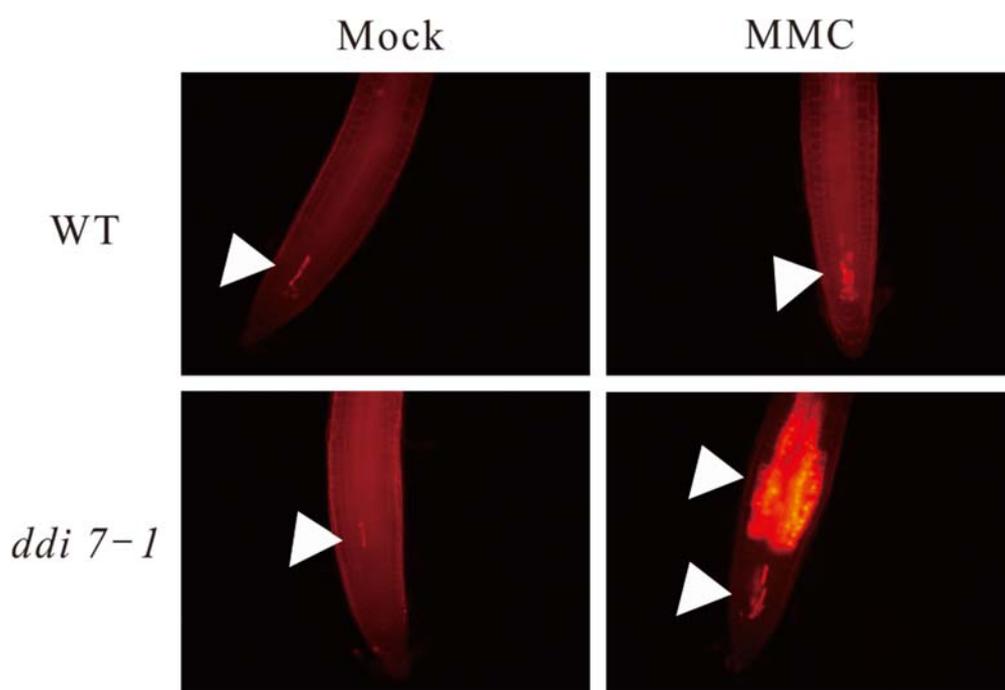
Fig. 36: Representative PI-stained root tips of WT and *ddi 7-1* after treatment with or without 10 µM MMC for 24 hours. Arrowheads indicate the region of dead cells.

However, very strong signals were detected at the transition zone (located between RAM and basal elongation region) in MMC-treated *ddi 7-1* (Fig. 38), much stronger than in RAM. This suggests that the mutation in *CHR31* might affect the cell growth

and development at transition zone, or it might be required specifically to regulate the repairing process or genome stability for the cells at this area.
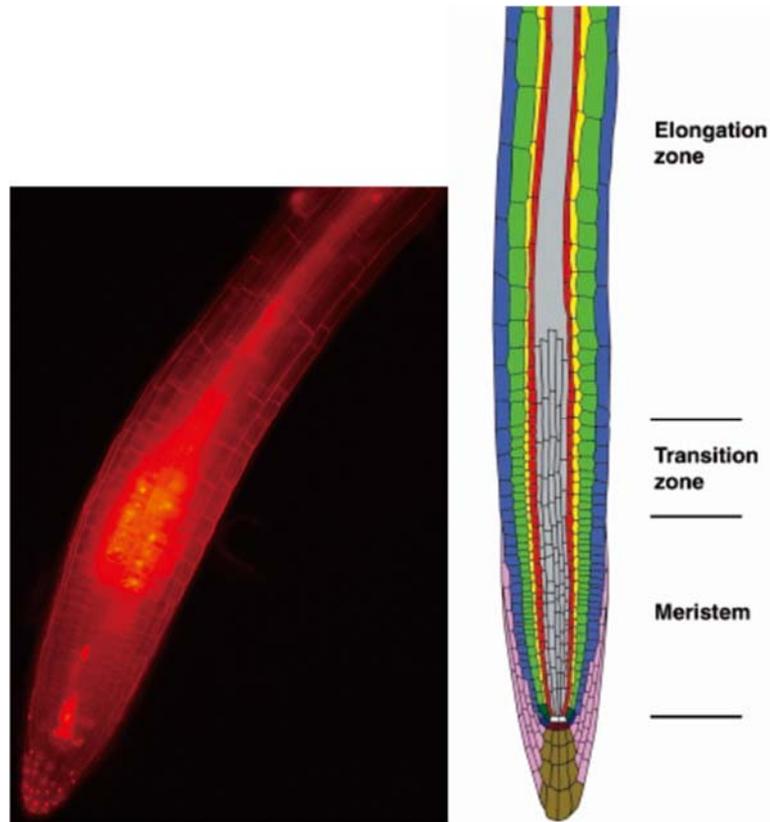
Fig. 37: The root of MMC-treated *ddi 7-1* showing that the red fluorescence detected (dead cells) was mainly at the transition zone. The figure at right was modified from Ubeda-Tomás et al.(2012).

5. Discussion

5.1 Genotoxic stress, DNA repair, and the application in improving crop productivity

Plants are routinely exposed to environmental stresses, such as temperature extremes, water availability, ion toxicity, UV radiation, or infection by pathogen and insects. The generation of reactive oxygen species (ROS) represents a universal response to most stress conditions (Tuteja et al., 2009). ROS are highly active molecules that induce

damages to all cellular macromolecules including DNA. They alter DNA bases and damage sugar residues, leading to DNA strand breaks (Roldán-Arjona and Ariza, 2009). UV radiation from sunlight also induces DNA damage by forming pyrimidine dimers, resulting in DNA cross-links (Kimura and Sakaguchi, 2006). The understanding and the expanding knowledge of DNA repair mechanism becomes more crucial in current breeding technologies for improving stress tolerance in crops (Balestrazzi et al., 2011). For example, transgenic Arabidopsis plants with reduced *PARP* levels showed enhanced tolerance to a wide range of abiotic stresses (Vanderauwera et al., 2007). Enhanced transcription of the gene encoding the detoxifying enzyme of pyrimidine dimers, photolyase, improved UV-tolerance in both *A. thaliana* and rice (Tanaka et al., 2002; Kimura and Sakaguchi, 2006). Overexpression of stress-induced DNA helicases (the enzyme that catalyzes the unwinding of DNA double helix to favor the accessibility of the damaged DNA region) gene has also shown higher tolerance to salinity in plants (Sanan-Mishra et al., 2005; Vashisht and Tuteja, 2006).

With increasing demand for high-quality food production due to expanding world population size, there is greater need for breeders to select stress-tolerant crops to cope with the changing climate. The role of DNA damage repairs genes in enhancing stress tolerance and the expanding knowledge in the link between DDR functions and protective mechanism under stress condition could provide an effective strategy for crop improvement and productivity (Balestrazzi et al., 2011; Balestrazzi et al., 2013).

5.2 Natural variation of Arabidopsis to HU

5.2.1 GWA studies of HU sensitivity

In this study, HU was applied as an inducer to dissect the genetic variation of

Arabidopsis natural accession in response to DNA replication stress. The natural accessions revealed a diverse response to HU from RS of 100% to 0%, while only few accessions were hypersensitive (most of the plants were dead) upon the high dose of HU treatment. HU was known to inhibit the activity of RNR and reduced the concentration of available nucleotides for DNA replication, thus inducing replication fork-stalling type of DNA damage and cell cycle arrest before entering into S-phase (Wang and Liu, 2006). It was also shown to produce nitric oxide (NO) both *in vitro* and *in vivo* studies (Huang et al., 2004; King, 2004) and in Arabidopsis to inhibit catalase-mediated H_2O_2 decomposition and the trigger for H_2O_2 formation (Juul et al., 2010). The generation of H_2O_2 was proposed to activate ATM-dependent signaling pathway and result in cell-cycle checkpoint in response to ROS-induced DNA damage (Yi et al., 2014).

To dissect the genetic basis of the Arabidopsis natural accessions to HU, GWA studies were performed. No strong peaks associated with functionally known genes was detected as shown in the previous GWA studies (Atwell et al., 2010; Li et al., 2010), probably due to very low number of sensitive accessions or complex genetic architecture of this trait. Several peaks of association contain SNP makers with *p*-values above the threshold were identified, and the most significant peak of association appeared at the bottom arm of chromosome 5. This region consists of genes encoding *CYCLIN-DEPENDENT KINASE* and *OXIDOREDUCTASE*, which are potentially related to HU-induced cell cycle arrest and ROS production, respectively. However, combining with other selected candidates, the screening of T-DNA insertional mutants on HU suggests the genetic basis of HU sensitivity might come from several loci rather one, single major locus.

According to Bergelson and Roux (2010), GWA studies in *A. thaliana* suffers from two major limitations: first, the standard GWA studies using single-locus analysis ignores the polygenic background and other unobserved variables, treating the trait of interest as it were due to a single locus. If the background variables are correlated with the SNP in the model, population structure will lead to inflation of test statistics and a high false positive rate. Second, different combination of genes in two different populations can lead to the same phenotype, which is called "genetic (allelic) heterogeneity." The alternative genes (or alleles at the same gene) leading to the same phenotype can impede the detection of genes that underlie the natural phenotypic variation. The mapping population with strong representation from specific area (for example, many of the accessions in this study are from UK, Germany, and France) are geographically limited and will reveal such extent of genetic and allelic heterogeneity.

Schott (2012) has shown in his thesis that all the HU-sensitive accessions don't belong to the same region (maybe some are geographically related, e.g. Ts-1 and Bla-1 are both from coastal regions of Spain). The complementation tests for the sensitive accessions also revealed that there are several genes rather than one responsible for HU sensitivity in this mapping population, suggesting the genetic basis of HU sensitivity in the natural accessions is a more complex trait than monogenic and with genetic or allelic heterogeneity.

5.2.2 QTL mapping of HU sensitivity – from QTL to QTG

GWAS was proven to be a powerful tool to identify single, major locus when it comes to monogenic traits, which are, in most of the cases, validated genes or loci (Atwell et al., 2010). Tradition linkage mapping, however, is useful to identify rare alleles, but the

genetic basis are specific to the parental lines of the mapping population and may not be representative of the genetic variation of the natural selections. A combination of traditional linkage mapping and GWAS was proposed (Bergelson and Roux, 2010). This might allow the true positives and the false negatives to be distinguished from false positives in GWA studies after population structure corrections. Ideally, the true positives, the causative SNP detected by GWA studies, will be overlapped by QTL regions, and the false negative, the causative SNP lost after corrections of population structure, can be validated by QTL regions.

QTL mapping identified two major QTLs at chromosome 3 and 5. QTL5 overlaps with the most significant associated region in GWA studies, but no peaks detected by GWA studies overlap with the region in QTL3. This suggests the genetic basis identified in QTL mapping might be more representative to Bla-1 and Col-0 as the parental lines in the mapping population rather than the genetic variation of the 391 natural accessions.

Fine mapping of QTL3 revealed eight candidate genes underlie this region. *RECA3* is the only candidate directly related to DDR. It was shown to be involved to maintain the genome stability and recombination-dependent repair in mitochondria and plant fitness under stresses (Miller-Messmer et al., 2012). The gene encoding P-loop NTPase was identified as coil-coiled protein with microtubule motor activity during cell division (Rose et al., 2004). *DFC* encodes a protein with tetrahydrofolylpolyglutamate synthase activity that is located in the mitochondrial matrix. It was shown to be involved in nitrogen utilization during early seedling development in Arabidopsis (Jiang et al., 2013). The gene encoding calcium-binding protein was proposed to be involved in plant immune responses (Le et al., 2014). The other four candidate genes are still functionally unidentified.

Since none of the candidate genes have been shown to be directly related to HU-induced phenotypes or physiological responses, the quantitative complementation experiment proposed by Long et al. (1996) could be an approach to identify the candidate gene. Crosses between a selected RIL (with Bla-1 genotype at QTL3) and the mutant (with Col-0 background) of the candidate genes could test whether the mutant allele is allelic to QTL3 and thus identify the gene responsible for HU sensitivity.

5.2.3 The potential additive and epistatic effects between QTLs for HU sensitivity

Two-dimensional genome scan showed an additive interaction between the two main loci (QTL3 and 5) associated with HU sensitivity, combining with the results from GWA studies, suggesting the genetic basis for HU sensitivity is more complex than a single, major locus but due to the genetic interaction between QTL3 and 5. However, when searching for proper HIFs for QTL confirmation, both for QTL3 and 5, separately, segregating progenies of several HIFs did not show the expected phenotypes, e.g. plants with Bla-1 genotype on both QTL3 and 5 didn't show sensitivity to HU, suggesting there might be a 3rd QTL. QTL1 was identified by two-locus effectplot, showing strong epistatic interaction with QTL5, and the alignment of the genotypes of the top 20 sensitive RILs also showed that the genetic bases of HU sensitivity requires Col-0 genotype at QTL1.

The epistasis between QTLs is estimated by fitting a statistical model that includes both the main effects of each QTL and the effects of the QTL-QTL interactions. Thus, the analysis of a single locus can detect a QTL that has no main effect but interacts epistatically with another locus and the power to formally detect epistasis is very

limited (Purcell and Sham, 2004), which both explain the epistatic effect between QTL1 and 5 was not identified in the two-dimensional scan.

Epistasis occur when the sum of the effects of introgressed fragments are significantly greater than or less than the mean difference in the phenotype between the parental lines (Mackay, 2014). To confirm the epistatic effect of QTL 1 on both QTL 3 and 5, proper HIFs with segregating QTL regions are needed. The available RILs were limited especially when selecting proper HIFs for QTL5. Developing NILs introgressed with Bla-1 for each of the QTLs in Col-0 background and making multiple crosses for cumulating isogenic lines (those cumulating several QTL regions) (Causse et al., 2007) could be an alternative approach to validate both single QTLs and the potential epistatic effects.

5.3 The effect of zebularine treatment on Arabidopsis genome stability

5.3.1 Zebularine: A nucleoside analogue, a DNA methylation inhibitor, and an inducer of DNA damage response in *Arabidopsis*

Nucleoside analogues inhibiting DNA methylation have been widely used in epigenetic therapy for cancer treatment. Since many tumors have the characteristic of aberrant hypermethylation of tumor suppressor genes, demethylating cytidine analogues were shown to reverse regional hypermethylation and restore the expression of tumor suppressor genes (Egger et al., 2004; Lim et al., 2011). Comparing to azanucleosides (5-aza-cytidine and 5-aza-2'-deoxycytidine), zebularine was proven to be a superior choice in terms of lower cytotoxicity, increased stability in aqueous solution, and high

specificity for cancer cells (Cheng et al., 2004; Marquez et al., 2005).

