Functional characterization of NSE4 paralogs in *Arabidopsis thaliana*

I n a u g u r a l - D i s s e r t a t i o n

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Mariana Andrea Diaz Smoje

aus Santiago, Chile

Köln, 2018

Diese Arbeit wurde am Max-Planck-Institute für Pflanzensüchtungsforschung in Köln in der Arbeitsgruppe von Dr. habil. Ales Pecinka in der Abteilung für Pflanzenzüchtungs und Genetik von Prof. Dr. Maarten Koornneef durchgeführt.

Prüfungsvorsitzender: Prof. Dr. Maria Albani

Berichtstatter: Prof. Dr. George Coupland

Prof. Dr. Björn Schumacher

Tag der mündlichen Prüfung: 02.03.18.

ABSTRACT

The control of genome integrity throughout cellular generations is critical for viable plant development and the correct transmission of genetic information to the progeny. Key factor involved in this process is the STRUCTURAL MAINTAINANCE OF CHROMOSOME 5/6 (SMC5/6) complex, related to cohesin and condensin complexes controlling sister chromatid alignment and chromosome condensation, respectively. In my Ph.D project I characterized *NON-SMC ELEMENT 4* (*NSE4*) paralogs of SMC5/6 complex in Arabidopsis. *NSE4A* is widely expressed in meristems and accumulates during DNA damage repair. Partial loss of function *NSE4A* mutants are viable, but hypersensitive to DNA damage induced by zebularine. In addition, *NSE4A* mutants show aberrant seed development, with embryo failing at organogenesis and non-cellularized endosperm. This resembles defects in cohesin and condensin mutants and thus suggests a role for all three SMC complexes in differentiation during seed development in plants. In contrast, *NSE4B* is expressed only in a few cell types and loss-of-function mutants do not have any obvious phenotype. In summary, this study shows that NSE4A subunit of SMC5/6 complex is essential for DNA damage repair in somatic tissues and plays a role during reproductive development.

ZUSAMMENFASSUNG

Die Kontrolle der Genomintegrität über die gesamte Zellgeneration hinweg ist entscheidend für die Entwicklung einer lebensfähigen Pflanze und die korrekte Übertragung der genetischen Information an die Nachkommenschaft. Schlüsselfaktor in diesem Prozess ist der STRUCTURAL MAINTENANCE OF CHROMOSOME 5/6 (SMC5/6)–Komplex. Der SMC5/6 –Komplex steuert, in Verbindung mit den Cohesin- und Condensin-Komplexen, die Schwesterchromatidausrichtung bzw. die Chromosomenkondensation. In meiner Doktorarbeit habe ich die paralogenen Proteine der Untereinheit NON-SMC ELEMENT 4 (NSE4) des SMC5/6 - Komplexes in *Arabidopsis thaliana* charakterisiert. NSE4A wird häufig in Meristemen exprimiert und akkumuliert dort während der Reparatur von DNA-Schäden. NSE4A-Mutanten mit partiellem Funktionsverlust sind lebensfähig, aber hypersensitiv gegenüber Zebularin-induzierten DNA-Schäden. Darüber hinaus weisen NSE4A-Mutanten eine gestörte Samenentwicklung bezüglich der Organogenese im Embryo, sowie eine ausbleibende Zellularisierung des Endosperms auf. Dies ähnelt Defekten in Cohesin- und Condensin-Mutanten und legt somit eine Rolle für alle drei SMC-Komplexe bei der Differenzierung während der Samenentwicklung in Pflanzen nahe. Im Gegensatz dazu wird NSE4B nur in wenigen Zelltypen exprimiert, und *loss-of-function*-Mutanten haben keinen offensichtlichen Phänotyp. Zusammenfassend zeigt diese Studie, dass die NSE4A-Untereinheit des SMC5 /6-Komplexes essentiell für die Reparatur von DNA-Schäden in somatischen Geweben ist und eine Rolle während der reproduktiven Entwicklung spielt.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

A. thaliana	*Arabidopsis thaliana*
3-AT	3-Amino-1,2,4-triazole
4a-1	*nse4a-1*
4a-1 swap	*nse4a-1 pNSE4A::gNSE4B*
4a-2	*nse4a-2*
4a-2 4b-2	*nse4a-2 nse4b-2*
4a-2 com4A	*nse4a-2 pNSE4A::gNSE4A*
4a-2 swap	*nse4a-2 pNSE4A::gNSE4B*
4A-VENUS	*nse4a-2 pNSE4A::VENUS::NSE4A::tNSE4A*
ABA	Absicic acid
AD	Activation domain
AGL	AGAMOUS-LIKE
AP	Apurinic/apyrimidinic
ARM	Armadillo
ARR	ARABIDOPSIS RESPONSE REGULATORS
ASAP1	ARABIDOPSIS SNI ASSOCIATED PROTEIN 1
ATM	ATAXIA TELANGIECTASIA MUTATED
ATP	Adenosine triphosphate
ATR	ATAXIA TELANGIECTASIA MUTATED AND RAD3-RELATED
BC	Bent cotyledon
BD	Binding domain
BER	Base excision repair
BIFC	Bi-molecular fluorescence complementation
BIR	Break-induced replication
BRCA1	BREAST CANCER 1
cDNA	Complementary DNA
ce	Chalazal endosperm

CHK2	CHECK POINT KINASE2
Col-0	Columbia-0
CSA	*COCKAYNE SYNDROME, TYPE A*
CSB	*COCKAYNE SYNDROME, TYPE B*
CYCB1;1	MITOTIC CYCLIN-DEPENDENT KINASE CYCB1;1
DAG	Days after germination
DAP	Days after pollination
DAPI	4',6-diamidino-2-phenylindole
DDR	DNA damage repair
DIC	Diferencial contrast
DNA	Deoxyribonucleic acid
DO	Dropout supplement
DPC	DNA-protein cross-link
DR	Differentiated root
DSB	Double strand break
dsDNA	Double stranded DNA
e	Embryo
EMS	Ethyl methanesulfonate
ERCC1	EXCISION REPAIR COMPLEMENTING DEFECTIVE REPAIR IN CHINESE HAMSTER 1
FEN1	5-FLAP ENDONUCLEASE
FLC	FLOWERING LOCUS C
FM	Functional megaspore
FT	FLOWERING TIME
G1 phase	Gap 1 phase
G2 phase	Gap 2 phase
GGR	Global genome repair
GI	Globular
GUS	β-glucuronidase
H	Heart
H3K27me3	Histone 3 lysine 27 tri-methylation

HAP	Hours after pollination
HJ	Holliday junction
HP1	HETEROCHROMATIN PROTEIN 1
HPY2	High ploidy 2
HR	Homologous recombination
ICLs	Inter-strand cross-links
LHP1	LIKE-HETEROCHROMATIN PROTEIN 1
LIG1	DNA LIGASE 1
LIG3	DNA LIGASE III
LIG4	LIGASE 4
LR	Lateral root
LW	Leucine minus/Tryptophan minus
LWH	Leucine minus/Tryptophan minus/Histidine minus
MAGE	MELANOME ANTIGEN
MMC	Mitomycin C
MMC	Megaspore mother cell
MMS	Methyl methane sulfonate
MS	Murashige and Skoog
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
NASC	Nottingham Arabidopsis Stock Center
NBS	NIBRIN
NCBI	National Center for Biotechnology Information
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NPR1	NONEXPRESSER OF PR GENES 1
NSE	NON-SMC-ELEMENT
p4A::GUS	*pNSE4A::GUS*
p4B::GUS	*pNSE4B::GUS*
p53	TUMOR PROTEIN 53
PCD	Programmed cell death
PCNA	PROLIFERATING CELL NUCLEAR ANTIGEN

PCR	Polymerase chain reaction
PI	Propidium iodide
PIAS	Protein Inhibitor of Activated STAT
PR	PATHOGEN RESISTANCE
PRC	POLYCOMB REPRESSIVE COMPLEX
RACE	Rapid amplification of cDNA-ends
RAD51	RADIATION SENSITIVE 51
RAM	Root apical meristem
rDNA	Ribosomal DNA
RFB	Replication fork barrier
RFC	REPLICATION FACTOR C
RING	Really Interesting New Gene
RNA	Ribonucleic acid
ROS	Reactive oxigen species
ROS1	REPRESSOR OF SILENCING 1
RPA	REPLICATION PROTEIN A
RT	Reverse transcription
SA	Salisilyc acid
SAR	SA-mediated systemic acquired resistance
SCC1	RAD21 homolog
SCI	SISTER CHROMATID INTERTWINING
SDSA	Synthesis-dependent strand annealing
SGS1	SLOW GROWTH SUPRESSOR 1
SLF1	SMC5-SMC6 COMPLEX LOCALIZATION FACTOR
SMC	STRUCTURAL MAINTENANCE OF CHROMOSOMES
SMC6B	STRUCTURAL MAINTENANCE OF CHROMOSOMES 6B
SNI	SUPPRESSOR OF NPR1-1, INDUCIBLE
SOC1	SUPPRESSOR OF CONSTANS
SOG1	SUPRESSOR OF GAMA RESPONSE 1
SSA	Single-strand annealing
SSB	Single strand break

SSD	Single stranded DNA
SSN	SUPPRESSOR OF SNI
STAT	Signal Transducer and Activator of Transcription
STR complex	SGS1-TOP3-RMI1 complex
SUMO	Small-Ubiquitin-Modifier
T	Torpedo
TCR	Transcription-coupled repair
T-DNA	Transfer-DNA
TFIIH	TRANSCRIPTION FACTOR IIH
TLS	TRANSLESION DNA SYNTHESIS
TOP1	TOPOISOMERASE 1
TOP2	TOPOISOMERASE 2
TOP3	TOPOISOMERASE 3
TRAX	TRANSLIN-ASSOCIATED FACTOR-X
UV-DDB	UV-DAMAGED DNA BINDING PROTEIN
WT	Wild type
XPA	XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP A
XPB	XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP B
XPC	XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP C
XPD	XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP D
XPF	XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP F
XPG	XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP G
XRCC1	X-RAY REPAIR COMPLEMENTING DEFECTIVE IN CHINESE HAMSTER 1
Y2H	Yeast-two-hybrid
YFP	Yellow fluorescent protein
ZEB	Zebularine

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Genome stability and DNA damage repair (DDR)

Eukaryotic nuclear genome is organized into linear chromosomes. Chromosomal DNA is wrapped around histone octamers forming nucleosomes, the primary chromatin units, which are folded into chromatin fibers and domains of different density and accessibility (Li et al., 2002; Kouzarides, 2007). Chromosome and chromatin stability can be challenged by endogenous factors including oxidative stress, replication errors and/or topological stress (Branzei et al., 2008) or adverse environmental conditions such as UV radiation and chemical pollutants (Roy, 2014). These (and other) factors may challenge genome stability by a wide range of toxic effects including base oxidation, alkylation, DNA single and double strand breaks (SSBs and DSBs) and formation of non-native bonds between and/or within DNA strands e.g intra-, interstrand cross-links (Hu et al., 2016). Unrepaired lesions result in mutations, which compromise gene functionality, cause loss/gain of genetic information and induce chromosome instability. This problem may be particularly pronounced in obligatory phototrophic sessile organisms such as plants, which are often exposed to challenging environmental conditions without possibility for escape (Balestrazzi et al., 2011; Willing et al., 2016).

1.2. DNA damage

1.2.1. Sources of DNA damage and damage products

Endogenous sources

Reactive oxygen species (ROS) are particularly important due to their reactivity. These include singlet oxygen ($^1O_2$), superoxide (O_2^-) hydrogen peroxide (H_2O_2), and hydroxyl radicals ($HO\cdot$), among others. These substances are produced through the life of the plant, due to metabolic processes such as photosynthesis, respiration and all kind of stresses such as high light, high or low temperature, salinity, drought, nutrient deficiency and pathogen attack (Tripathy and Oelmuller, 2012). They produce oxidative stress, oxidizing DNA bases, which results in single- and double-strand breaks (Hu et al., 2016).

An example for endogenous base alkylation, is DNA methylation (reviewed in Britt, 1994). The majority of the bonds in all four bases are susceptible to methylation, but most frequent ones occur at purine bases. If uncorrected, they can be premutagenic and/or lethal. The most recurrent generated alkylation product, 7-methyladenine. This product is able to pair normally and is considered as neither mutagenic nor toxic. In contrast, 3-methyladenine cannot serve as a template for DNA synthesis and therefore acts as a block to DNA replication. The effects of base alkylation will be discussed more in the next section.

On the other hand, hydrolysis of the glycosylic bond between purine bases and the DNA backbone can produce an abasic site (Lindahl, 1993). If left unrepair, similarly to base alkylation, this could block DNA replication and/or transcription and cause lethality. One consequence of hydrolysis is the formation of uracil and thymine due to the hydrolytic deamination of cytosine and 5-methylcytosine. These two deamination products are potentially mutagenic (reviewed in Britt, 1996). Whereas uracil is recognized as an incorrect base in DNA, and it is excised by uracil glycosylase, thymine cannot be recognized as a DNA damage product and subsequently is highly mutagenic, producing C:G to T:A transitions.

Exogenous sources

Exogenous agents also play an important role as DNA damage inducers. In fact, UV-B (280–320 nm) and UV-A (320–400 nm) present in sunlight are the most important sources of epidermal DNA damage in plants and animals (reviewed in Britt, 1996). The most recurrent UV-induced DNA damage products are cyclobutane pyrimidine dimers (CPD) and the pyrimidine (6-4) pyrimidinone dimer (the 6-4 photoproduct), but UV can also produce free radicals in a minor amount (Chen et al., 1994; Britt, 1996). These products cause cross-links between adjacent cytosine and thymine, blocking replication and/or transcription (Setlow et al., 1963). For plants this type of damage is particularly important as they need sunlight in order to perform photosynthesis.

In contrast to UV-radiation, ionizing radiation produces a broad range of damage products. In plants cells, the most frequent target of ionizing radiation is water, and the interaction results in formation of hydroxyl radicals, producing thus, oxidative damage (reviewed in Britt, 1996). If the sugar phosphate backbone absorbs ionizing radiation this can generate a nick, which increases the probability of DSBs. Furthermore, ionizing radiation is often used in research to

generate chromosomal breaks, inversions, duplications, and translocations, but also this damage can generate oxidative damage, resulting in point mutations (reviewed in Britt, 1996).

Finally, industrial- environmental chemicals, fungal and bacterial toxins can produce a wide spectra of DNA damage, such as base oxidation and alkylation, DNA cross-links, DNA-protein cross-links, among others (reviewed in Balestrazzi et al., 2011)

Note that there is a broad spectrum of types of DNA damage and repair mechanisms. In the following paragraphs I will described only the most relevant ones for this study.

1.2.2. Types of DNA damage

Base substitutions

Base substitutions can rise spontaneously in the cell from hydrolytic events like deamination or base loss, oxidative damage, methylation of ring nitrogens by endogenous agents, or due to exposure to exogenous agents such as X-rays (reviewed in Britt, 1996). In plant research, base modification by alkylating agents is widely used. Two examples for monofunctional alkylating agents are Ethyl methanesulfonate (EMS) and Methyl methanesulfonate (MMS). EMS has been extensively used in forward genetics for generation of mutagenized populations (Koornneef et al., 1982; Greene et al., 2003). It ethylates the N7- or O 6 -position of guanine, which leads to G/C → A/T transitions (Krieg, 1963). On the other hand, MMS can cause formation of toxic apurinic sites and/or blocking of DNA polymerases. MMS predominantly methylates DNA on N7-guanine and N3-adenine (reviewed in Pecinka and Liu, 2014), producing T/A → G/C transversions and A/T → G/C transitions (Beranek, 1990). DNA damage caused by this and other alkylating agents is predominantly repaired by the base excision repair (BER) pathway and DNA alkyltransferases (Lundin et al., 2005).

Double strand breaks (DSBs)

DSBs occur endogenously as part of crossing over during meiosis, due to mechanical stress or due to a nick in a single stranded region (reviewed in Britt, 1996). In addition, they can be induced by exogenous agents such as ROS, which are caused by various oxidative processes, e.g. imbibition prior to seed germination or UV radiation, transgenic DNAs, and radiomimetic chemicals, such as bleomycin/bleocin or zeocin (reviewed in Britt, 1999). DSBs can be repaired

by homologous recombination (HR) and non-homologous-end-joining, both discussed in detail in the next chapter.

DNA cross-links

DNA cross-linking damage arises when cross-linking agents covalently connect two nucleotide from the same DNA strand (intrastrand cross-link) or from opposite strands [interstrand cross-link (ICL)]. Pyrimidine dimers induced by UV-B and DNA adducts induced by cisplatin are few examples of this damage. Intrastrand cross-links can be easily removed by nucleotide excision repair (NER) mechanism (Huang and Li, 2013).

However, because Interstrand cross-links (ICLs) covalently link the two strands of the DNA, they can block replication and transcription (Raschle et al., 2008). Thus, they are considered more toxic. ICLs can be formed by endogenous metabolites but also exogenous agents such as mitomycin C (MMC). In eukaryotes, ICLs are repaired by a mix of different pathways, including HR and NER.

DNA-protein cross-links (DPC)

In order for a cell to perform all essential functions, interaction between DNA and proteins are vital. This highly dynamic interaction is most of the times transient. However, they can be trapped by stable covalent cross-linking of proteins to the DNA (Stingele and Jentsch, 2015). DPCs can be formed via two types of events, enzymatic and non‑enzymatic cross-linking. The first one consists on the permanent trapping of normally transient covalent protein–DNA intermediates, which could be caused spontaneously or due to enzyme poisons (e.g. camptothecin, which causes cross-linking of TOPOISOMERASE 1 (TOP1)).

On the other hand, non-enzymatic DPCs can be caused by certain agents that cause unspecific chemical cross-linking of random proteins to DNA (Stingele and Jentsch, 2015). Those agents can be ionizing radiation, UV light, or metabolically produced reactive aldehydes created from ethanol oxidation or histone demethylation (Ide et al., 2011; Stingele et al., 2015).

DPCs can cause stalling of replication forks, which can inhibit successful replication, cell division and induce cell death (Stingele and Jentsch, 2015).

In plants DPC has not been much studied. Nevertheless, our research group recently found that in plants the cytidine analogue zebularine, causes enzymatic DPCs (A. Finke and A. Pecinka, unpublished data). Plant zebularine-induced DPC repair pathway is currently unknown,

however an initial study revealed that the presence of DPC is signaled redundantly by both *ATAXIA TELANGIECTASIA MUTATED (ATM)* and *ATAXIA TELANGIECTASIA MUTATED AND RAD3 RELATED (ATR)* kinases (Liu et al., 2015).

1.2.3. DNA damage response

The response to DNA damage involves a complex pathway, which depending on the severity of the damage, in plants it can range from cell cycle arrest, activation of DNA repair genes, endoreplication[1] to apoptosis (Barnum and O'Connell, 2015; Hu et al., 2016). This response consists of sensors, transducers and effectors (Tuteja et al., 2009). Firstly recognition occurs, where the DNA-damage intermediates are sensed by *ATM* and *ATR*. In animals, activated ATM phosphorylates downstream components of checkpoint pathways that include TUMOR PROTEIN 53 (p53), BREAST CANCER 1 (BRCA1), NIBRIN (NBS1), and CHECK POINT KINASE2 (CHK2), initiating G1-, S-, or G2-phase arrest and/or apoptosis (Shiloh, 2001). On the other hand, activated ATR phosphorylates CHECK POINT KINASE1 (CHK1) to initiate G2-phase arrest and BRCA1 at other points of the cell cycle (Tibbetts et al., 2000). In plants however, phosphorylation products from ATM and ATR remain mostly unknown, as plant homologs for *p53, CHK1 and CHK2* have not been found. Recent studies showed that the DDR transcription factor *SUPRESSOR OF GAMA RESPONSE 1 (SOG1),* acts dependent of both *ATM* and *ATR*, is phosphorylated by ATM and is required for the induction of hundreds of transcripts in response to gamma radiation (Yoshiyama et al., 2009; Yoshiyama et al., 2013). The role of *SOG1* as master regulator in DDR suggests that despite the lack of sequence similarity, *SOG1* could be the functional homolog of p53 in plants. Furthermore, the lack of conservation of transcription factors involved in DDR may indicate that the DNA damage response could vary depending on the life strategy of the organism in question.

[1] Process in which DNA replication occurs in the absence of mitosis, resulting in cellular polyploidy.

1.3.DDR mechanisms

1.3.1. Photoreactivation

The damage induced by UV-B radiation is repaired preferentially by photoreactivation and only to a minor extent by NER. In plants, photoreactivation is carried out by photolyases using the energy of UV-A radiation. Photolyases bind cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts and upon absorption of a photon of a wavelength of 350–450 nm, directly reverse the damage in an error-free manner (Chen et al., 1994).

1.3.2. Base excision repair (BER)

BER is a mechanism that protects against non-bulky DNA base damage caused by alkylation, oxidation, deamination or replication errors (Cordoba-Canero et al., 2011). The first step is performed by DNA glycosylases and includes the recognition and removal of damaged base, resulting in an apurinic/apyrimidinic (AP) site (Tuteja et al., 2009).

After base incision, in mammals there are two pathways known for the DNA synthesis: long-patch BER, which is DNA polymerase δ/ε- dependent, and short-patch BER, in which synthesis is catalyzed by DNA polymerase β. In long-patch BER, PCNA and (5-FLAP ENDONUCLEASE) FEN-1 act in coordination with DNA polymerase and DNA ligase I for DNA synthesis and ligation. In short-patch BER, the 5′ -dRP residue is removed by the dRP lyase activity of DNA polymerase, DNA chains are then joined to complete the repair by DNA LIGASE III (LIG3) and X-RAY REPAIR COMPLEMENTING DEFECTIVE IN CHINESE HAMSTER 1 (XRCC1) (reviewed in Kimura and Sakaguchi, 2006). In plants however, there is no direct DNA polymerase β homolog, but instead a putative homolog for mammalian DNA polymerase λ, which also possesses dRP lyase activity and contributes to BER (Garcia-Diaz et al., 2001). One additional difference is the absence of direct homolog for mammalian LIG3, although it was shown that LIG1 is able to catalyze ligation in both short- and long-patch sub-pathways in plants (Cordoba-Canero et al., 2011).

1.3.3. Nucleotide excision repair (NER)

This mechanism can repair a wide spectra of lesions from a single base to a big bulky adduct (reviewed in Brit 1999). NER is subdivided in two pathways: global genome repair (GGR), which repairs damage everywhere in the genome and transcription-coupled repair (TCR), which repairs damage during transcription (Kimura and Sakaguchi, 2006). Recognition differs in the pathway of choice. In GGR, recognition is mediated by XERODERMA PIGMENTOSUM COMPLEMENTATION GROUP C/RAD23 (XPC/RAD23) or by UV-DAMAGED DNA BINDING PROTEIN (UV-DDB). Instead, in TCR the contact of RNA polymerase II with damaged bases triggers recognition by COCKAYNE SYNDROME, TYPE A (CSA) or COCKAYNE SYNDROME, TYPE B (CSB). Next, in both pathways the unwinding is mediated by TRANSCRIPTION FACTOR IIH (TFIIH) together with XERODERMA PIGMENTOSUM COMPLEMENTATION GROUPS B, D, A (XPB, XPD, XPA) and REPLICATION PROTEIN A (RPA). After unwinding, XPF, XPG and EXCISION REPAIR COMPLEMENTING DEFECTIVE REPAIR IN CHINESE HAMSTER 1 (ERCC1) excise about 20 to 30 bases (reviewed in Dronkert and Kanaar, 2001; Kimura and Sakaguchi, 2006). The gap formed is filled by PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) and REPLICATION FACTOR C (RFC)-dependent DNA synthesis carried out by DNA polymerase. Finally the repaired DNA strand is rejoined by DNA LIGASE 1 (LIG1) synthesis (reviewed in Kimura and Sakaguchi, 2006). If two nicks in complementary strands are too close to each other or a single stranded DNA (SSD) replicates, a DSB can be formed. Thus, synergism of two or more repair pathways often occurs (Dronkert and Kanaar, 2001).

1.3.4. Non-homologous end joining (NHEJ) and homologous recombination (HR)

DBSs can be repaired via two alternative pathways: homologous recombination (HR) and non-homologous end joining (NHEJ). HR uses an intact homolog sequence as a template (sister chromatid or less frequently homolog chromosome) and depending on genomic architecture, the DSB can be repaired via single-strand annealing (SSA) and synthesis-dependent strand annealing (SDSA) (Puchta et al., 1996; Roth et al., 2012). After a DSB induction, in both SSA and SDSA pathways an exonuclease-catalyzed end resection is performed, which leads to the formation of single-stranded overhangs. However, the difference between both pathways is that in SSA, direct annealing occurs between complementary sequences, which could result in

trimming of 3´overhangs and loss of information (Roth et al., 2012). In contrast, SDSA is characterized by strand invasion, in which the 3′ end invades a homologous double strand forming a D-loop. This permits repair using the newly paired strand as a template. Thus, as a result of the reaction, there is no loss of DNA sequence.

NHEJ, on the other hand, reseals the broken ends by activity of LIGASE 4 (LIG4). In this pathway, the broken ends are protected by donut-like KU70-KU80 dimer to prevent their digestion by nucleases. However, this protection is not absolute leading to occasional loss of sequences during NHEJ (Tamura et al., 2002). In addition, ligation of non-DSBs can lead to formation of large chromosomal rearrangements. Therefore, NHEJ is considered as an error-prone pathway. Unlike yeasts, which mainly use HR as a mechanism for repair of DSBs, plants repair DSBs primarily by NHEJ (Puchta et al., 1996). It is known that the DNA damage response induced by DSBs is signaled primarily by ATM kinase (Garcia, 2003).

1.3.5. Protein-cross-link repair (DPC)

In eukaryotes, two different models of DPC repair have been proposed. One of them involves proteolysis by DPC-specific metalloproteases, followed by TRANSLESION DNA SYNTHESIS (TLS) polymerases for replication completion, which can be mutagenic (Stingele et al., 2015). The second model occurs when cells lack a DPC protease. Here, persistent fork stalling can cause fork collapse, which is then cleaved by an endonuclease that generates a single-ended DSB (Stingele and Jentsch, 2015). This can then be repaired by break-induced replication (BIR) or homologous recombination (HR). Alternatively, convergence with an incoming replication fork can also repair a DPC. In this pathway the DNA past the DPC can be unwind, which allows replication completion or initiate a repair mechanism similar to ICL repair (Stingele and Jentsch, 2015).
In plants DPC has not been much studied. Nevertheless, our research group recently found that in plants the cytidine analogue zebularine, causes enzymatic DPCs (A. Finke and A. Pecinka, unpublished data).

