Stress-induced mobilization of retrotransposons for plant breeding.

Inauguraldissertation

zur Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Michael Oliver Thieme

aus Starnberg, Deutschland

Basel, 2017

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Thomas Boller (Fakultätsverantwortlicher) and Dr. Jean Molinier (Korreferent)

Basel, den 14.11.2017

____

Prof. Dr. Martin Spiess (Dekan)

Table of contents

1 LIST OF ABBREVIATIONS IV

2 SUMMARY VI

3 GENERAL INTRODUCTION 1

3.1 TRANSPOSABLE ELEMENTS IN PLANT GENOMES 2

3.1.1 Life cycle and general classification 2

3.1.2 Impact of transposable elements on genome size and architecture 6

3.1.3 Retrotransposons in the model plant Arabidopsis thaliana 7

3.2 REGULATION OF RETROTRANSPOSITION IN PLANTS 8

3.2.1 DNA methylation 8

3.2.1.1 Maintenance methylation 8

3.2.1.2 RNA-directed DNA methylation (RdDM) 9

3.2.2 Chromatin state 12

3.3 RETROTRANSPOSONS: LINKING THE GENOME TO THE ENVIRONMENT 14

3.3.1 Stress-dependent activation of retrotransposons in plants 14

3.3.1.1 Examples for environmental stimuli 14

3.3.1.2 Mechanisms underlying the stress response of LTR-retrotransposons 15

3.3.2 Impact on gene structure and expression 15

3.3.2.1 Insertional mutagenesis 15

3.3.2.2 Impact on splicing 16

3.3.2.3 Expressional changes in retrotransposon-flanking regions 16

3.3.2.4 Gene movement 17

3.4 RETROTRANSPOSONS AS A TOOL FOR CROP IMPROVEMENT AND BREEDING 17

3.4.1 Detection of retrotransposition events 17

3.4.2 Genetic engineering 18

3.4.3 Transient inhibition of retrotransposon silencing 18

3.5 POLICY RELATED ASPECTS OF BRIDGING SCIENCE AND INNOVATION 19

3.6 AIMS OF THE THESIS 20

4 RNA POLYMERASE II REGULATES RETROTRANSPOSON MOBILITY IN *ARABIDOPSIS* 21

4.1 ABSTRACT 21

4.2 INTRODUCTION 21

4.3 RESULTS 22

4.3.1 RNA polymerase II represses the heat-dependent activation of ONSEN 22

4.3.2 Inhibition of Pol II reduces global DNA-methylation 27

4.3.3 Pol II transcripts feed into DCL3-RdDM 29

4.4 DISCUSSION 33
4.5 MATERIAL AND METHODS 36
4.5.1 ChopPCR for CHH-methylation at the ONSEN-LTR 36
4.5.2 Detection of ONSEN-derived antisense transcripts 37

5 EVOLUTIONARY CONSEQUENCES OF AN INDUCED BURST OF *ONSEN* IN *ARABIDOPSIS* 38

5.1 ABSTRACT 38
5.2 INTRODUCTION 39
5.3 RESULTS 40
5.3.1 Pol II-deficiency and inhibition of DNA-methylation leads to retrotransposition of ONSEN 40
5.3.2 ONSEN high-copy lines show phenotypic diversity 47
5.3.3 Genetic stability of ONSEN high-copy lines under HS 50
5.4 DISCUSSION 52
5.5 MATERIAL AND METHODS 55

6 INDUCED ACTIVATION OF RETROTRANSPOSONS IN CROPS 56

6.1 ABSTRACT 56
6.2 INTRODUCTION 57
6.3 RESULTS 59
6.3.1 Simultaneous inhibition of Pol II and DNMtases mobilizes a copia-like retrotransposon in rice ... 59
6.3.2 Testing A&Z-treatments to induce retrotransposition in soybean 64
6.4 DISCUSSION 67
6.5 MATERIALS AND METHODS 71
6.5.1 Heat-stress treatment of rice 71
6.5.2 IRAP-analysis to detect Houba-copy number variation 71
6.5.3 Seed material and cultivation of soybean 71
6.5.4 In vitro culture of soybean 71
6.5.5 Cold treatment of soybean 72
6.5.6 Mobilome-analysis of soybean 72

7 POLICY-RELATED ASPECTS OF INTRODUCING A NOVEL BREEDING TECHNOLOGY TO THE MARKET 73

7.1 PATENT 73
7.2 DISSEMINATION 73
7.2.1 Article for the Newsletter of the PSC 73
7.2.2 Article pflanzenforschung.de 74
7.2.3 Open access publication 74
7.3 STAKEHOLDER DIALOG ON THE ACCEPTANCE OF NEW BREEDING TECHNOLOGIES 75
7.3.1 Identification and characterization of stakeholders 75
7.3.2 Fachtagung Dialog Grün 2016 77

7.3.3 Presentation and World-Café at FiBL ... 78
7.3.3.1 Aims an planning ... 78
7.3.3.2 Results and discussion ... 80
7.3.3.3 Synopsis and key findings from the World-Café and open questions from the questionnaire ... 82
7.3.3.4 Summary and outlook ... 87