In plants, zebularine was used as a DNA methylation inhibitor for epigenetic studies. It was shown to partially reduce DNA methylation and heterochromatin condensation, and cause transcriptional reactivation of TGS targets (Baubec et al., 2009; Baubec et al., 2014).

In this study, genome-wide expression analysis revealed that approximately 50% of the genes up-regulated by short-zebularine treatment were associated with DNA damage repair and additional DDR genes were induced after long-zebularine treatment. Among the common genes up-regulated after short- and long-zebularine treatment, those encoding *BRCA1*, *RAD51*, and *TK1A* were considered as the transcriptional hallmarks of DDR (regardless of the type of DNA damages), and *SMR7* was proposed to regulate DNA damage-induced cell-cycle checkpoint (Yi et al., 2014). *RNR1* and *TSO2* genes are more specific to the changes in dNTP pools for DNA replication and repair (Roa et al., 2009). This indicates that zebularine, as a nucleoside analogue, may interfere with the incorporation of dNTP into DNA during DNA replication and induce DDR. Comparing to the transcriptome analysis after 16 days of 5-aza-2'-deoxycytidine treatment, a functionally diverse set of genes were up-regulated but with no significant association to DNA damage repair (Chang and Pikaard, 2005). This might be due to the differences between the metabolism of nucleoside analogue (zebularine) and the deoxy-triphosphate form of 5-aza-cytidine (5-aza-2'-deoxycytidine), the treatment duration, the stability, and the biological effect of the drugs. In this study, the comparison between 5-aza-cytidine and zebularine treatment also suggests that the two nucleoside analogues have different effects on Arabidopsis. In contrast, transcriptional changes induced by short zebularine treatment overlapped more than 90% with those of the alkylating agent

MMC, suggesting similar stimuli and common signaling pathways induced by MMC and zebularine.

Zebularine was shown as a potent inhibitor of DNMT and DNA methylation *in vitro* and *in vivo* mammalian studies (Marquez et al., 2005; Champion et al., 2010; Meador et al., 2010; Lim et al., 2011). The genome-wide effects of zebularine on transcriptional DNA methylation targets was analyzed, and only less than 1% of zebularine-activated genetic elements were among the genes controlled by key DNA methylation factor *DDM1* (Zemach et al., 2013). Bisulfite sequencing analysis of several zebularine-activated genes showed no significant changes in DNA methylation, which could be potentially due to fast *de novo* DNA methylation activity in apical meristems (Baubec et al., 2014). This suggests the up-regulation of these genes by zebularine, especially the core DDR genes, were not due to the stable loss of DNA methylation.

The metabolic activation of zebularine requires it to be phosphorylated and incorporated into DNA (Egger et al., 2004). *In vivo* study of zebularine metabolism in human cancer cells proposed that the incorporation of this nucleoside analogue into RNA was 7-fold higher than that into DNA, and a considerable amount (25% to 50%) of phosphorylated zebularine were sequestered into diphosphocholine adduct. The complex metabolism of zebularine and its limited incorporation into DNA explain why higher doses are required for equivalent inhibition of DNMT comparing to azanucleosides (Ben-Kasus et al., 2005). The doses applied in this study for zebularine treatment were relatively low and the duration of drug treatment was short, suggesting the maintenance of DNA methylation of zebularine-up regulated genes might also be due to its limited incorporation into DNA and inhibition of DNMT.

5.3.2 The induction of cell cycle arrest and endoreduplication by zebularine

The DNA repair pathways are tightly coordinated with cell cycle progression through the activation of cell cycle checkpoint. A remarkable exception among the cell-cycle genes is the mitotic cyclin *CYCB1;1* which is induced rather than repressed upon DNA damage or inhibition of DNA replication (De Veylder et al., 2007). In this study, GUS reporter construct of *CYCB1;1* showed a delay of cell cycle progression from G2 to M phase after zebularine treatment.

To prevent the transmission of DNA lesions or to have higher potential for cell expansion upon DNA damages, several studies have shown that the presence of DNA damages by specific gene mutations or chemical inducers resulted in endoreduplication (Yoshiyama et al., 2013a). In this study, zebularine treatment was shown to induce endoreduplication, combining with the activation of cell cycle arrest, both results suggest the presence of DNA damages and the activation of DDR induced by zebularine.

5.3.3 The DNA repair mechanism induced by zebularine
5.3.3.1 The activation of DDR signaling transducers: ATM and ATR protein kinases

The signaling of DNA damages as DSBs are primarily transduced by ATM, and the signaling of stalled replication forks are mainly by ATR. But once these DNA lesions are generated, they can be processed into alternative forms which then will be recognized by the alternative damage transducers (Jiang et al., 1997).

The reverse genetic screen of *atm*, *atr*, and *atm(+/-) atr(-/-)* suggests both of the DNA damage transducers were involved in repairing zebularine-induced DDR. MMC was proposed to mainly activates ATR-dependent repair signaling by forming inter-strand cross-links (Culligan et al., 2004), whereas both of the protein kinases were involved in the response to gamma-irradiation (Culligan et al., 2006). The nucleoside-like nature of zebularine allows its interference with genome stability only during DNA replication, thus, zebularine-induced DNA damage may occur specifically after DNA strand separation and activate the DNA damage repair machinery by additive functions of ATR and ATM kinases.

Endoploidy analysis showed a significantly increase in CV in *atm*(+/-) *atr*(-/-) after all three drug-treatments. ATM and ATR were proposed to be involved in the induction of cell-cycle arrest upon gamma-irradiation, in which ATM was responsible for immediate transcriptional response to DSB, and ATR was to maintain the persistence of cell-cycle arrest (Culligan et al., 2006). The increase in root cell size upon DSB inducer required both ATM and ATR (Adachi et al., 2011). The increase in CV in *atm*(+/-) *atr*(-/-) upon all drug-treatments indicates a hyperactivation of cell-cycle arrest and reduced capability for repairing DNA damages, also further suggesting ATM and ATR kinases were both involved in zebularine-induced cell-cycle arrest.

5.3.3.2 Homologous recombination plays a crucial role in detoxifying zebularine-induced DNA damages

Several DDR mutants were tested on zebularine to dissect their involvement in detoxifying zebularine-induced DNA damages and repair mechanisms. A stronger reduced growth and higher induction of endoreduplication were observed in *fas1* after

zebularine treatment comparing to mock-treated plants. The reduction of cell number and the increase in endoreduplication in *fas1* was proposed to be a result of ATM-dependent DDR (Hisanaga et al., 2013). Both growth inhibition and endoreduplication induction were more severe after zebularine treatment, indicating an additive effect of ATM-dependent DDR of chemical (by zebularine) and genetic interference (by the mutation).

Mutants defective in NER (*xpf*) showed slightly reduced growth under zebularine treatment, suggesting only a minor role of NER in detoxifying zebularine-induced DDR. While growing under MMC treatment, *xpf* were hypersensitive to the drug and showed great induction of endoreduplication. The similar phenomenon was observed upon other genotoxic stresses in *xpf*, such as UV and gamma-irradiation, the mutation caused a hyperactivation of cell cycle arrest and reduced capability for repairing DNA damage (Preuss and Britt, 2003). Cross-talks between NER and HR was proposed in *XPF* in Arabidopsis, in which the endonuclease might play a role in the removal of non-homologous 3'-ended overhangs from SSA-HR intermediates (Dubest et al., 2002). The hypersensitivity of *xpf* to MMC and the high frequencies of SSA-HR induced by MMC might be another indication of *XPF* in facilitating SSA-HR.

An extremely sensitive phenotype upon zebularine treatment was observed in *smc6b*, which was more specific to zebularine than other two drug-treatments. *SMC6B* is the core component of the SMC5-SMC6 complex and was proposed to be involved in DNA repair and efficient HR upon genotoxic stress (Mengiste et al., 1999; Hanin et al., 2000; Kozak et al., 2009). The specific hypersensitivity of *smc6b* to zebularine treatment indicates an important role of SMC5-SMC6 complex in detoxifying zebularine-induced toxicity, and HR might be crucial to this repair mechanism. The high frequencies of

somatic HR induced by zebularine comparing to other stresses (Pecinka et al., 2009) further supports that HR is essential in repairing zebularine-induced damages. However, the detailed molecular mechanisms remain unknown.

5.3.3.3 Zebularine-induced DNA damages was preferentially repaired by synthesis-dependent strand annealing homologous recombination

HR is one of the pathways to repair DSBs. Depending on the genomic architecture, SSA and SDSA can be used to repair DSBs by HR (Roth et al., 2012). In this study, the analysis of specific HR pathways revealed that SSA is a preferred HR pathway for repairing of bleocin-and MMC-induced damages, while SDSA seems to be more important for repairing of zebularine-induced damages. SSA can occur at both non-replicated and replicated DNA, whereas SDSA occurs only at replicated DNA. The deoxy-triphosphate form of zebularine can only be incorporated into DNA during DNA replication in order to replace cytosine, i.e. it is only active during S-phase (Egger et al., 2004), thus the higher frequency of SDSA induced by zebularine indicates that the zebularine-induced damages occur during DNA replication and are mainly repaired by SDSA.

5.3.4 Zebularine-DNMT NPAs: the trigger for DDR and cytotoxicity?

The formation of nucleoside analogues-DNMT NPAs have been proposed to be the cause for DNA damage induction and cytotoxicity to cancer cells (Kiziltepe et al., 2007; Lim et al., 2011).

Zebularine was shown to induce DNA damage induction, cell cycle arrest, and

increased level of deoxy-zebularine adduct in human glioblastoma cells (Meador et al., 2010), and the growth inhibition by zebularine of lung cancer cell was related to the elevated ROS production and reduction in antioxidants (You and Park, 2014). In this study, plant growth inhibition was observed after zebularine treatment, and the inhibition was partially restored in *ddc* mutant. The up-regulation of DDR genes by zebularine were significantly reduced in *ddc*. Both of the results suggest the zebularine-induced cytotoxicity and DDR activation might be at least partially due to the formation of deoxy-zebularine-DNMT NPAs (Lim et al., 2011). The increase in ROS and the changes in redox state could be another source of zebularine-induced cytotoxicity and DDR (Ruiz-Magaña et al., 2012; You and Park, 2014), which needs to be further investigated.

The exposure to ionizing radiation, UV light, formaldehyde (FA), or compromised topoisomerase action could all lead to the formation of NPAs (or DNA-protein crosslinks). The formation and the repair mechanism of NPAs were proposed in yeast (Stingele et al., 2014). In the case of FA exposure, both NER and HR contribute to the repair of NPAs. While NER cannot act on large NPAs (it removes 24 to 32 base oligonucleotide from a damage strand), and the repairing by HR is independent of the NPAs size (Nakano et al., 2007). In this study, the zebularine-activated DDR was shown to be related to the formation of NPAs, and the repairing mechanism was highly dependent on HR, indicating by the hypersensitivity of *smc6b* and the high induction of SDSA-HR frequencies. However, the types of zebularine-induced DNA damages (whether they are large NPAs or smaller) and the coordination between HR and NER for their repair needs further investigation.

5.3.5 The differential effects of nucleoside analogues on Arabidopsis:

the comparison of zebularine and 5-aza-cytidine

Zebularine and 5-aza-cytidine were proposed to have identical mode of action in DNA demethylation and differ mainly by chemical stability and thus duration of treatment (Ben-Kasus et al., 2005). The cytotoxicity of zebularine and 5-aza-cytidine to cancer cells were proposed to be due to the induced-DDR and cell cycle arrest by forming NPAs with DNMT (Lim et al., 2011). In this study, the comparison of the effect of the two drugs suggests that the two nucleoside analogues may have different biological functions in Arabidopsis. The activation of DDR and high HR frequencies induced by zebularine were not observed after 5-aza-cytidine treatment, and 5-aza-cytidine seemed to be much more toxic to Arabidopsis than zebularine. The hypomethylation activity of 5-aza-cytidine was shown to be more potent than zebularine, and their effects on cancer cell transcriptome were shown to be very distinct (Flotho et al., 2009). Nonetheless, how they are metabolized and incorporated into DNA and RNA and their molecular activity in Arabidopsis require further studies.

5.4 Identifying a novel DNA damage repair gene based on the co-expression with known DNA damage repair genes

The DDR and repair mechanism are well studied in yeast and mammals, and orthologs of DDR-related genes in plants have been identified based on the conserved sequence of protein domain from the animal or yeast models (Culligan et al., 2004; De Schutter et al., 2007). The major limitation of this approach is the low potential to identify novel DDR genes that are specific to plants. An alternative method is forward genetic screening to find mutants showing altered sensitivity to genotoxic agents (Sweeney et

al., 2009) or defects in development and fertility (Wang and Liu, 2006). However, the screening of mutant population can be time-consuming with potentially high risk of identifying known candidates, and the subtle effects from the mutation could complicate the identifying process. In this study, a new approach was applied in identifying DDR genes in Arabidopsis by co-expression analysis, in which genes that are functionally related tend to be co-expressed.

5.4.1 *CHROMATIN REMODELING 31* – the connection between chromatin remodeling and DNA repair

Chromatin remodeling is the alteration of chromatin structure through the movement of the nucleosomes on DNA string in an ATP-dependent manner. These changes in high-order chromatin structure contribute to gene expression regulation by facilitating the access to e.g. transcription factors binding sites (Vriet et al., 2015).

Studies in yeast and mammals indicate that chromatin remodeling complexes are important and may remodel nucleosomes during DNA damage repair using the energy from ATP hydrolysis (Weiwei Lai et al., 2013). Several evidences in Arabidopsis have shown the link between chromatin structure stability and DNA repair. Hypersensitivity to genotoxic stresses and increased DDR were observed in the mutants of CAF-1 subunits, *BRUSHY1* (*BRU1*), and *HAM1* and *HAM2* (the histone acetyltransferases) and were related to the changes in chromatin structure (Takeda et al., 2004; Campi et al., 2012; Hisanaga et al., 2013), indicating that DNA repair is regulated both at genetic and epigenetic levels.