1.4. The SMC complexes

Structural maintenance of chromosomes (SMC) complexes are the key regulators of chromosome dynamics, structure and function in eukaryotes (reviewed in Losada and Hirano, 2005; Hirano, 2006; De Piccoli et al., 2009; Jeppsson et al., 2014a; Uhlmann, 2016). They operate from the scale of whole chromosomes in e.g. chromosome segregation to a few base pairs in DNA damage repair. The core subunits of SMC complexes are SMC proteins, which are large polypeptides (1,000–1,300 amino acids) containing Walker A and Walker B motifs at the N- and C- terminal globular domains. The prerequisite for SMC function is folding at the hinge domain and coiling of the arms. This brings the C- and N-terminal globular domains together and constitutes heads with ATP-dependent DNA binding activity (Kanno et al., 2015). The most characterized SMC complex is cohesin (SMC1-SMC3), which controls dynamics of sister chromatid cohesion and thus affects various chromosomal processes (reviewed in Losada and Hirano, 2005; Hirano, 2006; Jeppsson et al., 2014a). Condensin complex (SMC2-SMC4) plays a pivotal role in chromosome folding and condensation during interphase and nuclear division. Finally, the third complex consisting of SMC5-SMC6 heterodimer backbone (SMC5/6; (reviewed in De Piccoli et al., 2009) is famous for its role in maintaining genome stability. Besides of SMC5 and SMC6, this complex contains six additional NON-SMC ELEMENT (NSE) subunits (Figure 1A,B). The SMC5/6 subunits are organized into three sub-complexes: NSE2-SMC5-SMC6, NSE1-NSE3-NSE4 and NSE5-NSE6 acting as specialized functional modules performing particular biochemical functions (Sergeant et al., 2005; Palecek et al., 2006; Duan et al., 2009).

1.5. Architecture of the SMC5/6 complex

1.5.1. NSE1-NSE3-NSE4 sub-complex

NSE1-NSE3-NSE4 trimer is a highly conserved part of the SMC5/6 complex responsible for binding DNA and bridging SMC heads. NSE1 contains a RING-like domain necessary for the NSE1-NSE3-NSE4 trimer formation and recruitment of NSE4 and SMC5 to the sites of DNA damage (Fujioka et al., 2002; McDonald et al., 2003; Pebernard et al., 2008; Taylor et al., 2008). Mutations in the RING-like domain lead to DNA damage hypersensitivity and full deletion of *NSE1* is lethal in *S. cerevisiae*, *S. pombe* and Arabidopsis (Fujioka et al., 2002; McDonald et

al., 2003; Li et al., 2017). In multiple organisms it was shown that NSE1 interacts with the N-terminus of NSE3 subunit and strengthens its binding to dsDNA (Palecek et al., 2006; Pebernard et al., 2008; Duan et al., 2009; Hudson et al., 2011; Zabrady et al., 2016; Li et al., 2017). NSE3 occurs as single copy gene in fungi and plants, but has homology to *MAGE* (*MELANOMA ANTIGEN GENE*) family with over 60 members in humans. However, only MAGE-G1 and MAGE-F1 have been found to associate with NSE1 and only MAGE-G1 co-immunoprecipitated with the SMC5/6 holocomplex in human cells (Chomez et al., 2001; Taylor et al., 2008; Doyle et al., 2010). MAGEs interact *in vitro* with E3 RING-type ubiquitin ligases, are aberrantly expressed in a wide variety of cancer types and play a critical role in tumorigenesis (Barker and Salehi, 2002; Weon and Potts, 2015). The presence of MAGE-G1 and UBCH13/MMS2 E2 ubiquitin-conjugating enzyme significantly enhances NSE1 E3 ubiquitin ligase activity (Doyle et al., 2010). The C-terminal domain of NSE3 interacts with NSE4 (Hudson et al., 2011). NSE4 is a structural protein containing a winged helix motif, which forms a RING-like structure through interaction with SMC proteins (Schleiffer et al., 2003; Palecek et al., 2006). NSE4 is an essential protein and its functions include: interaction between NSE1-NSE3-NSE4 sub-complex and SMC5 as shown in *S. pombe* (Duan et al., 2009) and bridging SMC5 and SMC6 heads as found in *S. cerevisiae* (Palecek et al., 2006).

1.5.2. NSE2-SMC5-SMC6 sub-complex

NSE2-SMC5-SMC6 represents the core sub-complex, which serves as a scaffold and via NSE2/MMS1 enzymatic activity most likely regulates dynamics of the whole SMC5/6 complex throughout its target processes. *NSE2*/*MMS21* (*METHYL METHANE SULFONATE SENSITIVITY 21*) was initially identified via genetic screen as hypersensitive to MMS, X-rays and UV radiation in budding yeast (Prakash and Prakash, 1977), and was associated with SMC5/6 complex only about two decades later (Zhao and Blobel, 2005). NSE2/MMS21 is covalently bound to the SMC5 protein (Figure 1A), and this association appears to be conserved in fungi, animals and plants (Potts and Yu, 2005; Sergeant et al., 2005; Duan et al., 2009; Xu et al., 2013). NSE2/MMS21 contains a putative SIZ/PIAS (Protein Inhibitor of Activated STAT[2]) RING domain characteristic of Small Ubiquitin-like Modifier (SUMO) ligase (Zhao and Blobel, 2005). In general, SUMO modification is involved in various cellular processes, such as nuclear-cytosolic transport, transcriptional regulation, apoptosis, protein stability,

[2] STAT - Signal Transducer and Activator of Transcription

response to stress, and progression through the cell cycle (Hay, 2005; Jalal et al., 2017). However, which of these (and potentially other) processes can be assigned to NSE2/MMS21 is largely unknown. *In vitro* studies revealed that NSE2/MMS21 SUMOylates[3] to numerous proteins, some possibly in a species-specific manner. So far identified NSE2/MMS21 targets include SMC6, NSE3 and NSE4 in fission yeast; SMC5 and KU70 in budding yeast; SMC6, cohesin subunits SA2 and SCC1, TRANSLIN-ASSOCIATED FACTOR-X (TRAX) and several members of the SHELTERIN/TELOSOME complex in humans (Andrews et al., 2005; Potts and Yu, 2005; Zhao and Blobel, 2005; Potts et al., 2006; Potts and Yu, 2007; Pebernard et al., 2008; McAleenan et al., 2012). Surprisingly, in fungi, animals and plants NSE2 auto-SUMOylates at the C-terminal region and thus most likely auto-regulates its own function. Catalytically dead human NSE2/MMS21 is not able to alleviate hypersensitivity to DNA damage, suggesting that NSE2/MMS21 SUMO ligase activity is required for proper cellular response to DNA damage

1.5.3. NSE5-NSE6 sub-complex

The NSE5-NSE6 sub-complex is most likely responsible for loading, localization or multimerization of SMC5/6 complex (Figure 1B,C). Pull-down experiments using fungi, plants and *Xenopus* egg extracts identified a pair of previously unknown SMC5/6 associated proteins, which shared surprisingly little primary sequence similarity. It was YML023c (alias NSE5) and KRE29 in budding yeast (Zhao and Blobel, 2005), and NSE5 and NSE6 in fission yeast (Pebernard et al., 2006). In Arabidopsis, these unknown subunits included ARABIDOPSIS SNI ASSOCIATED PROTEIN 1 (ASAP1) and SUPPRESSOR OF NPR1, INDUCIBLE (SNI1) (Yan et al., 2013). Recently SMC5-SMC6 COMPLEX LOCALIZATION FACTORs 1 and 2 (SLF1 and SLF2) were found in vertebrates (Raschle et al., 2015). KRE29, NSE6, SNI1 and SLF2 contain armadillo (ARM) repeats (Huber et al., 1997), which are supposed to form a superhelix of α-helices resulting in a spiral structure. 3D modeling suggested that all these factors have a highly similar protein structure, where several essential residues of the ARM repeats create a binding surface not apparent from the linear sequence (Pebernard et al., 2006; Yan et al., 2013). Therefore, KRE29, SLF2 and SNI1 are considered as putative functional orthologs of NSE6, and their interaction partners YML023c, SLF1 and ASAP1 as the putative functional orthologs of NSE5, respectively. Beside little conserved protein sequence, NSE5 and

[3] Addition of SUMOs (small ubiquitin-like modifiers)

NSE6 differ also with respect to their position in the complex. In budding yeast NSE6 and NSE5 bind to the hinges of SMC5 and SMC6 (Duan et al., 2009), while in fission yeast they bind to SMC5 and SMC6 heads, without directly interacting with the NSE1-NSE3-NSE4 trimer (Figure 1B) (Palecek et al., 2006; Pebernard et al., 2006). The function of NSE5-NSE6 is unknown, but an earlier study (Palecek et al., 2006) proposed that this dimer could mediate SMC5/6 complex multimerization (Figure 1C). In vertebrates, SLF1-SLF2 sub-complex mediates interaction of SMC5/6 with RAD18 E3 ubiquitin protein-ligase during the process of DNA damage repair at stalled replication forks (Raschle et al., 2015) . Location of SNI1 and ASAP1 in Arabidopsis remains unknown (Figure 1B). Both NSE5 and NSE6 were shown to be essential in budding yeast, but not in fission yeast. In *A. thaliana*, loss of *SNI1* function leads to smaller and poorly looking plants with strongly reduced fertility. Homozygous *ASAP1* mutant plants are not able to develop beyond the cotyledon stage and die (Yan et al., 2013).

1.6. SMC5/6 complex molecular functions

1.6.1. DNA damage repair

Repair of damaged DNA represents the canonical function of SMC5/6 complex and multiple complex subunits were identified in genetic screens based on mutant hyper-sensitivity to genotoxic stress (Lehmann et al., 1995; Mengiste et al., 1999; Fousteri and Lehmann, 2000; Andrews et al., 2005; Santa Maria et al., 2007; Raschle et al., 2015).

Arabidopsis *SMC6B* and *NSE2*/*MMS21*/*HPY2* (*HIGH PLOIDY 2*) mutants show moderate hypersensitivity to UV, x-rays and MMC and strong hypersensitivity to methyl methane sulfonate (MMS) and zebularine (Mengiste et al., 1999; Ishida et al., 2009; Watanabe et al., 2009; Yuan et al., 2014; Liu et al., 2015). While genotoxic effects of most of these treatments are generally well understood (reviewed in Pecinka and Liu, 2014), effects of zebularine remain less clear. Our group found that beside its (relatively weak) DNA demethylating effects (Baubec et al., 2009; Baubec et al., 2014; Liu et al., 2015), it acts as a potent inducer of enzymatic DNA-protein cross-links (A. Finke and A. Pecinka, unpublished data). Collectively, the DNA damage assays indicate that the SMC5/6 complex participates in (post-)replicative repair of mainly complex or bulky lesions (Figure 1D) and has only a negligible role in non-homologous end joining repair of DSBs in Arabidopsis.

In animals and fungi, both cohesin and SMC5/6 complexes are recruited to DSB sites (De Piccoli et al., 2006; Lindroos et al., 2006; Potts et al., 2006; Outwin et al., 2009; Jeppsson et al., 2014b). Initial observations in human cells revealed that the SMC5/6 complex recruits cohesin, facilitating repair by HR, and this recruitment was dependent on NSE2/MMS21 mediated SUMOylation (Potts et al., 2006). Recent study using *Xenopus* eggs and human cells revealed that the recruitment of SMC5/6 to the sites of DNA damage is dependent on RAD18 and newly identified NSE5 and NSE6-like subunits SLF1 and SLF2 (Raschle et al., 2015). In contrast, the yeast SMC5/6 complex requires loading to chromosomes via the SCC1 subunit of cohesin complex (Strom et al., 2007; Outwin et al., 2009; Jeppsson et al., 2014b). Hence, loading of the SMC5/6 complex may be species-specific. Data using flow-sorted Arabidopsis G2 nuclei revealed an SMC5/6-dependent increase in sister chromatid alignment upon induction of DNA damage, which depends on correct S phase-mediated cohesion (Watanabe et al., 2009).

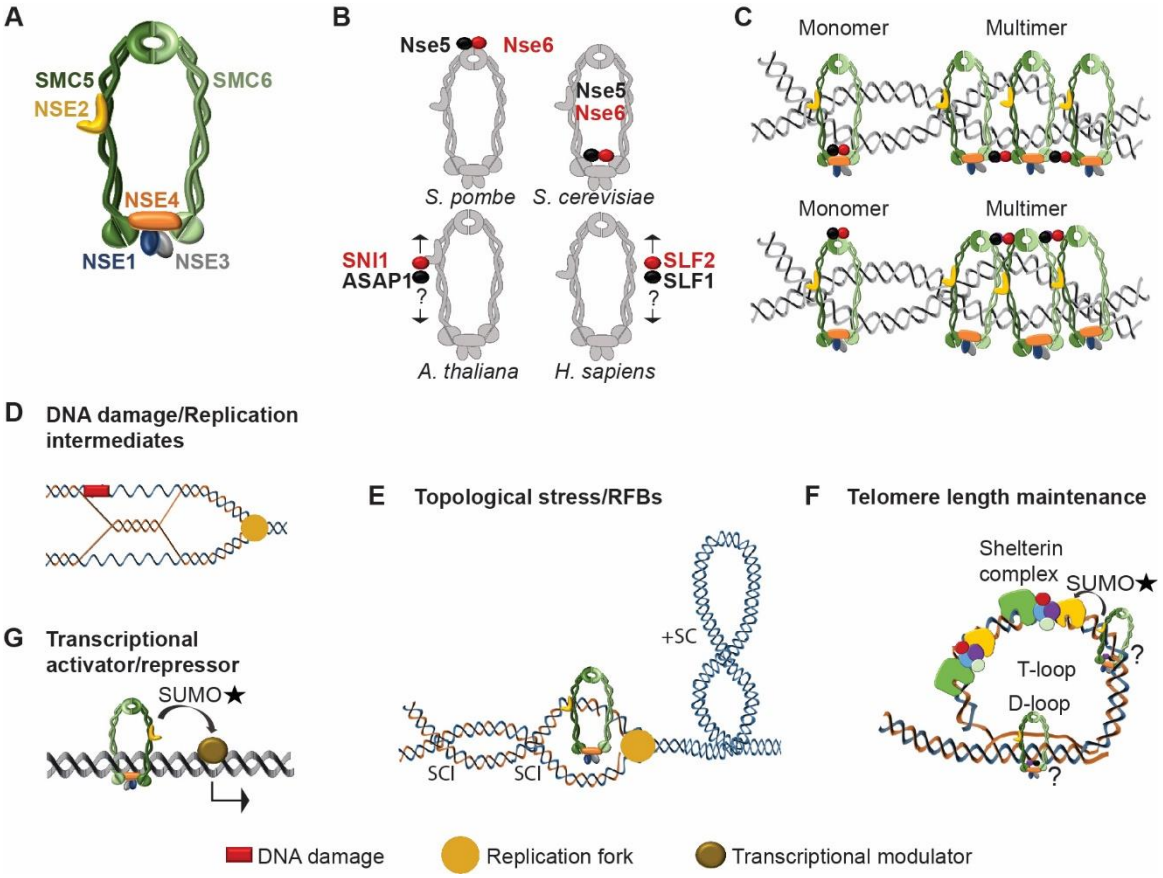


Figure 1. SMC5/6 complex composition and functions. (A) Consensual model of SMC5/6 complex without and **(B)** with species-specific positions of NSE5(-like) and NSE6(-like) subunits. **(C)** Hypothetical function of NSE5-NSE6 dimer in multimerizing SMC5/6 complexes via their heads (top) or hinges (bottom). **(D)** Replication intermediate structure bypassing DNA

damage site (red square). **(E)** Topological stress occurring during DNA replication and at replication fork barriers (RFBs) represented by the positive supercoil (+SC) ahead of the replication fork and sister chromatid intertwinings (SCIs) between the nascent chromatids. **(F)** Role of SMC5/6 complex in telomere length maintenance. **(G)** Speculative model for SUMOylation of transcriptional modulations by SMC5/6 complex. Note that the position of SMC5/6 complex in images (C), (E), (F) and (G) is only speculative.

Using transgenic reporter systems it was shown that Arabidopsis *SMC6A*, *SMC6B* and *NSE2*/*MMS21* mutants have reduced frequency of HR under control conditions (Mengiste et al., 1999; Watanabe et al., 2009; Yuan et al., 2014). Upon genotoxic treatments, wild-type and also *SMC6A* and *SMC6B* mutants (*NSE2*/*MMS21* was not tested) showed similar fold-increase, but the total number of HR events still remained much lower in the mutants. In contrast, expression of *SMC6B* under the control of a strong constitutive viral promoter doubled HR frequency (Mengiste et al., 1999; Hanin et al., 2000). This indicates that the plant SMC5/6 complex functions as positive regulator of HR in a regulatory network, where several pathways compete for processing lesions by HR or other repair mechanisms. This observation is consistent with the fungal and animal models (McDonald et al., 2003; De Piccoli et al., 2006; Potts et al., 2006; Stephan et al., 2011).

Two evolutionary conserved kinases ATM and ATM AND RAD3-RELATED (ATR) are involved in signaling presence of DNA strand breaks and single stranded DNA (typically at stalled replication forks), respectively, within the HR pathway (Garcia et al., 2003; Culligan et al., 2006; Cimprich and Cortez, 2008). Processing of spontaneous damages (presumably induced by DNA replication) is controlled by ATR, while both kinases are involved in signaling the presence of zebularine-induced DNA damage (Yan et al., 2013; Liu et al., 2015). Whether the Arabidopsis SMC5/6 complex is directly phosphorylated by ATM and/or ATR is unknown. A recent study analyzing phosphoproteomic targets of ATM and ATR did not reveal any SMC5/6 members (Roitinger et al., 2015). However, this could be due to the treatment with gamma-radiation producing mainly DSBs, i.e. substrate which is not a typical target of SMC5/6 complex mediated repair in Arabidopsis. Alternatively, SMC5/6 complex members could be activated at transcriptional level. However this scenario is less likely because none of the complex subunits was detected up- or down-regulated in genome-wide studies using wild-type and ATM and ATR mutant plants exposed to a variety of DNA damaging treatments (Culligan et al., 2006; Kilian et al., 2007; Liu et al., 2015).

Homology based repair is particularly challenging in tandemly repeated genome regions and replication fork barriers (RFBs). High similarity of individual repeat units increases the risk of HR between ectopic copies, which can lead to loss of genetic information (Devos et al., 2002; Ampatzidou et al., 2006). Data from yeasts and animals suggest that the SMC5/6 complex is recruited to RFBs during G2/M and controls HR at least in ribosomal DNA (rDNA) and telomeres (Figure 1E,F) (Torres-Rosell et al., 2005; De Piccoli et al., 2006; Hwang et al., 2008; De Piccoli et al., 2009; Menolfi et al., 2015). Presence of the SMC5/6 complex at rDNA loci (and telomeres) reduces activity of the recombination proteins like RAD51 (Chiolo et al., 2011; Ryu et al., 2015), while its loss is accompanied by formation of Rad52 foci, indicative of error-prone repair, increased frequency of holliday junctions, HR and chromosomal rearrangements (Torres-Rosell et al., 2005; Torres-Rosell et al., 2007). An interesting mechanism reducing the risk of ectopic recombination at repetitive DNA was described in insects and fungi (Torres-Rosell et al., 2007; Chiolo et al., 2011; Ryu et al., 2015). Here, SMC5/6 complex interaction with HETEROCHROMATIN PROTEIN 1 (HP1) blocks HR in heterochromatin until its expansion and relocation of damage sites into (repetitive DNA poor) euchromatic nuclear space. SMC5/6 dependent heterochromatin remodeling upon DNA damage has not been observed in plants so far, but one study showed that the kinetics of DSB repair is slower in the *SMC6B* mutant (Kozak et al., 2009). This may indicate that the SMC5/6 complex affects repair kinetics in plants, but whether this is accompanied with heterochromatin relaxation has to be analyzed.

1.6.2. Removal of replication-derived toxic structures

Assisting DNA replication machinery, removal of toxic replication structures and relief from DNA topological stress represent potentially highly conserved but only recently discovered SMC5/6 functions. Hypersensitivity of SMC5/6 mutants to the replication blocking agents, such as hydroxyurea (HU) and MMS, led to assumption that the complex may be involved in detoxifying toxic structures arising during DNA replication (Branzei et al., 2006). A recent study in budding yeast revealed that SMC5/6 complex functions are essential during (late) G2 phase, but not in the other cell cycle stages including S-phase, under non-damaging conditions (Menolfi et al., 2015). Absence of SMC5/6 during S-phase allows normal replication initiation and fork speed, suggesting that the SMC5/6 function is post-replicative (Menolfi et al., 2015). To date, two major post-replicative functions of SMC5/6 complex have been described: (i)

removal of DNA supercoils and sister chromatid intertwinings (SCIs) and (ii) resolving toxic DNA replication intermediates (Figure 1D,E).

Progressive separation of the parental DNA strands by replication machinery leads to the accumulation of positively supercoiled DNA ahead of the replication fork and formation of SCIs, i.e. coiled dsDNA strands, behind the fork (Figure 1E). Both structures are problematic as they cause topological stress and hinder sister chromatid separation during mitosis, respectively, and need to be removed in order to allow normal cellular functions (DiNardo et al., 1984). Experiments in budding yeast showed that DNA supercoils are resolved by the coordinated action of type I TOPOISOMERASE 1 (TOP1) and type II TOPOISOMERASE 2 (TOP2), while SCIs are removed by the activity of TOP2 as shown in budding yeast (Bermejo et al., 2007). However, recently also the SMC5/6 complex was found to play a role in removal of DNA supercoils and formation of SCIs in the same model system (Kegel et al., 2011; Jeppsson et al., 2014b). It is assumed that SMC5/6 facilitates fork rotation by sequestering nascent SCIs that form behind the replication machinery, thus decreasing the level of replication-induced supercoiling (Kegel et al., 2011; Jeppsson et al., 2014b). The SMC5/6 complex is loaded to the sites of DNA topological stress by the cohesin complex during S-phase as indicated by the absence of chromosome bound SMC5/6 in cohesin mutant *scc1*, but loss of SMC5/6 function does not affect cohesin localization (Jeppsson et al., 2014b). Based on the experiments with circular DNA molecules, it was suggested that the SMC5/6 and TOP2 complex functions as ATP-dependent DNA linker, which facilitates intermolecular interactions through their topological entrapment (Kanno et al., 2015). In addition, TOP2 causes SMC5/6 to dissociate from chromosome arms under non-stress conditions (Kegel et al., 2011; Jeppsson et al., 2014b), possibly by efficient removal of SCIs, upon which the presence of SMC5/6 is no longer required. Depletion of human SMC5 and SMC6 results in abnormal distribution of TOPOIIα, a homolog of the yeast TOP2, which probably leads to accumulation and/or abnormal distribution of SCIs and aberrant chromosome segregation (Gallego-Paez et al., 2014). In budding yeast, SMC5/6 may oppose the SCI stabilizing activity of the cohesin complex in the absence of TOP2 activity, and thus allow easier passive sister chromatid resolution at the end of chromosomes (Jeppsson et al., 2014b). This suggests that the SMC5/6 complex controls the TOP2-independent SCI resolution pathway. This model is based on budding yeast where sister chromatids remain paired with each other after DNA replication (reviewed in Cohen-Fix, 2001). In Arabidopsis, where centromeres and telomeres show the highest degree of cohesion, in spite of a generally low degree of sister chromatid association during interphase (Schubert et al.,

2006b), the SMC5/6 complex activity may be stimulated on demand after e.g. occurrence of DNA damage (Watanabe et al., 2009).

Another important SMC5/6 function linked to the post-replicative phase is a rescue of collapsed replication forks and repair of replication-derived toxic HR intermediates (Branzei et al., 2006; Bermudez-Lopez et al., 2010; Menolfi et al., 2015). These are typically represented by X-shaped holliday junction structures formed during template switch in HR events (Figure 1D). They arise during bypass synthesis, when DNA polymerase encounters a block during DNA synthesis, switches the template to the newly replicated strand and returns to the original template after the damage. It is known that HR intermediates are repaired synergistically by the SMC5/6 complex and the STR complex. The latter consists of RECQ type helicase SGS1, type I topoisomerase TOP3 and domain of unknown function containing protein RMI1 in budding yeast (Branzei et al., 2006; Bermudez-Lopez et al., 2016; Bonner et al., 2016; Bermudez-Lopez and Aragon, 2017). The SMC5/6 complex associates to SGS1 and SUMOylates the STR complex, which decreases the presence of recombination structures (Branzei et al., 2006; Bonner et al., 2016). The resolution of branched structures seems to be dependent on the SUMOylation ability of NSE2/MMS21, as SGS1 mutants, impaired in recognition of SUMOylated SMC5/6 complex, exhibited unprocessed holliday junctions at damaged replication forks, increased exchange frequencies between double helices during double-strand break repair, and severe impairment in DNA end resection (Bermudez-Lopez et al., 2010; Bermudez-Lopez et al., 2016). Furthermore, there is an alternative (non-canonical) HR intermediate resolution pathway represented by MPH1, MMS2 and the SHU complex in budding yeast. It was proposed that the SMC5/6 complex acts antagonistically to MPH1, in a pathway distinct from that of SGS1, preventing accumulation of toxic intermediate structures (Chen et al., 2009; Choi et al., 2010).

1.7. SMC5/6 complex functions specific to plants

SMC5/6 complex controls number of processes, which are unique to higher plants. Many of these phenotypes appear to be critical for successful plant development and affect also economically important traits such as yield or stress resistance. Given the current state of the knowledge, for many of these phenotypes it cannot be unambiguously decided whether they are mediated by the SMC5/6 DNA damage repair or other activities.

Plant SMC5/6 complex includes six evolutionarily conserved and two plant-specific (ASAP1 and SNI1) SMC5/6 subunits (Supplemental Table 1). In spite of frequent polyploidization events during evolution of seed plants, most subunits are represented by a single copy gene in extant species. The only exception, is *NSE4*, which is represented by two or more copies in almost all analyzed seed plants (Supplemental Table 1).

1.7.1. Developmental regulator

Multiple studies showed that the SMC5/6 complex regulates specific developmental processes including e.g. meristem and stem cell niche size, flowering time, meiosis, gametophyte and seed development in Arabidopsis (Huang et al., 2009; Ishida et al., 2009; Ishida et al., 2012; Liu et al., 2014; Kwak et al., 2016; Li et al., 2017). The unifying theme of the affected tissues and biological processes is that they contain a high proportion of replicating nuclei and (rapidly) dividing cells, which could be associated with higher amounts of naturally occurring DNA damage and/or topological stress (Figure 1D,E). Moreover, some of these tissues represent germline cells, which appear to be under a strict control concerning genome and epigenome stability in plants (Kimura and Sakaguchi, 2006; Yadav et al., 2009; Baubec et al., 2014b; Willing et al., 2016; Diaz and Pecinka, 2017).