8 GENERAL DISCUSSION AND OUTLOOK ... 88

9 REFERENCES ... 97

10 ACKNOWLEDGEMENTS ... 116

11 LIST OF APPENDICES ... 117

11.1 APPENDIX I: SCIENTIFIC PUBLICATION: THIEME, M., LANCIANO, S., BALZERGUE, S., DACCORD, N., MIROUZE, M. AND BUCHER, E. (2017) INHIBITION OF RNA POLYMERASE II ALLOWS CONTROLLED MOBILISATION OF RETROTRANSPOSONS FOR PLANT BREEDING. *GENOME BIOL* 18: 134. (17 PAGES) ... 118
11.2 APPENDIX II: PATENT: BUCHER*, E. AND THIEME*, M. (2017) MOBILIZATION OF TRANSPOSABLE ELEMENTS TO ENHANCE GENETIC AND EPIGENETIC VARIABILITY IN A POPULATION. *PATENT* WO2017/093317A1 (50 PAGES) ... 139
11.3 APPENDIX III: NEWSLETTER ARTICLE: THIEME, M. (2017) PUTTING PLANTS IN SHOOL: ON THE POTENTIAL OF EPIGENTIC MEMORY IN CROP BREEDING. *PLANT SCIENCE NEWS* 32: 4-5. (2 PAGES) ... 190
11.4 APPENDIX IV: TRANSCRIPT OF PRESENTATION FOR “FACHTAGUNG DIALOG GRÜN 2016”: THIEME, M. (2017) DER MOBILE TEIL DES PFLANZENGENOMS ALS RESSOURCE FÜR DEN ZUKÜNFTIGEN PFLANZENSCHUTZ. *NEUE TECHNOLOGIEN IN DER PFLANZENFORSCHUNG-EINE ALTERNATIVE ZU PFLANZENSCHUTZMITTELN?*: 36-39 (4 PAGES) ... 193
11.5 APPENDIX V: MULTIPLE CHOICE TEST (WORLD-CAFÉ) CORRECT ANSWERS MARKED IN RED (2 PAGES) ... 198
11.6 APPENDIX VI: QUESTIONNAIRE (WORLD-CAFÉ) (1 PAGE) ... 201
11.7 APPENDIX VII QUESTIONS FOR GUIDING THE DISCUSSION OF WORLD-CAFÉ (1 PAGE) ... 203
11.8 APPENDIX VIII: PRESENTATION GIVEN PRIOR TO THE DISCUSSION IN THE WORLD-CAFÉ AT FIBL (PICTURES WERE REMOVED) (5 PAGES) ... 205
11.9 APPENDIX IX ARTICLE PUBLISHED BY “PFLANZENFORSCHUNG.DE” “HITZE LÄSST *ONSEN* HÜPFEN-RETROTRANSPOSONEN KONTROLLIERT ZUM SPRINGEN BRINGEN” (2 PAGES) ... 211

1 List of abbreviations

5`UTR	5`untranslated region
A. thaliana	*Arabidopsis thaliana*
Ac/Ds	*Activator/Dissociation*
AGO	ARGONOUTE
AP	aspartic proteinase
ATRX5 or 6	ARABIDOPSIS TRITHORAX-RELATED PROTEIN 5 or 6
AZA	5-azacytidine
BARE-1 or *2*	Barley Retro Element-1 or 2
CMT3	CHROMOMETHYLASE3
CS	control stress (24h @ 6°C + 24h @ 24°C)
CTD	c-terminal domain
DCL	DICER-LIKE
DDM1	DECREASE IN DNA METHYLATION 1
DDT	dichlorodiphenyltrichloroethane
DNMtase	DNA methyltransferase
DRM	DOMAINS REARRANGED METHYLASE
dsRNA	double stranded RNA
E. coli	*Escherichia coli*
eccDNA	extrachromosomal circular DNA
ecDNA	extrachromosomal complementary DNA
EMS	ethylmethansulfonat
epiRIL	epigenetic recombinant inbred line
EU	European Union
EuGH	Europäischer Gerichtshof (European Court of Justice)
FWA	*FLOWERING WAGENINGEN*
GMO	genetically modified organism
H3K27m	histone 3 lysine 27 methylation
H3K9	histone 3 lysine 9
H3K9m2	histone 3 lysine 9 di-methylation
HDA6	HISTONE DEACETYLASE 6
HRE	heat response element
HS	heat stress (24h @ 6°C + 24h @ 37°C)
IN	integrase
IRAP	inter retrotransposon amplified polymorphism

KYP	KRYPTONITE
LARD	large retrotransposon derivative
LINE	long interspersed nuclear element
M. domestica	*Malus domestica*
^{m}C	5-methylcytosin
MET1	METHYLTRANSFERASE1
NRPB (Pol II)	RNA polymerase II
NRPB1 or 2	RNA polymerase II subunit 1 or 2
NRPD (Pol IV)	RNA polymerase IV
NRPE (Pol V)	RNA polymerase V
O. sativa	*Oryza sativa*
ORF	open reading frame
PBS	primer binding site
PSC	Zürich-Basel Plant Science Center
PTGS	post transcriptional gene silencing
PTT	polypurine tract
RdDM	RNA-directed DNA methylation
RDR	RNA-DEPENDENT RNA POLYMERASE
retroTE	retrotransposon
RH	RNaseH
RT	reverse transcriptase
sidRNA	dicer independent small RNA
SINE	short interspersed nuclear element
siRNA	small interfering RNA
SUVH	SUPPRESSOR OF VARIEGATION 3–9 HOMOLOGUE
TALEN	transcription activator-like effector nuclease
TE	transposable element
TIR	terminal inverted repeat
TRIM	terminal-repeat retrotransposons in miniature
TSD	target site duplication
TSS	transcriptional start site
VLP	virus-like particle
Z	zebularine
ZFN	zinc finger nuclease

2 Summary

Environmental stresses can trigger the activation and amplification of retrotransposons (retroTEs) in eukaryotes. The mobilization of retroTEs via their characteristic copy and paste mechanism, known to induce (epi)genetic diversity, is one the major drivers of genome evolution in plants and possibly most eukaryotes. Due to their potential as powerful mediators of adaptation to different environmental conditions, retroTEs are increasingly gaining attention as an endogenous genetic resource that could be harnessed for plant breeding. However, as transposition can also lead to detrimental mutations, retroTE mobility is normally limited by sophisticated silencing pathways implemented by their hosts.

In this work, I uncover the role of RNA polymerase II (Pol II) as a key regulator of retroTE-activity. By using two different inhibitors that specifically target the initiation of TE-silencing in plants, I was able to mobilize a stress responsive retroTE in *Arabidopsis* and thus opened up the so far sealed (epi)genetic resource of retroTEs for plant breeding. Due to the observation that the treatment with the two inhibitors also mobilized a retroTE in the distantly related rice crop, I conclude that this approach could in principle be used to mobilize retroTEs in virtually any plant.

Besides major advances in the field of epigenetics and retroTE- silencing, this work indeed provides the basis for a completely new approach that has also been patented and already commercialized in a start-up company. Thus, the second part of this thesis is thematically situated at the interface of science & innovation and science & policy. The described measures taken to communicate and discuss this new breeding approach form the basis for the futures public debate on the topic.

3 General introduction

Prior to the groundbreaking discovery of transposable elements (TEs) by Barbara McClintock (McClintock 1950) in the mid of the 20th century, the genetic information encoding for a certain phenotype had been assumed to be organized in a unidimensional and static way. Challenging this generally accepted theory, McClintocks observation of unmappable “mutable” loci underlying the highly diverse variegated kernel color of maize, initially did not attract a great deal of attention. Today it is known that the presence of potentially mobile genetic elements, like the *Ac/Ds* system discovered by McClintock, is the rule rather than the exception. Indeed, TEs have been detected in all organisms tested so far and in some extreme cases like maize and barley they make out more than 80% of the genome (Schnable *et al.* 2009; Wendel *et al.* 2016).