Reverse genetic screen of the candidate genes suggested by co-expression analysis

revealed *CHR31* as a potential candidate of a novel DDR-associated gene. *CHR31* was identified as a yeast *RAD54*-like gene and of the SWITCH 2 (SWI2) / SUCROSE NON-FERMENTING 2 (SNF2) chromatin remodeling gene family that contain the characteristic ATPase / helicase motifs (Shaked et al., 2006). The mutant of yeast *RAD54* was shown to be hypersensitive to ionizing radiation and involved in HR by interacting with *RAD51*. As part of the SWI2/SNF2 chromatin remodeling family, it was proposed to play a role in altering accessibility of template DNA via chromatin remodeling during synapsis (Raoul Tan et al., 2003). Several SWI2/SNF2 chromatin remodeling genes in Arabidopsis was shown to be sensitive to gamma-irradiation and their defects resulted in a dramatic reduction in HR frequencies (Shaked et al., 2006; Roth et al., 2012). In this study, increased hypersensitivity and SSA-HR frequencies upon MMC treatment were observed in *chr31*. Expression analysis showed enhanced DDR in *chr31* and up-regulation of *CHR31* in WT after MMC treatment. All the results point to that a stronger DDR or maybe more severe damages upon MMC treatment were observed in *chr31*, indicating *CHR31* as a putative chromatin remodeler (speculated from molecular phylogenetics) involved in DDR. Nonetheless, as a putative chromatin remodeler, it requires further investigation to test whether the involvement of *CHR31* in DDR is related to the changes in chromatin remodeling activities.

Activation of cell cycle checkpoint (indicated by the cyclin-GUS reporter line) and the induction of endoreduplication upon MMC were observed in WT, and in *chr31* they were significantly reduced, indicating the involvement of *CHR31* in activating MMC-induced G2/M checkpoint and regulating the mitosis-to-endocycle transition. In yeast, the chromatin remodeling complex were shown to be recruited to centromeres and implicated in chromosome segregation (Papamichos-Chronakis and Peterson, 2013). However, the detailed function of *CHR31* in cell-cycle control and the induction of

endoreduplication will need further studies.

CHR31 was also identified as *CLASSY3* (*CLY3*) which may interact with RNA polymerase IV (Pol IV). Pol IV is thought to initiate biogenesis of small interfering RNAs (siRNAs) in RNA-directed DNA methylation, and thus playing a role in DNA methylation (Law et al., 2011). The interaction between chromatin remodeling and DNA methylation was proposed in one of the SWI2/SNF2 genes, *DDM1*. Mutations in *DDM1* caused a rapid loss of cytosine methylation and hypersensitivity toward gamma-irradiation. However, the altered DDR in *ddm1* was proposed to be caused by disruption of chromatin-remodeling functions rather than the alterations in cytosine methylation (Shaked et al., 2006). Hence, it requires further investigation to test whether the MMC sensitivity of *chr31* is caused by disruption of chromatin-remodeling functions or by alterations in cytosine methylation.

5.4.2 *CHR31* and the transition zone in root apex

In Arabidopsis, chemical inducers of DNA damages usually induce cell death at meristematic tissues where there is active cell division going on to prevent deleterious chromosome transmission. In this study, MMC-induced cell death was observed at RAM in both WT and *chr31*. However, very strong red fluorescence was detected in *chr31* at the root transition zone.

The root transition zone concept, in its original sense, states that root cells leaving the apical meristem need to accomplish a transitional stage of cyto-architectural rearrangement, especially of the actin cytoskeleton, in order to perform rapid cell elongation. It is the most sensitive zone of the root apex as it integrates diverse inputs

from endogenous (hormonal) and exogenous (sensorial) stimuli and translates them into signaling and motoric outputs as adaptive differential growth responses. Though negligible with respect of cell growth, transition zone is the most active zone in the whole root apex with respect of oscillating electric spike activities, auxin flux, vesicle recycling activity, and oxygen demands (Baluška et al., 2010; Baluska et al., 2013).

The high induction of cell death at the transition zone in *chr31* indicates that defective in *CHR31* might affect cell development at this area and make them the target to MMC-induced toxicity. However, it needs further investigation for the connection between *CHR31* and cell development in the transition zone and their response to MMC toxicity or even to general stress response.

Summary

In conclusion, this thesis revealed the natural variation of *A. thaliana* to DNA replication stress using hydroxyurea (HU). Combining the result of Genome-wide association studies and quantitative trait locus mapping suggest that HU sensitivity is a complex and multigenic trait underlain with additive and epistatic effects. One of the genetic elements has been localized to a genomic region of 22 kb.

The study of non-methylable cytidine analogue zebularine and its cytotoxicity revealed an induction of DNA damage responses (DDR), which is at least partially contributed by the formation of nucleo-protein adducts between zebularine and DNA methyltransferases. Molecular and genetic studies suggest that zebularine-induced DNA damages mainly occur after DNA strand separation during replication and require the repair by synthesis-dependent strand-annealing type of homologous recombination.

Co-expression analysis of DDR genes identified *CHR31* as a potential candidate of novel DNA damage repair-associated genes. Characterizing *CHR31* functions provides novel evidence on the connection between chromatin remodeling and DNA damage repair.

References

Adachi, S., Minamisawa, K., Okushima, Y., Inagaki, S., Yoshiyama, K., Kondou, Y., Kaminuma, E., Kawashima, M., Toyoda, T., Matsui, M., Kurihara, D., Matsunaga, S., and Umeda, M. (2011). Programmed induction of endoreduplication by DNA double-strand breaks in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America **108,** 10004-10009.

Allis, C.D., Jenuwein, T., and Reinberg, D. (2007). Epigenetic regulation in plants. In Epigenetics, M. Matzke and O.M. Scheid, eds (Cold Spring Harbor Laboratory Press), pp. 175-176.

Alonso-Blanco, C., and Koornneef, M. (2000). Naturally occurring variation in *Arabidopsis*: an underexploited resource for plant genetics. Trends in Plant Science **5,** 22-29.

Atwell, S., Huang, Y.S., Vilhjalmsson, B.J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A.M., Hu, T.T., Jiang, R., Muliyati, N.W., Zhang, X., Amer, M.A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., de Meaux, J., Ecker, J.R., Faure, N., Kniskern, J.M., Jones, J.D.G., Michael, T., Nemri, A., Roux, F., Salt, D.E., Tang, C., Todesco, M., Traw, M.B., Weigel, D., Marjoram, P., Borevitz, J.O., Bergelson, J., and Nordborg, M. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature **465,** 627-631.

Balestrazzi, A., Confalonieri, M., Macovei, A., Donà, M., and Carbonera, D. (2011). Genotoxic stress and DNA repair in plants: emerging functions and tools for improving crop productivity. Plant Cell Reports **30,** 287-295.

Balestrazzi, A., Macovei, A., Donà, M., Carbonera, D., and Confalonieri, M. (2013). Genotoxic stress, DNA repair, and crop productivity. In Crop Improvement Under Adverse Conditions, N. Tuteja and S.S. Gill, eds (Springer New York), pp. 153-169.

Baluska, F., Baluska, F., and Mancuso, S. (2013). Root apex transition zone as oscillatory zone. Frontiers in Plant Science **4,** 1-15.

Baluška, F., Mancuso, S., Volkmann, D., and Barlow, P.W. (2010). Root apex transition zone: a signalling–response nexus in the root. Trends in Plant Science **15,** 402-408.

Barow, M., and Meister, A. (2003). Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. Plant, Cell & Environment **26,** 571-584.

Baubec, T., Pecinka, A., Rozhon, W., and Mittelsten Scheid, O. (2009). Effective, homogeneous and transient interference with cytosine methylation in plant genomic DNA by zebularine. The Plant Journal **57,** 542-554.

Baubec, T., Finke, A., Mittelsten Scheid, O., and Pecinka, A. (2014). Meristem-specific expression of epigenetic regulators safeguards transposon silencing in *Arabidopsis*. EMBO reports **15,** 446-452.

Ben-Kasus, T., Ben-Zvi, Z., Marquez, V.E., Kelley, J.A., and Agbaria, R. (2005). Metabolic activation of zebularine, a novel DNA methylation inhibitor, in human bladder carcinoma cells. Biochemical Pharmacology **70,** 121-133.

Bergelson, J., and Roux, F. (2010). Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. Nature Reviews Genetics **11,** 867-879.

Billam, M., Sobolewski, M., and Davidson, N. (2010). Effects of a novel DNA methyltransferase inhibitor zebularine on human breast cancer cells. Breast Cancer Resear Treat **120,** 581-592.

Bleuyard, J.-Y., Gallego, M.E., and White, C.I. (2006). Recent advances in understanding of the DNA double-strand break repair machinery of plants. DNA Repair **5,** 1-12.

Bleuyard, J.-Y., Gallego, M.E., Savigny, F., and White, C.I. (2005). Differing requirements for the *Arabidopsis* RAD51 paralogs in meiosis and DNA repair. The Plant Journal **41,** 533-545.

Böhmdorfer, G., Schleiffer, A., Brunmeir, R., Ferscha, S., Nizhynska, V., Kozák, J., Angelis, K.J., Kreil, D.P., and Schweizer, D. (2011). GMI1, a structural-maintenance-of-chromosomes-hinge domain-containing protein, is involved in somatic homologous recombination in *Arabidopsis*. The Plant Journal **67,** 420-433.

Bray, C.M., and West, C.E. (2005). DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity. New Phytologist **168,** 511-528.

Britt, A.B. (1996). DNA damage and repair in plants. Annual Review of Plant Physiology and Plant Molecular Biology **47,** 75-100.

Campi, M., D'Andrea, L., Emiliani, J., and Casati, P. (2012). Participation of chromatin-remodeling proteins in the repair of ultraviolet-B-damaged DNA. Plant Physiology **158,** 981-995.

Cao, J., Schneeberger, K., Ossowski, S., Gunther, T., Bender, S., Fitz, J., Koenig, D., Lanz, C., Stegle, O., Lippert, C., Wang, X., Ott, F., Muller, J., Alonso-Blanco, C., Borgwardt, K., Schmid, K.J., and Weigel, D. (2011). Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nature

Genetics **43,** 956-963.

Causse, M., Chaïb, J., Lecomte, L., Buret, M., and Hospital, F. (2007). Both additivity and epistasis control the genetic variation for fruit quality traits in tomato. Theor Appl Genet **115,** 429-442.

Champion, C., Guianvarc'h, D., Sénamaud-Beaufort, C., Jurkowska, R.Z., Jeltsch, A., Ponger, L., Arimondo, P.B., and Guieysse-Peugeot, A.-L. (2010). Mechanistic insights on the inhibition of C5 DNA methyltransferases by zebularine. PLoS ONE **5,** e12388.

Chan, S.W.L., Henderson, I.R., and Jacobsen, S.E. (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. Nature Reviews Genetics **6,** 351-360.

Chang, S., and Pikaard, C.S. (2005). Transcript profiling in *Arabidopsis* reveals complex responses to global inhibition of DNA methylation and histone deacetylation. Journal of Biological Chemistry **280,** 796-804.

Cheng, J.C., Yoo, C.B., Weisenberger, D.J., Chuang, J., Wozniak, C., Liang, G., Marquez, V.E., Greer, S., Orntoft, T.F., Thykjaer, T., and Jones, P.A. (2004). Preferential response of cancer cells to zebularine. Cancer Cell **6,** 151-158.

Christman, J., K. (2002). 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. Oncogene **21,** 5483-5495.

Christopherson, R.I., Lyons, S.D., and Wilson, P.K. (2002). Inhibitors of de novo nucleotide biosynthesis as drugs. Accounts of Chemical Research **35,** 961-971.

Clark, R.M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., Warthmann, N., Hu, T.T., Fu, G., Hinds, D.A., Chen, H., Frazer, K.A., Huson, D.H., Schölkopf, B., Nordborg, M., Rätsch, G., Ecker, J.R., and Weigel, D. (2007). Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. Science **317,** 338-342.

Collard, B., Jahufer, M., Brouwer, J., and Pang, E. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica **142,** 169-196.

Colón-Carmona, A., You, R., Haimovitch-Gal, T., and Doerner, P. (1999). Spatio-temporal analysis of mitotic activity with a labile cyclin–GUS fusion protein. The Plant Journal **20,** 503-508.

Cools, T., Iantcheva, A., Maes, S., Van den Daele, H., and De Veylder, L. (2010). A replication stress-induced synchronization method for *Arabidopsis thaliana* root meristems. The Plant Journal **64,** 705-714.

Cools, T., Iantcheva, A., Weimer, A.K., Boens, S., Takahashi, N., Maes, S., Van

den Daele, H., Van Isterdael, G., Schnittger, A., and De Veylder, L. (2011). The *Arabidopsis thaliana* checkpoint kinase *WEE1* protects against premature vascular differentiation during replication stress. The Plant Cell **23,** 1435-1448.

Culligan, K., Tissier, A., and Britt, A. (2004). ATR regulates a G2-phase cell-cycle checkpoint in *Arabidopsis thaliana*. The Plant Cell **16,** 1091-1104.

Culligan, K.M., Robertson, C.E., Foreman, J., Doerner, P., and Britt, A.B. (2006). ATR and ATM play both distinct and additive roles in response to ionizing radiation. The Plant Journal **48,** 947-961.

Da Ines, O., Degroote, F., Amiard, S., Goubely, C., Gallego, M.E., and White, C.I. (2013). Effects of *XRCC2* and *RAD51B* mutations on somatic and meiotic recombination in *Arabidopsis thaliana*. The Plant Journal **74,** 959-970.

De Schutter, K., Joubès, J., Cools, T., Verkest, A., Corellou, F., Babiychuk, E., Van Der Schueren, E., Beeckman, T., Kushnir, S., Inzé, D., and De Veylder, L. (2007). *Arabidopsis WEE1* kinase controls cell cycle arrest in response to activation of the DNA integrity checkpoint. The Plant Cell **19,** 211-225.

De Veylder, L., Beeckman, T., and Inze, D. (2007). The ins and outs of the plant cell cycle. Nature Reviews Molecular Cell Biology **8,** 655-665.

De Veylder, L., Larkin, J.C., and Schnittger, A. (2011). Molecular control and function of endoreplication in development and physiology. Trends in Plant Science **16,** 624-634.

Deans, A.J., and West, S.C. (2011). DNA interstrand crosslink repair and cancer. Nature Reviews Cancer **11,** 467-480.

Dubest, S., Gallego, M.E., and White, C.I. (2002). Role of the AtRad1p endonuclease in homologous recombination in plants. EMBO reports **3,** 1049-1054.

Egger, G., Liang, G., Aparicio, A., and Jones, P.A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. Nature **429,** 457-463.

Ehlting, J., Provart, N.J., and Werck-Reichhart, D. (2006). Functional annotation of the Arabidopsis P450 superfamily based on large-scale co-expression analysis. Biochem Soc Trans **34,** 1192-1198.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America **95,** 14863-14868.

Eshed, Y., and Zamir, D. (1996). Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics **143,** 1807-1817.

Flotho, C., Claus, R., Batz, C., Schneider, M., Sandrock, I., Ihde, S., Plass, C.,

Niemeyer, C.M., and Lubbert, M. (2009). The DNA methyltransferase inhibitors azacitidine, decitabine and zebularine exert differential effects on cancer gene expression in acute myeloid leukemia cells. Leukemia **23,** 1019-1028.