Most developmental phenotypes controlled by SMC5/6 complex are described for *NSE2*/*MMS21*, which is (together with *SNI1*) the only non-duplicated subunit producing viable homozygous mutants. Arabidopsis *NSE2*/*MMS21* mutants were identified based on the short roots with increased nuclear endoploidy (therefore named as *HIGH POIDY 2* - *HPY2*), abnormally developed shoots with small leaves, irregular phylotaxy, occasional fasciations and partial sterility (Huang et al., 2009; Ishida et al., 2009). Cells within *NSE2*/*MMS21* mutant root apical meristems are disorganized and display an increased frequency of cell death. Molecular and genetic studies in Arabidopsis revealed that NSE2/MMS21 promote G1/S and G2/M transitions by destabilizing E2Fa/DPa transcription factor complex and promoting CYCLINB1;1, respectively (Ishida et al., 2009; Liu et al., 2016). However, NSE2/MMS21 affects other pathways during root development. The NSE2/MMS21 mutants show reduced response to exogenous cytokinin and down-regulation of transcription factors from cytokinin-induced ARABIDOPSIS RESPONSE REGULATORS (ARR) family (Huang et al., 2009). There is also missregulation of stem cell niche-defining transcription factors (Xu et al., 2013)

and a recent study revealed that NSE2/MMS21 activity is required for high levels of BRAHMA chromatin remodeling factor and thus normal root development (Zhang et al., 2017). The phenotypes of *NSE2*/*MMS21* mutants are strengthened by application of exogenous DNA damaging factors, suggesting that inability to process particular types of toxic DNA structures represents another challenge (Xu et al., 2013).

Recently, NSE2/MMS21 was identified as floral repressor (Kwak et al., 2016). The *mms21* mutant flowered earlier under both long and short day conditions, it had reduced amount of transcript and protein of the key floral repressor FLOWERING LOCUS C (FLC) and an increased transcript amount of the floral inducers SUPPRESSOR OF CONSTANS (SOC1) and FLOWERING LOCUS T (FT). FLC is the direct upstream regulator of FT, which then regulates SOC1 (Andres and Coupland, 2012). This indicates that the SMC5/6 complex promotes FLC transcription. This could occur via interaction or competition with Polycomb Repressive Complexes and/or LIKE-HETEROCHROMATIN PROTEIN 1 (LHP1), which are important modulators of FLC transcription (Bastow et al., 2004; Mylne et al., 2006; Schubert et al., 2006a). Beside of altering FLC transcription, NSE2/MMS21 also SUMOylates FLC protein.
Collectively, this suggests that NSE2/MMS21 prevents precocious flowering in Arabidopsis.

Once the decision to flower is reached, plants undergo a series of complex developmental events including production of gametes and seeds. The SMC5/6 complex plays critical role during multiple stages of generative development. *NSE2*/*MMS21* mutants showed lagging chromosomes and occasional anaphase bridges at meiotic metaphase I, indicating genome instability in male meiosis (Liu et al., 2014). In addition, several transcripts for meiotic genes related to chromosome maintenance and recombination were altered in *NSE2*/*MMS21* mutants (Liu et al., 2014). At the end of meiosis, *NSE2*/*MMS21* mutant plants developed not only tetrads but also diads with large nuclei, which produced a smaller number of pollen, with poor germination and abnormal tube growth. Hence, NSE2/MMS21 activity is required for successful male gametogenesis. The role of SMC5/6 complex in plant meiosis is far from being understood. Currently it is unknown whether other subunits take part in this process and there are also no data concerning a potential female meiotic phenotype of SMC5/6 mutants.
Fully developed micro- and mega-gametophytes, represented by pollen grains and ovules with mature embryo sac, respectively, allow double fertilization (egg and central cell) and give rise to seeds. Seeds are important propagation units of plants and an important source of nutrition for humans (Bewley, 1997; Tanksley and McCouch, 1997). There is accumulating evidence

that the SMC5/6 complex plays key role in seed development. Homozygous mutants in multiple complex subunits: *SMC5* (alias *EMBRYO DEFECTIVE 2782*), *NSE1* (alias *EMBRYO DEFECTIVE 1379*), *NSE3* and *SMC6A SMC6B* double mutant do not produce viable seeds (McElver et al., 2001; Watanabe et al., 2009; Yan et al., 2013; Li et al., 2017). However, *NSE2*/*MMS21* and partially complemented *NSE1* and *NSE3* homozygous mutants produce 50% to 75% of wild-type seed set, but seed viability is reduced (Liu et al., 2014; Li et al., 2017). These seeds had typically an underdeveloped embryo, arrested in early stages of development, and a more or less over-proliferated endosperm (Li et al., 2017). Although the mechanism of SMC5/6 complex involvement in seed development is currently unknown, its similarity with the *TITAN* seed phenotypes of cohesin and condensin mutants (Liu et al., 2002; Tzafrir et al., 2002) makes it tempting to speculate that the underlying mechanism arises via combinatorial action of cohesin and SMC5/6 complexes.

1.7.2. Modulator of abiotic stress responses

It was reported that NSE2/MMS21 mutant plants show improved resistance to drought, while NSE2/MMS21 over-expressors are drought hypersensitive (Zhang et al., 2013). NSE2/MMS21 works as a negative regulator of proline biosynthesis, and drought tolerance is associated with higher proline concentrations, which could explain, at least in part, the phenotype observed. One of the responses to drought stress is abscisic acid (ABA) accumulation. NSE2/MMS21 expression is reduced upon ABA treatment. Mutations in *NSE2*/*MMS21* lead to upregulation of ABA-mediated stress responsive genes and to hypersensitivity to ABA, as indicated by stomatal aperture, seed germination and cotyledon greening assays. Finally, ABA-induced accumulation of SUMO-protein conjugates was reduced in *NSE2*/*MMS21* mutant. Altogether, this indicates that NSE2/MMS21 plays a role as negative regulator of ABA-mediated stress response, by SUMOylating ABA responsive gene products (or more likely their transcriptional activators/repressors in a mechanism proposed above for FLC regulation) and thus influences stomata opening (Zhang et al., 2013).

1.7.3. Suppressor of immune responses

Arabidopsis NONEXPRESSER OF PR GENES 1 (NPR1) is a key positive regulator of salicylic acid (SA)-mediated systemic acquired resistance (SAR) pathway essential for defense against microbial pathogens (Cao et al., 1997; Ryals et al., 1997). NPR1 function is critical for

expression of PATHOGEN RESISTANCE (PR) genes. Among suppressors of *npr1* phenotype (i.e. PR genes are up-regulated), an unknown protein named SUPPRESSOR OF NPR1-1, INDUCIBLE (SNI1) was identified (Li et al., 1999). Recently, purification of the SNI1 complex in Arabidopsis revealed that it is associated with an unknown factor termed ARABIDOPSIS SNI1 ASSOCIATED PROTEIN 1 (ASAP1) and the two SMC5/6 subunits SMC5 and SMC6B (Yan et al., 2013). Although ASAP1 and SNI1 do not share significant sequence homology with any proteins outside of the plant kingdom, modeling of their structure revealed that they are structurally highly similar to the yeast NSE5 and NSE6, respectively. Hence, ASAP1 and SNI1 are the putative functional orthologs of yeast NSE5 and NSE6 in plants acting as suppressors of SAR by unknown mechanism(s). Screening for *SUPPRESSOR OF SNI* (*SSN*), i.e. for the mutations reverting smaller size *sni1* mutant plants to a wild-type like phenotype, revealed the following genes: *SSN1* (*RAD51D*), *SSN2* (*SWS1*/*SHU2*), *SSN3* (*BRCA2A*), *SSN4* (*RAD17*) and *ATR* (Durrant et al., 2007; Wang et al., 2010; Song et al., 2011; Yan et al., 2013). This suggests that either excessive production of SA and/or reduced genome stability in the absence of functional SMC5/6 complex lead to an increased frequency of repair via a (possibly error prone) pathway represented by the *SSN* genes. After disrupting the function of the repair signaling components ATR and RAD17 and their putative downstream SSN effectors, the balance may be reestablished, allowing for normal plant growth.

1.8. Angiosperm development

1.8.1. Vegetative development and flowering

The angiosperm[4] *A. thaliana* has separate vegetative and reproductive growth phases. Following germination, the seedling goes through a phase of vegetative growth, named organogenesis, in which begins to produce additional organs (leaves, stems, and roots), starting from the root and shoot meristems (Jurgens, 1991; Mayer et al., 1991). Next, the shoot apical meristem produces leaves with little elongation forming a rosette. Then the apical meristem that produced leaves in the vegetative phase switches to producing flowers in the reproductive phase, generating the primary inflorescence shoot. Furthermore, after first flower primordia is

[4] A plant of a large group that comprises those that have flowers and produce seeds enclosed within a carpel

produced, the plant bolts due to increased internode elongation between the uppermost leaves and between flowers. Hence, the basal positions on the primary inflorescence shoot possess only a few cauline (stem) leaves, while the apical positions have an outsized number of flowers (Bowman, 1994).

The inflorescence is an open raceme with the development of individual flowers proceeding acropetally (from basal to apical positions). The structure of the flower consists on a calyx of four sepals and a corolla of four petals, six stamens (androecium) and one pistil (Müller, 1961). The male part (stamen) consists of an anther and a filament. The anthers dehisce longitudinally, resulting in self-fertilization. The gynoecium/pistil, occupies the center of the flower, is composed by a stigma, style and the ovary, which is the structure that contains the ovules or seeds precursors.

1.8.2. Pollen development

Male development in *Arabidopsis thaliana* involves the initiation and elaboration of the male structural organs, the development of normal gametes inside the locules of the anthers, and the interaction of mature pollen with the stigma, leading to pollen tube growth.

Mature pollen grains are formed inside anther locules by a process known as microsporogenesis. During microsporogenesis the diploid sporogenous cells differentiate as microsporocytes (pollen mother cells or meiocytes) which will undergo two meiotic divisions to form four haploid microspores. (Bowman, 1994). Thus, each diploid meiocyte gives rise to a tetrad of four haploid microspores and microsporogenesis is complete with the formation of distinct single-celled haploid microspores.

Next, each haploid microspore will develop further in a process called microgametogenesis. Here the formation of a large vacuole will cause a rapid increase in size and cause the displacement the microspore nucleus against the microspore wall (Bowman, 1994; Regan and Moffatt, 1990). The vacuole is reabsorbed and first mitotic division occurs, which results in the formation of a large vegetative cell and a small generative cell. Subsequently the generative cell divides once more by mitosis to form the two sperm cells. Maturation is accompanied by dehydration of both anther tissues and pollen grains. The mature pollen grains are tricolpate and tricellular.

1.8.3. Ovule development

In most angiosperms, the megaspore mother cell (MMC) undergoes meiotic divisions to form four megaspores. Three of these degenerate, and only one will constitute a functional megaspore (FM). Afterwards, the FM will undergo three mitotic divisions to form an 8-nucleate embryo sac, consisting in seven cells: the egg, two synergids, a single diploid central cell and three antipodals (Mansfield, 1991; Bowman et al., 1994). The egg and synergids are arranged in the micropylar chamber of the embryo sac, and form the egg apparatus. The central cell is located between the egg apparatus and the antipodals in the chalazal portion of the embryo sac. Each type of cell will perform a distinct function during embryo and seed development: the synergids play a role in pollen tube guidance, while egg cell and central cell will give rise to the embryo and endosperm, respectively. However, the role of the antipodals during seed development still remains elusive.

1.8.4. Fertilization, embryogenesis and seed development

The gametophytic development concludes in fertilization and the beginning of development of the next sporophytic generation. Pollination is the result of a series of events, which include pollen contact and adhesion to the stigma, hydration, pollen tube germination, pollen tube growth and travel through the stigma, style and ovary and double fertilization (Bowman, 1994). Here, one sperm cell fuses with the egg cell and develops into the embryo –in a process called embryogenesis, while the second sperm cell fuses with the central cell and will proliferate into the endosperm (Sargant, 1900). From this moment on, both tissues follow very distinct pathways, where embryo represents the next generation, while endosperm nourishes and protects the embryo, controls its growth and acts as inter-ploidy reproductive isolation barrier (Johnston et al., 1980; Erilova et al., 2009).

All major parts form in growing embryo during early seed development. In parallel, endosperm stimulated by the maternally expressed AGAMOUS-LIKE (AGL) transcription factors rapidly divides and proliferates into a multi-nucleate syncytium. Soon after, AGLs become epigenetically silenced by the activity of maternally expressed Polycomb Repressive Complex 2 (PRC2) and syncytium will cellularize. This is an important step necessary for further development, in which the embryo absorbs almost the entire endosperm in many dicotyledons. By the end of embryogenesis, dicotyledon plants made two meristems, one that will give rise to the above ground shoot system and another that will form the root system (Bowman, 1999).

Once the embryo is fully developed, orthodox seeds undergo desiccation, in which a strong reduction of water takes place. Finally, when environmental conditions are favorable, orthodox seeds increase their water content, restart the cellular machinery and embryo will germinate from its seed (Bewley, 1997).

2. AIMS OF THE PROJECT

The variety of plant SMC5/6 complex mutant phenotypes suggests that this complex participates in multiple developmental and cellular pathways possibly linked to stress responses. Currently, the composition of plant SMC5/6 complex, the roles of individual subunits and their functional requirement in cellular and developmental processes (beside DNA damage repair) are poorly characterized. In an effort to obtain a comprehensive understanding of Arabidopsis SMC5/6 complex, this project aims to characterize the functional role of putative *NSE4A* and *NSE4B* subunits.

3. MATERIALS AND METHODS

3.1. Plant Material

A. thaliana wild-type and mutants were in Col-0 background: *nse4a-1* (SALK_057130), *nse4a-2* (GK-768H08), *nse4b-1* (SAIL_296_F02), *nse4b-2* (GK-175D10), *smc6b-1* (SALK_ SALK_101968C), *hpy2-2* (SAIL_77_G06), *atr-2* (SALK_032841C) and *lig4* (SALK_044027C). We also used cyclin-GUS containing the *pCYCB1;1::CYCB1;1:GUS* construct (Colón-Carmona et al., 1999). For promoter reporter constructs, regions 18943545 to 18941640 and 7260588 to 7258919 bp upstream of *NSE4A* and *NSE4B* transcription start sites, respectively, were PCR amplified, cloned into pDONOR221 and recombined into binary gateway vector pGWB553 containing β-glucuronidase gene. The final plasmids were transformed into *Agrobacterium tumefaciens* strain GV3101 and then into Arabidopsis Col-0 by floral dip method (Zhang et al., 2006). T1 generation seeds were screened on ½ MS plates containing hygromycin B 25 µg/L (Duchefa Biochemie) and resistant plants were transferred to soil. T2 generation with approximately 75% resistant seedlings, indicating single locus T-DNA insertions, were considered for further analysis. For promoter swap analysis, the *NSE4A* promoter and genomic region of *NSE4B* were PCR amplified and cloned into pGWB550 vector by MultiSite Cloning Gateway (Thermo Fisher Scientific). Plant transformation and selection was carried out as described above, which the exception that the construct was transformed into *nse4a-2* background. For NSE4A-fluorescent protein translational fusion, the *NSE4A* promoter, coding sequence, terminator and *VENUS* N-terminal tag and a BASTA resistance cassette were cloned using Gibson assembly (New England Biolabs) into pGGA000, pGGC000, pGGE000, pGGB000, pGGF000 respectively, to generate entry clones. The Greengate cloning reaction was performed as described (Lampropoulos et al., 2013) and the multi entry cassette was assembled into pAGM4723 backbone. The *nse4a-2* mutant plants were transformed with this construct as described above. For *nse4a-2* complementation analysis, the *NSE4A* promoter and genomic region of *NSE4A* were PCR amplified and cloned into pGWB550 vector by MultiSite Cloning Gateway (Thermo Fisher Scientific). Plant transformation and screening of transformants was carried out exactly as for the promoter swap experiment.

3.2. Phylogenetic analysis

NSE4 protein sequences were retrieved from NCBI and Phytozome (Supplemental Table 2). Protein alignment was performed using MUSCLE algorithm (Edgar, 2004) and the resulting alignment was submitted to Gblocks (Castresana, 2000). Curation and selection of aligned blocks was performed in Gblocks using less astringent parameters.

3.3. Drug Treatments

Sterilized seeds were evenly spread on sterile 1/2 Murashige and Skoog (1/2 MS) medium with or without zebularine (Sigma-Aldrich), MMC (Duchefa Biochemie), and bleocin (Calbiochem) in concentrations specified in the text and grown at 16 h light : 8 h dark at 21°C. Seven days old plants were used for root length measurements. For MMS experiment, sterilized seeds were grown in ½ MS medium for 5 days and then transferred to liquid ½ MS medium with and without 100 ppm MMS and grown for 26 days. Roots from five seedlings per genotype were straightened and in total three replicates were performed.

3.4. Nucleic acids isolation

For DNA isolation, leaf material of plants in rosette stage was harvested and DNA was isolated using DNeasy Plant Mini Kit (QIAGEN), following manufacturer's instructions. For RNA isolation, floral buds were collected, shock frozen in liquid nitrogen and kept at -80°C until use. Total RNA isolation was performed with QIAzol (QIAGEN), and the RNA integrity was assessed by formaldehyde agarose gel electrophoresis. cDNA synthesis was performed from 1 µg of total RNA as starting material, using RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific) with oligo dT primer according to manufacturer's instructions.

3.5. Primers

Primers used in this study are provided in Supplemental Table 4.

3.6. 3´RACE PCR

Four nested PCR reactions were performed using the primer combinations listed in Supplemental Table 5. The first PCR was performed using 1/100 dilution of cDNA synthesized

from *nse4a-2* mutant. Afterwards the PCR product was gel purified and used for the following nested PCR reaction. This step was repeated until the fourth reaction. PCR product obtained from the fourth reaction was cloned into a pJET1.2 vector and sequenced.

3.7. GUS histochemical staining

The staining protocol was adapted according to different tissues. Somatic tissues were stained as described (Liu et al., 2015). Inflorescences were dissected under a Leica MZ16FA stereomicroscope, fixed for 30 min in ice cold 4% formaldehyde in 1× PBS buffer, washed for 3×5 min in 1× PBS and infiltrated with GUS staining solution (Stangeland and Salehian, 2002) under vacuum. After 10-15 min. vacuum was released and samples were incubated at 37°C for 3 days, followed by overnight clearing in ethanol 70%. Subsequently, inflorescences were rinsed with water and mounted in petri dishes containing agarose and water. For staining of ovules and young seeds, developing siliques were first opened and fixed in 90% cold acetone at -20°C for 45 min. Afterwards, they were rinsed 3× with phosphate buffer 100 mM, transferred to GUS staining solution and stained at 37°C for 48h. After staining, siliques were quickly rinsed with phosphate buffer and mounted in 8:2:1 chloral hydrate solution. In order to avoid loss of signal where we observed weaker GUS staining, we performed a less severe clearing. We dissected pistils and immediately transferred them to GUS solution. Staining of ovules was performed as described (Vielle-Calzada et al., 2000). After clearing, mounted ovules where immediately imaged using Zeiss Zeiss AxioImager M2 microscope.

3.8. Hoyer's clearing

For DIC microscopy of seeds, 14 DAP siliques were dissected in a microscopy slide on 50 µL of Hoyer's solution (Anderson, 1954). Afterward, seeds were covered with a coverslip, and slides were stored in at 4°C for 16 hours. Imaging was performed using a Zeiss AxioImager M2 microscope.

3.9. Cell cycle arrest

The double-homozygous *nse4a-2* cyclin-GUS and *nse4b-2* cyclin-GUS plants were grown for 5 d in liquid ½ MS medium, transferred to liquid ½ MS supplemented with 10 µM zebularine

for 0, 1, 3, 6, 12, 24, 48 h, GUS-stained overnight, cleared in 70% ethanol and imaged using a Leica MZ16FA stereomicroscope.

3.10. Confocal Microscopy

For cell death analysis, seeds from transgenic lines were grown on vertically positioned plates with ½ MS medium for 4 days and then transferred for 1 day to liquid ½ MS media with 20 µM zebularine. Seedlings were stained with 10 mg mL^{-1} of Propidium Iodide (PI) solution (Sigma) for 3 min, followed by a rinsing step with sterilized water, and were placed on slides in a drop of water and then evaluated with a laser scanning confocal microscope Carl Zeiss LSM700. For subcellular localization of NSE4A in roots, transgenic lines expressing *pNSE4A::VENUS::NSE4A::tNSE4A* were grown for 5 days in either solid ½ MS or ½ MS supplemented with 10 µM zebularine. Afterwards, seedlings were stained with PI, following the same procedure as above and imaged with a Leica TCS SP8 confocal microscope (Leica Microsystems). For imaging of ovules, pistils were quickly dissected in a drop of water and ovules from different stages mounted in slide with a drop of water and placed on ice. After few minutes preparations were observed using a Leica TCS SP8 confocal microscope (Leica Microsystems)

3.11. Yeast-two-hybrid assay and bimolecular fluorescence complementation

The full-length coding sequence of Arabidopsis *SMC5*, *SMC6A*, *SMC6B*, *NSE1* and *NSE3* were PCR-amplified from cDNA and cloned into vector pGADT7 (Clontech), to produce a protein fusion with the GAL4 DNA AD in N-terminal orientation. The SMC5, *NSE4A*, *NSE4B, NSE1* and *NSE3* PCR fragments were cloned into pGBKT7, to produce a protein fusion with the GAL4 DNA BD. In order to avoid negative results due to interference of BD or AD domain with possible interactors, all genes were cloned into both C-terminal pGBKCg and pGADCg Y2H vectors, to produce both GAL4 DAN AD and BD fusion proteins C-terminally tagged; with exception of NSE4B which was only cloned in pGADCg vector.

The hinge and fragments of coils of SMC5 (corresponding to amino acids 415 – 699), SMC6A (amino acids 367 – 670) and SMC6B (amino acids 358 – 691) were cloned into pGBKCg and pGADCg vectors for testing interaction of the core subunits. The GAL4-based interaction was tested in the yeast strain AH109 (Clontech). Co-transformed yeast strains were selected on

synthetic defined (SD)/–Leu/–Trp (-LW) medium. Protein–protein interactions were tested using stringent (SD)/–Leu/–Trp/–His (-LWH) selection media supplemented with defined concentrations of 3-AT (Supplemental Table 3). The interaction between pGADT7-T and pBKT7-53 was used as the positive control, and that between pGADT7-T and pBKT7-LamC was used as the negative control.

For BiFC we used the same coding sequences as for Y2H experiments. The *SMC5*, *SMC5* (hinge) and *NSE3* were cloned into pBATL-nYFP, and *NSE4A*, *NSE4B*, *NSE1* and *SMC6B* hinge were cloned into the pBaTL-cYFP. Both plasmids produce C-terminal fusion proteins. *N. benthamiana* leaves were transiently transformed as described (Tian et al., 2011). YFP fluorescence was observed using Confocal Microscope Zeiss LSM 700.

4. RESULTS

4.1. *NSE4* is a duplicated gene in *A. thaliana*

Arabidopsis genome contains two putative uncharacterized NSE4 homologs: NSE4A (At1g51130; 403 amino acids) and NSE4B (At3g20760; 383 amino acids) sharing 65.1% protein identity (Figure 2). To identify the age of this duplication, we built NSE4 phylogeny across the green plants using *Schizosaccharomyces pombe* and human NSE4 proteins as outgroups (Figure 3).

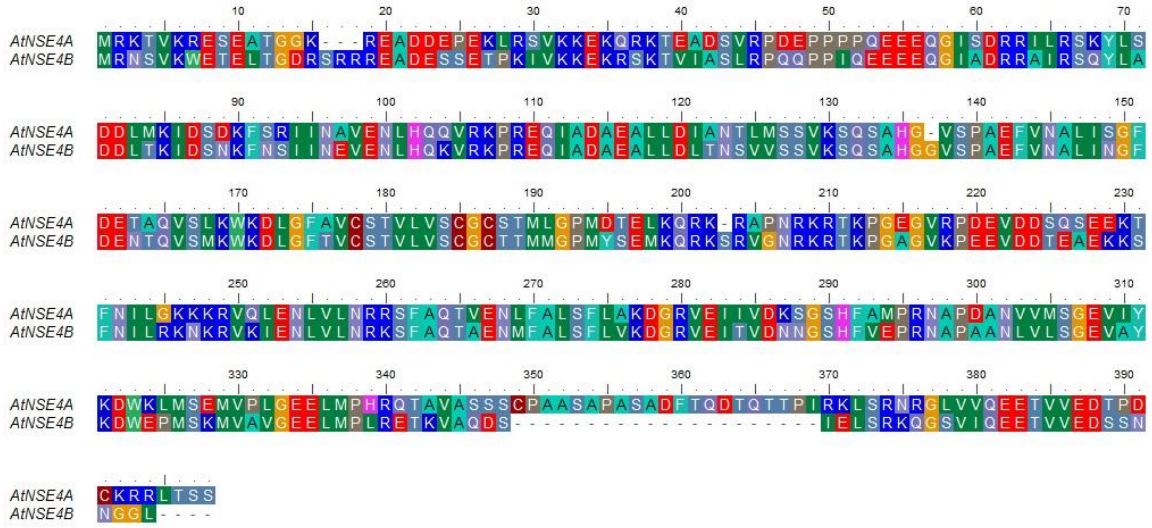


Figure 2. Alignment of Arabidopsis NSE4A and NSE4B proteins

Except for mosses carrying a single copy of NSE4, all other plant genomes contained at least two NSE4 copies. Orthologs of *A. thaliana* NSE4A and NSE4B occurred in *Arabidopsis lyrata, Capsella rubella* and *Eutrema salsugineum*. Both NSE4 copies of *Brassica rapa* were derived from NSE4A, while NSE4B was missing. This suggests that the Brassicaceae-specific NSE4A and NSE4B originate from the whole genome duplication event.

There was another set of conserved NSE4 paralogs in Poaceae, including *Brachypodium distachyon*, barley and maize. These paralogs most likely appeared during the whole genome duplication event ~70 million years ago (Paterson et al., 2009). We found in total six *NSE4* copies in rice and four in tomato. Some of these copies were short and grouping with more

distantly related species (Figure 2), raising a question on their origin and functionality. The high frequency of multiple *NSE4* copies per genome may indicate rapid *NSE4* sub- or neo-functionalization in different plant groups.

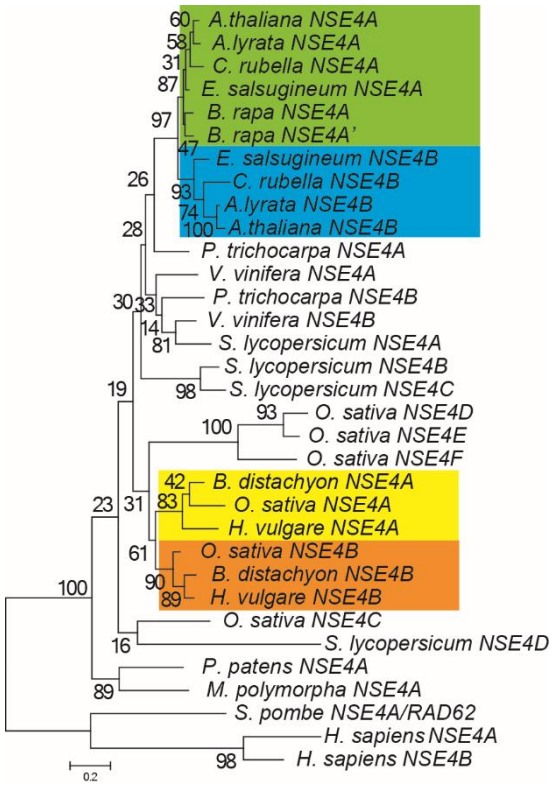


Figure 3. Phylogenetic three of *NSE4* homologs in green plants. *S. pombe NSE4*/*RAD62* and human NSE4 paralogs were used as outgroups. Brassicaceae and Poaceae *NSE4* duplications are indicated by the colored squares.