In stark contrast to their previous image as “parasitic elements” or “junk DNA” (Ohno 1972), increasing knowledge in the field of TE-research has rather revealed a multitude of cases illustrating their indispensable function during development, response to environmental triggers and as drivers of evolution. Apart from a multitude of examples found in plants, prominent findings from other kingdoms of life including humans are completing the latest image of TEs as vital basic modules of life (Chuong *et al.* 2017). For instance, the analysis of the genome of the black truffle (*Tuber melanosporum)* genome revealed a high number of unsilenced, active TEs in this species. Based on this observation the authors concluded this to be a mechanism to actively promote genome plasticity that allows truffle to adapt to sudden environmental changes (Montanini *et al.* 2014). Another striking example for the adaptive potential of mobile elements was found in *Drosophila*, where the insertion of a TE upstream of a cytochrome P450-gene underlies the emergence of insecticide (DDT)-resistant individuals (Chung *et al.* 2007). In humans there is increasing evidence for the pivotal role of TEs during fundamental developmental processes such as neurogenesis that ultimately allowed for the evolution of the enormous complexity of the human brain (Erwin *et al.* 2014). Consistent with the importance of TEs in developmental processes, more and more studies report on for the fatal consequences of misregulated TEs in humans. Thus, TEs and particularly LINE-1-elements are currently attracting great attention in cancer research (Burns 2017).

Against the background of the versatile functions of TEs that are partially based on highly conserved principles, I will hereafter focus on TEs and more specifically retrotransposons (retroTEs) in plants.

3.1 Transposable elements in plant genomes

3.1.1 Life cycle and general classification

Depending on the mechanism of transposition, TEs are assigned to two major classes. Class I TEs are evolutionary closely related to retroviruses and also referred to as retrotransposons (retroTEs). These elements fully depend on transcription by the host RNA polymerases II (Pol II) - or in special cases Pol III - and amplify themselves via an RNA intermediate resulting in a “copy and paste” transposition. In contrast, class II elements, also known as DNA-transposons, transpose without producing an RNA-intermediate, usually resulting in a “cut and paste” movement. Because of the increasing taxonomical complexity of known TEs, a regularly updated classification system is required (Piegu *et al.* 2015). However, in general, both main classes of TEs can be further subdivided according to their distinct structural and enzymatic characteristics.

For retroTEs, the main criterion for differentiation is the presence or absence of two long terminal repeats (LTRs) with the same orientation at the 5`and 3`end of the transposing unit. Plant LTR-retroTEs can range in size from less than 1 kb to up to 22 kb (*Ogre* element in pea) (Neumann *et al.* 2003). The LTRs, 0.085 kb- 5 kb in length (Zhao *et al.* 2016), contain both regulatory regions and the transcriptional start site (TSS) needed for transcription of the TE by the host Pol II (Fig. 1) (Grandbastien 2015). Transcripts originating from the 5´ LTR play two important roles in the retroTE life cycle. On one hand, they code for the retroTE-replication machinery or polyprotein (POL) consisting of the aspartic proteinase (AP), the reverse transcriptase (RT), the ribonuclease RNaseH (RH), the integrase (INT) and for the structural GAG capsid proteins that form a virus-like particle (VLP) (Grandbastien 2015). On the other hand, they serve as the template for reverse transcription resulting in extrachromosomal complementary retroTE-DNA (ecDNA) that is potentially capable to enter the nucleus and to integrate into genomic DNA (Schulman 2013). As transcription of LTR-retroTEs starts and ends within the LTRs, the region upstream of the TSS in the 5` LTR and

downstream the terminator site in the 3´ LTR is missing in the initial transcript that serves as a template for reverse transcription (Fig. 1). Hence, these regions would get lost unless restored in a complex mechanism during cDNA synthesis. Based on the homology of the two LTRs, retroTEs have evolved a mechanism to restore missing sequence information at both extremities of the LTRs using the homolog sequence that adjoins the internal domain of the TE. Two specific domains are crucial during this process: primer binding site (PBS) and polypurine tract (PPT) (Schulman 2013) (Fig. 1).

Depending on the sequential arrangement of the protein domains in the TE-body, LTR-TEs are further subdivided into Ty1-*copia* (Pseudoviridae) and Ty3-*gypsy* (Metaviridae)- retroTEs (Wicker *et al.* 2007) (Fig. 1).

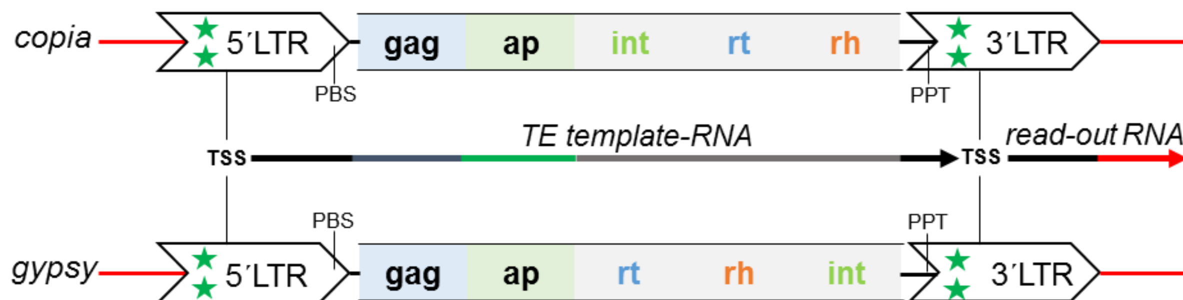


Figure 1 **Schematic representation of the structure of *copia* and *gypsy*- like LTR-retrotransposons in plants**. Upon activation, often mediated by regulatory motifs (green asterisks) located in the LTRs, transcription by the host Pol II starts at the transcriptional start sides (TSSs) that are located within in the 5´- and the 3´-LTR of the TE. Transcription from the 5´ LTR generates template transcripts that are either translated into the GAG capsid protein (gag), the polyprotein consisting of the aspartic proteinase (ap), the integrase (int), the reverse transcriptase (rt) and the RNaseH (rh) or used as a template for reverse transcription into TE- cDNA. The primer binding site (PBS) and the polypurine tract (PPT) are indicated. Transcription originating in the 3´ LTR generates read-out RNA of flanking genomic regions. Depending on the sequential arrangement of int, rt and rh LTR retrotransposons are subdivided into *copia* and *gypsy*-like elements. Adapted from (Grandbastien 2015; Wicker et al. 2007).