Fu, D., Calvo, J.A., and Samson, L.D. (2012). Balancing repair and tolerance of DNA damage caused by alkylating agents. Nature Reviews Cancer **12,** 104-120.

Garcia, V., Bruchet, H., Camescasse, D., Granier, F., Bouchez, D., and Tissier, A. (2003). AtATM is essential for meiosis and the somatic response to DNA damage in plants. The Plant Cell Online **15,** 119-132.

Hanin, M., Mengiste, T., Bogucki, A., and Paszkowski, J. (2000). Elevated levels of intrachromosomal homologous recombination in *Arabidopsis* overexpressing the *MIM* gene. The Plant Journal **24,** 183-189.

Henderson, I.R., and Jacobsen, S.E. (2008). Tandem repeats upstream of the *Arabidopsis* endogene *SDC* recruit non-CG DNA methylation and initiate siRNA spreading. Genes & Development **22,** 1597-1606.

Hisanaga, T., Ferjani, A., Horiguchi, G., Ishikawa, N., Fujikura, U., Kubo, M., Demura, T., Fukuda, H., Ishida, T., Sugimoto, K., and Tsukaya, H. (2013). The ATM-dependent DNA damage response acts as an upstream trigger for compensation in the *fas1* mutation during *Arabidopsis* leaf development. Plant Physiology **162,** 831-841.

Huang, J., Kim-Shapiro, D.B., and King, S.B. (2004). Catalase-mediated nitric oxide formation from hydroxyurea. Journal of Medicinal Chemistry **47,** 3495-3501.

Jiang, C.Z., Yee, J., Mitchell, D.L., and Britt, A.B. (1997). Photorepair mutants of *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America **94,** 7441-7445.

Jiang, L., Liu, Y., Sun, H., Han, Y., Li, J., Li, C., Guo, W., Meng, H., Li, S., Fan, Y., and Zhang, C. (2013). The mitochondrial folylpolyglutamate synthetase gene is required for nitrogen utilization during early seedling development in *Arabidopsis*. Plant Physiology **161,** 971-989.

Juul, T., Malolepszy, A., Dybkær, K., Kidmose, R., Rasmussen, J.T., Andersen, G.R., Johnsen, H.E., Jørgensen, J.-E., and Andersen, S.U. (2010). The in vivo toxicity of hydroxyurea depends on its direct target catalase. Journal of Biological Chemistry **285,** 21411-21415.

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.-y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nature Genetics **42,** 348-

354.

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., D'Angelo, C., Bornberg-Bauer, E., Kudla, J., and Harter, K. (2007). The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. The Plant Journal **50,** 347-363.

Kim, S., Plagnol, V., Hu, T.T., Toomajian, C., Clark, R.M., Ossowski, S., Ecker, J.R., Weigel, D., and Nordborg, M. (2007). Recombination and linkage disequilibrium in *Arabidopsis thaliana*. Nature Genetics **39,** 1151-1155.

Kimura, S., and Sakaguchi, K. (2006). DNA repair in plants. Chemical Reviews **106,** 753-766.

King, S.B. (2004). Serial review: Mechanisms and novel directions in the biological applications of nitric oxide donors. Free Radical Biology & Medicine **37,** 737-744.

Kiziltepe, T., Hideshima, T., Catley, L., Raje, N., Yasui, H., Shiraishi, N., Okawa, Y., Ikeda, H., Vallet, S., Pozzi, S., Ishitsuka, K., Ocio, E.M., Chauhan, D., and Anderson, K.C. (2007). 5-Azacytidine, a DNA methyltransferase inhibitor, induces ATR-mediated DNA double-strand break responses, apoptosis, and synergistic cytotoxicity with doxorubicin and bortezomib against multiple myeloma cells. Molecular Cancer Therapeutics **6,** 1718-1727.

Klaue, D., Kobbe, D., Kemmerich, F., Kozikowska, A., Puchta, H., and Seidel, R. (2013). Fork sensing and strand switching control antagonistic activities of RecQ helicases. Nature Communications **4**.

Konieczny, A., and Ausubel, F.M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. The Plant Journal **4,** 403-410.

Kooke, R., Wijnker, E., and Keurentjes, J.J. (2012). Backcross populations and nearisogenic lines. In Quantitative Trait Loci (QTL) : methods and protocols, pp. 3-16.

Koornneef, M., Alonso-Blanco, C., and Vreugdenhil, D. (2004). Natural occurring genetic variation in *Arabidopsis Thaliana*. Annual Review of Plant Biology **55,** 141-172.

Koornneef, M., Reymond, M., and Alonso-Blanco, C. (2011). Natural variation in *Arabidopsis thaliana* In Genetics and Genomics of the *Brassicaceae*, R. Schmidt and I. Bancroft, eds (Springer New York), pp. 123-151.

Kozak, J., West, C.E., White, C., da Costa-Nunes, J.A., and Angelis, K.J. (2009). Rapid repair of DNA double strand breaks in *Arabidopsis thaliana* is dependent on proteins involved in chromosome structure maintenance. DNA

Repair **8,** 413-419.

Lang, J., Smetana, O., Sanchez-Calderon, L., Lincker, F., Genestier, J., Schmit, A.-C., Houlné, G., and Chabouté, M.-E. (2012). Plant γH2AX foci are required for proper DNA DSB repair responses and colocalize with E2F factors. New Phytologist **194,** 353-363.

Law, J.A., Vashisht, A.A., Wohlschlegel, J.A., and Jacobsen, S.E. (2011). SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. PLoS Genetics **7,** e1002195.

Le, M.H., Cao, Y., Zhang, X.-C., and Stacey, G. (2014). LIK1, a CERK1-interacting kinase, regulates plant immune responses in *Arabidopsis*. PLoS ONE **9,** e102245.

Li, Y., Huang, Y., Bergelson, J., Nordborg, M., and Borevitz, J.O. (2010). Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America **107,** 21199-21204.

Lim, S., Neilsen, P., Kumar, R., Abell, A., and Callen, D. (2011). The application of delivery systems for DNA methyltransferase inhibitors. BioDrugs **25,** 227-242.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods **25,** 402-408.

Long, A.D., Mullaney, S.L., Mackay, T.F.C., and Langley, C.H. (1996). Genetic interactions between naturally occurring alleles at quantitative trait loci and mutant alleles at candidate loci affecting bristle number in *Drosophila Melanogaster*. Genetics **144,** 1497-1510.

Ma, J.-L., Kim, E.M., Haber, J.E., and Lee, S.E. (2003). Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. Molecular and Cellular Biology **23,** 8820-8828.

Mackay, T.F.C. (2014). Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nature Reviews Genetics **15,** 22-33.

Mannuss, A., Dukowic-Schulze, S., Suer, S., Hartung, F., Pacher, M., and Puchta, H. (2010). RAD5A, RECQ4A, and MUS81 have specific functions in homologous recombination and define different pathways of DNA repair in *Arabidopsis thaliana*. The Plant Cell **22,** 3318-3330.

Marquez, V.E., Kelley, J.A., Agbaria, R., Ben-Kasus, T., Cheng, J.C., Yoo, C.B., and Jones, P.A. (2005). Zebularine: A unique molecule for an epigenetically

based strategy in cancer chemotherapy. Annals of the New York Academy of Sciences **1058,** 246-254.

Meador, J.A., Su, Y., Ravanat, J.-L., and Balajee, A.S. (2010). DNA-dependent protein kinase (DNA-PK)-deficient human glioblastoma cells are preferentially sensitized by zebularine. Carcinogenesis **31,** 184-191.

Meinke, D.W., Cherry, J.M., Dean, C., Rounsley, S.D., and Koornneef, M. (1998). *Arabidopsis thaliana*: a model plant for genome analysis. Science **282,** 662-682.

Mengiste, T., Revenkova, E., Bechtold, N., and Paszkowski, J. (1999). An SMC-like protein is required for efficient homologous recombination in *Arabidopsis*. The EMBO Journal **18,** 4505-4512.

Miller-Messmer, M., Kühn, K., Bichara, M., Le Ret, M., Imbault, P., and Gualberto, J.M. (2012). RecA-dependent DNA repair results in increased heteroplasmy of the *Arabidopsis* mitochondrial genome. Plant Physiology **159,** 211-226.

Molinier, J., Ries, G., Bonhoeffer, S., and Hohn, B. (2004). Interchromatid and interhomolog recombination in *Arabidopsis thaliana*. The Plant Cell **16,** 342-352.

Murashige, T., and Skoog, F. (1962). A Revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiologia Plantarum **15,** 473-497.

Nakano, T., Morishita, S., Katafuchi, A., Matsubara, M., Horikawa, Y., Terato, H., Salem, A.M.H., Izumi, S., Pack, S.P., Makino, K., and Ide, H. (2007). Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks. Molecular Cell **28,** 147-158.

Nezames, C.D., Sjogren, C.A., Barajas, J.F., and Larsen, P.B. (2012). The *Arabidopsis* cell cycle checkpoint regulators TANMEI/ALT2 and ATR mediate the active process of aluminum-dependent root growth inhibition. The Plant Cell **24,** 608-621.

Nordborg, M., and Weigel, D. (2008). Next-generation genetics in plants. Nature **456,** 720-723.

Nordborg, M., Hu, T.T., Ishino, Y., Jhaveri, J., Toomajian, C., Zheng, H., Bakker, E., Calabrese, P., Gladstone, J., Goyal, R., Jakobsson, M., Kim, S., Morozov, Y., Padhukasahasram, B., Plagnol, V., Rosenberg, N.A., Shah, C., Wall, J.D., Wang, J., Zhao, K., Kalbfleisch, T., Schulz, V., Kreitman, M., and Bergelson, J. (2005). The pattern of polymorphism in *Arabidopsis thaliana*. PLoS Biology **3,** e196.

Obayashi, T., and Kinoshita, K. (2010). Coexpression landscape in ATTED-II:

usage of gene list and gene network for various types of pathways. J Plant Res **123,** 311-319.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H., and Kinoshita, K. (2009). ATTED-II provides coexpressed gene networks for *Arabidopsis*. Nucleic Acids Research **37,** D987-D991.

Obayashi, T., Okamura, Y., Ito, S., Tadaka, S., Aoki, Y., Shirota, M., and Kinoshita, K. (2014). ATTED-II in 2014: evaluation of gene coexpression in agriculturally important plants. Plant and Cell Physiology **55,** e6.

Page, D.R., and Grossniklaus, U. (2002). The art and design of genetic screens: *Arabidopsis thaliana*. Nature Reviews Genetics **3,** 124-136.

Pan, W.H., Houben, A., and Schlegel, R. (1993). Highly effective cell synchronization in plant roots by hydroxyurea and amiprophos-methyl or colchicine. Genome **36,** 387-390.

Papamichos-Chronakis, M., and Peterson, C.L. (2013). Chromatin and the genome integrity network. Nature Reviews Genetics **14,** 62-75.

Pecinka, A., Rosa, M., Schikora, A., Berlinger, M., Hirt, H., Luschnig, C., and Scheid, O.M. (2009). Transgenerational stress memory is not a general response in *Arabidopsis*. PLoS ONE **4,** e5202.

Povirk, L.F. (1996). DNA damage and mutagenesis by radiomimetic DNA-cleaving agents: bleomycin, neocarzinostatin and other enediynes. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **355,** 71-89.

Preuss, S.B., and Britt, A.B. (2003). A DNA-damage-induced cell cycle checkpoint in *Arabidopsis*. Genetics **164,** 323-334.

Puchta, H., Swoboda, P., and Hohn, B. (1995). Induction of intrachromosomal homologous recombination in whole plants. The Plant Journal **7,** 203-210.

Purcell, S., and Sham, P. (2004). Epistasis in quantitative trait locus linkage analysis: interaction or main effect? Behav Genet **34,** 143-152.

Ransbotyn, V., Yeger-Lotem, E., Basha, O., Acuna, T., Verduyn, C., Gordon, M., Chalifa-Caspi, V., Hannah, M.A., and Barak, S. (2014). A combination of gene expression ranking and co-expression network analysis increases discovery rate in large-scale mutant screens for novel Arabidopsis thaliana abiotic stress genes. Plant Biotechnology Journal**,** 1-13.

Raoul Tan, T.L., Kanaar, R., and Wyman, C. (2003). RAD54, a Jack of all trades in homologous recombination. DNA Repair **2,** 787-794.

Roa, H., Lang, J., Culligan, K.M., Keller, M., Holec, S., Cognat, V., Montané, M.-H., Houlné, G., and Chabouté, M.-E. (2009). Ribonucleotide reductase regulation in response to genotoxic stress in *Arabidopsis*. Plant Physiology

151, 461-471.

Roldán-Arjona, T., and Ariza, R.R. (2009). Repair and tolerance of oxidative DNA damage in plants. Mutation Research/Reviews in Mutation Research **681,** 169-179.

Rosa, M., Von Harder, M., Cigliano, R.A., Schlögelhofer, P., and Scheid, O.M. (2013). The *Arabidopsis* SWR1 chromatin-remodeling complex is important for DNA repair, somatic recombination, and meiosis. The Plant Cell**,** 1-12.

Rose, A., Manikantan, S., Schraegle, S.J., Maloy, M.A., Stahlberg, E.A., and Meier, I. (2004). Genome-wide identification of *Arabidopsis* coiled-coil proteins and establishment of the ARABI-COIL database. Plant Physiology **134,** 927-939.

Roth, N., Klimesch, J., Dukowic-Schulze, S., Pacher, M., Mannuss, A., and Puchta, H. (2012). The requirement for recombination factors differs considerably between different pathways of homologous double-strand break repair in somatic plant cells. The Plant Journal **72,** 781-790.

Ruiz-Magaña, M.J., Rodríguez-Vargas, J.M., Morales, J.C., Saldivia, M.A., Schulze-Osthoff, K., and Ruiz-Ruiz, C. (2012). The DNA methyltransferase inhibitors zebularine and decitabine induce mitochondria-mediated apoptosis and DNA damage in p53 mutant leukemic T cells. International Journal of Cancer **130,** 1195-1207.

Sanan-Mishra, N., Pham, X.H., Sopory, S.K., and Tuteja, N. (2005). Pea DNA helicase 45 overexpression in tobacco confers high salinity tolerance without affecting yield. Proceedings of the National Academy of Sciences of the United States of America **102,** 509-514.

Sancar, A. (2003). Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. Chemical Reviews **103,** 2203-2238.

Santi, D.V., Norment, A., and Garrett, C.E. (1984). Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. Proceedings of the National Academy of Sciences of the United States of America **81,** 6993-6997.