To assess the effects of NSE4 genes on plant growth, we isolated *NSE4A* and *NSE4B* mutant alleles carrying exonic insertions (Figure 4A). The *nse4a-1* allele carried T-DNA in the second exon and was most likely lethal, because we did not find any homozygous mutants in the progeny of *nse4a-1* heterozygous parent. However, we could recover viable homozygous *nse4a-2* plants with the T-DNA in the last exon, 56 amino acids before the stop codon (Figure 4B). 3' RACE revealed that in *nse4a-2* the *NSE4A* transcript continues into the T-DNA sequence and remains in the reading frame for 201 nt and, based on *in silico* reconstruction, adds 67 alien amino acids (Figure 4C). Therefore, *nse4a-2* most likely represents only a partial loss of function mutant allele with trimmed and modified C-terminus.

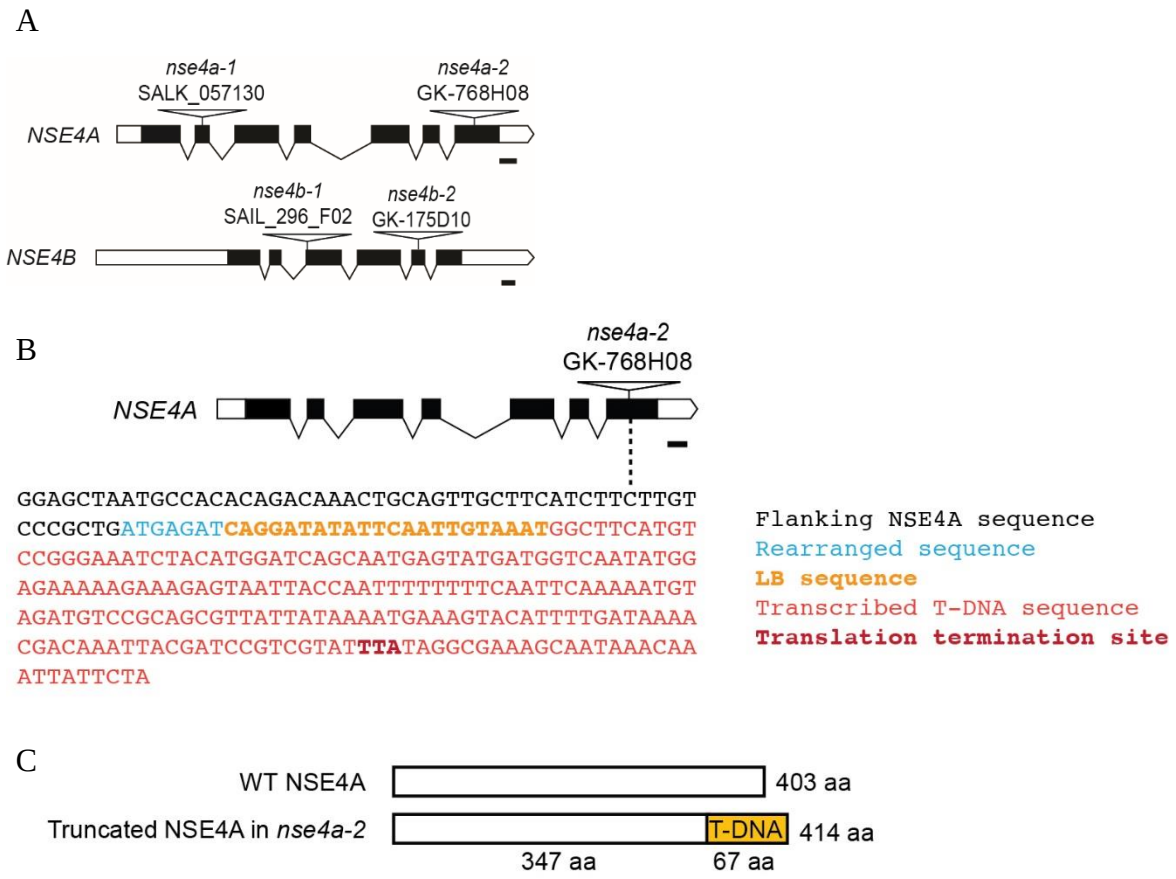


Figure 4. *NSE4A* and *NSE4B* mutant alleles.

(A) Gene structure of *A. thaliana NSE4A* and *NSE4B* with indicated positions of the mutations used in this study. Scale bars: 100 bp. (**B**) Scheme of *NSE4A* with indicated T-DNA position and 3' RACE recovered sequence in *nse4a-2*. (**C**) Schematic representation of NSE4A protein in wild-type and *nse4a-2*. In the later were the 56 C-terminal amino acids (aa) replaced by 67 amino acids derived from T-DNA.

Juvenile and non-flowering *nse4a-2* plants were smaller than wild-type plants (Figure 5A-B), but later recovered and were indistinguishable from the control plants during adult stage (Figure 5C). In contrast, both *nse4b* mutant alleles, carrying T-DNA insertions in the second intron (*nse4b-1*) and the fifth exon (*nse4b-2*), were viable and showed wild-type-like phenotypes (Figure 5A-C). Combining *nse4a-2* and *nse4b-2* alleles into a homozygous double mutant resulted in the *nse4a-2*-like phenotype suggesting that *NSE4A* and *NSE4B* do not act redundantly during vegetative stages.

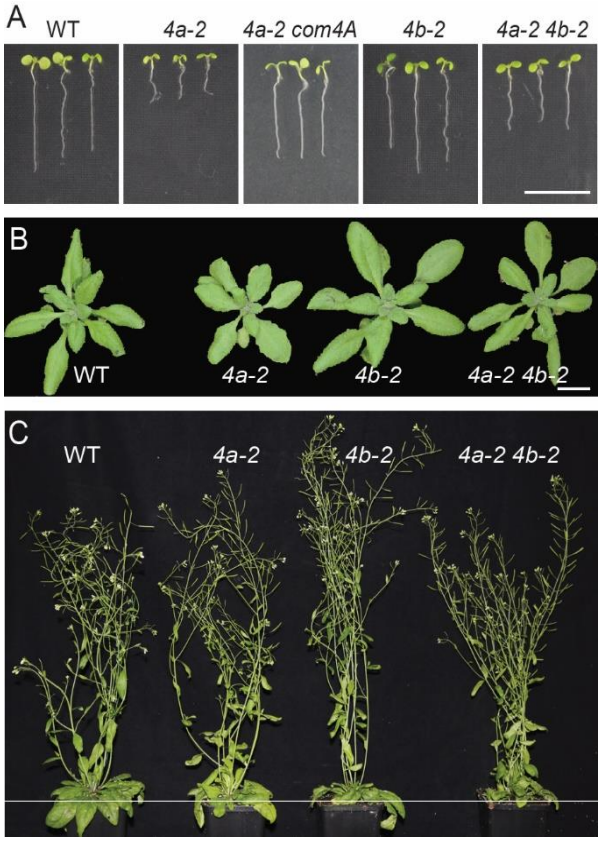


Figure 5. Phenotypes of homozygous wild-type (WT), *nse4a-2* (*4a-2*), *nse4a-2* complemented with *pNSE4A::gNSE4A* (*4a-2 com4A*), *nse4b-2* (*4b-2*) and *nse4a-2 nse4b-2* (*4a-2 4b-2*) plants. (A) Seedlings after one week in control *in vitro* conditions. (B) Three weeks old plants in soil and (C) six weeks old mature plants.

To characterize NSE4 expression pattern during Arabidopsis development, we generated stable reporter lines carrying NSE4A and NSE4B promoters fused with β-glucuronidase (GUS) coding sequence (*pNSE4A::GUS* and *pNSE4B::GUS*). The *NSE4A* promoter was strongly active in emerging true leaves and weakly in the cotyledon vasculature at seven days after germination (DAG; Figure 6A). In addition we observed signal in stele tissues within differentiated root zone, but there was no *pNSE4A* activity in the root meristems. At 14 DAG, *pNSE4A* was weakly active all over the aerial tissues (Figure 6B).

Flowers had GUS signals in sepals, upper half of the fully elongated anther filaments, pistils (ovules) and anthers (Figure 6C,D). In contrast, *pNSE4B::GUS* signals were restricted to the leaf stipules and a small domain within the root apical meristem at 7 DAG (Figure 6A, insets and red arrowheads), and this pattern remained unchanged during the whole vegetative stage

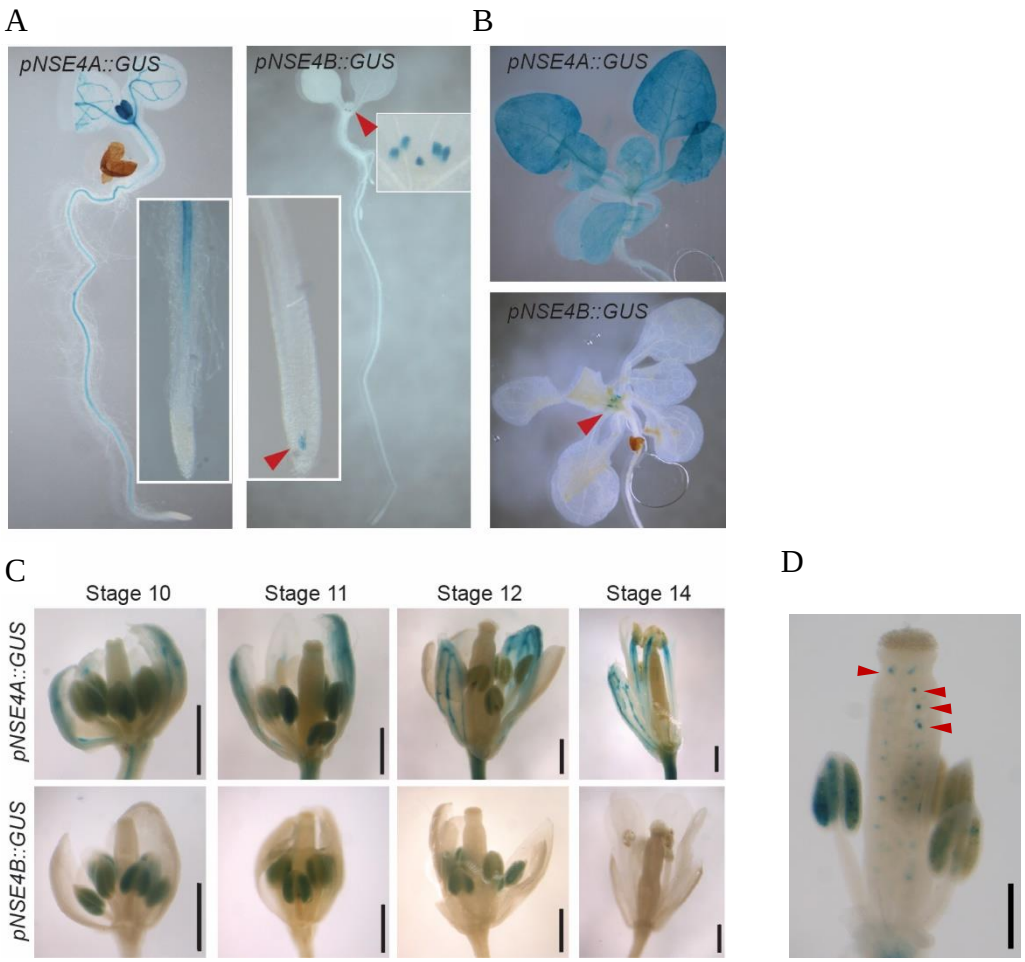
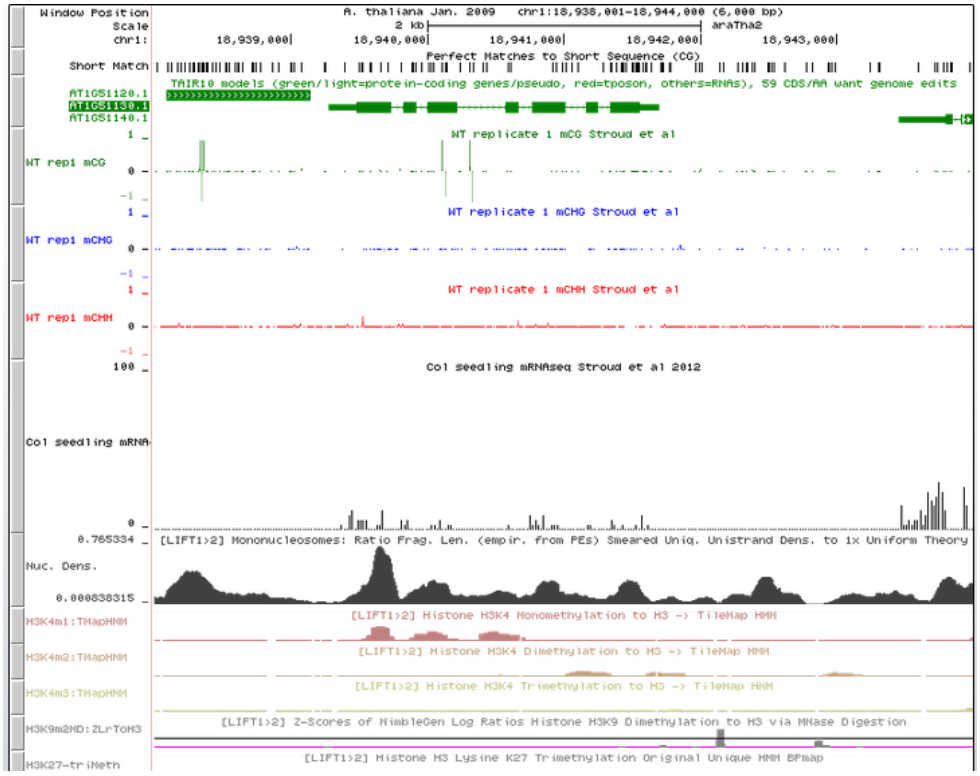


Figure 6. Analysis of *NSE4A* and *NSE4B* promoter activity using GUS reporter system. (A) One week old plants. Red arrowheads indicate signals in root meristematic zone and leaf stipules (upper inset). (B) Plants after 14 days of *in-vitro* culture. (C) Flowers at different stages. (D) Flower bud at stage 11 with signals in ovules indicated by red arrowheads.

(Figure 6B). During generative stage, *pNSE4B* was active in anthers between stages 10 and 12 (Figure 6C). A possible reason for the sparse *pNSE4B* transcription could be the significant association of the locus with a transcriptionally repressive environment, marked by histone H3 lysine 27 tri-methylation (Figure 7B).

A



B

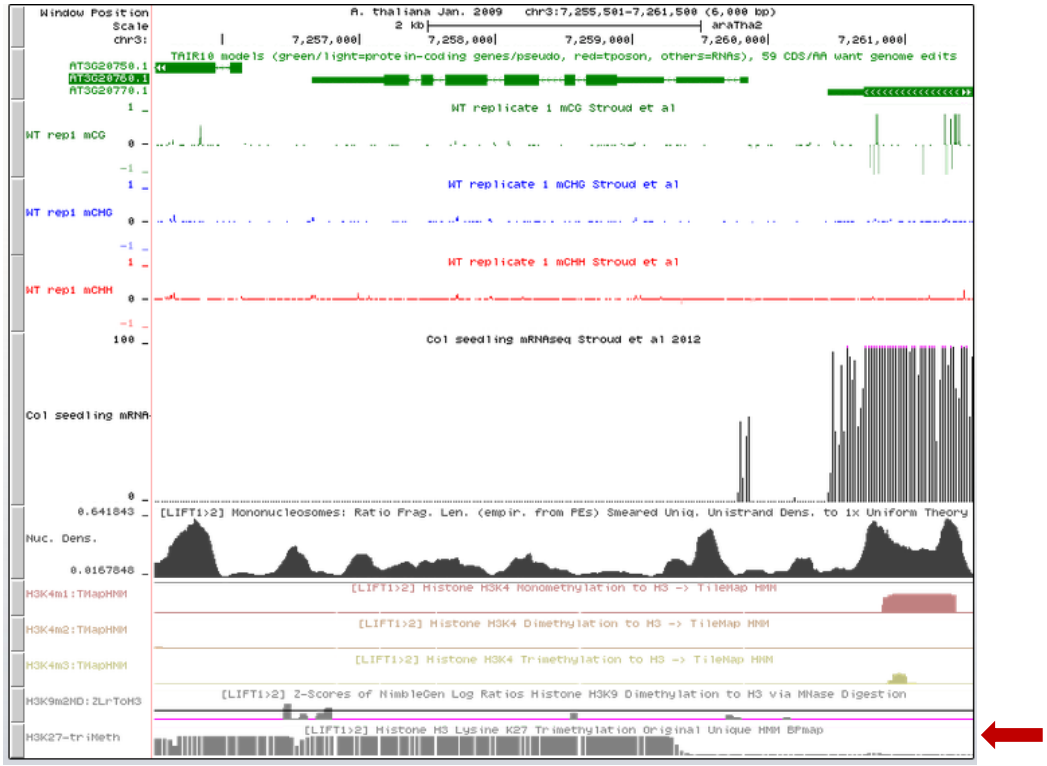


Figure 7. Chromatin environment of (A) *NSE4A* and (B) *NSE4B* genomic regions. The snapshots were made from Epigenome browser created by Jacobsen lab as UCLA, Davis. Both genes are displayed in complement strand. Red arrow indicates H3K27me3 island.

4.2. NSE4A expression analysis during ovule and seed development

The activity of *pNSE4A* and *pNSE4B* in flowers prompted us to analyze the generative stages in more detail. To get better insight into the presence of NSE4A protein, we expressed a translational fusion of NSE4A with VENUS under the control of its native promoter (*pNSE4A::NSE4A:VENUS*) in the *nse4a-2* background. Based on the full complementation of *nse4a-2* hypersensitivity to zebularine (Figure 8), we conclude that the addition of VENUS does not interfere with NSE4A functions.

.

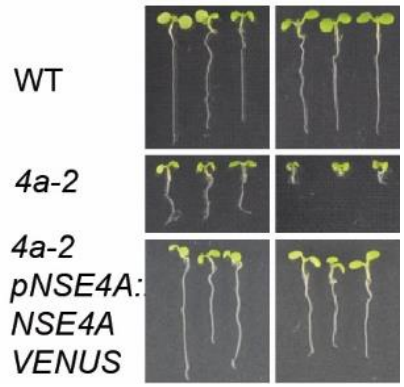


Figure 8. VENUS tag does not interfere with NSE4A functions. Wild-type (WT), *nse4a-2* and *nse4a-2* complemented with *pNSE4A::NSE4A:VENUS* were germinated and grown on the control and zebularine containing media for seven days.

Survey of the pollen developmental series revealed a strong and a weak activity of *pNSE4A* and *pNSE4B*, respectively (Figures 9A,B). The microspores (flower stage 10; (Bowman, 1994)) showed on average the strongest signals, which were decreasing in subsequent developmental stages, and there were practically no GUS signals in mature pollen from open pollen sacs in *pNSE4B::GUS* (flower stage 14). At the protein level, NSE4A was present in all pollen stages in the cell lineage leading to the generative cells as indicated by the VENUS signals in the single nucleus of the microspore, one nucleus of the bicellular pollen (flower stage 11) and two sperm nuclei of the stages 12-14 (Figure 9A). No NSE4A-VENUS signal in vegetative nucleus could be observed.

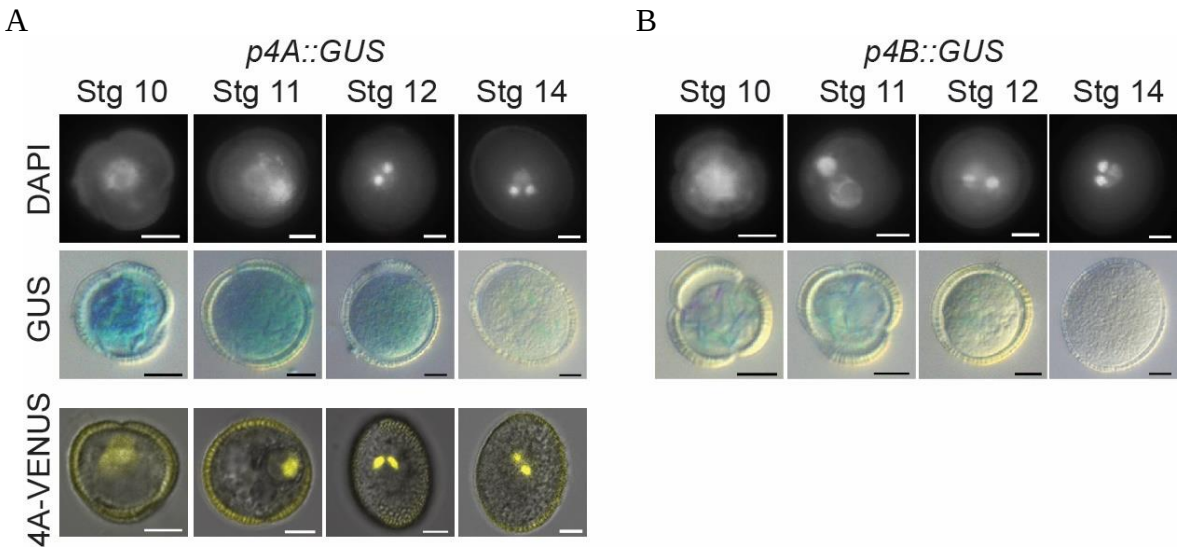


Figure 9. NSE4 expression analysis during pollen development

(A) The first two stacks show DAPI stained and GUS stained pollen of *pNSE4A::GUS* (*p4A::GUS*) reporter line. Stage (Stg) 10 corresponds to microspore, stage 11 to bicellular pollen, stage 12 to tricellular pollen and stage 14 to mature pollen from open anthers. The last stack presents *pNSE4A::NSE4A:VENUS* (4A-VENUS) signal in same stages of pollen development. Bar = 5 µm. **(B)** Transcription from *pNSE4B::GUS* (*p4B::GUS*) reporter line organized in the same way as (A). Bar = 5 µm.

During ovule development (Figures 10A,B), we observed *pNSE4A::GUS* signals in ovule primordia at flower stage 10, nucellus at stage 11 and embryo sac in stages 12 to 14. The transcriptional profile was largely in agreement with NSE4A protein accumulation (Figure 10B). NSE4A-VENUS gave strong signals in almost all cells of the nucellus, except for the megaspore mother cell, where the tagged protein was barely detectable (Figure 10B, flower stage 10).

However, NSEA-VENUS accumulated strongly in female meiocytes initiating meiotic prophase I (Figure 10B, flower stage 11, arrowhead). After pollination, *pNSE4A* is transcriptionally active in the embryo and the chalazal endosperm, and later (at 96 h after pollination) also in the syncytial endosperm (Figure 11). This corresponded well with the strong NSE4A-VENUS signals in developing embryos (Figure 10C), and also prominent localization to the nuclei of the syncytial endosperm (Figure 10D). By contrast to *NSE4A*, *pNSE4B* activity during early ovule development remained below threshold detection (Figure 10A).

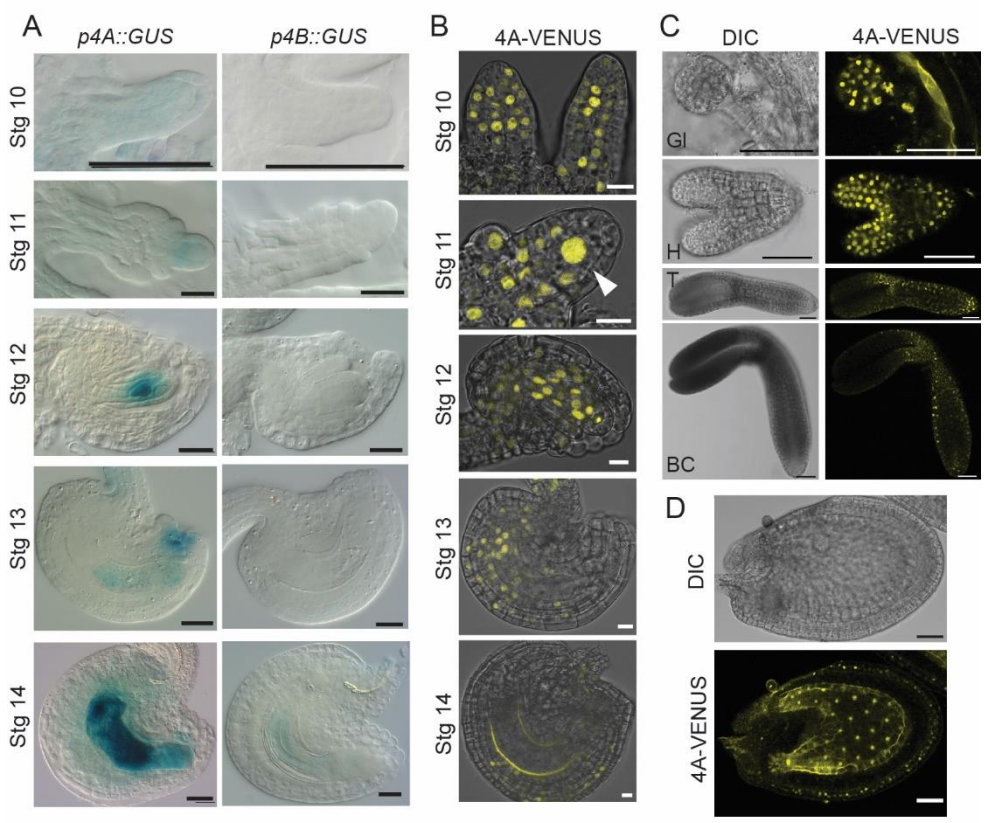


Figure 10. NSE4 expression analysis during ovule and seed development**. (A)** Activity of *NSE4A* (left) and *NSE4B* (right) promoters from ovule primordia to early post-fertilization. Stage 10 shows ovule primordia, stage 11 nucellus and stages 12 to 14 developing embryo sac. Bars = 50 µm. **(B)** NSE4A-VENUS signals in the same stages as described in (A). In ovule primordia of stage 11 is megaspore mother cell (MMC) almost without NSE4A-VENUS signal (arrowhead). However, is expression is greatly increased in the female meiocyte of the stage 11. Bar = 10 µm. **(C)** Accumulation of NSE4A-VENUS in globular (Gl), heart (H), torpedo (T) and bent cotyledon (BC) stage embryos. Left images represent differential interference contrast (DIC) and the right ones VENUS signal. Bars = 50 µm. **(D)** NSE4A-VENUS signals in endosperm 72 HAP. Regularly spaced dots represent nuclei of the syncytial endosperm. Bars = 50 µm.

Nevertheless, we detected weak *pNSE4B::GUS* activity in young embryo. *pNSE4B* activity reached its maximum at the globular stage (Figure 11).