While intact *copia* and *gypsy*– like retroTEs code for all domains needed for their transposition, several additional non-autonomous LTR-retroTEs are known in plants. These include LARDs (Large Retrotransposon Derivatives) (Kalendar *et al.* 2004) TRIMs (Terminal-Repeat Retrotransposons In Miniature) (Witte *et al.* 2001) and SMARTs (Small LTR Retrotransposons) (Gao *et al.* 2012). All three families lack the coding regions needed to produce their own replication machinery. Instead, they carry an internal non-coding region that ranges in size from less than 0.3 kb (SMARTs and TRIMS) to up to 3.5 kb (LARDS). For their replication, these elements depend on proteins encoded by related autonomous elements. A descriptive case of this kind of functional trans-compensation of domains between related retrotransposons was shown for the *BARE-1 and 2 copia*-like LTR-retroTEs in barley. Due to a mutation in the open reading frame (ORF) of its structural GAG protein, *BARE-2* itself is considered to be non-autonomous. However, thanks to a high degree of homology, the *BARE-2*-retroTE can utilize *BARE-1*-GAG to fulfill its own lifecycle (Tanskanen *et al.* 2007). Although possibly biased by different class-dependent approaches of transposon annotations, retroTEs and more precisely LTR-retroTEs usually represent the largest class of TEs in currently known plant genomes (Fig. 2) (Vitte *et al.* 2014).

Non-LTR retroTEs in plants are represented by autonomous LINEs (long interspersed nuclear elements) and non-autonomous SINEs (short interspersed nuclear elements (Schmidt 1999). Transcription of LINEs is driven by an internal Pol II promoter located within in the 5´ UTR, making the process of self-amplification less complex. LINEs generally don`t code for structural capsid proteins such as the GAG. Rather, they use an RNA-binding protein that assembles together with the RT and RH to form a ribonucleoprotein instead of a VLP. Furthermore, the mechanism of integration differs between LTR and non-LTR elements. While reverse transcription of LTR-retroTEs takes place in the cytoplasm and is hence spatially separated from integration, LINE-elements integrate by target-primed reverse transcription and without making use of an integrase (Schulman 2013). In contrast to all previously described TEs, SINEs are derived from RNA-polymerase III transcripts such as 5S RNAs, 7SL RNA and tRNAs. They carry an internal Pol III promoter for transcription but depend on autonomous LINE-elements for their proliferation (Schulman 2013; Sun *et al.* 2007; Wenke *et al.* 2011).

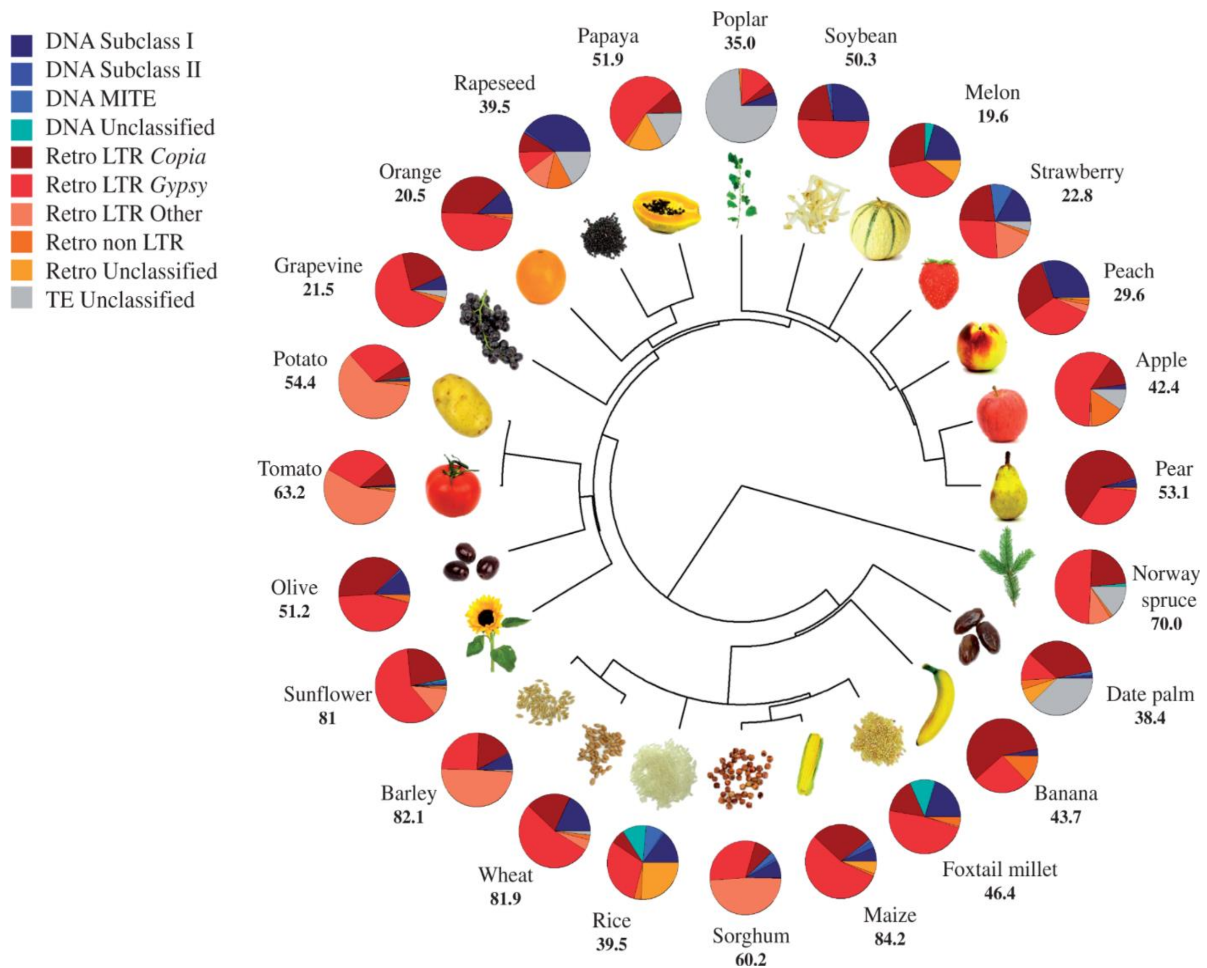


Figure 2 **Transposon content and diversity in 24 sequenced crop genomes in the context of their phylogenetic relationship**. The number below the species name indicates the total fraction of the genome annotated as TE-derived. The pie charts indicate the determined proportions of TE-subclasses. Adapted from (Vitte *et al.* 2014).