Schott, A. (2012). Natural variation of *Arabidopsis thaliana* in response to hydroxyurea. In Plant Breeding and Genetics, Max Planck Institute for Plant Breeding Research (MPIPZ) (Cologne: Fachhochschule Aachen, Fachbereich Chemie und Biotechnologie), pp. 51.

Shaked, H., Avivi-Ragolsky, N., and Levy, A.A. (2006). Involvement of the *Arabidopsis* SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination. Genetics **173,** 985-994.

Simon, M., Loudet, O., Durand, S., Bérard, A., Brunel, D., Sennesal, F.-X.,

Durand-Tardif, M., Pelletier, G., and 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.

Singh, S., Roy, S., Choudhury, S., and Sengupta, D. (2010). DNA repair and recombination in higher plants: insights from comparative genomics of arabidopsis and rice. BMC Genomics **11,** 443.

Smetana, O., Široký, J., Houlné, G., Opatrný, Z., and Chabouté, M.-E. (2012). Non-apoptotic programmed cell death with paraptotic-like features in bleomycin-treated plant cells is suppressed by inhibition of ATM/ATR pathways or *NtE2F* overexpression. Journal of Experimental Botany **63,** 2631-2644.

Spadafora, N.D., Doonan, J.H., Herbert, R.J., Bitonti, M.B., Wallace, E., Rogers, H.J., and Francis, D. (2011). *Arabidopsis* T-DNA insertional lines for *CDC25* are hypersensitive to hydroxyurea but not to zeocin or salt stress. Annals of Botany **107,** 1183-1192.

Stam, L.F., and Laurie, C.C. (1996). Molecular dissection of a major gene effect on a quantitative trait: The level of alcohol dehydrogenase expression in *Drosophila Melanogaster*. Genetics **144,** 1559-1564.

Stingele, J., Schwarz, Michael S., Bloemeke, N., Wolf, Peter G., and Jentsch, S. (2014). A DNA-dependent protease involved in DNA-protein crosslink repair. Cell **158,** 327-338.

Sweeney, P.R., Britt, A.B., and Culligan, K.M. (2009). The Arabidopsis ATRIP ortholog is required for a programmed response to replication inhibitors. The Plant Journal **60,** 518-526.

Takeda, S., Tadele, Z., Hofmann, I., Probst, A.V., Angelis, K.J., Kaya, H., Araki, T., Mengiste, T., Scheid, O.M., Shibahara, K.-i., Scheel, D., and Paszkowski, J. (2004). *BRU1*, a novel link between responses to DNA damage and epigenetic gene silencing in *Arabidopsis*. Genes & Development **18,** 782-793.

Tanaka, A., Sakamoto, A., Ishigaki, Y., Nikaido, O., Sun, G., Hase, Y., Shikazono, N., Tano, S., and Watanabe, H. (2002). An ultraviolet-B-resistant mutant with enhanced DNA repair in *Arabidopsis*. Plant Physiology **129,** 64-71.

Thiel, T., Kota, R., Grosse, I., Stein, N., and Graner, A. (2004). SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. Nucleic Acids Research **32,** e5.

Tuinstra, M.R., Ejeta, G., and Goldsbrough, P.B. (1997). Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ

at quantitative trait loci. TAG Theoretical and Applied Genetics **95,** 1005-1011.

Tuteja, N., Ahmad, P., Panda, B.B., and Tuteja, R. (2009). Genotoxic stress in plants: Shedding light on DNA damage, repair and DNA repair helicases. Mutation Research/Reviews in Mutation Research **681,** 134-149.

Ubeda-Tomás, S., Beemster, G.T.S., and Bennett, M.J. (2012). Hormonal regulation of root growth: integrating local activities into global behaviour. Trends in Plant Science **17,** 326-331.

Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F.M., Bassel, G.W., Tanimoto, M., Chow, A., Steinhauser, D., Persson, S., and Provart, N.J. (2009). Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. Plant, Cell & Environment **32,** 1633-1651.

Vanderauwera, S., De Block, M., Van de Steene, N., van de Cotte, B., Metzlaff, M., and Van Breusegem, F. (2007). Silencing of poly(ADP-ribose) polymerase in plants alters abiotic stress signal transduction. Proceedings of the National Academy of Sciences of the United States of America **104,** 15150-15155.

Vashisht, A.A., and Tuteja, N. (2006). Stress responsive DEAD-box helicases: A new pathway to engineer plant stress tolerance. Journal of Photochemistry and Photobiology B: Biology **84,** 150-160.

Vriet, C., Hennig, L., and Laloi, C. (2015). Stress-induced chromatin changes in plants: of memories, metabolites and crop improvement. Cell. Mol. Life Sci. **72,** 1261-1273.

Wang, C., and Liu, Z. (2006). *Arabidopsis* ribonucleotide reductases are critical for cell cycle progression, DNA damage repair, and plant development. The Plant Cell **18,** 350-365.

Weigel, D. (2012). Natural variation in *Arabidopsis*: From molecular genetics to ecological genomics. Plant Physiology **158,** 2-22.

Weiwei Lai, Hongde Li, Shuang Liu, and Tao, Y. (2013). Connecting chromatin modifying factors to DNA damage response. International Journal of Molecular Sciences **14,** 2355–2369.

Yang, Q., Zhang, D., and Xu, M. (2012). A sequential quantitative trait locus fine-mapping strategy using recombinant-derived progeny. Journal of Integrative Plant Biology **54,** 228-237.

Yi, D., Alvim Kamei, C.L., Cools, T., Vanderauwera, S., Takahashi, N., Okushima, Y., Eekhout, T., Yoshiyama, K.O., Larkin, J., Van den Daele, H., Conklin, P., Britt, A., Umeda, M., and De Veylder, L. (2014). The *Arabidopsis* SIAMESE-RELATED cyclin-cependent kinase inhibitors SMR5

and SMR7 regulate the DNA damage checkpoint in response to reactive oxygen species. The Plant Cell **26,** 296-309.

Yoshiyama, K., Conklin, P.A., Huefner, N.D., and Britt, A.B. (2009). Suppressor of gamma response 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA damage. Proceedings of the National Academy of Sciences **106,** 12843-12848.

Yoshiyama, K.O., Sakaguchi, K., and Kimura, S. (2013a). DNA damage response in plants: conserved and variable response compared to animals. Biology **2,** 1338-1356.

Yoshiyama, K.O., Kobayashi, J., Ogita, N., Ueda, M., Kimura, S., Maki, H., and Umeda, M. (2013b). ATM-mediated phosphorylation of SOG1 is essential for the DNA damage response in *Arabidopsis*. EMBO reports **14,** 817-822.

You, B.R., and Park, W.H. (2014). Zebularine inhibits the growth of A549 lung cancer cells via cell cycle arrest and apoptosis. Molecular Carcinogenesis **53,** 847-857.

Zeller, G., Clark, R.M., Schneeberger, K., Bohlen, A., Weigel, D., and Rätsch, G. (2008). Detecting polymorphic regions in *Arabidopsis thaliana* with resequencing microarrays. Genome Research **18,** 918-929.

Zemach, A., Kim, M.Y., Hsieh, P.-H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, Stacey L., and Zilberman, D. (2013). The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell **153,** 193-205.

Zhong, S., Fei, Z., Chen, Y.-R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., Shao, Y., and Giovannoni, J.J. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nature Biotechnology **31,** 154-159.

Zhu, C., Gore, M., Buckler, E.S., and Yu, J. (2008). Status and prospects of association mapping in plants. The Plant Genome **1,** 5-20.

Appendix

Supplemental Table 1. The list of *A. thaliana* accessions and their relative survival rate (RS) on HU.

Name	Country	Mean RS	SD RS
Ag-0	FRA	88.5	5.2
Ak-1	GER	95.9	1.8
Bay-0	GER	22.6	4.3
Bla-5	ESP	91.8	3.9
Bar-1	SWE	96.3	2.3
Bor-4	CZE	46.7	3.3
Bur-0	IRL	97.3	3.8
Bur-1	IRL	97.9	2.9
C24	POR	82.6	2.0
Can-0	ESP	N.A.	N.A.
Col-0	USA	95.7	3.4
Ct-1	ITA	49.0	10.2
N13	RUS	90.8	6.6
Oy-0	NOR	79.7	0.3
Eden-2	SWE	96.0	5.7
Ei-2	GER	84.1	4.5
Er-0	GER	28.8	2.5
Fäb-4	SWE	98.7	1.3
Fei-0	POR	95.6	3.1
Fr-2	GER	82.1	6.0
Kas-1	IND	96.8	3.2
Kin-0	USA	97.0	2.4
Kno-18	USA	64.5	4.5
Kz-1	Kazakhstan	82.1	7.1
Ga-0	GER	34.7	1.4
Got-7	GER	91.3	2.4
Got-22	GER	92.5	6.9
Gu-0	GER	95.6	4.4
Gy-0	FRA	79.3	6.6
Hk-3	N.A.	79.5	5.5
Hp2-5	N.A.	97.8	2.2
Hr-5	UK	76.5	4.7

Hr-10	UK	100.0	0.0
Is-0	GER	95.3	3.9
Kz-9	Kazakhstan	77.7	4.3
Ler-1	GER	31.1	31.1
LL-0	ESP	46.1	7.6
Lov-1	DEN	96.1	2.8
Lp2-2	CZE	85.0	4.3
Lp2-6	CZE	100.0	0.0
Lov-5	N.A.	98.1	2.7
Lz-0	FRA	89.2	4.8
Mr-0	SWI	60.2	21.7
Ms-0	RUS	60.7	7.7
Mt-0	LIB	55.7	5.7
Nd-1	SUI	99.2	1.1
NFA-10	UK	96.2	5.4
Nok-1	NED	73.0	18.0
Nw-0	GER	96.4	3.5
Nw-1	GER	99.8	0.3
Old-2	GER	97.0	4.2
Ömö2-1	SWE	71.4	6.0
Ove-0	GER	98.0	1.5
Per-1	RUS	73.3	10.3
Pna-10	USA	74.2	5.0
Pna-17	USA	79.8	10.9
Pro-0	ESP	73.7	2.2
Pu2-7	GER	80.4	7.7
Pu2-23	CZE	84.3	11.1
Ra-0	FRA	90.3	4.4
Ren-1	FRA	62.2	8.3
Ren-11	FRA	98.3	1.7
Rmx-Ps02	USA	89.3	15.1
Rmx-A180	USA	94.7	4.3
RRS-7	USA	69.9	7.3
RRS-10	USA	64.8	19.3
Rsch-4	RUS	98.2	2.5
Se-0	ESP	100.0	0.0
Shahdara	TJK	36.1	2.8

Sorbo	TJK	95.7	2.3
Spr1-6	SWE	61.5	18.2
Sq-1	UK	100.0	0.0
Sq-8	UK	94.7	1.2
Ta-0	CZE	74.5	4.5
Tamm-2	FIN	97.2	0.8
Tamm-27	FIN	98.5	2.2
Ts-5	ESP	94.7	0.0
N4	RUS	90.4	5.0
N5	RUS	96.4	2.6
N6	RUS	6.2	5.1
N7	RUS	75.6	1.9
N8	RUS	91.4	6.2
N9	RUS	98.9	1.5
Uod-1	AUT	97.4	2.6
Uod-7	AUT	86.9	2.9
Van-0	CAN	98.2	2.5
Vår2-1	SWE	98.5	1.5
Vår2-6	DEN	92.9	3.4
Wei-0	SUI	18.7	12.5
Wil-1	LTU	95.4	3.8
Ws-0	RUS	86.7	4.6
Ws-2	BEL	92.3	8.0
Wt-5	GER	79.3	0.0
Yo-0	USA	91.1	3.6
Zdr-1	CZE	78.5	5.5
Zdr-6	CZE	63.1	3.5
Fas1	N.A.	94.7	0.0
Hod	CZE	N.A.	N.A.
Kas-1	IND	93.5	0.0
KBS-Mac-8	USA	N.A.	N.A.
Köln	GER	91.2	5.0
LAC-3	FRA	93.2	1.8
LDV-58	FRA	81.7	5.0
Liarum	SWE	93.1	6.4
Lip-0	POL	92.2	8.7
Lm-2	FRA	86.3	1.3

Ha-0	GER	93.7	2.1
Hn-0	GER	93.8	6.3
Jm-1	CZE	28.8	7.4
Kr-0	GER	100.0	0.0
Kelsterbach-2	GER	N.A.	N.A.
Kl-5	GER	94.3	4.1
Kn-0	LTU	97.5	2.5
Kro-0	GER	100.0	0.0
Krot-2	GER	90.0	0.0
Li-7	GER	94.8	1.0
Mc-0	UK	89.6	5.2
Mh-0	POL	N.A.	N.A.
Nc-1	FRA	95.8	5.9
No-0	GER	100.0	0.0
Nw-0	GER	100.0	0.0
Nw-2	GER	87.1	1.8
Aa-0	GER	90.7	4.0
Alst-1	UK	97.0	2.3
An-2	BEL	99.3	0.7
Ang-0	BEL	88.5	5.0
Arby-1	SWE	93.9	8.6
Benk-1	NED	100.0	0.0
Boot-1	UK	94.5	4.1
Bs-2	SUI	97.2	2.3
Bsch-0	GER	100.0	0.0
Bu-8	GER	100.0	0.0
Ca-0	GER	75.9	3.2
Chat-1	FRA	96.6	3.4
Cit-0	FRA	82.4	12.7
Cnt-1	UK	94.9	4.5
Co-2	POR	92.0	0.8
CSHL-5	USA	75.8	5.0
Com-1	FRA	98.1	1.9
Da-0	GER	90.3	8.3
Da(1)-12	CZE	94.3	5.7
Db-0	GER	85.2	2.8
Di-1	FRA	95.3	0.0

Do-0	GER	91.5	6.3
Dra-2	CZE	80.8	0.0
Ede-1	NED	83.5	3.1
Ep-0	GER	97.2	3.9
Es-0	FIN	46.3	0.4
Fi-1	GER	93.7	0.8
Fr-4	GER	84.6	0.0
Ga-2	GER	82.1	12.1
Ge-1	SUI	92.2	5.6
Gel-1	NED	93.8	6.3
Gie-0	GER	60.6	19.3
Go-0	GER	20.4	6.2
Gr-5	AUT	90.0	6.5
Gu-1	GER	81.4	8.9
Hau-0	DEN	74.1	0.9
Hey-1	NED	81.6	9.5
Je-0	GER	N.A.	N.A.
Jl-3	CZE	20.4	14.1
KNO-11	USA	79.4	0.0
Li-3	GER	78.2	4.2
Li-5:2	GER	88.0	6.0
Li-6	GER	77.0	16.8
Mnz-0	GER	91.2	8.8
N7	RUS	82.1	6.0
NFC-20	UK	51.6	31.7
Nok-1	NED	48.6	19.8
Nz1	NZL	90.0	3.3
Ob-1	GER	73.4	4.7
Old-1	GER	78.5	11.8
Or-0	GER	100.0	0.0
Pa-2	ITA	81.7	1.7
PHW-13	UK	84.6	15.4
PHW-14	UK	35.3	0.3
PHW-20	UK	96.7	3.3
PHW-22	UK	97.4	2.6
PHW-26	UK	88.5	11.5
PHW-28	UK	92.7	4.3