To assess whether both maternal and paternal *NSE4A* copy can contribute to early embryo development, we analyzed *NSE4A* transcriptional reporter line (*pNSE4A::GUS*) in reciprocal crosses to wild-type plants in: (i) not-pollinated plants, (ii) self-pollinated plants, (iii) maternally-inherited reporter gene and (iv) paternally-inherited reporter gene (Figure 12). All

plants inheriting reporter locus maternally (i, ii and iii) had GUS signals in the central cell and synergids during the first 16 hours. This pattern was maintained in non-pollinated ovules (i), while the signals expanded to entire embryo sac after pollination (ii and iii). Paternal allele (iv) showed the first signals in embryo at 24 h after pollination and maternal signal remained stronger also at later stages, with dominating transcription in the globular to heart stage embryos and endosperm (Figure 12). This suggests preferential transcription from maternal NSE4A allele.

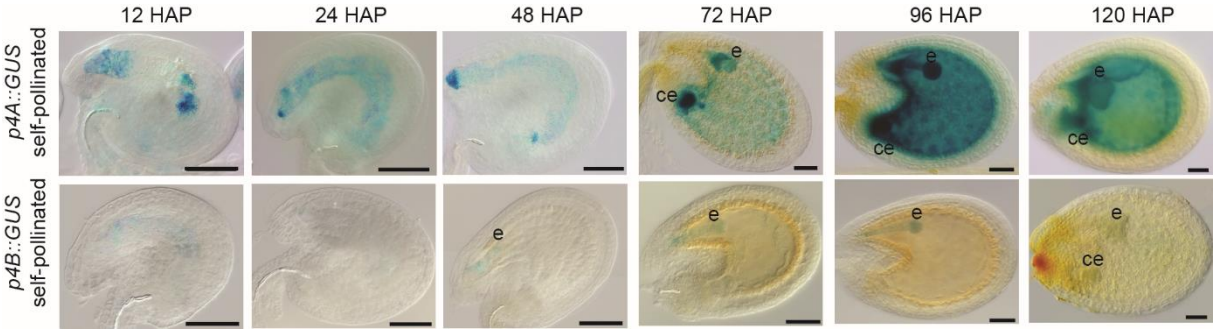


Figure 11. Transcription from *NSE4A* and *NSE4B* promoters at indicated number of hours after pollination (HAP). Reporter lines were pollinated with their own pollen 48h after emasculation. e = embryo, ce = chalazal endosperm. Bars = 50 µm.

In summary, this series of analyses confirmed NSE4A as nucleus-localized protein, as expected for a protein involved in genomic DNA repair, and revealed a dynamic expression pattern of NSE4A during sporogenesis, gametogenesis, embryogenesis and endosperm development. Possibly, high levels of NSE4A during meiosis and in the proliferating fertilization products might be linked with a DNA repair function e.g. in cross overs during meiosis or in genomic integrity processes in response to fast mitosis in the embryo and endosperm).

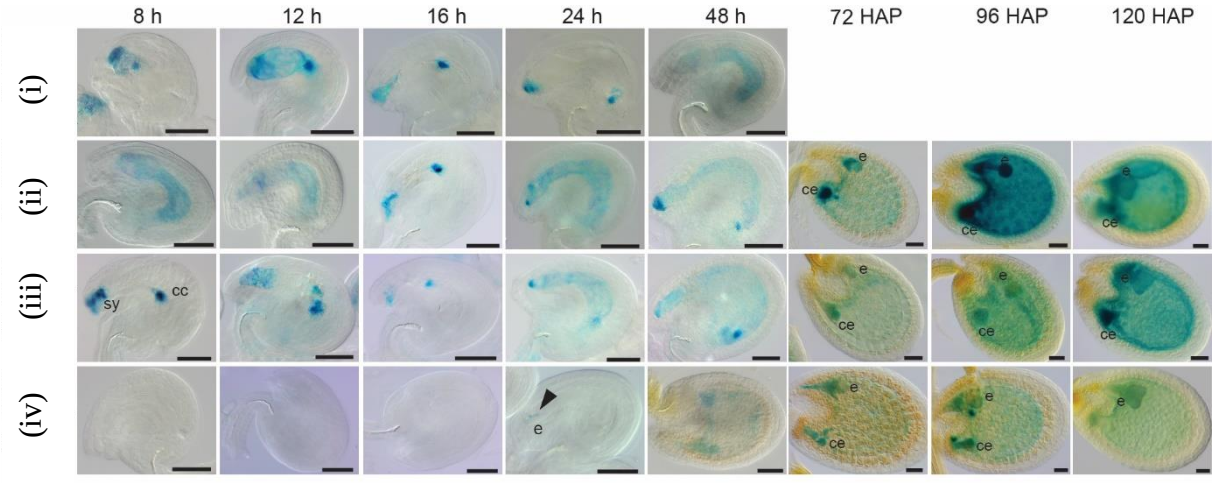


Figure 12. Parental contribution of *NSE4A* to seed development. *pNSE4A::GUS* was crossed reciprocally to wild-type plants. *NSE4A* activity in ovules from (i) non-pollinated plants, (ii) self-pollinated plants, (iii) maternally-inherited reporter gene and (iv) paternally-inherited reporter gene. Numbers indicate hours after pollination (HAP), starting 48h after emasculation. e – embryo (at 24h paternal transcription indicated by black arrowhead), ce – chalazal endosperm. Bars = 50 µm.

4.3. *NSE4A* plays role in seed development

Based on the NSE4 expression in seeds, we analyzed fertility of *nse4a* and *nse4b* mutants two weeks after pollination (Figure 13A,B). In contrast to wild-type plants, siliques from *nse4a-1/NSE4A* heterozygous plants produced 28.8% abnormal seeds (pale seeds underscoring delayed embryos and/or aborted seeds; n =1402, Figure 13A,B). Fertility was even more impaired in homozygous *nse4a-2* plants, with approximately half (53.4%) of the seeds developing normally, 22% early aborted ovules and 24.6% abnormally large seeds with glossy surface and liquid endosperm (n = 1008). Clearing of abnormal *nse4a-1* and *nse4a-2* seeds revealed embryo arrest at the heart and heart-to-torpedo stage, respectively (Figure 13C, black arrows). A *NSE4A* genomic construct could fully rescue *nse4a-2* mutant seed phenotype (up to 96.5% normal seeds, n = 949) confirming that seed lethality is a consequence from *NSEA* loss-of-function (Figure 13A,B). By contrast, and in agreement with *NSE4B* expression pattern, *nse4b-1* and *nse4b-2* single mutants were fully fertile, while the *nse4a-2 nse4b-2* double mutant phenocopied *nse4a-2* single mutant (Figure 13A,B). Hence, NSE4A is necessary for normal seed development while *NSE4B* is dispensable.

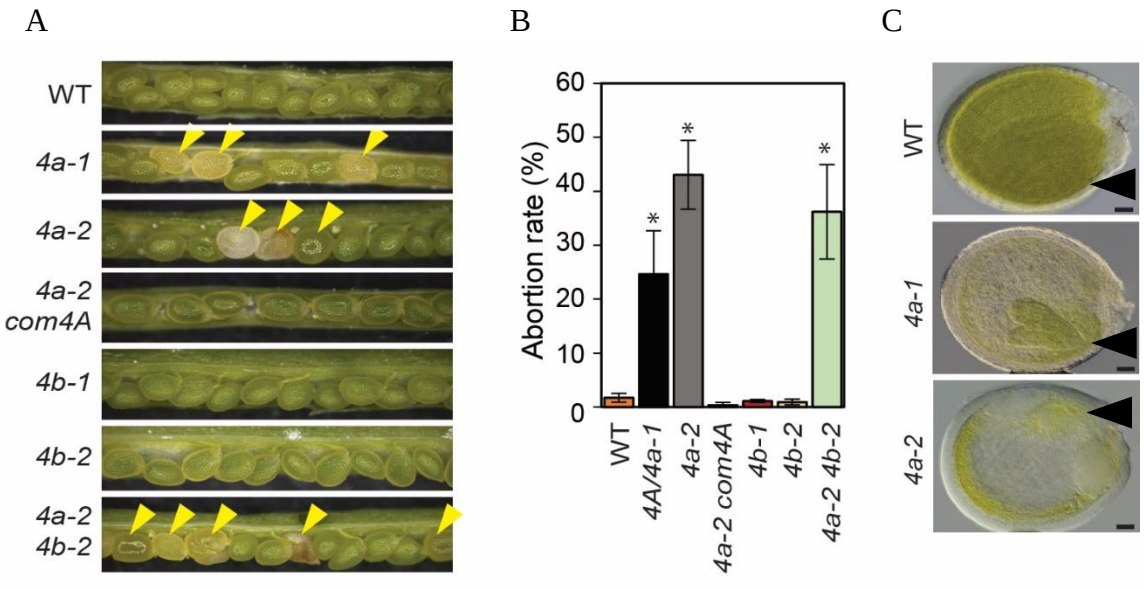


Figure 13. NSE4A is necessary for seed development. **(A)** Seed phenotypes in wild-type (WT), heterozygous self-pollinated *NSE4A*/*nse4a-1* (*4a-1*), and homozygous *nse4a-2* (*4a-2*), homozygous *nse4a-2* complemented with genomic *NSE4A* locus (*4a-2 com4A*), *nse4b-1* (*4b-1*), *nse4b-2* (*4b-2*), and homozygous *4a-2 4b-2* double mutant. Abnormally developing seeds are indicated by yellow arrowheads. **(B)** Quantification of the late aborted seeds in genotypes listed in (A). Error bars indicate standard deviation between means of three biological replicates. At least 140 seeds from a minimum of four siliques were analyzed in one replicate. Asterisks indicate significant differences relative to WT in Student's T-test at $P < 0.05$. **(C)** Equally old cleared WT, pale self-pollinated *NSE4A*/*nse4a-1* (*4a-1*) and large *nse4a-2* (*4a-2*) seeds. The *nse4-2* seed phenotype was variable and therefore we show multiple examples. Embryos were pointed by a black arrow. Bar = 50 µm.

4.4. *NSE4A* is involved in somatic DNA damage repair

Next, we tested which of the Arabidopsis *NSE4* paralogs is involved in DNA damage repair. First, we scored for the transcriptional response of *NSE4A* and *NSE4B* to drug treatment using the promoter-GUS reporter lines (Figure 14A). No induction was observed for *pNSE4B::GUS* upon treatment with DNA damaging agent zebularine (10 µM) generating enzymatic DNA-protein cross-links (A. Finke and A. Pecinka, unpublished data) (Figure 13A). In contrast, *pNSE4A* became activated throughout the entire meristematic zone and also in the emerging

lateral roots (Figure 14A). This transcriptional activation was accompanied by NSE4A protein accumulation as indicated by NSE4A-VENUS signals within the larger area of root apical meristems of the stressed reporter plants (Figure 14B).

Subsequently, we analyzed role of *NSE4* genes in DNA damage repair by growing wild-type, *nse4a-2* single mutant, *nse4a-2* complemented with *NSE4A* genomic construct (*pNSE4A::NSE4A::tNSE4A*), *nse4b* (both alleles) and *nse4a-2 nse4b-2* double mutant plants on media containing 10 µM MMC, 50 nM bleocin and 10 µM zebularine and monitored their growth (Figure 14C). In separate assay we applied also DNA alkylating agent methyl methane sulfonate (MMS; Figure 14E), which caused poor growth of Arabidopsis *smc6b-3* (*mim-1*) mutant (Mengiste et al., 1999). As positive controls for the drug sensitivity we used *ATR* signaling kinase mutant (*atr-2*), *LIG4* mutant (*lig4-1*), and mutants in the two SMC5/6 complex subunits *SMC6B* (*smc6b-1*) and *HPY2* (*hpy2-2*) (Ishida et al., 2009; Yuan et al., 2014; Liu et al., 2015). The *nse4b-1* and *nse4b-2* single mutant plants were indistinguishable from wild-type plants under all genotoxic treatments (Figures 14C,D). The *nse4a-2* single and *nse4a-2 nse4b-2* double mutants showed wild-type-like phenotypes under MMC and bleocin stress, but were strongly hypersensitive to zebularine and MMS (Figures 14C-E). The sensitivity of *nse4a-2* to zebularine was largely suppressed by its complementation with *NSE4A* genomic construct (Figure 14C,D).

Inhibition of root growth in response to DNA damage is frequently accompanied by increased cell death. Therefore, we monitored the amount of dead cells using the Propidium Iodide (PI) assay in control and 20 µM zebularine-treated plants (Figure 15A). While there was no or few dead cells in wild-type and *nse4b-2* plants *nse4a-2* and *nse4a-2 nse4b-2* double mutant plants showed a drastic increase upon zebularine treatment. The drug sensitivity phenotype (growth and cell death) of *nse4a-2* to zebularine is directly due to *NSE4A* loss-of-function as shown by complementation using an *NSE4A* genomic construct (Figure 15A).

We noticed that the root meristem was partially disorganized in zebularine-treated *nse4a-2* plants during PI assay. Therefore, we estimated the meristem size by counting the number of cortex layer cells between the quiescent center and the differentiation zone (Figure 15B). Wild-type and *nse4b-2* contained 38 to 45 cells, and this number did not change significantly after 24 h of 20 µM zebularine treatment (Student's T-test, $P > 0.05$). In contrast, *nse4a-2* showed a significant 31% reduction to 26 cells upon zebularine treatment ($P < 0.001$ in Student's T-test).

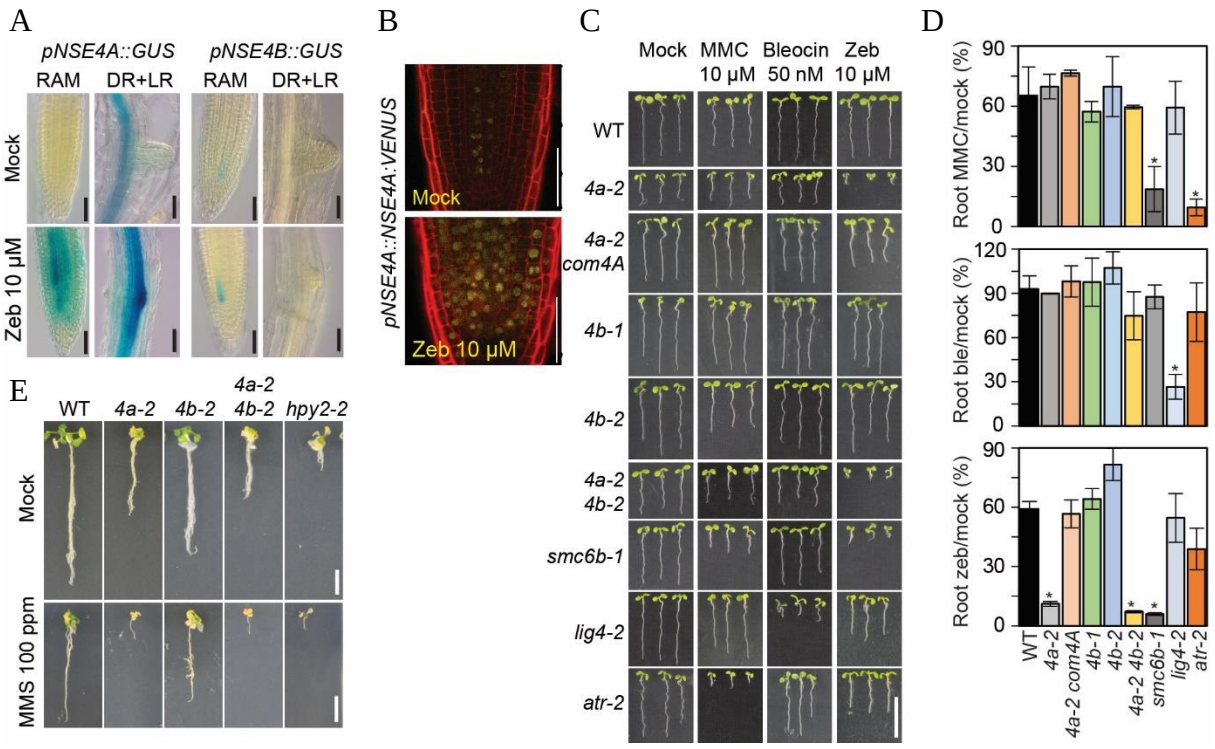


Figure 14. *NSE4A* is involved in somatic DNA damage repair in Arabidopsis. **(A)** Transcriptional response of *pNSE4A* and *pNSE4B* promoters after 24 h of 10 µM zebularine treatment in main root apical meristem (RAM) and differentiated root (DR) section with emerging lateral root (LR). Error bars = 50 µM. **(B)** *pNSE4A::NSE4A:VENUS* accumulation in root apical meristem under control and 10 µM zebularine conditions. Error bars = 50 µM. **(C)** Sensitivity to genotoxic stress. Wild-type (WT), *nse4a-2* (*4a-2*), *nse4b-1* (*4b-1*), *nse4b-2* (*4b-2*), *nse4a-2 nse4b-2* (*4a-2 4b-2*), *nse4a-2* complemented with *pNSE4A::gNSE4A*, *smc6b-1*, *atr-2* and *lig4-2* plants were germinated and maintained for one week on 10 µM MMC, 50 nM bleocin and 10 µM zebularine (zeb). Bar = 10 mm. **(D)** Quantitative data for (C) calculated as the relative root length under zebularine versus control conditions. Error bars are standard deviation between means of three biological replicates, each containing at least 20 plants. * = statistically significant differences in Student's T-test at $P < 0.05$. **(E)** Sensitivity to MMS. Representative phenotypes of WT, *4a-2*, *4b-2*, *4a-2 4b-2* double mutant and *hpy2-2* plants grown for one week in control liquid media and then for three weeks in control and 100 ppm MMS containing media. Bar = 10 mm.

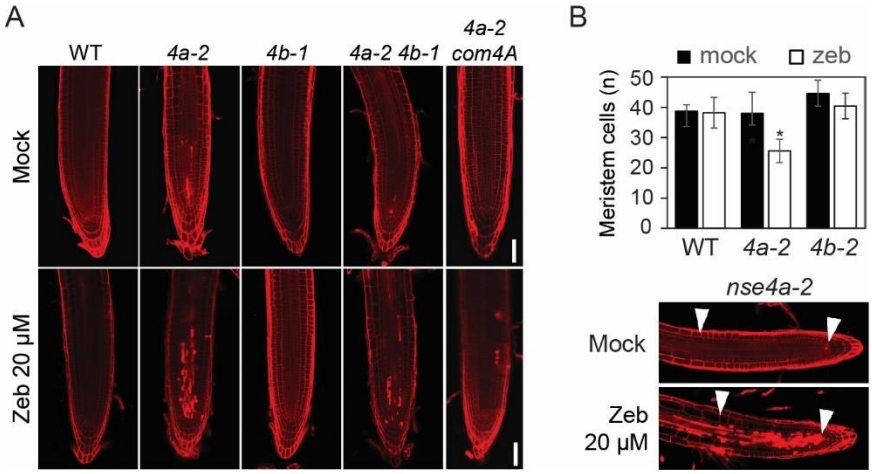


Figure 15. Cell death assay. **(A)** Propidium iodide stained roots from living Arabidopsis seedlings treated as control and by 20 µM zebularine for 24 h. **(B)** Meristem size estimation. Plants from (A) were used to estimate the number of cells within the root apical meristem (indicated by white arrowheads). Error bars in graph indicate standard deviation among 12 analyzed roots. * = statistically significant differences in T-test with $P < 0.05$ in comparison to mock-treated control.

To test an effect of the mutation on cell cycle regulation we introduced a G2/M phase reporter line, which utilizes a translational fusion between CyclinB1;1 and GUS (Colón-Carmona et al., 1999) into *nse4a-2* and *nse4b-2* mutant backgrounds. The chimeric protein accumulates specifically in G2 phase of cycling cells and is destroyed at the onset of mitosis resulting in loss of the signal. Double homozygous lines were exposed to 10 µM zebularine for up to 48 h, and the domain of GUS expression was monitored (Figure 16). The *nse4a-2* roots showed increased number of GUS positive cells already at 0 h, indicating a prolonged G2 phase. After 48 h treatment, the meristem of *nse4a-2* plants was largely damaged, as indicated by diffuse signals and root hairs close to the root tip. The response in *nse4b-2* and wild-type was slower, less severe and not different from each other (Figure 16).

Collectively, this shows that NSE4A responds to genotoxic stress and is likely involved in DNA repair of zebularine-induced DNA protein cross-links and is required to promote cell division in response to the genotoxic drug, possibly to actively propagate healed cells.

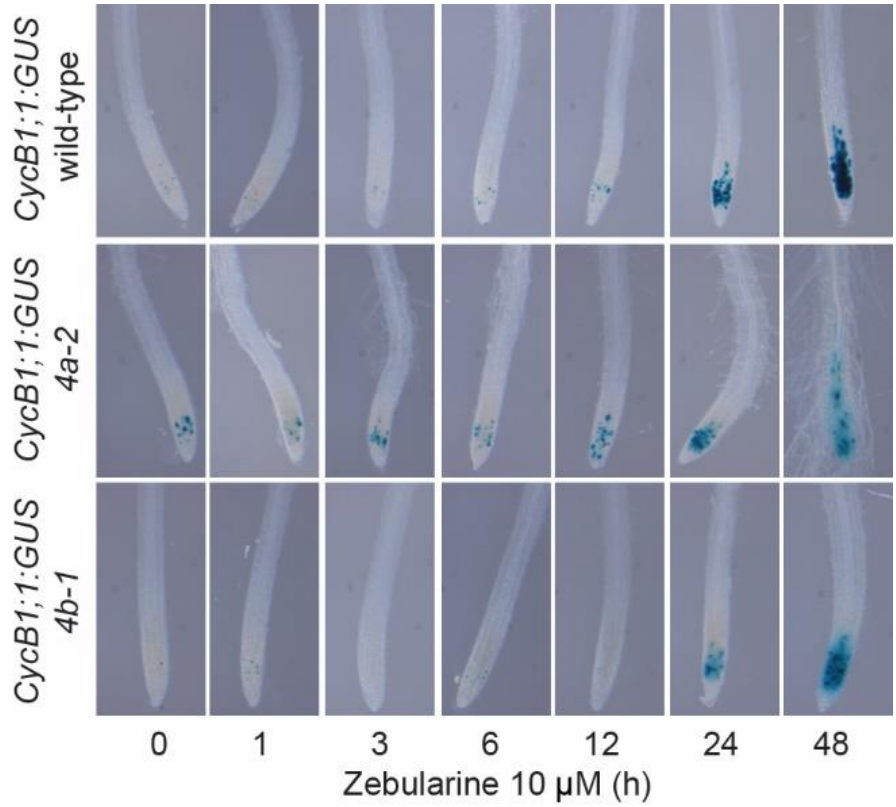


Figure 16. G2-M cell cycle progression in *nse4a-2* and *nse4b-2*. Analyzed by CycB1;1-GUS (CycB1;1::CycB1;1:GUS) after exposure to 10 µM zebularine for the given number of hours.

4.5. Both NSE4A and NSE4B interact with the same SMC5/6 complex subunits

The architecture of SMC5/6 complex remains unknown in plants. Based on fungal and animal models, we assume that NSE4 may act as a central subunit interacting with SMC5 and SMC6 and possibly also several NSEs (Duan et al., 2009; Hudson et al., 2011). To test whether this hypothesis holds true for both NSE4 paralogs we performed yeast-two-hybrid (Y2H) assays. The assay conditions were optimized using the positive (T+53) and the negative (T+lam C) controls, and we also suppressed protein auto-activation by adjusting the 3-Amino-1,2,4-triazole (3-AT) concentrations (Figure 16A; Figure 18; Supplemental Table 3). To test functionality of Y2H in our hands we first tested interaction between SMC5 and SMC6 paralogs. This confirmed our expectation that both SMC6A and SMC6B hinges can interact with SMC5 hinge.

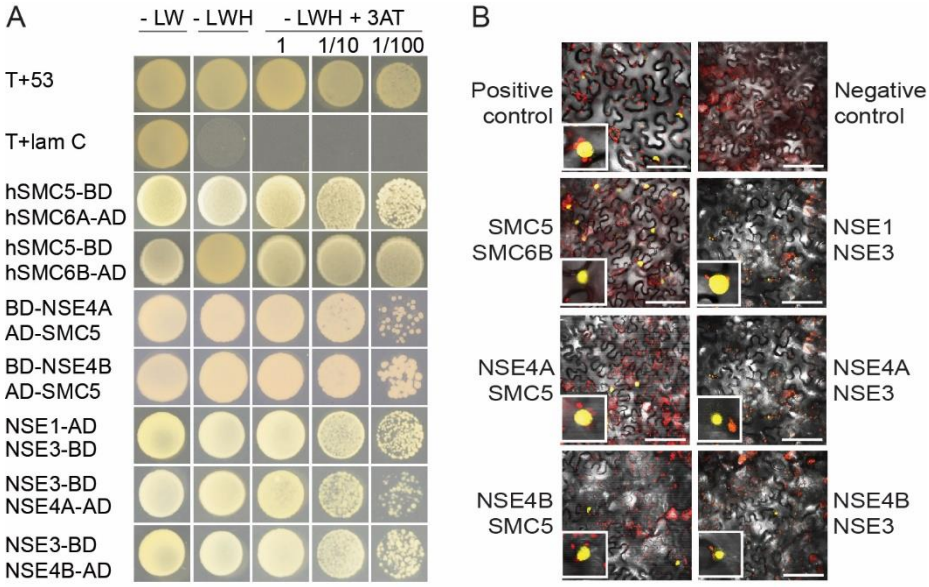


Figure 17. Analysis of protein-protein interactions.

(A) Yeast-two-hybrid (Y2H) assays. T+53 = positive control and T + lamC = negative control; BD = binding domain and AD = activation domain. Domain position before/after the gene name indicates N- or C-terminal fusions, respectively. Autoactivation controls, negatively tested combinations and used 3-AT concentration are provided in Figure 18 and Supplemental Table 3. **(B)** Bimolecular fluorescence complementation (BiFC) validation of interactions predicted by Y2H. Insets show nuclei with positive signals. Error bars = 50 µm.

Subsequently, we tested for interactions of the full length SMC5 or SMC6 with NSE4A and NSE4B. While the interaction between both NSE4 paralogs and SMC5 was positive (Figure 17A), we did not observe yeast growth in the tests for interactions with SMC6A and SMC6B. This remained true even after switching the tag positions (N- and C-terminal) and extensive optimization (Figure 18). Within the NSE1-NSE3-NSE4 sub-complex, we positively tested interactions of both NSE4 paralogs with NSE3, and confirmed (after Li et al., 2017) the interaction of NSE1 with NSE3 (Figure 17A).

However, we did not detect interactions between NSE4A or NSE4B and NSE1. To confirm the interactions predicted by Y2H, we performed Bimolecular Fluorescence Complementation (BiFC) assays in *Nicotiana benthamiana* and analyzed signals using confocal microscopy (Figure 17B).