According to the current classification, DNA transposons generally comprise elements that do not rely on an RNA-intermediate for their transposition. With few exceptions such as *Helitrons* that transpose in a rolling circle-like mechanism (Kapitonov and Jurka 2001) the vast majority of class II elements in plants move through “cut and paste”, hence in a conservative mode of transposition. For this purpose, they solely encode for a single protein called the transposase. Analogous to the LTRs of class I elements, the internal region of DNA transposons is flanked by two terminal inverted repeat (TIR) sequences that are essential for transposition. The transposase catalyzes both the excision and the ligation of the DNA transposon into the target site. After integration, the sticky ends that are generated during the excision, are filled up a by a

DNA polymerase and a DNA ligase. This results in an element-specific target site duplication (TSD) at both termini of the TE. Both length of the TSD and the sequence of TIR and the sequence similarities of their transposases can be used to further subdivide class II elements. Accordingly, plant DNA-TEs are classified into six different subgroups: *Tc1-Mariner*, *hAT*, *Mutator*, *P*, *PIF-Harbinger* and *CACTA* (Wicker *et al.* 2007). Analogical to retroTEs, there are examples for non-autonomous DNA-transposons in plants. These include miniature inverted-repeat transposable elements (MITEs) that can be highly abundant in crop genomes (Vitte *et al.* 2014). In rice it was shown that non-autonomous MITEs depend on autonomous *Tc1-Mariner* elements for their mobility (Feschotte *et al.* 2003).

Although DNA-TEs typically transpose in a conservative process and without using an RNA intermediate, they can occasionally increase in copy number if transposition happens during the S phase of cell replication (Chen *et al.* 1987; Zhang *et al.* 2014).

3.1.2 Impact of transposable elements on genome size and architecture

As TEs represent substantial portions of plant genomes, they also play important roles as structural elements that shape genome architectures (Bennetzen and Wang 2014). The structural analysis and comparison of various sequenced angiosperm genomes including *Arabidopsis*, soybean and rice, has revealed that there is a strong positive correlation between the TE content and the actual genome size (Tenaillon *et al.* 2010). Besides drastic events such as whole genome duplications, TE-bursts are the main contributors to increases in genome size. With few exceptions, such as a reported burst of MITEs in rice (Naito *et al.* 2006), retroTE proliferation through copy and paste is the main driver of genome expansions in plants (Lee and Kim 2014; Piegu *et al.* 2006; Zedek *et al.* 2010). Such expansions are often triggered by a massive amplification of a few TE-families or individual elements (Bennetzen and Wang 2014). The potentially unlimited growth in genome size resulting from the copy and paste activity of active class I elements would ultimately lead to potentially adverse effects to the host (Bennett 1972; Bennetzen and Kellogg 1997; Bennetzen *et al.* 2005; Diez *et al.* 2013; Schubert and Oud 1997). Hence, unlimited accumulation of transposable elements in plant genomes is counteracted by illegitimate and homologous recombination events resulting in loss of DNA and genome shrinking (Devos *et al.*

2002; Hawkins *et al.* 2009). Sufficient sequence homology needed for homologous recombination can either exist between the two LTRs of one element but also between LTRs of two closely related retroTEs. This type of inter-element recombination can result in chromosome rearrangement and deletions of larger genomic fragments (Bennetzen and Wang 2014). Isolated LTRs, also called soloLTRs are abundant remnants of such events in plant genomes (Devos *et al.* 2002; Ma *et al.* 2004; Shirasu *et al.* 2000). Besides their substantial role as regulators of genome size, plant TEs are pivotal for global genome architecture. For instance, there is evidence from rice and maize that TE insertions have been driving the evolution of the centromeric regions of chromosomes that are crucial during cell replication (Gao *et al.* 2015; Sharma *et al.* 2013; Wolfgruber *et al.* 2009). On a smaller scale, active TEs can contribute to the evolution of new genes by mediating gene fragment transduction or exon shuffling (Elrouby and Bureau 2010; Jiang *et al.* 2004). It has further been reported that mobile *Helitrons* and MULE-elements (so called Pack-MULEs) can capture and amplify functional genes in maize and rice (Jameson *et al.* 2008; Jiang *et al.* 2004).

3.1.3 Retrotransposons in the model plant *Arabidopsis thaliana*

Depending on the annotation methods used to detect TEs (Joly-Lopez and Bureau 2014) current estimations of TE-derived DNA in the *Arabidopsis thaliana* genome range from 15 % (de la Chaux *et al.* 2012) to up to 23.7 % (Hu *et al.* 2011). Confirming the general trend in plants, class I elements represent the larger share of TE-derived DNA in *Arabidopsis* although their overall copy number (9 021) is assumed to be smaller compared to that of more abundant but shorter class II elements (12 631) (de la Chaux *et al.* 2012). As also reported for many crop genomes (Vitte *et al.* 2014), *gypsy* and *copia*-like TEs are the predominant retroTEs in *Arabidopsis* (Buisine *et al.* 2008; Quadrana *et al.* 2016; The_Arabidopsis_Genome_Initiative 2000).

Genome wide analyses suggest that family-specific patterns of TE-distributions exist in eukaryotes (Sultana *et al.* 2017). In *Arabidopsis*, the three main types of class I TEs, *copia*, *gypsy* and *LINE* differ in abundancy along the chromosomes (Underwood *et al.* 2017). While *gypsy* elements are more likely to be found in close proximity to centromeres, it appears that *copia* and *LINE* elements are more frequently observed in pericentromeric regions. Furthermore, there is evidence that *copia* elements in

Arabidopsis tend to be more often linked to coding regions (Ito *et al.* 2011; Quadrana *et al.* 2016). Although similar family specific patterns of TE-distribution have also been reported for maize (Baucom *et al.* 2009) *Arabidopsis* seems to be an exception considering the equal TE-distributions in other crops such as apple (Daccord *et al.* 2017) and rice (Mirouze and Vitte 2014; Nobuta *et al.* 2007).

3.2 Regulation of retrotransposition in plants

3.2.1 DNA methylation

DNA methylation at cytosines in the form of 5-methylcytosine (^{m}C) is a key feature of epigenetic regulation and silencing of potentially mobile genetic elements in plants (Miura *et al.* 2001). The abundancy of ^{m}C has been shown to correlate with the presence of repetitive elements and transposons in the genome of *Arabidopsis* substantiating the role of DNA-methylation in regulating the activity of these elements (Cokus *et al.* 2008; Zhang *et al.* 2006). Depending on the sequence context (CG, CHG or CHH where H can be any base but G), it has been proposed that plants have evolved different families of specialized DNA-methyltransferases (DNMtases) to establish or maintain DNA-methylation (Henderson and Jacobsen 2007). Global DNA-methylation levels in *Arabidopsis* have previously been determined to be approximately 24 % (CG), 6.7.% (CHG) and 1.7 % (CHH), respectively (Cokus *et al.* 2008).