PHW-31	UK	97.1	4.1
PHW-33	NED	90.8	3.2
PHW-35	FRA	85.0	6.0
PHW-36	FRA	94.4	4.0
PHW-37	FRA	52.4	5.3
Pla-0	ESP	90.7	4.4
Pn-0	FRA	98.6	2.0
Pog-0	CAN	99.1	1.3
Pr-0	GER	61.8	6.2
Pu2-24	CZE	68.7	1.3
Rhen-1	NED	92.8	7.2
Rou-0	FRA	93.6	1.8
S96	UNK	96.2	3.0
Sapporo-0	JPN	91.2	6.3
Sav-0	CZE	72.7	14.7
Sei-0	ITA	87.5	0.0
Sg-1	GER	30.6	6.1
Sh-0	GER	61.0	14.7
Si-0	GER	92.8	2.7
Sp-0	GER	96.8	3.2
Ste-0	GER	96.3	3.7
Tha-1	NED	95.2	4.1
Ting-1	SWE	98.1	1.9
Tiv-1	ITA	88.8	4.9
Tscha-1	AUT	96.8	2.6
Tsu-0	JPN	94.4	1.3
Uk-1	GER	88.8	2.9
Uk-1	GER	93.8	4.4
Uk-2	GER	100.0	0.0
Utrecht	NED	100.0	0.0
Ven-1	NED	92.2	5.7
Wa-1	POL	99.2	0.8
Wag-3	NED	37.8	10.5
Wag-4	NED	70.1	2.7
Wag-5	NED	98.4	1.6
WAR	USA	50.0	0.0
Wc-2	GER	69.2	7.0

Wl-0	GER	99.4	0.8
Ws	RUS	98.1	1.9
Wt-3	GER	87.7	3.6
Zu-1	SUI	25.2	19.4
Ors-1	ROU	66.7	0.0
11ME1.32	USA	99.0	1.4
11PNA4.101	USA	99.4	0.6
328PNA054	USA	93.6	0.9
627ME-4Y1	USA	96.1	2.3
Alc-0	ESP	22.4	8.3
ALL1-2	FRA	97.2	2.0
ALL1-3	FRA	N.A.	N.A.
An-1	BEL	89.3	0.0
App1-16	SWE	98.3	2.4
Bå1-2	SWE	96.6	3.0
Belmonte-4-94	ITA	100.0	0.0
Bg-2	USA	95.5	2.2
Bla-1	ESP	23.6	5.2
Blh-1	CZE	50.9	3.3
Bor-1	CZE	N.A.	N.A.
Br-0	CZE	86.1	1.4
Brö1-6	SWE	39.5	0.0
Bu-0	GER	58.8	1.1
BUI	FRA	92.2	5.6
CAM-16	FRA	100.0	0.0
CAM-61	FRA	93.2	6.8
Can-0	ESP	96.9	0.0
Cen-0	FRA	N.A.	N.A.
CIBC-17	UK	22.4	24.7
CLE-6	FRA	100.0	0.0
Col-0	USA	86.1	2.8
CUR-3	FRA	N.A.	N.A.
Cvi-0	CPV	96.1	5.5
DraIV	CZE	68.0	3.3
DraIV	CZE	91.1	2.6
DraIV	CZE	55.2	2.8
DraIV	CZE	83.6	9.7

DraIV	CZE	77.8	4.8
Duk	CZE	89.6	3.1
Edi-0	UK	20.6	0.0
Est-1	RUS	94.6	5.6
Fjä1-2	SWE	69.2	5.8
Fjä1-5	SWE	N.A.	N.A.
Gd-1	GER	80.8	0.0
Ge-0	SUI	98.6	1.4
Gr-1	AUT	89.5	4.5
Gul1-2	SWE	25.0	25.0
Hi-0	NED	66.7	N.A.
Hov4-1	SWE	76.3	13.8
Hovdala-2	SWE	85.5	2.0
Hs-0	GER	98.0	1.8
HSm	CZE	94.1	3.2
In-0	AUT	55.9	0.3
JEA	FRA	96.3	3.8
Ka-0	AUT	93.3	6.7
Kelsterbach-4	GER	100.0	0.0
Nordborg	SWE	N.A.	N.A.
LAC-5	FRA	96.8	2.6
Lc-0	UK	31.5	8.0
LDV-14	FRA	83.2	4.3
LDV-25	FRA	90.9	6.9
LDV-34	FRA	100.0	0.0
LI-OF-095	USA	91.1	6.8
Lillö-1	SWE	80.0	6.7
Lis-2	SWE	92.3	0.0
Lisse	NED	64.2	7.1
Lom1-1	SWE	96.9	4.4
Löv-5	SWE	91.6	3.7
Map-42	USA	95.2	6.8
MIB-15	FRA	94.5	5.0
MIB-22	FRA	84.1	3.7
MIB-28	FRA	95.3	0.1
MIB-84	FRA	96.6	4.8
MNF-Che-2	USA	90.3	2.8

MNF-Jac-32	USA	100.0	0.0
MNF-Pot-48	USA	91.2	6.2
MNF-Pot-68	USA	71.6	1.8
MOG-37	FRA	10.6	0.0
Mrk-0	GER	89.4	3.7
Mz-0	GER	95.2	6.7
N13	RUS	94.0	3.1
Na-1	FRA	90.3	9.7
NC-6	USA	77.8	10.5
NFA-8	UK	70.0	6.9
Ör-1	SWE	95.0	5.0
Ost-0	SWE	88.4	8.4
PAR-3	FRA	95.6	6.2
PAR-4	FRA	N.A.	N.A.
PAR-5	FRA	69.7	32.4
Paw-3	USA	89.9	6.5
Pent-1	USA	N.A.	N.A.
Per-1	RUS	89.1	1.8
Petergof	RUS	82.4	0.0
PHW-34	FRA	91.9	5.3
Rak-2	CZE	N.A.	N.A.
Rev-2	SWE	97.5	2.5
ROM-1	FRA	N.A.	N.A.
Sanna-2	SWE	100.0	0.0
Sap-0	CZE	88.4	11.1
Sav-1	CZE	35.0	0.0
SLSP-30	USA	37.4	4.0
Sparta-1	SWE	97.8	0.0
St-0	SWE	97.8	3.1
Ste-3	USA	24.9	5.8
T1040	SWE	98.4	2.3
T1060	SWE	72.1	5.5
T1080	SWE	95.1	3.5
T1110	SWE	86.6	0.9
T510	SWE	92.3	7.7
T540	SWE	100.0	0.0
T620	SWE	N.A.	N.A.

Ta-0	CZE	66.7	16.8
TÅD	SWE	89.0	3.3
TDr-1	SWE	81.1	0.0
TDr-18	SWE	91.7	8.3
TDr-3	SWE	95.9	5.8
TDr-8	SWE	50.2	2.4
Tomegap-2	SWE	53.5	0.5
Tottarp-2	SWE	68.0	4.7
TOU-A1-115	FRA	91.0	2.3
TOU-A1-116	FRA	93.2	6.8
TOU-A1-12	FRA	44.4	0.0
TOU-A1-43	FRA	90.9	2.0
TOU-A1-62	FRA	50.2	27.0
TOU-A1-67	FRA	94.4	0.3
TOU-A1-96	FRA	85.7	0.0
TOU-C-3	FRA	9.2	0.8
TOU-E-11	FRA	51.3	18.0
TOU-H-12	FRA	98.1	1.9
TOU-H-13	FRA	97.2	2.8
TOU-I-17	FRA	92.9	0.0
TOU-I-2	FRA	N.A.	N.A.
TOU-I-6	FRA	28.8	8.8
TOU-J-3	FRA	60.9	4.2
TOU-K-3	FRA	97.5	2.5
Ts-1	ESP	4.5	6.3
UduI	CZE	73.9	3.5
UKID101	UK	100.0	0.0
UKID37	UK	27.3	0.0
UKID48	UK	92.2	2.5
UKID80	UK	57.7	22.9
UKNW06-059	UK	97.6	2.4
UKNW06-060	UK	N.A.	N.A.
UKNW06-386	UK	82.8	4.2
UKNW06-436	UK	28.1	0.0
UKNW06-460	UK	73.4	1.6
UKSE06-062	UK	95.8	1.1
UKSE06-192	UK	59.7	13.0

UKSE06-272	UK	82.9	2.9
UKSE06-278	UK	93.8	4.6
UKSE06-349	UK	76.2	1.2
UKSE06-351	UK	55.4	30.1
UKSE06-414	UK	96.2	0.0
UKSE06-429	UK	90.9	6.4
UKSE06-466	UK	80.4	1.6
UKSE06-482	UK	100.0	0.0
UKSE06-520	UK	72.6	3.6
UKSE06-628	UK	98.9	1.1
UKSW06-202	UK	99.2	0.8
Ull2-3	SWE	94.7	5.3
Ull3-4	SWE	93.5	4.7
VOU-1	FRA	N.A.	N.A.
VOU-2	FRA	84.6	0.0
Wil-1	LTU	93.4	2.6
ZdrI	CZE	52.1	26.2
ZdrI	CZE	85.6	4.0

Supplemental Table 2. The list of Bla-1 x Col-0 RILs and their RS on HU.

RIL No.	RS	RIL No.	RS	RIL No.	RS	RIL No.	RS
4	93.8	146	97.1	268	11.8	397	93.5
6	88.7	147	NA	270	82.9	398	83.4
7	58.6	151	24.6	271	61.0	404	86.4
13	100.0	153	100.0	274	91.3		
18	100.0	156	97.1	275	NA		
24	84.6	159	93.8	276	48.6		
30	90.7	160	88.0	278	NA		
38	90.0	163	54.6	279	NA		
39	87.3	164	88.9	286	91.1		
42	72.8	165	96.9	289	NA		
45	66.7	166	30.8	293	95.2		
47	93.4	169	22.6	295	97.2		
48	98.9	178	77.5	301	81.3		
50	94.6	182	84.8	305	90.7		
51	43.8	193	65.0	306	35.0		
52	100.0	194	38.5	308	79.4		
54	50.0	195	50.8	316	100.0		
57	87.5	196	86.1	317	69.4		
62	34.3	197	90.9	323	70.2		
64	67.5	198	85.1	324	NA		
67	89.8	199	97.2	326	75.0		
76	99.0	201	100.0	328	95.0		
79	100.0	203	30.4	329	90.7		
82	46.9	204	91.7	334	NA		
84	94.4	205	100.0	340	72.0		
90	97.6	208	44.7	343	94.1		
91	96.9	209	88.1	347	NA		
93	60.8	212	19.2	348	100.0		
98	87.8	216	91.1	350	16.3		
101	89.2	217	20.5	353	95.4		
102	70.8	218	23.4	354	29.0		
105	91.8	219	61.8	355	55.2		
107	82.3	222	73.7	358	81.3		
108	84.4	225	58.4	363	8.3		
110	100.0	228	96.3	364	41.3		
118	NA	231	51.7	365	98.9		

120	75.1	234	94.1	366	74.0
123	85.1	236	100.0	369	94.4
124	61.0	237	98.0	370	NA
126	0.0	238	93.0	371	66.1
130	94.3	245	71.1	377	86.1
131	30.7	246	20.2	383	8.0
136	97.6	251	96.8	387	NA
140	42.0	253	72.6	389	15.4
142	94.9	255	61.3	394	NA
145	83.9	257	95.8	395	60.9

Supplement Table 3. Rosette area (mm^2) in response to zebularine-, MMC- and bleocin-treatment. Significant differences (*t*-test, $P < 0.05$) relative to wild-type are indicated in bold. M = mean values of three to five replicates, SD = standard deviation of the replicate means.

	mock		zebularine						MMC						bleocin					
			10 µM		20 µM		40 µM		10 µM		15 µM		20 µM		25 nM		50 nM		100 nM	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
WT	18.6	2.6	15.1	3.7	14.1	1.5	12.6	2.2	18.8	4.2	15.1	4.1	13.1	3.9	20.4	2.9	17.4	2.7	13.7	1.9
atr	17.1	2.3	10.9	3.6	**10.1**	**3.0**	**9.0**	**1.3**	**6.5**	**2.2**	**5.5**	**1.3**	**4.6**	**0.8**	**17.5**	**2.0**	17.3	2.1	12.8	2.3
atm	18.2	3.8	13.3	3.7	12.4	2.8	11.4	3.9	16.0	4.8	11.6	3.2	11.7	3.0	18.3	3.0	16.5	3.5	13.4	4.8
atm/atr	16.9	2.1	10.3	2.1	**8.0**	**1.5**	**7.2**	**2.3**	**7.8**	**3.4**	**6.5**	**2.7**	**5.2**	**1.0**	17.6	2.8	16.9	1.5	12.6	1.2
rad1	17.0	5.0	12.1	1.9	**10.5**	**1.6**	**8.8**	**1.2**	**4.7**	**0.5**	**3.9**	**0.9**	**3.9**	**0.7**	17.0	3.3	16.2	2.4	12.8	2.2
brca1	17.8	1.0	12.4	2.4	13.9	0.1	11.8	2.0	15.8	4.3	13.4	2.8	12.6	3.1	18.6	4.2	18.0	1.2	12.2	3.3
rad51B	16.6	3.8	12.8	4.4	13.8	0.8	9.9	1.2	16.7	3.8	12.2	3.8	12.0	3.7	19.6	0.3	17.5	1.6	11.4	2.9
mim1	19.7	3.8	**7.4**	**2.9**	**4.8**	**1.4**	**4.6**	**0.2**	18.1	4.7	14.4	3.0	14.2	4.3	21.3	2.2	19.0	2.8	15.6	2.5
fas1	19.5	3.5	10.5	2.0	**10.6**	**1.8**	**8.7**	**0.9**	14.9	4.4	10.1	2.9	9.6	1.9	19.9	1.5	**14.0**	**1.5**	12.5	1.9
ku70	18.6	3.4	11.6	4.1	12.7	3.1	**8.6**	**1.1**	14.7	5.9	12.8	4.3	10.3	4.4	**12.5**	**0.8**	**10.1**	**4.3**	**5.8**	**2.9**
lig4	17.6	4.7	14.2	2.9	13.9	2.8	11.5	1.5	16.3	2.6	13.3	4.3	12.3	4.3	**15.3**	**1.6**	**10.8**	**1.7**	**8.3**	**0.8**

Supplement Table 4. Mean CV of nuclei isolated from cotyledons of 15-day-old seedlings in response to MMC, zebularine (zeb), and bleocin (Ble) treatment. Treatment-specific significant changes (*t*-test, $P < 0.05$) relative to WT are indicated in bold. Std = standard deviation.