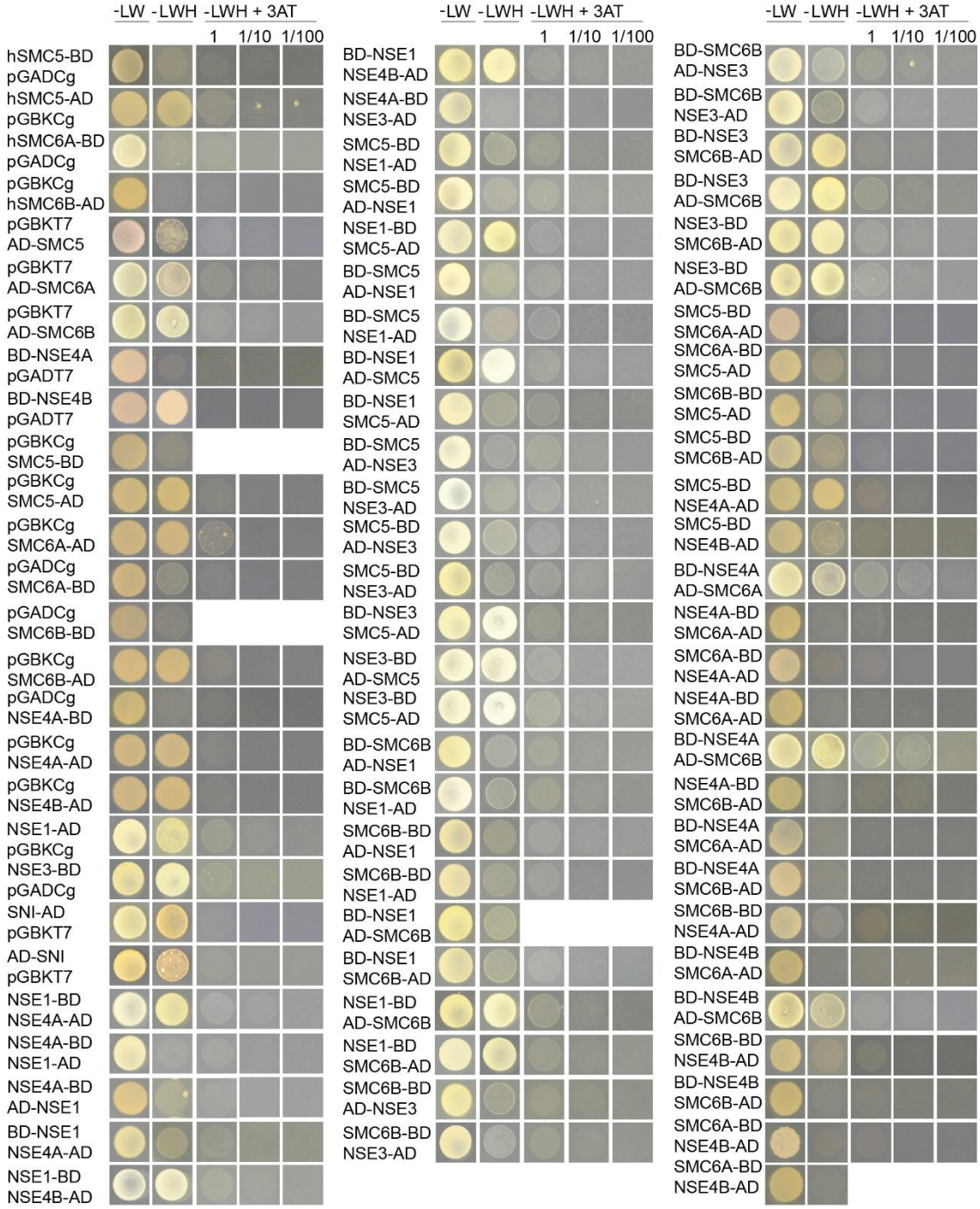


Figure 18. Yeast-two-hybrid autoactivation and negative results.

In all cases the signals showed nuclear localization and confirmed that both NSE4A and NSE4B are able to interact with SMC5 and NSE3. Based on the combination of Y2H and BiFC assays together with published data we propose a model for interactions between individual SMC5/6 subunits in Arabidopsis (Figure 19).

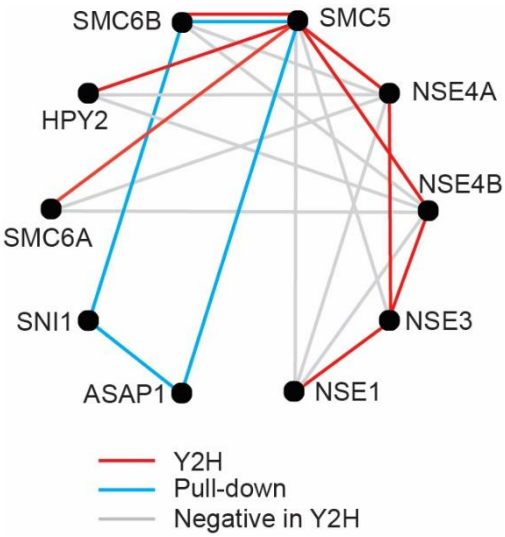


Figure 19. The model of protein-protein interactions within Arabidopsis SMC5/6 complex. Based on Y2H and BiFC (red lines) or pull down (blue lines; Yan et al., 2013) experiments. Negatively tested combinations are indicated by gray lines. Interaction between HPY2 and SMC5 was published previously (Xu et al., 2013).

4.6. NSE4B is only partially redundant with NSE4A

NSE4A and *NSE4B* paralogs show little overlap in their transcription patterns and loss-of-function phenotypes. To test whether *NSE4A* and *NSE4B* also diverged functionally, we developed a promoter swap construct consisting of *NSE4B* genomic coding sequence under the control of *NSE4A* promoter (*pNSE4A::gNSE4B::tNSE4B*). This construct was transformed into the homozygous *nse4a-2* plants and individuals heterozygous or homozygous for the promoter swap were selected in T2 and tested for zebularine sensitivity in T3 generation. While the control *nse4a-2* plants were strongly hypersensitive, several independent promoter swap lines showed a rescue, albeit incomplete, of the drug sensitivity phenotype with roots of intermediate average length between *nse4a-2* and wild-type plants (Figure 20A,B).

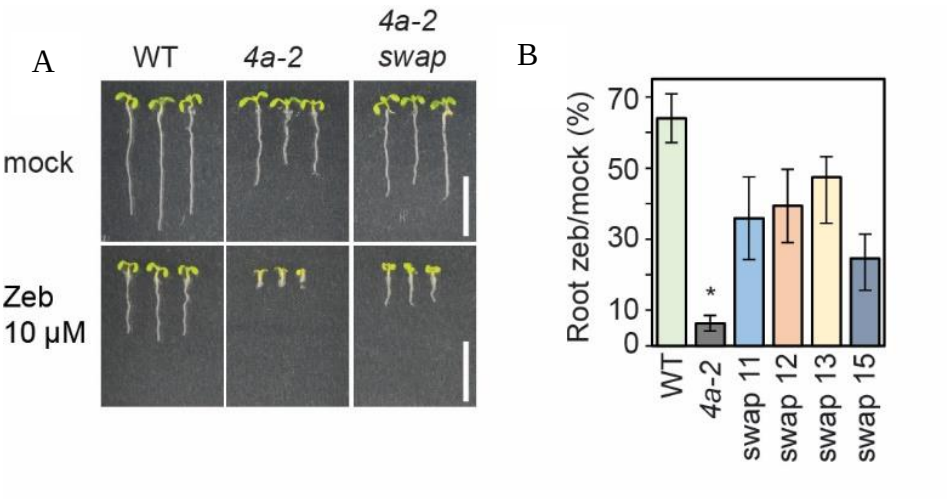


Figure 20. Analysis of NSE4B in DNA damage repair. **(A)** Zebularine hypersensitivity assay. Wild-type (WT), nse4a-2 (4a-2) and nse4a-2 complemented with pNSE4A::NSE4B::tNSE4B (4a-2 swap) line 13 were germinated and kept on control and 10 µM zebularine (Zeb) containing media for one week. Bar = 10 mm. **(B)** Quantitative data for root length of zebularine-treated versus control plants as described in (A). Lines 11, 12, 13 and 15 represent independent promoter swap transgenics. Error bars indicate standard deviation of the means from two biological replicates, each with at least 20 plants. * = statistically significant differences in Student's T-test at $P < 0.05$ relative to WT and nse4a-2.

In addition, analysis of seed phenotypes in the *nse4a-2* promoter swap line revealed absence of the typical seed defects, indicating rescue by more broadly expressed NSE4B. In addition, we introduced promoter swap into the heterozygous *NSE4A*/*nse4a-1* background and analyzed genotypes of 92 individual F2 plants by PCR (Tables 1 and 2). We found 22 out of 92 (24%) plants, which contained at least one copy of the promoter swap construct and were homozygous for *nse4-1* mutation, corresponding well to our theoretical expectation for promoter swap-mediated rescue (Table 2). This suggests that the promoter swap alleviates lethality of homozygous *nse4-1* genotype. In addition, promoter swap suppressed formation of the large pale seeds typical for *nse4a-1* mutant (Figure 21A).

Table 1. Punnett square indicating frequencies of genotypes in F2 generation of self-pollinated F1 hybrid *nse4a-1/nse4a-2 T/0*. Symbols: *nse4a-1 (a-1)*, *nse4a-2 (a-2)*, promoter swap transgene (T), no promoter swap (0). Colors indicate groups of plants with expected similar phenotype. Note that the combination *a-1/a-1 0/0* is lethal.

Gametes	***a-1 T***	***a-1 0***	***a-2 T***	***a-2 0***
a-1 T	*a-1/a-1 T/T*	*a-1/a-1 T/0*	*a-1/a-2 T/T*	*a-1/a-2 T/0*
a-1 0	*a-1/a-1 T/0*	*a-1/a-1 0/0 lethal*	*a-1/a-2 T/0*	*a-1/a-2 0/0*
a-2 T	*a-1/a-2 T/T*	*a-1/a-2 T/0*	*a-2/a-2 T/T*	*a-2/a-2 T/0*
a-2 0	*a-1/a-2 T/0*	*a-1/a-2 0/0*	*a-2/a-2 T/0*	*a-2/a-2 0/0*

Table 2. Promoter swap rescues *nse4a-1* lethality. Expected and observed frequencies of genotypes defined in Table 1. Expected lethal combinations were excluded from frequency calculations to allow for easier comparison with the frequencies among 92 PCR genotyped F2 plants. Dark green and red fields indicate congruence and discrepancy between the observed versus rescue or no rescue frequencies, respectively.

Genotype	**Expected**		**Observed (n = 92)**
	No rescue	**Rescue**	
a-1a-1 T/T or a-1a-1 T/0	0% (lethal)	20% (3/15)	24% (22/92)
a-1a-1 0/0	0% (lethal)	0% (lethal)	0% (0/92)
a-1a-2 T/0	50% (6/12)	40% (6/15)	36% (33/92)
a-1a-2 0/0	17% (2/12)	13% (2/15)	15% (14/92)
a-2a-2 T/0 or a-2a-2 T/T	25% (3/12)	20% (3/15)	16% (15/92)
a-2a-2 0/0	8% (1/12)	7% (1/15)	9% (9/92)

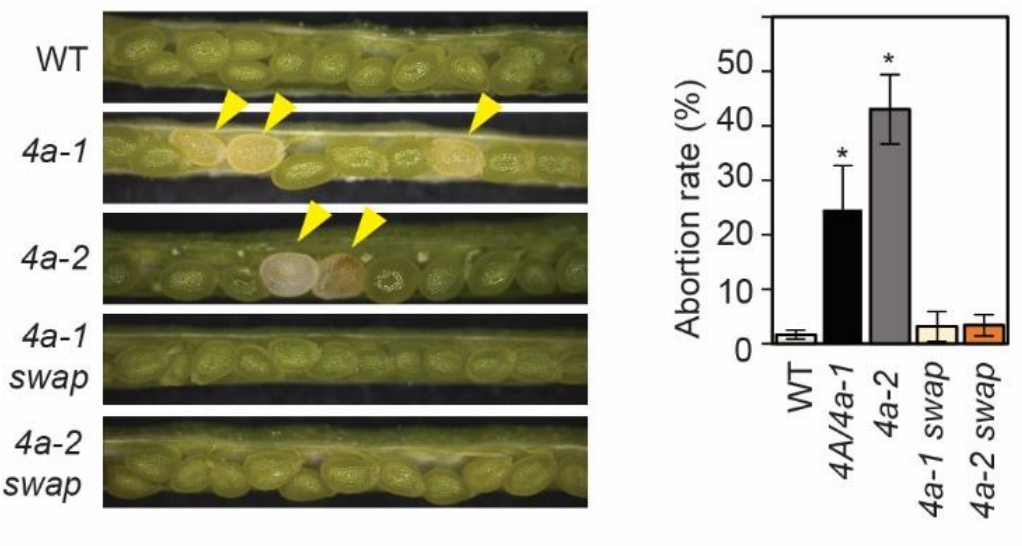


Figure 21. NSE4B promoter swap line alleviates seed developmental problems in *nse4a-1* and *nse4a-2* mutant background. **(A)** Analysis of seed development phenotypes in WT, heterozygous *NSE4A/nse4a-1 (4a-1)* and *4a-2*. The two bottom pictures show homozygous *nse4a-1* and *nse4a-2* containing homozygous swap 13 construct *(4a-1 swap* and *4a-2 swap).* Yellow arrowheads indicate aberrantly developing seeds. **(B)** Quantification of abortion rates in genotypes described in (A). Error bars indicate standard deviation between means of two biological replicates, each with at least 300 seeds. * = statistically significant differences in Student's T-test at $P < 0.05$ relative to WT.

5. DISCUSSION

SMC5/6 complex plays a crucial role in the maintenance of genome stability in eukaryotes (de Piccoli, Torres-Rosell, and Aragón 2009; Kegel and Sjögren 2010; Jeppsson et al. 2014; Diaz and Pecinka 2018). Some of its subunits remain poorly characterized in plants, including the two NSE4 homologs. In my project I demonstrated that *NSE4A* is involved in preserving genome stability and control of seed development, while *NSE4B*, despite sharing functional redundancy with *NSE4A* is barely active under normal development and drug stress, possibly via transcriptional suppression.

5.1. NSE4A is an essential gene in Arabidopsis

The *NSE4* paralogs of *Brassicaceae* species originate from the ~47 million years old whole genome (alpha) duplication event (Kagale et al., 2014). Surprisingly, there were at least two *NSE4A* copies in all analyzed vascular plants and the highest number of six copies in rice. The NSE4 amplifications are family-specific and much more frequent than the duplications of any other SMC5/6 complex member in plants (Diaz and Pecinka 2018). Our data from Arabidopsis and published data from humans (Hudson et al., 2011) suggest that at least some of these duplicated copies differ as to their transcriptional domains. We found that both NSE4A and NSE4B can interact with the core subunit SMC5 and NSE3, but not NSE1, where the last two represent the members of the NSE1-NSE3-NSE4 sub-complex (Palecek and Gruber, 2015). However, in spite of an extensive optimization we did not observe interactions of NSE4 proteins with SMC6B. This interaction seems particularly difficult to confirm as indicated by studies in *S. cerevisiae* and *S. pombe* (Duan et al. 2009; Palecek et al. 2006; J. Palecek, pers. comm.), possibly due to mutual organization of SMC6 and NSE4, steric disturbances by the tags or absence of an activating and/or stabilizing component.

Strong *NSE4A* mutation was homozygous lethal and self-pollinated heterozygotes showed 28.8% seed abortion. This resembles phenotypes of *SMC5*, *NSE1*, *NSE3*, and *ASAP1* mutants and *SM6A SMC6B* double mutant, which cause embryonic or early somatic lethality in Arabidopsis (Watanabe et al., 2009; Xu et al., 2013; Yan et al., 2013; Li et al., 2017). However, we found also a hypomorphic *nse4a-2* mutant allele, which likely produces a protein with a modified C-terminal part. This allele alleviate the problem of homozygous lethality

encountered in the loss-of-function allele *nse4a-1* hence enabling analysis of *NSE4A* functions during plant development and genotoxic stress. Its phenotypes partially resemble phenotypes of *HPY2* and *SNI1* mutants, which survive, but are strongly affected in development and fertility (Li et al., 1999; Huang et al., 2009; Ishida et al., 2009).

5.2. NSE4A is involved in sporogenesis, gametogenesis and the development of the fertilization products

We observed prominent and dynamic expression of *NSE4A* during Arabidopsis reproductive development. In the male gametophyte, NSE4A signal was tightly associated with the generative cells lineage and fully absent in the vegetative cell. This is consistent with the observation that the sperm nucleus is rich in the components of active chromatin control, while vegetative cell nucleus has lost multiple repressive chromatin modifications and will no longer divide (Schoft et al., 2009; Slotkin et al., 2009; Abdelsamad and Pecinka, 2014). However, the function of NSE4A in pollen development remains currently unknown as even the loss-of-function *nse4a* allele produced viable pollen. Possibly, *NSE4A* may secure faster or more accurate response, which is not detected in our laboratory conditions, upon environmental challenges affecting genome integrity in the germline.

NSE4A is broadly expressed in the ovule primordia, with a notable accumulation in the female meiocyte. Thus, beside its role in male meiosis (Liu et al., 2014), the SMC5/6 complex may play a role during female meiosis, possibly in the process of DNA replication, meiotic recombination or DNA damage repair. During embryo sac development and early fertilization, *NSE4A* was expressed in synergids and central cell and later in the embryo, the syncytial and the chalazal endosperm. *NSE4A* activity at these stages may be interpreted as the functional requirement for genome integrity safeguarding processes involving DNA repair as a consequence of the challenges posed by massive DNA replication and chromatin dynamics in these tissues (Baroux et al., 2007; Baroux and Autran, 2015). Genome integrity is necessary to ensure a proper differentiation and functioning of the new plant progeny individual and avoid the propagation of genetic mutations. In addition and not exclusively, the high levels of *NSE4A* in the syncytial endosperm may be interpreted as a requirement for detoxification from the endogenously occurring replication-derived toxic DNA structures. DNA replication produces high frequency of intertwinings between nascent chromatids, DNA supercoils and X-shaped

toxic DNA replication intermediates, which all require (to different extent) SMC5/6 functions for resolution (Diaz and Pecinka 2018; Jeppsson et al. 2014; Menolfi et al. 2015).

While SMC5/6 null mutations lead to an early seed abortion, hypomorphic NSE4 mutant produced large and ”glossy” seeds with a ”liquid” endosperm, which turned brown at later stages and died. Seed phenotype similar to *nse4a-1* or *nse4a-2* was observed in *NSE1*, *NSE3* and *MMS21*/*HPY2* mutants (Liu et al., 2014; Li et al., 2017). Hence, all SMC complexes participate in early seed development in Arabidopsis. Study in *S. cerevisiae* revealed that SMC5/6 complex is loaded by the SCC1 subunit of the cohesin complex to the specific sites during DNA replication (Jeppsson et al. 2014). This could explain the similarity of SMC5/6 and cohesin mutant seed phenotypes, and indicates that both complexes cooperate during seed development.

5.3. NSE4A but not NSE4B is required for resistance to genotoxic stress

SMC5/6 complex functions are widely associated with maintenance of genome stability (Kegel and Sjögren, 2010; Wu and Yu, 2012; Jeppsson et al., 2014a); however, it was not clear which of the Arabidopsis NSE4 paralogs confers this function. We observed activation of *NSE4A*, but not *NSE4B* in response to genotoxic treatments with drugs inducing various types of DNA damage. In addition, viable and phenotypically almost wild-type-like *NSE4A* mutant was hypersensitive to the cytidine analog zebularine and alkylating agent MMS, but not to other treatments. Lack of sensitivity to bleocin and MMC could be caused by the fact that the used allele is not null and/or that such damages can be processed by SMC5/6-independent pathways.

We previously showed that *SMC6B* mutants are hypersensitive to zebularine-induced damage (Liu et al., 2015). This suggests that SMC5/6 complex is essential for detoxification from complex toxic structures such as zebularine-induced enzymatic DNA-protein cross-links (A. Finke and A. Pecinka, unpublished data). DNA repair in response to zebularine treatment is mediated by both *ATM* and *ATR* kinases (Liu et al., 2015), which are known to phosphorylate proteins at SQ/TQ motifs (Awasthi et al., 2015). NSE4A contains two adjacent TQ motifs at amino acids 361 to 365 (TQDTQ), which makes it a good candidate for direct target of phosphorylation by ATM and/or ATR.

Recent studies from the non-plant models suggest that SMC5/6 complex acts as an ATP-dependent inter-molecular linker, which helps resolving toxic DNA structures at late replicating sites and also prevents recombination between non-homologous sequences (Chiolo et al., 2011; Kanno et al., 2015; Menolfi et al., 2015). In Arabidopsis, SMC5/6 promotes association of sister chromatids and is required for normal levels of homologous recombination in Arabidopsis (Mengiste et al., 1999; Hanin et al., 2000; Watanabe et al., 2009; Yuan et al., 2014). In addition to its role in somatic DNA damage repair, there is emerging evidence that SMC5/6 complex plays also a role in meiosis (Liu et al., 2014) and immune responses (Yan et al., 2013). Our data indirectly support this association by observation that NSE4A amount strongly increases during female meiosis. However, the exact molecular mechanism of genome maintenance by SMC5/6 complex remains unknown.

5.4. NSE4B and NSE4A are diversified primarily transcriptionally and NSE4B is not responsive to DNA damage

The functions of *NSE4B* are less clear than those of *NSE4A* in Arabidopsis. *NSE4B* single mutants are morphologically indistinguishable from wild-type and do not worsen the phenotype of a weak *NSE4A* mutant. We provided ample evidence that *NSE4B* is silenced throughout most of the development, except for a small domain within the root apical meristem, leaf stipules and embryo up to globular stage. Based on the results of *in silico* analysis which revealed extensive coverage of *NSE4B* locus by histone H3 lysine 27 tri-methylation we hypothesize that *NSE4B* is controlled by the Polycomb Repressive Complex 2 (reviewed in Mozgova and Hennig, 2015).

To test the *NSE4B* function in the non-silenced state, we swapped its promoter with that of *NSE4A* and tested whether the *NSE4A*-like transcribed *NSE4B* complements the *nse4a* phenotypes. The seed developmental phenotype was complemented to a full extent, but we found only a weak rescue under the DNA damaging conditions. This may point to the dual functions of the SMC5/6 complex as described in budding yeast (Menolfi et al. 2015). The DNA damage-independent function during DNA replication and DNA damage-dependent function in DNA damage repair. Both *NSE4A* and *NSE4B* seem capable of performing the first function, while DNA damage repair can be done only by *NSE4A* in Arabidopsis.

SUMMARY

In conclusion, I have identified *NSE4A* as a novel gene involved in somatic DNA damage repair and seed development in Arabidopsis, and have provided evidence that NSE4A functions are essential and its full loss is lethal. Weak mutant allele displayed sensitivity to zebularine-induced DNA-protein cross-links and aberrant endosperm development resulting in frequent seed abortion. How SMC5/6 complex mechanistically functions in these processes remains to be addressed in future studies.

REFERENCES

Abdelsamad, A., and Pecinka, A. (2014). Pollen-specific activation of Arabidopsis retrogenes is associated with global transcriptional reprogramming. Plant Cell **26,** 3299-3313.

Alabert, C., and Groth, A. (2012). Chromatin replication and epigenome maintenance. Nat Rev Mol Cell Biol **13,** 153-167.

Anderson, L.E. (1954). Hoyer's Solution as a rapid mounting medium for bryophytes. Bryologist 57: 242–244.

Andres, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. Nat Rev Genet **13,** 627-639.

Andrews, E.A., Palecek, J., Sergeant, J., Taylor, E., Lehmann, A.R., and Watts, F.Z. (2005). Nse2, a component of the Smc5-6 complex, is a SUMO ligase required for the response to DNA damage. Mol Cell Biol **25,** 185-196.

Ampatzidou, E., Irmisch, A., O'Connell, M.J., and Murray, J.M. (2006). Smc5/6 is required for repair at collapsed replication forks. Mol Cell Biol **26,** 9387-9401.

Awasthi, P., Foiani, M., and Kumar, A. (2015). ATM and ATR signaling at a glance. J Cell Sci **128,** 4255-4262.

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

Barker, P.A., and Salehi, A. (2002). The MAGE proteins: emerging roles in cell cycle progression, apoptosis, and neurogenetic disease. J Neurosci Res **67,** 705-712.

Barnum, K.J., and O'Connell, M.J. (2015). Molecular mechanisms involved in initiation of the DNA damage response. Mol Cell Oncol 2, e970065.

Baroux, C., and Autran, D. (2015). Chromatin dynamics during cellular differentiation in the female reproductive lineage of flowering plants. Plant J **83,** 160-176.

Baroux, C., Pecinka, A., Fuchs, J., Schubert, I., and Grossniklaus, U. (2007). The triploid endosperm genome of Arabidopsis adopts a peculiar, parental-dosage-dependent chromatin organization. Plant Cell **19,** 1782-1794.

Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. Nature **427,** 164-167.

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. Plant J **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 Rep **15,** 446-452.

Beranek, D.T. (1990). Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutat Res 231, 11-30.

Bermejo, R., Doksani, Y., Capra, T., Katou, Y.M., Tanaka, H., Shirahige, K., and Foiani, M. (2007). Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. Genes Dev **21,** 1921-1936.

Bermudez-Lopez, M., Ceschia, A., de Piccoli, G., Colomina, N., Pasero, P., Aragon, L., and Torres-Rosell, J. (2010). The Smc5/6 complex is required for dissolution of DNA-mediated sister chromatid linkages. Nucleic Acids Res **38,** 6502-6512.

Bermudez-Lopez, M., Villoria, M.T., Esteras, M., Jarmuz, A., Torres-Rosell, J., Clemente-Blanco, A., and Aragon, L. (2016). Sgs1's roles in DNA end resection, HJ dissolution, and crossover suppression require a two-step SUMO regulation dependent on Smc5/6. Genes Dev **30,** 1339-1356.

Bermudez-Lopez, M., and Aragon, L. (2017). Smc5/6 complex regulates Sgs1 recombination functions. Curr Genet **63,** 381-388.

Bewley, J.D. (1997). Seed Germination and Dormancy. Plant Cell **9,** 1055-1066.

Bonner, J.N., Choi, K., Xue, X., Torres, N.P., Szakal, B., Wei, L., Wan, B., Arter, M., Matos, J., Sung, P., Brown, G.W., Branzei, D., and Zhao, X. (2016). Smc5/6 mediated sumoylation of the Sgs1-Top3-Rmi1 complex promotes removal of recombination intermediates. Cell Rep **16,** 368-378.

Bowman, J.L. (1994). Ovules. In Arabidopsis (New York:Springer), pp. 297-331.

Bowman, J.L., Mansfield, S.G., and Koorneef, M. (1994). Embryogenesis. In Arabidopsis (New York: Springer), pp. 349-401.

Branzei, D., Sollier, J., Liberi, G., Zhao, X., Maeda, D., Seki, M., Enomoto, T., Ohta, K., and Foiani, M. (2006). Ubc9- and mms21-mediated sumoylation counteracts recombinogenic events at damaged replication forks. Cell **127,** 509-522.

Branzei, D., Vanoli, F., and Foiani, M. (2008). SUMOylation regulates Rad18-mediated template switch. Nature **456,** 915-920.