3.2.1.1 Maintenance methylation

Following DNA replication that results in hemimethylated DNA, the specific ^{m}C-pattern of the newly synthetized daughter strand has to be restored. In case of symmetric methylation in the CG-context, METHYLTRANSFERASE1 MET1 (Finnegan *et al.* 1996; Kankel *et al.* 2003) which is a homologue of the mammalian maintenance methyltransferase Dnmt1 (Finnegan and Dennis 1993) adds methyl groups to the new unmethylated daughter strand. Symmetric CHG methylation in plants is catalyzed by the plant specific CHROMOMETHYLASE3 (CMT3) and in a reinforcing loop triggered by the activity of the histone 3 lysine 9 (H3K9) methyltransferase KRYPTONITE (KYP, also known as SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 4, SUVH4), SUVH5 and 6 (Ebbs and Bender 2006; Jackson *et al.* 2004; Jackson *et al.* 2002; Lindroth *et al.* 2001). Recently, MET2a, a less well characterized homolog of MET1

has also been reported to be important for CHG methylation exclusively at active TE families in *Arabidopsis* (Quadrana *et al.* 2016; Stroud *et al.* 2013). So far, two different mechanisms leading to specific methylation at asymmetric CHH sites are known in plants. While CMT2 recognizes H3K9me2, a heterochromatic mark co-localizing with long TEs (Gouil and Baulcombe 2016), activity of the *de novo* methyltransferase DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) is guided by sequence specific small interfering (si)RNAs (Cao and Jacobsen 2002).

3.2.1.2 RNA-directed DNA methylation (RdDM)

The potential of retroTEs to produce and integrate intact and unmethylated copies of themselves along the genome triggered the evolution of specialized silencing machineries in plants as well as in animals and fungi (Wheeler 2013). Generally, these silencing mechanisms are based on the production of highly specific small non-coding RNA molecules that guide the silencing machinery to the corresponding locus in the genome. In plants, this process known as RNA-directed DNA methylation (RdDM) (Matzke *et al.* 2015; Wassenegger *et al.* 1994) is based on the activity of two additional plant specific RNA-polymerases, NRPD (Pol IV) and NRPE (Pol V) that have evolved as specialized paralogs of NRPB (Pol II) (Herr *et al.* 2005; Kanno *et al.* 2005; Matzke *et al.* 2015; Onodera *et al.* 2005; Pontier *et al.* 2005; Ream *et al.* 2009). In contrast to Pol II which is the main source of gene-transcripts in plants, mutants of Pol IV and V do not show any developmental defects in *Arabidopsis* (Pontier *et al.* 2005). The core of this complex silencing pathway, also called the “canonical RdDM“, is now well investigated in *Arabidopsis* (Wendte and Pikaard 2017).

Current models suggest that in a first step, RNA-DEPENDENT RNA POLYMERASE 2 (RDR2), converts non-coding RNAs produced by Pol IV into double stranded RNAs (dsRNAs) (Haag *et al.* 2012; Law *et al.* 2011). These non-coding dsRNAs are subsequently processed by DCL3 (DICER-LIKE 3) into 24 nt siRNAs, stabilized by methylation of their 3`-OH groups (Yang *et al.* 2006) and loaded onto AGO4 (Zilberman *et al.* 2003) or AGO6 (Zheng *et al.* 2007). Finally, presumably mediated by specific base pairing of AGO4-bound siRNAs with a nascent Pol V transcript (Wierzbicki *et al.* 2009), DRM2 is recruited resulting in a highly specific methylation of the Pol V transcribed loci in all three sequence contexts (Böhmdorfer *et al.* 2014).

Besides this well studied Pol IV-RDR2-dependent branch of the “canonical” RdDM pathway, there is emerging evidence for the importance of alternative routes of siRNA biogenesis resulting in Pol V-dependent TE-silencing (Cuerda-Gil and Slotkin 2016) in plants. With few exceptions including the Pol IV-NERD RdDM pathway (Pontier *et al.* 2012), these so-called “non-canonical” RdDM pathways have in common that they are directly dependent on transcriptional activity of Pol II. Although reverse genetic approaches for studying Pol II-dependent silencing mechanisms are limited by a drastic loss of viability of Pol II-mutants (Zheng *et al.* 2009) there is an increasing evidence describing silencing pathways acting downstream of Pol II (Fig. 3). Accordingly, it was shown that Pol II transcripts can be processed into primary 21-22 nt siRNAs in a RDR6-DCL2-DCL4-dependent manner (Nuthikattu *et al.* 2013). There is recent evidence, that this pathway also termed RDR6-RdDM targets especially transcriptionally active, full length TEs in *Arabidopsis* (Panda *et al.* 2016). Importantly, primary siRNAs whose production can also be triggered by the DCL1-dependent formation of microRNAs (Creasey *et al.* 2014) can feed into an RNAi loop leading to a strong RDR6-dependend accumulation of secondary siRNAs and post transcriptional gene silencing (PTGS) (Cuerda-Gil and Slotkin 2016). Besides the RDR6-DCL2-DCL4-dependent production of 21-22 nt siRNAs, Pol II transcripts can also be processes in a RDR6-DCL3-dependet manner resulting in 24 nt siRNAs, presumably reflecting partial redundancy and potential functional compensation between DCL2/4 and DCL3 in specifically silencing TEs present in high copy numbers in the genome (Gasciolli *et al.* 2005; Mari-Ordonez *et al.* 2013).

Only recently, another Pol II-dependent “non-canonical” RdDM pathway that also relies on DCL3 but that functions independently of RDRs resulting in the production of 24 nt siRNAs has been identified in *Arabidopsis* (Panda *et al.* 2016). It is presumed that this kind of RDR-independent TE silencing is triggered by the formation of imperfectly paired dsRNAs as they occur after transcription of inverted repeats. This scenario has been reported for the Mu killer locus that causes stable trans silencing of potentially active members of the the MuDR-family in maize (Slotkin *et al.* 2005). As transcription by Pol II directly precedes dicing by DCL3 it is self-evident that this pathway is also important in re-silencing TEs that are in a transcriptionally active state (Panda *et al.* 2016).