	Mock		10 µM MMC		10 µM zeb		50nM Ble	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
WT	1.354	0.08	**1.492**	**0.08**	**1.496**	**0.15**	1.324	0.19
atr	1.284	0.13	1.586	0.12	1.348	0.13	1.317	0.14
atm	1.311	0.10	1.403	0.11	1.549	0.19	**1.815**	**0.46**
atm/atr	1.354	0.13	**1.774**	**0.09**	**1.770**	**0.11**	**1.680**	**0.13**
rad1	1.334	0.12	**2.240**	**0.10**	1.554	0.12	1.466	0.08
rpa2	1.402	0.17	**1.686**	**0.06**	1.349	0.19	1.424	0.16
brca1	1.278	0.17	1.503	0.10	1.462	0.22	1.224	0.16
rad51B	1.254	0.12	1.442	0.14	1.418	0.20	1.268	0.14
mim1	1.332	0.12	1.587	0.08	1.570	0.16	1.216	0.32
fas1	**1.537**	**0.17**	**1.782**	**0.07**	**1.853**	**0.21**	1.487	0.14
ku70	1.276	0.18	1.584	0.12	1.518	0.19	1.317	0.15
lig4	1.329	0.12	1.455	0.07	1.440	0.14	1.372	0.21

Supplement Table 5. List of T-DNA insertional mutants used in this study.

Mutant line	AGI code	NASC ID	Name	Background
atm-2	At3g48193	N506953	SALK_006953	Col
atm-2 x atr-2	Kindly provided by Kevin Culligan			Col
atr-2	At5g40820	N669551	SALK_032841C	Col
brca1	At4g21070	N24951	SALK_014731	Col
ddi 1-1	At3g02400	N800701	SAIL_15_A10	Col
ddi 1-2	At3g02401	N595830	SALK_095830	Col
ddi 2-1	At1g09815	N676545	SALK_112640C	Col
ddi 2-2	At1g09816	N612638	SALK_112638	Col
ddi 3-1	At5g64060	N829349	SAIL_671_G04	Col
ddi 3-2	At5g64061	N610288	SALK_110288	Col
ddi 4-1	At2g46180	N675486	SALK_062053C	Col
ddi 4-2	At2g46181	N658097	SALK_014489C	Col
ddi 5-1	At3g42860	N500388	SALK_000388	Col
ddi 5-2	At3g42861	N524080	SALK_024080	Col
ddi 6-1	At1g68200	N681117	SALK_045897C	Col
ddi 6-2	At1g68201	N565040	SALK_065040	Col
ddi 6-3	At1g68202	N746757	GK-848H09	Col
ddi 7-1	At1g05490	N552751	SALK_052751	Col
ddi 7-4	At1g05493		GK-339H02	Col
ddi 8-1	At3g48770	N671308	SALK_015383C	Col
ddi 8-2	At3g48771	N672441	SALK_027476C	Col
ddi 9-1	At3g07810	N666782	SALK_094167C	Col
ddi 9-2	At3g07811	N507316	SALK_007316	Col
fas1	At1g65470	N828822	SAIL_662_D10	Col
ku70	At1g16970	N656936	SALK_123114C	Col
lig4	At5g57160	N656431	SALK_044027C	Col
nvh11	AT3G03620	N654360	SALK_113658C	Col
nvh13	AT5G54040	N523894	SALK_023894 (BA)	Col
nvh14	AT5G54050	N681316	SALK_075380C	Col
nvh18-1	AT5G63610	N674891	SALK_046407C	Col
nvh18-2	AT5G63610	N617306	SALK_117306 (BU)	Col
nvh19	AT5G63620	N662924	SALK_072101C	Col
nvh2	AT1G53160	N677087	SALK_137581C	Col
nvh3	AT1G70210	N675047	SALK_050467C	Col
nvh4	AT1G76540	N436323	GK-379C11	Col

			SALK_028948 (M)	
nvh6	AT2G01720	N528948	(AE)	Col
rad51B	At2g28560	N659269	SALK_024755C	Col
rpa2-4	At2g24490	N663677	SALK_111834	Col
smc6b	At5g61460	N601968	SALK_101968	Col
xpf	At5g41150	N682684	SALK_096156C	Col

Supplemental Table 6. CAPS markers used in this study.

Chr.	Position (bp)	Ecotype with restriction site	Enzyme	Fragments with Col-0 allele	Fragments with Bla-1 allele	Sequence(5´-3´) Left primer	Right primer
1	32210	Col-0	EcoRI	462+386	848	GTCAGTGTAGTGCTTATATTCAG	TTGTCAAGAAGATCCGTAGCTG
1	2148724	Col-0	EcoRV	394+461	855	tccaataatcaaccgagtcgtc	ACACCATTATTTTACGGCATCC
1	2708672	Bla-1	XbaI	710	284+426	cagATTAAATGTGGGGATCTGT	GTTCAGGACCACATGTGTTATC
1	3720074	Col-0	XbaI	477+297	774	cgtttgagaaccctcttgtttc	gttctacagccttacttaactatc
1	5093558	Bla-1	EcoRI	710	323+387	AAGAACTACGGTTGGGGAAAAG	ctttgtatgtcgcaattttggag
2	11897080	Bla-1	AhdI	828	463+365	ggattcattcgatttgagggaaac	gtcaccttttcgtatgctaac
2	13103424	Col-0	XmnI	206+501	707	ggtagtgatctgaagcattaac	GACAAGGGAGGAACTTCGATG
2	14277341	Bla-1	HpaI	718	227+491	gataacaactatggagctttgg	gcatgaacagatcacgaaattg
2	16620148	Bla-1	EcoRI	808	311+497	gacccgaattaaacaatccgac	gaaggatacAATGGCAAACCAC
2	18249608	Bla-1	EcoRI	720	285+435	CATGGACAGGAATGACTACTAC	GTCGTATTCTTGAACATAGCTG
3	202256	Col-0	AflII	797	471+326	gaaattagtggtccacgtgtc	GTCAATGACGAAGACGTTGAC
3	885721	Col-0	EcoRV	688	249+439	gttacCTGCAACTCTTTGTCC	CAAACTCACGTTATCTCGGTG
3	1459219	Col-0	DdeI	710	230+480	GACATAATGACCCAAGGTAGC	cagGCTTTTCAGAGAGACTTG
3	2081316	Col-0	EcoRV	727	231+496	TTTGGTCACTTTTCCCTGTTTC	CAAGCTGGAGCCTTTAAAATTG
3	2300524	Bla-1	PvuII	794	365+429	cagcgacaacaaacctgttatg	TCAAGTTTCTAACCGTGGTGTG
3	2456212	Col-0	BsrGI	807	495+312	GAAGAATGTTTCTCCCACAAGG	tgcacaaggaacaaccatattc
3	2650296	Bla-1	EcoRI	797	305+492	CTCAAAGTCCTCAATGCCTAC	taaaggatattcggtcggtgtc

3	2806106	Col-0	BsrI	806	322+484	ctactccaaagtcatctttggc	GATGTTCATATTCCCGAAACTC
3	2807822	Col-0	FokI	603	258+317+28	TGATAGTGGAGGTACAAACTAC	ACAATGGTTCGGATGGTTAGATC
3	2821813	Col-0	BsmI	1017	504+513	GTTCACGAGATGACTGAACAGA	GATTACAGTTCCTCTGCTTTAG
3	2857679	Bla-1	BsmAI	563	301+262	GACGTTTGAACCGCTAAGATTC	GTCAATAATGGAAGACCAGCG
3	2902656	Col-0	RsaI	786	405+381	tgaagGTTGGAGGAAAATATCG	ttcagcaaccaagttctattcg
3	3000025	Col-0	AciI	702	479+223	ctaaccactcgcacacactttc	ttttgggatcttacacaaagtacac
3	3025991	Col-0	BamHI	711	318+393	CTGTCCTGGAGATTTAGCCTTG	AAACCCAATTCGGAAATCAAAC
3	3043046	Col-0	BsaAI	757	331+426	CCGAGTAACCATGTCCAATC	GGGAAGTACATATGGTCAGATG
3	3087462	Bla-1	BsmAI	340+498	340+159+338	CTTGAAGGTCAAAGTCGAATCG	CTGTTATTGATCATACAGGAGG
3	3104907	Col-0	AciI	908	450+458	GCAAGTCGCTCTTTGTGAATATC	CAACAGGTATCCACTGACCTC
3	3131011	Col-0	ScrFI	795	345+450	CCTGAAACTGATTAGCCCATG	AGTGCTCGTGAAGCATCTAAC
3	3131769	Col-0	BstYI	794	329+465	gttagatgcttcacgagcact	GACAACTCCAGTGATAACTAGC
3	3134037	Bla-1	AflIII	615	269+346	gaacacaatattcgggatacgac	gatcctgcttgcataccaaaatc
3	3138936	Bla-1	BstYI	844	427+417	GATTGCTATGTTGTTAGCCGTC	GAATCGTATGAGCTGGACACTG
3	3147392	Col-0	PvuII	882	445+437	ctgtgagctgtaaaccaaacC	tcagaatgaactccctgtagc
3	3149640	Col-0	Sau96I	738	288+450	GGTGTTCTGAGAGCATCTAAC	CCAGTATGCCATTGCCTAATTG
3	3150841	Col-0	Sau96I	720	253+468	CATGAAGCACAGAGGCATTTC	GTGACAAGCAGAAAGAAACCATC
3	3154290	Bla-1	HaeIII	759	406+353	GAGCCAGCAAATCGTTAATATC	GAATCCATGTCTCAGTCAGAG
3	3155826	Col-0	Hpy188I	675	409+266	ATCTTCCGGCCAAGACTGTAA	GTTAGAACTACAGCTTATTCAGC
3	3158573	Bla-1	Tth111I	707	326+381	GTATGGAGAAGTAATGTGAACC	ACTAGTGCAATGTGCTGGAAG
3	3158723	Col-0	BsrI	762	368+394	AGCTGCAGAATACATCCAGAG	GTCCACTGATTTGCACACAATC
3	3161071	Bla-1	KflI	750	306+444	CGCCAGAGTTTATTAGATCATG	GAAGCGATGACGTTGATGAAC

3	3170323	Col-0	Hpy188I	519 (417)	140+379	AGCCTGCTCTAGTAAAGTATCC	GAGGTTTAGTTTGGGGTAAAATAG
3	3201803	Bla-1	BanI	750	290+460	GATATATCAGTGGCAACTGGAG	GGTAAACTTACACCAATGATGCG
3	3190295	Col-0	BstBI	147+887	147+322+565	CGAATCGGACGGATAAGATTAG	CACTCATGGTATGCCTTAAGTC
3	3212725	Bla-1	Fnu4HI	317+352	218+99+352	TGCAGCCGTACATTTGTAGAG	CTGGGTAACCTCAATGCTGCA
3	3249408	Col-0	TatI	680	211+469	CGGATTGGACGGTTTAATTCG	CCAGAGTAGAAGAGGATCTTC
3	3281850	Bla-1	PvuII	800	413+387	cgaaattaatatgcgcggttgc	agttacgtccattttcccttcc
3	3321269	Bla-1	BsrFI	777	261+516	GTTCACTCTGCTGAGATTCTC	TGCTAAGGTTGTAGCGGATC
3	3338013	Col-0	BsrFI	826	405+421	aagcaatcctaaacccataccg	CTCTTTTACCCTCAACCACTG
3	3357793	Col-0	BsrFI	903	542+361	gaatttgccgtaacggtaagg	tggaccataacttgttagtgtg
3	3448809	Bla-1	Sau96I	775	555+212	gcaagttggtcctacttagag	TGTCTGACCTCGATCTAGTTC
3	3449039	Col-0	DraI	646	313+333	GCTGCTAGTGAAGGAAAACTG	CATGTGACTCAGAGCTGATGA
3	3449868	Col-0	BsrFI	741	345+396	GTGAAGATGTCAAAGTGAATGC	GCCGAAGTATATCCCTTCTTG
3	3496923	Col-0	BsmFI	726	258+468	gcaagctgactaatttcgactc	gatcaattccgtaattcagtgc
3	3578604	Col-0	BstYI	703	458+245	ttggtagacaaaacgccaaaac	CTTCGGTAGAAACCATAGTGAG
3	3580021	Bla-1	AciI	651 (739)	372+279	aaacgggcaaagtttaagttcc	ATTGCTCTTTGAGCGTCTTCTC
3	3665352	Col-0	BsrI	779	354+425	ccttaaaccgcttttgaccttc	gactaaaatcccttacaaatacg
3	3702464	Col-0	BsrI	744	378+366	ctcgtctcccattcatataccc	CAATGCCTACACGGAGCTAAAC
3	3793485	Bla-1	BsrGI	774	280+494	caacttaaaggacggcagtttc	agtaccttggcctgagattttg
3	3826211	Bla-1	HpaI	746	246+500	caaaatcctttgccatcaactgc	cctaaacttacgcagtaggac
3	4109964	Col-0	BamHI	807	480+327	AAGCTTCCAGGGTTTAGGTTTC	ATATCCGGCATAACGTCCATAC
3	4524217	Bla-1	XbaI	721 (750)	389+332	GAACCTCCCCTAAATTCTCAGC	GTAAGCAATCCCTCAAAACGAC
3	4707008	Bla-1	DraI	736	299+437	gtgcggaaggtaaaattgtgag	GTTCTTTGTGAAGACTCAATTGG