Britt, A.B. (1996). DNA damage and repair in plants. Annu Rev Plant Biol 47, 75-100.

Britt, A.B. (1999). Molecular genetics of DNA repair in higher plants. Trends Plant Sci 4, 20-25.

Cao, H., Glazebrook, J., Clarke, J.D., Volko, S., and Dong, X. (1997). The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell **88,** 57-63.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol **17,** 540-552.

Chen, J.J., Mitchell, D.L., and Britt, A.B. (1994). A Light-Dependent Pathway for the Elimination of UV-Induced Pyrimidine (6-4) Pyrimidinone Photoproducts in Arabidopsis. Plant Cell **6,** 1311-1317.

Chen, Y.H., Choi, K., Szakal, B., Arenz, J., Duan, X., Ye, H., Branzei, D., and Zhao, X. (2009). Interplay between the Smc5/6 complex and the Mph1 helicase in recombinational repair. Proc Natl Acad Sci USA **106,** 21252-21257.

Chiolo, I., Minoda, A., Colmenares, S.U., Polyzos, A., Costes, S.V., and Karpen, G.H. (2011). Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. Cell **144,** 732-744.

Choi, K., Szakal, B., Chen, Y.H., Branzei, D., and Zhao, X. (2010). The Smc5/6 complex and Esc2 influence multiple replication-associated recombination processes in Saccharomyces cerevisiae. Mol Biol Cell **21,** 2306-2314.

Chomez, P., De Backer, O., Bertrand, M., De Plaen, E., Boon, T., and Lucas, S. (2001). An overview of the MAGE gene family with the identification of all human members of the family. Cancer Res **61,** 5544-5551.

Cimprich, K.A., and Cortez, D. (2008). ATR: an essential regulator of genome integrity. Nat Rev Mol Cell Biol **9,** 616-627.

Cohen-Fix, O. (2001). The making and breaking of sister chromatid cohesion. Cell **106,** 137-140.

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. Plant J **20,** 503-508.

Cordoba-Canero, D., Roldan-Arjona, T., and Ariza, R.R. (2011). Arabidopsis ARP endonuclease functions in a branched base excision DNA repair pathway completed by LIG1. Plant J **68,** 693-702.

Culligan, K., Tissier, A., and Britt, A. (2004). ATR regulates a G2-phase cell-cycle checkpoint in Arabidopsis thaliana. 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. Plant J **48,** 947-961.

De Piccoli, G., Cortes-Ledesma, F., Ira, G., Torres-Rosell, J., Uhle, S., Farmer, S., Hwang, J.Y., Machin, F., Ceschia, A., McAleenan, A., Cordon-Preciado, V., Clemente-Blanco, A., Vilella-Mitjana, F., Ullal, P., Jarmuz, A., Leitao, B., Bressan, D., Dotiwala, F., Papusha, A., Zhao, X., Myung, K., Haber, J.E., Aguilera, A., and Aragon, L. (2006). Smc5-Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination. Nat Cell Biol **8,** 1032-1034.

De Piccoli, G., Torres-Rosell, J., and Aragon, L. (2009). The unnamed complex: what do we know about Smc5-Smc6? Chromosome Res **17,** 251-263.

Devos, K.M., Brown, J.K., and Bennetzen, J.L. (2002). Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. Genome Res **12,** 1075-1079.

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

Diaz, M., and Pecinka, A. (2017). Seeds as emerging hotspot for maintenance of genome stability. Cytologia Focus **82**.

DiNardo, S., Voelkel, K., and Sternglanz, R. (1984). DNA topoisomerase II mutant of Saccharomyces cerevisiae: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. Proc Natl Acad Sci U S A **81,** 2616-2620.

Doyle, J.M., Gao, J., Wang, J., Yang, M., and Potts, P.R. (2010). MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. Mol Cell **39,** 963-974.

Dronkert, M.L., and Kanaar, R. (2001). Repair of DNA interstrand cross-links. Mutat Res **486,** 217-247.

Duan, X., Yang, Y., Chen, Y.H., Arenz, J., Rangi, G.K., Zhao, X., and Ye, H. (2009). Architecture of the Smc5/6 Complex of Saccharomyces cerevisiae Reveals a Unique Interaction between the Nse5-6 Subcomplex and the Hinge Regions of Smc5 and Smc6. J Biol Chem **284,** 8507-8515.

Durrant, W.E., Wang, S., and Dong, X.N. (2007). Arabidopsis SNI1 and RAD51D regulate both gene transcription and DNA recombination during the defense response Proc Natl Acad Sci USA **104,** 7307-7307.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32,** 1792-1797.

Erilova, A., Brownfield, L., Exner, V., Rosa, M., Twell, D., Mittelsten Scheid, O., Hennig, L., and Kohler, C. (2009). Imprinting of the polycomb group gene MEDEA serves as a ploidy sensor in Arabidopsis. PLoS Genet 5, e1000663.

Fousteri, M.I., and Lehmann, A.R. (2000). A novel SMC protein complex in Schizosaccharomyces pombe contains the Rad18 DNA repair protein. EMBO J **19,** 1691-1702.

Fujioka, Y., Kimata, Y., Nomaguchi, K., Watanabe, K., and Kohno, K. (2002). Identification of a novel non-structural maintenance of chromosomes (SMC) component of the SMC5-SMC6 complex involved in DNA repair. J Biol Chem **277,** 21585-21591.

Gallego-Paez, L.M., Tanaka, H., Bando, M., Takahashi, M., Nozaki, N., Nakato, R., Shirahige, K., and Hirota, T. (2014). Smc5/6-mediated regulation of replication progression contributes to chromosome assembly during mitosis in human cells. Mol Biol Cell **25,** 302-317.

Garcia-Diaz, M., Bebenek, K., Kunkel, T.A., and Blanco, L. (2001). Identification of an intrinsic 5'-deoxyribose-5-phosphate lyase activity in human DNA polymerase lambda: a possible role in base excision repair. J Biol Chem **276,** 34659-34663.

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. Plant Cell **15,** 119-132.

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L., and Henikoff, S. (2003). Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics **164,** 731-740.

Grossniklaus, U., and Paro, R. (2014). Transcriptional silencing by Polycomb-group proteins. Cold Spring Harbor Persp Biol **6,** a019331.

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

Hay, R.T. (2005). SUMO: a history of modification. Mol Cell **18,** 1-12.

Hirano, T. (2006). At the heart of the chromosome: SMC proteins in action. Nat Rev Mol Cell Biol **7,** 311-322.

Hu, Z.B., Cools, T., and De Veylder, L. (2016). Mechanisms used by plants to cope with DNA damage. Annu Rev Plant Biol **67,** 439-462.

Huang, L., Yang, S., Zhang, S., Liu, M., Lai, J., Qi, Y., Shi, S., Wang, J., Wang, Y., Xie, Q., and Yang, C. (2009). The Arabidopsis SUMO E3 ligase AtMMS21, a homologue of NSE2/MMS21, regulates cell proliferation in the root. Plant J **60,** 666-678.

Huang, Y., and Li, L. (2013). DNA crosslinking damage and cancer - a tale of friend and foe. Transl Cancer Res **2,** 144-154.

Hudson, J.J., Bednarova, K., Kozakova, L., Liao, C., Guerineau, M., Colnaghi, R., Vidot, S., Marek, J., Bathula, S.R., Lehmann, A.R., and Palecek, J. (2011). Interactions between the Nse3 and Nse4 components of the SMC5-6 complex identify evolutionarily conserved interactions between MAGE and EID Families. PLoS One **6,** e17270.

Hwang, J.Y., Smith, S., Ceschia, A., Torres-Rosell, J., Aragon, L., and Myung, K. (2008). Smc5-Smc6 complex suppresses gross chromosomal rearrangements mediated by break-induced replications. DNA Repair **7,** 1426-1436.

Ide, H., Shoulkamy, M.I., Nakano, T., Miyamoto-Matsubara, M., and Salem, A.M. (2011). Repair and biochemical effects of DNA-protein cross-links. Mutat Res 711, 113-122.

Ishida, T., Fujiwara, S., Miura, K., Stacey, N., Yoshimura, M., Schneider, K., Adachi, S., Minamisawa, K., Umeda, M., and Sugimoto, K. (2009). SUMO E3 ligase HIGH PLOIDY2 regulates endocycle onset and meristem maintenance in Arabidopsis. Plant Cell **21,** 2284-2297.

Ishida, T., Yoshimura, M., Miura, K., and Sugimoto, K. (2012). MMS21/HPY2 and SIZ1, two Arabidopsis SUMO E3 ligases, have distinct functions in development. PLoS One **7,** e46897.

Jalal, D., Chalissery, J., and Hassan, A.H. (2017). Genome maintenance in Saccharomyces cerevisiae: the role of SUMO and SUMO-targeted ubiquitin ligases. Nucleic Acids Res **45,** 2242-2261.

Jeppsson, K., Kanno, T., Shirahige, K., and Sjogren, C. (2014a). The maintenance of chromosome structure: positioning and functioning of SMC complexes. Nat Rev Mol Cell Biol **15,** 601-614.

Jeppsson, K., Carlborg, K.K., Nakato, R., Berta, D.G., Lilienthal, I., Kanno, T., Lindqvist, A., Brink, M.C., Dantuma, N.P., Katou, Y., Shirahige, K., and Sjogren, C. (2014b).

The chromosomal association of the Smc5/6 complex depends on cohesion and predicts the level of sister chromatid entanglement. PLoS Genet **10,** e1004680.

Johnston, S.A., den Nijs, T.P., Peloquin, S.J., and Hanneman, R.E., Jr. (1980). The significance of genic balance to endosperm development in interspecific crosses. Theor Appl Genet 57, 5-9.

Jurgens, G., Mayer, U., Torres, R., Berleth, T., Misera, S. (1991). Genetic analysis of pattern formation in the Arabidopsis embryo. Development **113,** 27-38.

Kagale, S., Robinson, S.J., Nixon, J., Xiao, R., Huebert, T., Condie, J., Kessler, D., Clarke, W.E., Edger, P.P., Links, M.G., Sharpe, A.G., and Parkin, I.A.P. (2014). Polyploid evolution of the *Brassicaceae* during the cenozoic era. Plant Cell **26,** 2777-2791.

Kanno, T., Berta, D.G., and Sjogren, C. (2015). The Smc5/6 complex is an ATP-dependent intermolecular DNA linker. Cell Rep **12,** 1471-1482.

Kegel, A., and Sjogren, C. (2010). The Smc5/6 complex: more than repair? Cold Spring Harb Symp Quant Biol **75,** 179-187.

Kegel, A., Betts-Lindroos, H., Kanno, T., Jeppsson, K., Strom, L., Katou, Y., Itoh, T., Shirahige, K., and Sjogren, C. (2011). Chromosome length influences replication-induced topological stress. Nature **471,** 392-396.

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. Plant J **50,** 347-363.

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

Koornneef, M., Dellaert, L.W., and van der Veen, J.H. (1982). EMS- and radiation-induced mutation frequencies at individual loci in Arabidopsis thaliana (L.) Heynh. Mutat Res **93,** 109-123.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell **128,** 693-705.

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.

Krieg, D.R. (1963). Ethyl methanesulfonate-induced reversion of bacteriophage T4rII mutants. Genetics **48,** 561-580.

Kwak, J.S., Son, G.H., Kim, S.I., Song, J.T., and Seo, H.S. (2016). Arabidopsis HIGH PLOIDY2 Sumoylates and Stabilizes Flowering Locus C through Its E3 Ligase Activity. Front Plant Sci **7,** 530.

Laflamme, G., Tremblay-Boudreault, T., Roy, M.-A., Andersen, P., Bonneil, É., Atchia, K., Thibault, P., D'Amours, D., and Kwok, B.H. (2014). Structural maintenance of chromosome (SMC) proteins link microtubule stability to genome integrity. J Biol Chem **289,** 27418-27431.

Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J.U., and Forner, J. (2013). GreenGate - A novel, versatile, and efficient cloning system for plant transgenesis. PLOS ONE **8,** e83043.

Lehmann, A.R., Walicka, M., Griffiths, D.J., Murray, J.M., Watts, F.Z., McCready, S., and Carr, A.M. (1995). The rad18 gene of Schizosaccharomyces pombe defines a new subgroup of the SMC superfamily involved in DNA repair. Mol Cell Biol **15,** 7067-7080.

Li, G., Hall, T.C., and Holmes-Davis, R. (2002). Plant chromatin: development and gene control. Bioessays **24,** 234-243.

Li, G., Zou, W., Jian, L., Qian, J., Deng, Y., and Zhao, J. (2017). Non-SMC elements 1 and 3 are required for early embryo and seedling development in Arabidopsis. J Exp Bot **68,** 1039-1054.

Li, X., Zhang, Y., Clarke, J.D., Li, Y., and Dong, X. (1999). Identification and cloning of a negative regulator of systemic acquired resistance, SNI1, through a screen for suppressors of npr1-1. Cell **98,** 329-339.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature **362,** 709-715.

Lindroos, H.B., Strom, L., Itoh, T., Katou, Y., Shirahige, K., and Sjogren, C. (2006). Chromosomal association of the Smc5/6 complex reveals that it functions in differently regulated pathways. Mol Cell **22,** 755-767.

Liu, C.-M., McElver, J., Tzafrir, I., Joosen, R., Wittich, P., Patton, D., Van Lammeren, A.A.M., and Meinke, D. (2002). Condensin and cohesin knockouts in Arabidopsis exhibit a titan seed phenotype. Plant J **29,** 405-415.

Liu, C.-H., Finke, A., Diaz, M., Rozhon, W., Poppenberger, B., Baubec, T., and Pecinka, A. (2015). Repair of DNA damage induced by the cytidine analog zebularine requires ATR and ATM in Arabidopsis. Plant Cell **27,** 1788-1800.

Liu, M., Shi, S.F., Zhang, S.C., Xu, P.L., Lai, J.B., Liu, Y.Y., Yuan, D.K., Wang, Y.Q., Du, J.J., and Yang, C.W. (2014). SUMO E3 ligase AtMMS21 is required for normal meiosis and gametophyte development in Arabidopsis. BMC Plant Biol **14**.

Losada, A., and Hirano, T. (2005). Dynamic molecular linkers of the genome: the first decade of SMC proteins. Genes Dev **19,** 1269-1287.

Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A.S., and Helleday, T. (2005). Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. Nucleic Acids Res 33, 3799-3811.

Mansfield, S., Briarty, L. (1991). Early embryogenesis in Arabidopsis thaliana. II. The developing embryo. Can J Bot **69** 461-476.

Mayer, U., Ruiz, R.A.T., Berleth, T., Misera, S., and Jurgens, G. (1991). Mutations Affecting Body Organization in the Arabidopsis Embryo. Nature **353,** 402-407.

McAleenan, A., Cordon-Preciado, V., Clemente-Blanco, A., Liu, I.C., Sen, N., Leonard, J., Jarmuz, A., and Aragon, L. (2012). SUMOylation of the alpha-kleisin subunit of cohesin is required for DNA damage-induced cohesion. Curr Biol **22,** 1564-1575.

McDonald, W.H., Pavlova, Y., Yates, J.R., 3rd, and Boddy, M.N. (2003). Novel essential DNA repair proteins Nse1 and Nse2 are subunits of the fission yeast Smc5-Smc6 complex. J Biol Chem **278,** 45460-45467.

McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., Thomas, C., Schetter, A., Zhou, Q., Cushman, M.A., Tossberg, J., Nickle, T., Levin, J.Z., Law, M., Meinke, D., and Patton, D. (2001). Insertional mutagenesis of genes required for seed development in Arabidopsis thaliana. Genetics **159,** 1751-1763.

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

Menolfi, D., Delamarre, A., Lengronne, A., Pasero, P., and Branzei, D. (2015). Essential roles of the Smc5/6 complex in replication through natural pausing sites and endogenous DNA damage tolerance. Mol Cell **60,** 835-846.

Mozgova, I., and Hennig, L. (2015). The polycomb group protein regulatory network. Annu Rev Plant Biol **66,** 269-296.

Müller, A. (1961). Zur Charakterisierung der Blüten und Infloreszenzen vonArabidopsis thaliana (L.) Heynh. Die Kulturpflanze **9,** 364–393.

Mylne, J.S., Barrett, L., Tessadori, F., Mesnage, S., Johnson, L., Bernatavichute, Y.V., Jacobsen, S.E., Fransz, P., and Dean, C. (2006). LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc Natl Acad Sci U S A **103,** 5012-5017.

Outwin, E.A., Irmisch, A., Murray, J.M., and O'Connell, M.J. (2009). Smc5-Smc6-dependent removal of cohesin from mitotic chromosomes. Mol Cell Biol **29,** 4363-4375.

Palecek, J., Vidot, S., Feng, M., Doherty, A.J., and Lehmann, A.R. (2006). The Smc5-Smc6 DNA repair complex. bridging of the Smc5-Smc6 heads by the KLEISIN, Nse4, and non-Kleisin subunits. J Biol Chem **281,** 36952-36959.

Palecek, Jan J., and Gruber, S. (2015). Kite proteins: a superfamily of SMC/kleisin partners conserved across bacteria, archaea, and eukaryotes. Structure **23,** 2183-2190.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H., Wang, X., Wicker, T., Bharti, A.K., Chapman, J., Feltus, F.A., Gowik, U., Grigoriev, I.V., Lyons, E., Maher, C.A., Martis, M., Narechania, A., Otillar, R.P., Penning, B.W., Salamov, A.A., Wang, Y., Zhang, L., Carpita, N.C., Freeling, M., Gingle, A.R., Hash, C.T., Keller, B., Klein, P., Kresovich, S., McCann, M.C., Ming, R., Peterson, D.G., Mehboob ur, R., Ware, D., Westhoff, P., Mayer, K.F.X., Messing, J., and Rokhsar, D.S. (2009). The *Sorghum bicolor* genome and the diversification of grasses. Nature **457,** 551-556.

Pebernard, S., Perry, J.J., Tainer, J.A., and Boddy, M.N. (2008). Nse1 RING-like domain supports functions of the Smc5-Smc6 holocomplex in genome stability. Mol Biol Cell **19,** 4099-4109.

Pebernard, S., Wohlschlegel, J., McDonald, W.H., Yates, J.R., 3rd, and Boddy, M.N. (2006). The Nse5-Nse6 dimer mediates DNA repair roles of the Smc5-Smc6 complex. Mol Cell Biol **26,** 1617-1630.

Pecinka, A., and Liu, C.H. (2014). Drugs for plant chromosome and chromatin research. Cytogenet Genome Res **143,** 51-59.

Potts, P.R., and Yu, H. (2005). Human MMS21/NSE2 is a SUMO ligase required for DNA repair. Mol Cell Biol **25,** 7021-7032.

Potts, P.R., Porteus, M.H., and Yu, H. (2006). Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks. EMBO J **25,** 3377-3388.

Potts, P.R., and Yu, H. (2007). The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. Nat Struct Mol Biol **14,** 581-590.

Potts, P.R. (2009). The Yin and Yang of the MMS21-SMC5/6 SUMO ligase complex in homologous recombination. DNA Repair **8,** 499-506.

Prakash, S., and Prakash, L. (1977). Increased spontaneous mitotic segregation in MMS-sensitive mutants of Saccharomyces cerevisiae. Genetics **87,** 229-236.

Puchta, H., Dujon, B., and Hohn, B. (1996). Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proc Natl Acad Sci U S A **93,** 5055-5060.

Raschle, M., Smeenk, G., Hansen, R.K., Temu, T., Oka, Y., Hein, M.Y., Nagaraj, N., Long, D.T., Walter, J.C., Hofmann, K., Storchova, Z., Cox, J., Bekker-Jensen, S., Mailand, N., and Mann, M. (2015). DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. Science **348,** 1253671.

Regan, S.M., and Moffatt, B.A. (1990). Cytochemical Analysis of Pollen Development in Wild-Type Arabidopsis and a Male-Sterile Mutant. Plant Cell **2,** 877-889.

Roitinger, E., Hofer, M., Kocher, T., Pichler, P., Novatchkova, M., Yang, J., Schlogelhofer, P., and Mechtler, K. (2015). Quantitative phosphoproteomics of the ataxia telangiectasia-mutated (ATM) and ataxia telangiectasia-mutated and rad3-related (ATR) dependent DNA damage response in Arabidopsis thaliana. Mol Cell Proteomics **14,** 556-571.

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. Plant J **72,** 781-790.

Roy, S. (2014). Maintenance of genome stability in plants: repairing DNA double strand breaks and chromatin structure stability. Front Plant Sci **5,** 487.

Ryals, J., Weymann, K., Lawton, K., Friedrich, L., Ellis, D., Steiner, H.Y., Johnson, J., Delaney, T.P., Jesse, T., Vos, P., and Uknes, S. (1997). The Arabidopsis NIM1 protein shows homology to the mammalian transcription factor inhibitor I kappa B. Plant Cell **9,** 425-439.

Ryu, T., Spatola, B., Delabaere, L., Bowlin, K., Hopp, H., Kunitake, R., Karpen, G.H., and Chiolo, I. (2015). Heterochromatic breaks move to the nuclear periphery to continue recombinational repair. Nat Cell Biol **17,** 1401-1411.

Santa Maria, S.R., Gangavarapu, V., Johnson, R.E., Prakash, L., and Prakash, S. (2007). Requirement of Nse1, a subunit of the Smc5-Smc6 complex, for Rad52-dependent postreplication repair of UV-damaged DNA in Saccharomyces cerevisiae. Mol Cell Biol **27,** 8409-8418.

Sargant, E. (1900). Recent work on the results of fertilization in angiosperms. Ann Bot 14, 689-712.

Schleiffer, A., Kaitna, S., Maurer-Stroh, S., Glotzer, M., Nasmyth, K., and Eisenhaber, F. (2003). Kleisins: a superfamily of bacterial and eukaryotic SMC protein partners. Mol Cell **11,** 571-575.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., and Lohmann, J.U. (2005). A gene expression map of Arabidopsis thaliana development. Nat Genet **37,** 501-506.

Schoft, V.K., Chumak, N., Mosiolek, M., Slusarz, L., Komnenovic, V., Brownfield, L., Twell, D., Kakutani, T., and Tamaru, H. (2009). Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. EMBO Rep **10,** 1015-1021.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., and Goodrich, J. (2006a). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. EMBO J **25,** 4638-4649.

Schubert, V., Klatte, M., Pecinka, A., Meister, A., Jasencakova, Z., and Schubert, I. (2006b). Sister chromatids are often incompletely aligned in meristematic and endopolyploid interphase nuclei of Arabidopsis thaliana. Genetics **172,** 467-475.

Schubert, V. (2009). SMC proteins and theirs multiple functions in higher plants. Cytogenet Genome Res **124,** 202-214.

Sergeant, J., Taylor, E., Palecek, J., Fousteri, M., Andrews, E.A., Sweeney, S., Shinagawa, H., Watts, F.Z., and Lehmann, A.R. (2005). Composition and architecture of the Schizosaccharomyces pombe Rad18 (Smc5-6) complex. Mol Cell Biol **25,** 172-184.

Setlow, R.B., Swenson, P.A., and Carrier, W.L. (1963). Thymine Dimers and Inhibition of DNA Synthesis by Ultraviolet Irradiation of Cells. Science **142,** 1464-1466.

Shiloh, Y. (2001). ATM and ATR: networking cellular responses to DNA damage. Curr Opin Genet Dev **11,** 71-77.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijo, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell **136,** 461-472.

Song, J., Durrant, W.E., Wang, S., Yan, S., Tan, E.H., and Dong, X. (2011). DNA repair proteins are directly involved in regulation of gene expression during plant immune response. Cell Host Microbe **9,** 115-124.

Stangeland, B., and Salehian, Z. (2002). An improved clearing method for GUS assay in Arabidopsis endosperm and seeds. Plant Mol Biol Reporter **20,** 107-114.

Stephan, A.K., Kliszczak, M., Dodson, H., Cooley, C., and Morrison, C.G. (2011). Roles of vertebrate Smc5 in sister chromatid cohesion and homologous recombinational repair. Mol Cell Biol **31,** 1369-1381.

Stingele, J., and Jentsch, S. (2015). DNA-protein cross-link repair. Nat Rev Mol Cell Biol 16, 455-460.

Stingele, J., Habermann, B., and Jentsch, S. (2015). DNA-protein cross-link repair: proteases as DNA repair enzymes. Trends Biochem Sci 40, 67-71.

Strom, L., Karlsson, C., Lindroos, H.B., Wedahl, S., Katou, Y., Shirahige, K., and Sjogren, C. (2007). Postreplicative formation of cohesion is required for repair and induced by a single DNA break. Science **317,** 242-245.

Tamura, K., Adachi, Y., Chiba, K., Oguchi, K., and Takahashi, H. (2002). Identification of Ku70 and Ku80 homologues in Arabidopsis thaliana: evidence for a role in the repair of DNA double-strand breaks. Plant J **29,** 771-781.

Tanksley, S.D., and McCouch, S.R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. Science **277,** 1063-1066.

Taylor, E.M., Copsey, A.C., Hudson, J.J., Vidot, S., and Lehmann, A.R. (2008). Identification of the proteins, including MAGEG1, that make up the human SMC5-6 protein complex. Mol Cell Biol **28,** 1197-1206.

Tian, G., Lu, Q., Zhang, L., Kohalmi, S.E., and Cui, Y. (2011). Detection of protein interactions in plant using a Gateway compatible Bimolecular Fluorescence Complementation (BiFC) system. J Visual Exp**,** 3473.

Tibbetts, R.S., Cortez, D., Brumbaugh, K.M., Scully, R., Livingston, D., Elledge, S.J., and Abraham, R.T. (2000). Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress. Genes Dev **14,** 2989-3002.

Torres-Rosell, J., Machin, F., Farmer, S., Jarmuz, A., Eydmann, T., Dalgaard, J.Z., and Aragon, L. (2005). SMC5 and SMC6 genes are required for the segregation of repetitive chromosome regions. Nat Cell Biol **7,** 412-419.

Torres-Rosell, J., Sunjevaric, I., De Piccoli, G., Sacher, M., Eckert-Boulet, N., Reid, R., Jentsch, S., Rothstein, R., Aragon, L., and Lisby, M. (2007). The Smc5-Smc6 complex and SUMO modification of Rad52 regulates recombinational repair at the ribosomal gene locus. Nat Cell Biol **9,** 923-931.

Tripathy, B.C., and Oelmuller, R. (2012). Reactive oxygen species generation and signaling in plants. Plant Signal Behav **7,** 1621-1633.

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. Mutat Res 681, 134-149.

Tuteja, N., Singh, M.B., Misra, M.K., Bhalla, P.L., and Tuteja, R. (2001). Molecular mechanisms of DNA damage and repair: progress in plants. Crit Rev Biochem Mol Biol 36, 337-397.

Tzafrir, I., McElver, J.A., Liu Cm, C.M., Yang, L.J., Wu, J.Q., Martinez, A., Patton, D.A., and Meinke, D.W. (2002). Diversity of TITAN functions in Arabidopsis seed development. Plant Physiol **128,** 38-51.