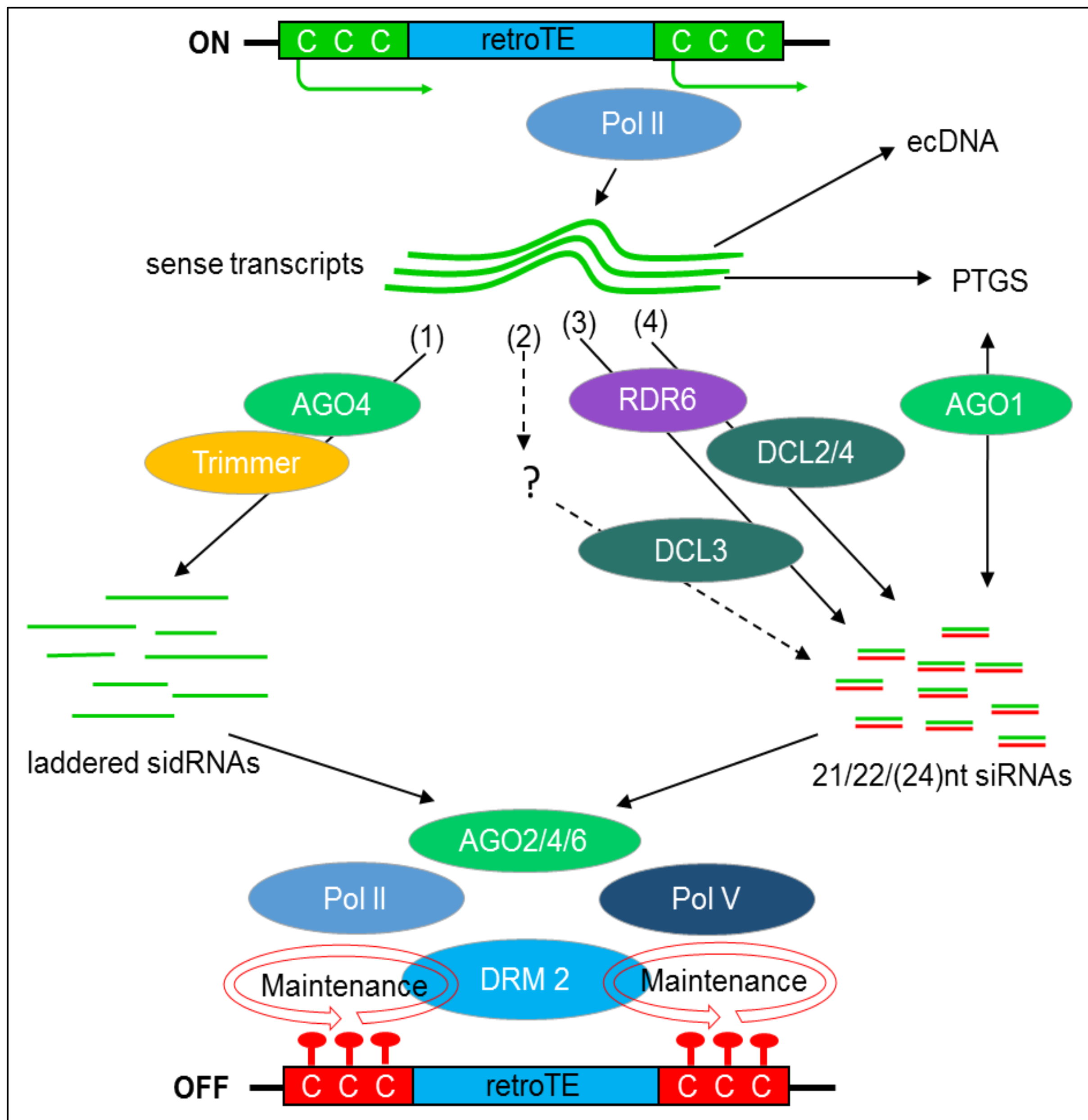


Figure 3 **Current model for the presumed mechanism of the “non-canonical” expression-dependent RdDM leading to retroTE silencing in plants.** Four known major routes for the processing of Pol II-derived sense-transcripts of retroTEs resulting in the production of retroTE-specific siRNAs are depicted. The interaction of AGO-bound siRNAs with scaffold transcripts produced by Pol II or Pol IV guides the DNMtase DRM2 and leads to highly specific and stable methylation (red lollipops) of the previously active retroTE and potentially other homologous copies in the genome. Numbers in brackets indicate key references that lead to this model as follows: (1) (Ye *et al.* 2016); (2) (Panda *et al.* 2016); (3) (Mari-Ordonez *et al.* 2013); (4) (Nuthikattu *et al.* 2013). See text and (Cuerda-Gil and Slotkin 2016) for details.

Another Pol II-dependent mechanism for the production of small RNAs triggering initial TE-silencing was found to be independent of DCL-enzymes (Ye *et al.* 2016). The authors of this study detected a new class of dicer independent small RNAs (sidRNAs) that ranged from 20-60 nt in size. Based on their observations they concluded that these sidRNAs are generated by stepwise trimming of the 3` end of AGO4-bound Pol II or Pol IV transcripts originating from active loci including retroTEs.

Although the various routes of siRNA biogenesis described above can be disentangled based on their specific mechanistic properties, their actual biological function and importance is challenging to be determined. For instance, Panda and colleagues found evidence that a functional loss of the RDR6-RdDM pathway can be compensated by Pol IV-dependent silencing. More importantly, it seems that the transcriptional state of a TE predetermines which of the above described RdDM pathways becomes dominant in (re)-silencing certain TEs (Panda *et al.* 2016).

3.2.2 Chromatin state

Chromatin density is a key regulatory element of gene regulation and TE-silencing in plants. Nucleosomes have approximatively 146 bp of DNA wrapped around a histone core complex and form the basic repeating units of eukaryotic chromatin. The core complex itself consists of eight histone proteins and more specifically of two dimers of H2A and H2B and a tetramer composed of two dimers of histones H3 and H4 (Luger *et al.* 1997).

Besides the reported distinct distribution of different histone variants (Stroud *et al.* 2012; Yelagandula *et al.* 2014), dynamic chemical modifications, such as acetylation, phosphorylation or methylation, of the N-terminal histone tails strictly correlate with the expressional state of chromosomal regions in *Arabidopsis* (Fuchs *et al.* 2006). There is evidence for a very close connection between DNA-methylation pathways and certain types of histone modifications in plants. For instance, it was reported that CHG-methylation correlates with histone H3 lysine nine di-methylation (H3K9me2) (Bernatavichute *et al.* 2008; Jackson *et al.* 2004) which is considered as an indicator for the presence of heterochromatin (Fuchs *et al.* 2006). A simultaneous knock-out of the histone methyltransferases SUVH4-6 resulted in the concomitant reduction of H3K9me2 and non-CG DNA methylation resulting in release of TE-transcription (Ebbs

and Bender 2006). Likewise, it was shown that TEs coincide with high levels of histone 3 lysine 27 monomethylation and that defects in the H3K27 monomethyltransferases ARABIDOPSIS TRITHORAX-RELATED PROTEIN 5 (ATXR5) and ATRX6 resulted in transcriptional up-regulation of TEs (Jacob *et al.* 2010).