3	4951046	Col-0	HpaI	717	492+225	TTGTAATGGTATGCCTGCTCAC	TCGCAGCTATAAAGCCTTCTTC
3	4954181	Bla-1	XbaI	877 (878)	397+480	aaagggaggctgctttaatttc	GCTTCGAGCAGAGATTGCTAAC
3	5567210	Bla-1	HpaI	711	55+283+373	TGAGATTGATGCGGATGACTAC	AACAAGAGAGGTGTGGAACGTC
3	5512488	Bla-1	AvaII	740	439+301	GTACCTTAGGCCAGCTTAAAC	CTCAAACGTCGCctgttataag
3	5527431	Col-0	FokI	804	360+444	CATTTCTTGGACCAGGAACATG	CTCCAGATGCTAAGTCTGACA
3	6174790	Col-0	XbaI	835	451+384	GAAAGAGAGAAGCACGCCATAC	AACATGAAGCAACCAAATCCAC
3	6399920	Bla-1	EcoRV	713 (719)	256+457	AACTTGGTGGTGGTAATTAGTCA	GGAAAGGCATTGCTTGTCTTG
3	6747743	Col-0	XbaI	794	496+298	GACACCAATTGAAACTGAGAACC	GACAAGAGATTGTATCCAGCTC
3	6932394	Bla-1	XbaI	705	426+279	AATCTTTGTCGGGACTCATACG	TGCTCTATGCTTTGCCACTAAC
3	7838291	Bla-1	HindIII	666	265+401	CTTATTCTGGTCATCCAGTTTC	CTTCAGGTGAGATAATGAGCG
3	8667633	Col-0	EcoRV	748	308+439	TAGGGGCAAATTCAAAGTAACG	TGTCCATCACGACTCATTTCTC
3	8876471	Col-0	AhdI	828	349+479	ccgaatccacgttagagaaatc	TCAGCCCTCAGATAGAGAGGA
3	9191017	Bla-1	XbaI	742	470+272	GTCATTTCTGCCCTGCACTTAC	GGCATACTCAGGGTCTATGGTC
3	9469186	Col-0	BsrGI	776	375+401	gtacccaaaacgactttagcac	gttggcagacgaaattggaca
5	17889843	Bla-1	PvuI	779	487+292	CATTCCTGAAGAAATTGGTTGG	AAGGAACCTACCGTTGACACAG
5	19102124	Col-0	HpaI	731	300+431	TTCATAAATTTTGCAGGTTGGAC	GAAAGGGGAATAGAAGGGTGAG
5	19672640	Col-0	HpaI	707	490+217	CTTTTGAGAATGCCGTCGTATG	CAAATGGTCAATCAACACAACC
5	20316047	Col-0	EcoRV	722	346+376	CACCAAGAACCTTCCTCTGAAC	CCTGGATAGAAGATTGGCAAAG
5	20762828	Bla-1	PvuI	744	493+251	GATGTGAGCAAGAATGTCAACG	ATGAGATACACCTCTGGGATCG
5	20837531	Bla-1	HpaI	750	455+295	GGGAATGGGAGTGACCATAAC	ATCCCGTCATGTGCCTAAATAC
5	20906855	Bla-1	EcoRV	739	492+247	ATCTACCCAAACATGAGGATGG	CTCTGTTGCCATTATTGGACTG
5	21210139	Col-0	EcoRV	754	472+282	CGACTAGGCATATCGAACACC	TGTTGTCGTCAATCGCAATATC
5	21719412	Col-0	HpaI	709	324+385	TTGGGAGTGAAAGGTTTACGAG	AGGCAGACTCAACTTCTTAGGC

5	22003686	Bla-1	EcoRV	739	385+354	ACACCGGAACATACATGAAAGG	TGCTTAAAGGCTCTTGTTTGTG
5	22593937	Col-0	PvuII	807	433+374	ATATCAAGATGGAATCCGAAGG	GGAACTTTCCTTGACACTGGAG
5	22990339	Col-0	BamHI	715	368+347	AAAGTCGTACACGCAAGAAGATAAC	TGAACCTCAAAGAAGGAGCATC
5	23048549	Col-0	EcoRV	714	376+338	TGATAGCAGATCTCCAAGAGACG	AACAATTCCAGAACTCAGAGCAG
5	23955563	Bla-1	PvuII	753	357+396	ctacgaaaatcatagtattggtttg	gacaatttacactaaagccacgc
5	24223366	Bla-1	XbaI	811	311+500	ccagtcggaatgttcatttgac	TATCTTCGCCTTGTTTGTAGCC
5	25216306	Col-0	HpaI	723	306+417	CTTGGTGTCGGAGGAATAAAAG	AGGGAGGAGGAGGTGGTTATAG
5	25710147	Col-0	XbaI	771	416+355	caccgacgaacgaaagagac	GCTTCATCCACGAATCCATAAAC
5	26239227	Col-0	PvuI	721	502+266	GTATCCGTGACTCGTGAACTTG	AAATCGAAGTCCTGCAAGAGAG
5	26662774	Col-0	AflII	797	348+449	ccaatttaatgcaatgttcgtctg	ataaagcatggttggactcagg
5	26654966	Bla-1	EcoRV	745	448+297	GAGAGTTGTCCCTGTGTCTGTG	GTTTCTTTTCGCCAATTTCAAG

Supplemental Table 7. Genotyping primers used in this study.

Locus	Name	5' to 3'
At3g48190	atm_2_F	GCAGGTTCTGAAGTGTTATCC
	atm_2_R	GTTCCAGTTGTTTCACACTACG
At5g40820	atr_F	CAGAATCATAGAGACTCCAGAG
	atr_R	GTCAGCCTGACAGATAGATTATC
At4g21070	brca1_F	GGATAGCTTGGTGAGCATTTAC
	brca1_R	CCTGACGATTCTTATCACTTGC
At2g28560	rad51B_F	GTGCATAGTCCTCAGAGAATG
	rad51B_R	CAACTGAGGTTGTCTCTGAG
At1g16970	ku70_F	CTTCTTCAGTTGGTCCTTCATG
	ku70_R	GCTAAGGTAGTTTAGCTTCAAC
At5g57160	lig4_F	CATCAAGGATACACCTATATGG
	lig4_R	CATTTCCTCGAGATACAATCGA
At5g41150	rad_F	CAGAGATACGGCATTGAAACTG
	rad_R	GGTTTCGATGTCGAATTACATG
	ddi1-1_F	GCTAAGAGGGTAGAGCAGGT
At3g02400	ddi1-1_R	CGCTTTGTGCACAATGTAGATC
	ddi1-2_F	GACACTATCGACTCAGATACG
	ddi1-2_R	TCACTGTTCTTACCTCTCCGA
At1g09815	ddiR2_F	GTGTTCATGAGTACGTGTAGTG
	ddi2_R	CATCGATCAAGTCTCGTCATAC
At5g64060	ddi3_F	CTACAGGTAAGACATCCTGATG
	ddi3_R	CATCCTCGTACTATGTTTGCG
	ddi4-1_F	TAAGGCACGGCAAGAACTAAAG
At2g46180	ddi4-1_R	CTTTAGCCACTGCAAGCTCTC
	ddi4-2_F	CAGAGGATCTTAGTAGACACC
	ddi4-2_R	GGTTTCCTGTTAGAATATAGCAAG
At3g42860	ddi5-1_F	GGTAATCCCAAGTATGAACCAG
	ddi5-1_R	GCAGAAGTTATCGGAGAGAAATG
	ddi6-1_F	CTGCGAATCCAAAGCCTGTAT
At1g68200	ddi6-1_R	GATCTCACAGAGATGCTCTTTG
	ddi6-2_F	CATTGACGCACTATAGAACTGG
	ddi6-2_R	GGTCATCTAAGCAGCTTGAAG
	ddi7-1_F	CAATGTTGTGTACACTTGTGCTC
At1g05490	ddi7-1_R	GAGTAGGATCAGATTCATATCCA
	ddi7-4_F	CACAGAAGAGCTCGGAACTG
	ddi7-4_R	CAGCAAGAGGAAGGTAGTTG

	ddi8-1_F	GGAGTCATACACGATACATCTC
At3g48770	ddi8-1_R	CTCAGCGATGAAGAGTAGAAC
	ddi8-2_F	GTTTGGGAGTTACAGTTGAGC
	ddi8-2_R	GTCGCTTATGCTAGATAACATG
	ddi9-1_F	GTTCACGGGAGGTACAAATTAC
At3g07810	ddi9-1_R	ACTAGTGATACGAAGGTGTTCT
	ddi9-2_F	GTAGTTGCACACCTTCATGTC
	ddi9-2_R	CAGATGCCGAATCTCGTAGA
At5g28490	recA_F	GAGAAACTTAGCTGGTTGTGATGAT
	recA_R	CTTTAACCATTGGTCACACTCTCTT
At2g27510	recB_F	TGCTGGTGAACACGTAAAGC
	recB_R	CAGTCGGATGGTTCGTTTCT
At5g25050	IC9C_FW	ACTGAGTTTGGTAACCTGTG
	IC9C_Rev	CCAGAAGTAAAAGTGAAGTTC
For SALK lines	LBb1.3	ATTTTGCCGATTTCGGAAC
	LB1	GCCTTTTCAGAAATGGATAAATAGC
For SAIL lines	LB2	CTTCCTATTATATCTTCCCAAATTAC
	LB3	GCATCTGAATTTCATAACCAATCTC
For GABI lines	o8474_m	ATAATAACGCTGCGGACATCTAC
For GUS	GUS_F	ACCCGACGCGTCCGATCACCT
	GUS_R	ATTGAGTGCAGCCCGGCTAACG

Supplemental Table 8. qPCR primers used in this study.

Gene	AGI ID	Name	5' to 3'
ACTIN7	AT5G09810	Act7qF	TGGATCGGAGGATCAATCCTTG
		Act7qR	GACTCATCGTACTCACTCTTTGAA
BRCA1	At4g21070	BRCA1qF	GTTACGTGTGCAAAACTCATACCAGAATG
		BRCA1qR	GATACTTGTTTAGGCTGAGAGTGCAGTGG
GMI1	AT5G24280	GMI1qF	AGCTAGCCTCGGACGATCTATCA
		GMI1qR	TATCATGTTCACAGCGTAGCCTTTGA
PARP2	At4g02390	PARP2qF	CGAACTATTGCTATGCCAACACT
		PARP2qR	CCACACCTTTTGTGCTTAGCT
RAD51	At5g20850	AtRAD51fwd	CTC CGA GGA AGG ATC TCT TGC AG
		AtRAD51rev	GCT CGC ACT AGT GAA CCC CAG AGG
RNR1	At2g21790	RNR1qF	CTGATATGGGACTTTGGACTCCAAC
		RNR1qR	GATCTATGTAGCATCCACGATCAGC
SDC	At2g17690	SDCqF	GAAGGTCTTACTTACGCTGTGGATAC
		SDCqR	CTACAATGTAAAGTTCTCCACGACATC
TSO2	At3g27060	TSO2qF	GAACGAATCATCGCTTTCGCTTGC
		TSO2qR	CGTCACGTGAGATCAATTCGTTTGAG

ACKNOWLEDGMENTS

First, I would like to express my gratitude to my supervisor, Dr. Ales Pecinka for giving me the opportunities to pursue my PhD degree in such great search environment and to participate in these three projects. He has been always open for discussion and encourages me to go on.

I would also like to show my gratitude to Prof. Dr. Maarten Koornneef for giving me the opportunities to do my PhD in his department, his supports for all my projects, and co-supervising me during my PhD time. He is always patient and caring to show his concerns for students.

I would like to thank the other members for my thesis committee for their positive replies and for their time. I would like to thank Prof. Dr. Björn Schumacher for his evaluation of this PhD thesis. I would like to thank Prof. Dr. Mirka Uhlirova for the chair and Dr. Wim Soppe as assessor.

I would like to show my appreciation to the coworkers in MPIZ. I would like to thank Dr. Pecinka for the RNA sequencing, Dr. Finke for bi-sulfite sequencing in the zebularine project, and Dr. Piofczyk for the German translation for abstract and CV. Barbara Eilts, Petra Pecinkova, and Regina Gentges for their technical support in countless experiments and taking care of the plants in the greenhouse. Björn Pietzenuk for a lot of German translations, and other members in the Pecinka group for their help and discussion in everyday-lab life.

Many thanks to the those who become my friends in the past few years, Ganga, Maggie, Mariana, and Ting-Ting, thank you for adding more good memories in my life in Cologne.

I am also grateful to my family and friends for their support during my time in Germany. Many thanks to Pei-An, Mu-An, Sheila, and Pinky, especially for their

support in the past few months, without you, I wouldn't be able to finish writing. Thank you S for everything you've done for me that sometimes I almost forgot that I am a stranger here.

Thank You Lord for choosing me before the world has begun and Your predestination for the sonship. Thank You for bringing me to Cologne, keeping me in the daily enjoyment of truth and life, and adding more faith in me. Thank You for letting me know Your sovereignty and showing me Your all-sufficiency during my PhD time. Surely goodness and lovingkindness will follow me all the days of my life, and I will dwell in the house of Jehovah for the length of my days (Psalms 23:6).

LEBENSLAUF

PERSÖNLICHE INFORMATIONEN

Name: Chun-Hsin Liu

Adresse: Herkulesstr. 76, 50823 Köln

Geburtsdatum: 20. September 1982

Geburtsort: Taipei, Taiwan

Staatsangehörigkeit: Taiwanesisch

AUSBILDUNG

seit Okt 2011	**Promotion**	International Max Planck Research School (IMPRS), Max Planck Institute for Plant Breeding Research

Titel der Abschlussarbeit: The study of DNA damage repair in *Arabidopsis thaliana*.

Sep 2008 - Jul 2010	**M.A** **Agronomy**	National Taiwan University

Titel der Abschlussarbeit: Studies on the relationship between potassium and cadmium stress of rice seedlings

Sep 2001 - Jun 2006	**B.A** **Agronomy**	National Taiwan University

Titel der Abschlussarbeit: The effects of polyamines on the oxidative damage of detached rice leaves caused by excess copper

WISSENSCHAFTLICHE VERÖFFENTLICHUNGEN

1. Chun-Hsin Liu, Wen-Dar Huang, and Ching-Huei Kao (2012) "The decline in potassium concentration is associated with cadmium toxicity of rice seedlings." Acta Physiologiae Plantarum. 34:495-502.

2. Chun-Hsin Liu, Yun-Yang Chao, and Ching-Huei Kao (2012) "Abscisic Acid Is an Inducer of Hydrogen Peroxide Production in Leaves of Rice Seedlings Grown under Potassium Deficiency." Botanical Studies. 53:229-237.

3. Chun-Hsin Liu, Yun-Yang Chao, and Ching-Huei Kao (2013) "Effect of potassium deficiency on antioxidant status and cadmium toxicity in rice seedlings." Botanical Studies. 54:1-10.

4. Ales Pecinka, Chun-Hsin Liu (2014) "Drugs for Plant Chromosome and Chromatin Research." Cytogenetic and Genome Research. 143:51-59.

ERKLÄRUNG

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit −einschließlich Tabellen, Karten und Abbildungen −, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie abgesehen von unten angegebenen Teilpublikationen noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Maarten Koornneef betreut worden.

Köln, August 2014

Liu, Chun-Hsin