Uhlmann, F. (2016). SMC complexes: from DNA to chromosomes. Nat Rev Mol Cell Biol **17,** 399-412.

Vielle-Calzada, J.P., Baskar, R., and Grossniklaus, U. (2000). Delayed activation of the paternal genome during seed development. Nature **404,** 91-94.

Wang, S., Durrant, W.E., Song, J., Spivey, N.W., and Dong, X. (2010). Arabidopsis BRCA2 and RAD51 proteins are specifically involved in defense gene transcription during plant immune responses. Proc Natl Acad Sci USA **107,** 22716-22721.

Watanabe, K., Pacher, M., Dukowic, S., Schubert, V., Puchta, H., and Schubert, I. (2009). The STRUCTURAL MAINTENANCE OF CHROMOSOMES 5/6 complex promotes sister chromatid alignment and homologous recombination after DNA damage in Arabidopsis thaliana. Plant Cell **21,** 2688-2699.

Weon, J.L., and Potts, P.R. (2015). The MAGE protein family and cancer. Curr Opin Cell Biol **37,** 1-8.

Willing, E.M., Piofczyk, T., Albert, A., Winkler, J.B., Schneeberger, K., and Pecinka, A. (2016). UVR2 ensures transgenerational genome stability under simulated natural UV-B in Arabidopsis thaliana. Nat Commun **7,** 13522.

Wu, N., and Yu, H. (2012). The Smc complexes in DNA damage response. Cell Biosci **2,** 5.

Xu, P., Yuan, D., Liu, M., Li, C., Liu, Y., Zhang, S., Yao, N., and Yang, C. (2013). AtMMS21, an SMC5/6 complex subunit, is involved in stem cell niche maintenance and DNA damage responses in Arabidopsis roots. Plant Physiol **161,** 1755-1768.

Yadav, R.K., Girke, T., Pasala, S., Xie, M., and Reddy, G.V. (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. Proc Natl Acad Sci USA **106,** 4941-4946.

Yan, S., Wang, W., Marques, J., Mohan, R., Saleh, A., Durrant, W.E., Song, J., and Dong, X. (2013). Salicylic acid activates DNA damage responses to potentiate plant immunity. Mol Cell **52,** 602-610.

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. Proc Natl Acad Sci U S A **106,** 12843-12848.

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

Yuan, D., Lai, J., Xu, P., Zhang, S., Zhang, J., Li, C., Wang, Y., Du, J., Liu, Y., and Yang, C. (2014). AtMMS21 regulates DNA damage response and homologous recombination repair in Arabidopsis. DNA Repair **21,** 140-147.

Zabrady, K., Adamus, M., Vondrova, L., Liao, C., Skoupilova, H., Novakova, M., Jurcisinova, L., Alt, A., Oliver, A.W., Lehmann, A.R., and Palecek, J.J. (2016). Chromatin association of the SMC5/6 complex is dependent on binding of its NSE3 subunit to DNA. Nucleic Acids Res **44,** 1064-1079.

Zhang, J., Lai, J., Wang, F., Yang, S., He, Z., Jiang, J., Li, Q., Wu, Q., Liu, Y., Yu, M., Du, J., Xie, Q., Wu, K., and Yang, C. (2017). A SUMO Ligase AtMMS21 Regulates the Stability of the Chromatin Remodeler BRAHMA in Root Development. Plant Physiol **173,** 1574-1582.

Zhang, S., Qi, Y., Liu, M., and Yang, C. (2013). SUMO E3 ligase AtMMS21 regulates drought tolerance in Arabidopsis thaliana(F). J Integr Plant Biol **55,** 83-95.

Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., and Chua, N.-H. (2006). Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat Protocols **1,** 641-646.

Zhao, X.L., and Blobel, G. (2005). A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization Proc Natl Acad Sci USA **102,** 9086-9086.

APPENDIX

Supplemental Table 1. Overview of SMC5/6 complex subunits in plants. The species are represented by spreading earthmoss *(Physcomitrella patens), Brachypodium distachyon, Oryza sativa* (rice) and *Hordeum vulgare* (barley), *Solanum lycopersicum* (tomato) and *Arabidopsis thaliana* (Arabidopsis). The number of gene identifiers indicates the number of copies per genome. *Functional (not protein sequence-based) homologs. Genes for *P. patens, B. distachyon, O. sativa and H. vulgare* were identified by BLAST searches in Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html#), for tomato in Sol Genomics Network database (https://solgenomics.net/) and for Arabidopsis in TAIR (https://www.arabidopsis.org/index.jsp).

Subunit	*P. patens*	*B. distachyon*	*O. sativa*	*H. vulgare*	*S. lycopersicum*	*A. thaliana*
SMC5	***PpSMC5***	***BdSMC5***	***OsSMC5***	***HvSMC5***	***SlSMC5***	***AtSMC5***
	Pp3c24_4940	Bradi2g14160	LOC_Os05g51790	HORVU1Hr1G095230	Solyc01g087720	At5g15920
SMC6	***PpSMC6***	***BdSMC6***	***OsSMC6***	***HvSMC6***	***SlSMC6***	***AtSMC6A***
	Pp3c11_11190	Bradi4g08527	LOC_Os09g03370	HORVU5Hr1G050720	Solyc05g051680	At5g07660
						***AtSMC6B* (*MIM*)**
						At5g61460
NSE1	***PpNSE1***	***BdNSE1***	***OsNSE1***	***HvNSE1***	***SlNSE1***	***AtNSE1***
			LOC_Os12g03360			AT5G21140
	Pp3c20_10070	Bradi4g43810	LOC_Os11g03590	HORVU0Hr1G010660	Solyc01g006210	
		Bradi2g12255				
NSE2	***PpNSE2***	***BdNSE2***	***OsNSE2***	***HvNSE2***	***SlNSE2***	***NSE2/MMS21/HPY2***
	Pp3c22_18560	Bradi2g16600	LOC_Os05g48880	HORVU1Hr1G087520	Solyc07g062780	
		Bradi2g16580				At3g15150
NSE3	***PpNSE3A***	***BdNSE3***	***OsNSE3***	***HvNSE3***	***SlNSE3***	***NSE3***
	Pp3c15_18480	Bradi1g58440	LOC_Os07g05650	HORVU2Hr1G060140	Solyc10g018870	At1g34770
NSE4	***PpNSE4***	***BdNSE4***	***OsNSE4***	***HvNSE4***	***SlNSE4A***	***NSE4A***
	Pp3c27_130	Bradi3g06970	LOC_Os02g10090	HORVU7Hr1G094270	Solyc10g078730	AT1G51130
				HORVU6Hr1G03375		***NSE4B***
				0		
		Bradi1g35930	LOC_Os06g41380		Solyc12g041890	
			LOC_Os08g40010		Solyc01g006460	At3g20760
			LOC_Os02g29620		Solyc04g025510	
			LOC_Os04g10870			
			LOC_Os07g01010			
NSE5*	***PpASAP1***	***BdASAP1***	***OsASAP1***	***HvASAP1***	***SlASAP1***	***ASAP1***
	Pp3c4_7040	Bradi2g08380	LOC_Os01g13940	HORVU3Hr1G032750	Solyc11g066340	At2g28130
NSE6*	***PpSNI1***	***BdSNI1***	***OsSNI1***	***HvSNI1***	***SlSNI1***	***SNI1***
	Pp3c13_1090	Bradi3g11450	LOC_Os02g20870	HORVU6Hr1G054340	Solyc02g077320	At4g18470

Supplemental Table 2. Protein sequences used in the phylogenetic analysis.

Species	**Gene**	**Accession number**	**Source**
Arabidopsis thaliana	*NSE4A*	NP_175525.1	NCBI
Arabidopsis thaliana	*NSE4B*	NP_188712.4	NCBI
Arabidopsis lyrata	*NSE4A*	XP_002894295.1	NCBI
Arabidopsis lyrata	*NSE4B*	XP_002885402.1	NCBI
Capsella rubella	*NSE4A*	XP_006307701.1	NCBI
Capsella rubella	*NSE4B*	XP_006299562.1	NCBI
Brassica rapa	*NSE4A*	XP_009147782.1	NCBI
Brassica rapa	*NSE4A'*	XP_009144924.1	NCBI
Eutrema salsugineum	*NSE4A*	XP_006393077.1	NCBI
Eutrema salsugineum	*NSE4B*	XP_006406352.1	NCBI
Vitis vinifera	*NSE4A*	CBI16452.3	NCBI
Vitis vinifera	*NSE4B*	CBI30028.3	NCBI
Solanum lycopersicum	*NSE4A*	XP_004249168.1	NCBI
Solanum lycopersicum	*NSE4B*	XP_004252465.1	NCBI
Solanum lycopersicum	*NSE4C*	XP_019067398.1	NCBI
Solanum lycopersicum	*NSE4D*	Solyc04g02510.2.1	Phytozome
Populus trichocarpa	*NSE4A*	XP_002304893.1	NCBI
P. trichocarpa	*NSE4B*	XP_002323286.2	NCBI
Hordeum vulgare	*NSE4A*	HORVU7Hr1G094270.3	Phytozome
Hordeum vulgare	*NSE4B*	HORVU6Hr1G033750.1	Phytozome
Brachypodium distachyon	*NSE4A*	XP_010229801.1	NCBI
Brachypodium distachyon	*NSE4B*	XP_003571216.1	NCBI
Oryza sativa	*NSE4A*	EEC80977.1	NCBI
Oryza sativa	*NSE4B*	XP_015623492.1	NCBI
Oryza sativa	*NSE4C*	EEC83860.1	NCBI
Oryza sativa	*NSE4D*	XP_015627110.1	NCBI
Oryza sativa	*NSE4E*	CAE02529.2	NCBI

Oryza sativa	*NSE4F*	EEE66414.1	NCBI
Marchantia polymorpha	*NSE4A*	OAE22228.1	NCBI
Physcomitrella patens	*NSE4A*	XP_001780495.1	NCBI
Schizosaccharomyces pombe	*NSE4A/Rad62*	NP_001018837.1	NCBI
Homo sapiens	*NSE4A*	NP_060085.2	NCBI
Homo sapiens	*NSE4B*	NP_001008395.1	NCBI

Supplemental Table 3. 3-AT concentrations used in yeast-two-hybrid experiments.

Autoactivation tests

Gene 1	Gene 2	3-AT (mM)
BD-SMC5	pGADT7	3
AD-SMC5	pGBKT7	3
AD-SMC6A	pGBKT7	3
AD-SMC6B	pGBKT7	3
BD-NSE4A	pGADT7	0
BD-NSE4B	pGADT7	0
SMC5-BD	pGADCg	0
SMC5-AD	pGBKCg	5
SMC6A-BD	pGADCg	5
SMC6A-AD	pGBKCg	10
SMC6B-BD	pGADCg	0
SMC6B-AD	pGBKCg	10
hSMC5-BD	pGADCg	5
hSMC5-AD	pGBKCg	10
hSMC6A-BD	pGADCg	5
hSMC6B-AD	pGBKCg	10
NSE4A-BD	pGADCg	5
NSE4A-AD	pGBKCg	25
NSE4B-AD	pGBKCg	25
BD-NSE1	pGADT7	5
AD-NSE1	pGBKT7	5
BD-NSE3	pGADT7	5
AD-NSE3	pGBKT7	10
NSE1-BD	pGADCg	5
NSE1-AD	pGBKCg	5
NSE3-BD	pGADCg	20
NSE3-AD	pGBKCg	5

Interaction tests

Gene 1	Gene 2	3-AT (mM)
BD-SMC5	AD-SMC6A	3
BD-SMC5	AD-SMC6B	3
BD-NSE4A	AD-SMC5	3
BD-NSE4B	AD-SMC5	3
BD-NSE4A	AD-SMC6A	3
BD-NSE4A	AD-SMC6B	3
BD-NSE4B	AD-SMC6B	3
BD-NSE4A	SMC6A-AD	10
BD-NSE4A	SMC6B-AD	10
BD-NSE4B	SMC6A-AD	10
BD-NSE4B	SMC6B-AD	10
NSE4A-BD	SMC6A-AD	10
SMC6A-BD	NSE4A-AD	25
NSE4A-BD	SMC6B-AD	10
SMC5-BD	SMC6A-AD	10
SMC5-BD	SMC6B-AD	10
SMC6A-BD	SMC5-AD	5
SMC6B-BD	SMC5-AD	5
SMC5-BD	NSE4A-AD	25
SMC5-BD	NSE4B-AD	25
SMC6B-BD	NSE4A-AD	25
SMC6B-BD	NSE4B-AD	25
hSMC5-BD	hSMC6B-AD	10
hSMC6A-BD	hSMC5-AD	10
NSE4A-BD	AD-NSE1	5
NSE4A-BD	NSE1-AD	5
BD-NSE1	NSE4A-AD	25
NSE1-BD	NSE4B-AD	25
NSE1-BD	NSE4A-AD	25
BD-NSE1	AD-SMC5	5
BD-SMC5	NSE1-AD	5
BD-SMC5	AD-NSE1	5
NSE1-BD	SMC5-AD	5
BD-NSE1	SMC5-AD	5
SMC5-BD	AD-NSE1	5
SMC5-BD	NSE1-AD	5
NSE4A-BD	NSE3-AD	5
BD-NSE1	NSE4B-AD	25
NSE3-BD	AD-SMC5	20
BD-SMC5	NSE3-AD	5

Interaction tests

Gene 1	Gene 2	3-AT (mM)
BD-SMC5	AD-NSE3	5
NSE3-BD	SMC5-AD	20
BD-NSE3	SMC5-AD	5
NSE1-BD	AD-SMC6B	5
BD-NSE1	AD-SMC6B	5
BD-SMC6B	NSE1-AD	5
BD-SMC6B	AD-NSE1	5
NSE1-BD	SMC6B-AD	10
SMC6B-BD	NSE1-AD	5
SMC6B-BD	AD-NSE1	5
SMC6B-BD	NSE3-AD	5
SMC6B-BD	AD-NSE3	5
SMC5-BD	NSE3-AD	5
BD-NSE1	SMC6B-AD	10
SMC5-BD	AD-NSE3	10
BD-NSE4A	SMC6B-AD	10
NSE4A-BD	SMC6A-AD	10
SMC6A-BD	NSE4A-AD	25
SMC6B-BD	NSE4B-AD	25
SMC6B-BD	NSE4A-AD	25
BD-NSE4B	SMC6B-AD	10
BD-NSE4B	SMC6A-AD	10
BD-NSE4A	SMC6A-AD	10
NSE4A-BD	SMC6B-AD	10
NSE4A-BD	SMC6A-AD	10
SMC5-BD	NSE4B-AD	25
SMC5-BD	NSE4A-AD	25
SMC6B-BD	SMC5-AD	25
SMC6A-BD	SMC5-AD	25
SMC5-BD	SMC6B-AD	10
SMC5-BD	SMC6A-AD	10
BD-SMC6B	AD-NSE3	10
NSE3-BD	AD-SMC6B	20
BD-NSE3	AD-SMC6B	5
BD-SMC6B	NSE3-AD	5
NSE3-BD	SMC6B-AD	20
BD-NSE3	SMC6B-AD	10
SMC6A-BD	NSE4B-AD	25
SMC6A-BD	NSE4B-AD	25

Supplemental Table 4. Primers used in this study. p – promoter, g – genomic sequence, t – terminator.

Target	Primer name	Sequence 5' to 3'	Application
nse4a-1	Nse4a1_F	ACCTGCCGTTAGGGTTTTCA	Genotyping
nse4a-1	Nse4a1_R	AGCAAACCCAAGATCCTTCCA	Genotyping
nse4a-2	Nse4a2_F	GCTCAACAGGCGGTCATTTG	Genotyping
nse4a-2	Nse4a2_R	ACAAAAGCCACTTAACTGCTACA	Genotyping
nse4b-1	Nse4b1_F	TCGGAAGCCTAGAGAGCAGA	Genotyping
nse4b-1	Nse4b1_R	ATAAGCGACCTCCCCTGACA	Genotyping
nse4b-2	Nse4b2_F	AGCAGAGACTTCGACGTTGG	Genotyping
nse4b-2	Nse4b2_R	AAAACCGTGGAACAGACCGT	Genotyping
LB Salk lines	LB-AP1	ACTGGAACAACACTCAACCCTATCT	Genotyping
LB Gabi-kat line	o8474_m	ATAATAACGCTGCGGACATCTAC	Genotyping
LB_SAIL line	SAIL-LB2m	tagcatctgaatttcataaccaat	Genotyping
pNSE4A	pNSE4A-attb1	GGGGACAAGTTTGTACAAAAAAGCAGGCTtagaatcatcaacaatttcata	Gateway cloning
pNSE4A	pNSE4A-attb2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGTGAATTTAGAACTGTCG	Gateway cloning

pNSE4B	pNSE4B2-attb1	ACAAGTTTGTACAAAAAAGCAGGCTCTCTCTTCTTCTGTGTGTTTTTG	Gateway cloning
pNSE4B	pNSE4B2-attb2	ACCACTTTGTACAAGAAAGCTGGGTATTTCACCGAGTTTCTCAT	Gateway cloning
pNSE4A	pNSE4A-attb5r	GGGGACCCAACTTTTGTATACAAAGTTGTGGTGAATTTAGAACTGTCG	Gateway cloning
gNSE4B	gNSE4B-attb5	GGGGACAACTTTGTATACAAAAGTTGGGTATGAGAAACTCGGTGAAATGGGA	Gateway cloning
Target	**Primer name**	**Sequence 5' to 3'**	**Application**
gNSE4B	gNSE4B-attb2	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTTATCAGATTTCAGTTTTTCA	Gateway cloning
pNSE4A	pNSE4A-F-Gr	agaagtgaagcttggtctcaacctTAGAATCATCAACAATTTCATAAATCTTG	Greengate cloning
pNSE4A	pNSE4A-R-Gr	gggcgagaattcggtctcaTGTTGGTGAATTTAGAACTGTCG	Greengate cloning
NSE4A CDS	NSE4A-F-Gr	agaagtgaagcttggtctcaggctgtATGAGGAAGACGGTGAAG	Greengate cloning
NSE4A CDS	NSE4A-R-Gr	gggcgagaattcggtctcactgaAGAAGAGGTGAGTCTCCG	Greengate cloning
tNSE4A	t-NSE4A-F-Gr	agaagtgaagcttggtctcactgcgtTAAGGAGTGTATAGAGAGAGAAG	Greengate cloning
tNSE4A	t-NSE4A-R-Gr	gggcgagaattcggtctcaTagtCTAGCTACTTGGTTCTTAGG	Greengate cloning
SMC5 CDS	smc5-attb1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCTGAACGTCGTGCTAAG	Gateway cloning, Y2H, BiFC

SMC5 CDS	smc5-attb2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGGAACATTGACTAGCTTCGGTT	Gateway cloning, Y2H, BiFC
*SMC6A*CDS	smc6a-attb1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATGAACATGGCGACCA	Gateway cloning, Y2H, BiFC
*SMC6A*CDS	smc6a-attb2	GGGGACCACTTTGTACAAGAAAGCTGGGTGAGAACGTGGAGCAGCCATTTGT	Gateway cloning, Y2H, BiFC
SMC6B CDS	smc6b-attb1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGTAAAATCTGGTGCTCG	Gateway cloning, Y2H, BiFC
SMC6B CDS	smc6b-attb2	GGGGACCACTTTGTACAAGAAAGCTGGGTGAGAACGAGGAGCAGCCATTT	Gateway cloning, Y2H, BiFC
SMC5 hinge	smc5-hge-attb1	GGGACAAGTTTGTACAAAAAAGCAGGCTATGGAGAATGCAAATAACAAACTCT	Gateway cloning, Y2H, BiFC
Target	**Primer name**	**Sequence 5' to 3'**	**Application**
SMC5 hinge	smc5-hge-attb2	GGGACCACTTTGTACAAGAAAGCTGGGTGCATGTCCTCCTCTTGCTCTAAA	Gateway cloning, Y2H, BiFC
SMC6A hinge	smc6a-hge-attb1	GGGACAAGTTTGTACAAAAAAGCAGGCTATGACTATAAGAAGCACACAGGT	Gateway cloning, Y2H, BiFC
SMC6A hinge	smc6a-hge-attb2	GGGACCACTTTGTACAAGAAAGCTGGGTGCTCGAGATTCATTTCTGCCTC	Gateway cloning, Y2H, BiFC
SMC6B hinge	smc6b-hge-attb1	GGGACAAGTTTGTACAAAAAAGCAGGCTATGAAGAACACACAGGCTGAACA	Gateway cloning, Y2H, BiFC
SMC6B hinge	smc6b-hge-attb2	GGGACCACTTTGTACAAGAAAGCTGGGTGCTTCAAATCGTGCATCTCAAGTTC	Gateway cloning, Y2H, BiFC

NSE4A CDS	nse4a-attb1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGGAAGACGGTGAAGCG	Gateway cloning, Y2H, BiFC
NSE4A CDS	nse4a-attb2	GGGGACCACTTTGTACAAGAAAGCTGGGTCAGAAGAGGTGAGTCTCCGCTTA	Gateway cloning, Y2H, BiFC
NSE4B CDS	nse4b-attb1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGAAACTCGGTGAAATGG	Gateway cloning, Y2H, BiFC
NSE4B CDS	nse4b-attb2	GGGGACCACTTTGTACAAGAAAGCTGGGTGCAAGCCACCATTTTTACTATCT	Gateway cloning, Y2H, BiFC
NSE1 CDS	NSE1_attb1	GGGACAAGTTTGTACAAAAAAGCAGGCTATGGCATCGCTAAGCTGGAA	Gateway cloning, Y2H, BiFC
NSE1 CDS	NSE1_attb2	GGGACCACTTTGTACAAGAAAGCTGGGTGCCTCGAGTGCCCAACGCT	Gateway cloning, Y2H, BiFC
NSE3 CDS	NSE3_attb1	GGGACAAGTTTGTACAAAAAAGCAGGCTATGGCCGACGAAGAAGATTCT	Gateway cloning, Y2H, BiFC
NSE3 CDS	NSE3_attb2	GGGACCACTTTGTACAAGAAAGCTGGGTGATCATTAAGCTCTACAACTGCTACA	Gateway cloning, Y2H, BiFC
Target	**Primer name**	**Sequence 5' to 3'**	**Application**
Adapter Primer	AP	GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTT	3' RACE PCR
Univ. amplificat. primer	UAP	GGCCACGCGTCGACTAGTAC	3' RACE PCR
NSE4A	NSE4-1-3F	CACCTGATGCAAACGTGGT	3' RACE PCR

NSE4A	NSE4-1-2F	AGGTGTTCGTCCAGATGAGG	3' RACE PCR
NSE4A	nse4a-3_F	TGCAGTTGAAAACTTGCACCA	3' RACE PCR

Supplemental Table 5. Primer combination used for 3' RACE PCR.

		Primer combinations used for 3' RACE PCR
RACE first reaction	nse4a-3_F	TGCAGTTGAAAACTTGCACCA
	UAP	GGCCACGCGTCGACTAGTAC
RACE second reaction	NSE4-1-2F	AGGTGTTCGTCCAGATGAGG
	UAP	GGCCACGCGTCGACTAGTAC
RACE third reaction	nse4a_2F	GCTCAACAGGCGGTCATTTG
	UAP	GGCCACGCGTCGACTAGTAC
RACE four reaction	NSE4-1-3F	CACCTGATGCAAACGTGGT
	UAP	GGCCACGCGTCGACTAGTAC

ACKNOWLEDGMENTS

First, I would like to express my gratitude to my supervisor Dr. Ales Pecinka for giving me the opportunity to pursue my Ph.D in his laboratory. He always motivated me to move the project forward and encouraged discussions and being ambitious. Moreover he acted as a mentor and guided me through my professional development during the entire period, helping me to figure out my plans in the future. I´m also thankful for all the support provided by Dr. Andreas Finke, who helped me to feel integrated in the group from my very first day of my Ph.D; being an essential pillar for the successful completion of my Ph.D; for the discussions with Dr. Anna Nowicka, Dr. Catarina Markus, the support from Dr. Ton Timmers and the Cemic facility and the help received from our technicians Petra Pecinkova, Barbara Eilts and Regina Gentges. I can´t forget the numerous students that passed through our lab and those who worked very hard in my Ph.D project.

In addition I´m also thankful to DAAD for awarding me the fellowship for conducting doctoral studies in Germany, Max Planck Institute for providing facilities and cutting edge technology, which allowed the completion of my Ph.D. Prof. Dr. George Coupland, Prof. Dr. Björn Schumacher and Prof. Dr. Maria Albani, for accepting being part of my committee defense.

Prof. Dr. Jaroslav Dolezel from the Centre of Plant Structural and Functional Genomics in Olomouc, who kindly provided me access to his facility in the final part of my studies.

Finally I want to express my gratitude to my family, my partner and friends Nathan, Jon, Alice, Ana and Lei who always believed in me and supported me whenever I needed it.

LEBENSLAUF

PERSÖNLICHE INFORMATIONEN

Name: Mariana Andrea Diaz Smoje
Adresse: Adamsstr. 67, 51063 Köln, Deutschland
Staatsbürgerschaft: Chilenisch und Kroatisch
Geburtsdatum: 4.11.1988
Geburtsort: Santiago, Chile

AUSBILDUNG

Seit Oktober 2013	Promotion	Max Planck Institut für Pflanzenzüchtungsforschung / Universität zu Köln

Titel der Abschlussarbeit: „Functional characterization of NSE4 paralogs in *Arabidopsis thaliana"*. Finanziert durch Deutscher Akademischer Austauschdienst (DAAD). Stipendium für eine Doktorarbeit in Deutschland

03/08/2012 03/28/2013	Pflanzenbiotechnologie Engineur	Universidad de Concepción, Chile

Titel der Abschlussarbeit: Relative expression of cold resistance genes in *Eucalyptus globulus*

03/17/2008-03/08/2012	Bachelor in Pflanzenbiotechnologie, Universidad de Concepcion, Chile	Universidad de Concepción, Chile

WISSENSCHAFTLICHE VERÖFFENTLICHUNGEN

Diaz, M., and Pecinka, A. (2018). Scaffolding for repair: Understanding molecular functions of the SMC5/6 complex. Genes (zur Veröffentlichung akzeptiert)

Diaz, M., and Pecinka, A. (2017). Seeds as emerging hotspot for maintenance of genome stability Cytologia Focus 82.

Liu, C. H., Finke, A., **Díaz, M.,** Rozhon, W., Poppenberger, B., Baubec, T., & Pecinka, A. (2015). Repair of DNA damage induced by the cytidine analog zebularine requires ATR and ATM in Arabidopsis. The Plant Cell, 27(6), 1788-1800.

Figueroa, C.R., Opazo, M.C., Vera, P., Arriagada, O., **Díaz, M**. and Moya-León, M.A. (2012). Effect of postharvest treatment of calcium and auxin on cell wall composition and expression of cell wall-modifying genes in the Chilean strawberry (Fragaria chiloensis) fruit. Food Chemistry 132(4) 2014-2022.

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 Dr. Ales Pecinka und Prof. Dr. George Coupland betreut worden.

Köln, 08.01.18

_____

Mariana Diaz Smoje