Besides the addition of repressive marks, it was also reported that the active removal of activating marks, such as histone acetylation plays an important role for TE-silencing. For instance, knocking-out HISTONE DEACETYLASE 6 (HDA6) led to a loss of heterochromatin formation. It is assumed that an interaction between HDA6 and MET1 results in locus specific methylation in the CG-context thereby forming the basis for subsequent non-CG methylation resulting in silent chromatin (To *et al.* 2011). Consequently, it was found that a subset of TEs are transcriptionally reactivated in the *hda6* background (Liu *et al.* 2012; Probst *et al.* 2004).

A key-factor that mediates the crosstalk between the chromatin state and DNA-methylation in *Arabidopsis* is the chromatin-remodeling ATPase DECREASE IN DNA METHYLATION 1 (DDM1). By providing DNMtases access to heterochromatin (Zemach *et al.* 2013), DDM1 has been reported to play an important general role in maintaining CG and non-CG methylation (Jeddeloh *et al.* 1999; Vongs *et al.* 1993) and TE-silencing (Lippman *et al.* 2003; Miura *et al.* 2001; Tsukahara *et al.* 2009). Mutating DDM1 results in a loss of H3K9me2, a global reduction of CG-DNA methylation and a decrease of heterochromatin condensation at TEs in *Arabidopsis* (Gendrel *et al.* 2002; Lippman *et al.* 2004). As a consequence, *ddm1* plants show a genome-wide upregulation of TEs (Lippman *et al.* 2004) and it was shown that inbreeding of *ddm1* mutant lines resulted in transposition of various class I and II TEs, including *gypsy* and *copia* elements (Tsukahara *et al.* 2009). Due to its central function in maintaining TE-silencing, several main discoveries elucidating expression-dependent TE-silencing pathways such as the RDR6- or the DCL3-dependent RdDM have been made in the *ddm1*-background where TEs are in a more “active state” (McCue *et al.* 2015; Nuthikattu *et al.* 2013; Panda *et al.* 2016)

3.3 Retrotransposons: Linking the genome to the environment

3.3.1 Stress-dependent activation of retrotransposons in plants

Already in the mid of the last century, when Barbara McClintock published her pioneering work on DNA-TEs in maize she described them as “controlling elements” (McClintock 1950; McClintock 1956) and linked their activation to the occurrence of a “challenge” or a stress (McClintock 1984). She suggested that the occurrence of “genomic shocks” as for example of global mutations caused by irradiation could trigger the activation of transposition in plants (McClintock 1984). Today her revolutionary hypotheses have been confirmed in a multitude of studies that clearly illustrate the role of both classes of TEs as potentially mobile regulatory elements and sensors of various stresses (Negi *et al.* 2016). Generally, retroTE-activity can be detected during different stages of their life cycle (Schulman 2013). As reviewed by Negi and colleagues, most of the currently known examples for the stress-dependent activation of class I elements are based on the detection of increased TE-transcript levels (Negi *et al.* 2016).

3.3.1.1 Examples for environmental stimuli

The list of conditions reported to induce retroTE-transcription comprises a broad panel of various biotic and abiotic stresses including tissue culture. In *Arabidopsis* there is evidence that elicitation with bacterial flagellin (Yu *et al.* 2013), heat (Pecinka *et al.* 2010; Tittel-Elmer *et al.* 2010), arsenic stress (Castrillo *et al.* 2013) salt stress, nutrition starvation as well as abscisic acid and cytokinin treatment (Duan *et al.* 2008; Zeller *et al.* 2009) can trigger the transcriptional activation of certain retroTEs. From what was found in other plant species including important crops, the list of activating environmental conditions can be expanded by treatment with the fungal elicitor cryptogein (*Nicotiana tabacum*) (Anca *et al.* 2014), wounding (*Quercus suber, Solanum chilense*) (Rocheta *et al.* 2012; Tapia *et al.* 2005), aphids infestation (*Pinus sylvestris*) (Voronova *et al.* 2014), cold (*Citrus sinensis*) (Butelli *et al.* 2012), water deficiency (*Zea mays*) (Lu *et al.* 2011), UV-light (*Avena sativa and Cucumis melo)(Kimura et al. 2001; Ramallo et al. 2008)* and herbicide treatment with atrazine (*Oryza sativa*) (Zhang *et al.* 2012). In rice, there is further evidence that inter-specific hybridization events can trigger retrotransposition of some class I elements (Wang *et al.* 2009).

3.3.1.2 Mechanisms underlying the stress response of LTR-retrotransposons

As described above (Fig. 1) transcription of *copia* and *gypsy*-type class I elements initiates at the TSSs of both 5`and 3` LTRs. As reviewed by Galindo-Gonzalez and colleagues, there are some well-studied cases in plants elucidating the molecular mechanism behind the observed stress-dependency of retroTE-activation (Galindo-Gonzalez *et al.* 2017). Hence, it was shown for different stress-responsive class I elements that *cis*-regulatory motifs in the LTRs play a pivotal role in converting an environmental trigger into transcriptional upregulation and retroTE-mobilization (Casacuberta and Grandbastien 1993; Takeda *et al.* 1998; Tapia *et al.* 2005). In this regard, a well-studied retroTE is the heat responsive *Ty1*/*copia*-like retroTE *ONSEN* (*AtCOPIA78*) in *Arabidopsis* (Pecinka *et al.* 2010; Tittel-Elmer *et al.* 2010). A closer examination of the promoter in the *ONSEN*-LTRs revealed that all eight copies in the Columbia ecotype of *Arabidopsis* contain a heat response element (HRE) with the consensus sequence nTTCnnGAAn (Cavrak *et al.* 2014). This study further provided evidence that binding of specific heat shock factors (HSFs) such as HSFA2 to these HREs mediates the heat-stress dependent activation of *ONSEN* transcription (Cavrak *et al.* 2014).

3.3.2 Impact on gene structure and expression

3.3.2.1 Insertional mutagenesis

Transposition of class I elements can affect gene expression in various ways (Lisch 2013). Loss of function mutations caused by insertions into regulatory regions or introns and exons of genes are commonly observed in plant genomes. Apart from obvious phenotypes caused by insertions of these elements mobilized in silencing mutants of *Arabidopsis* (Mirouze *et al.* 2009; Tsukahara *et al.* 2009) many examples are found among selected crop varieties. For instance, photoperiod-insensitivity in cultivars of *Glycine max* (Kanazawa *et al.* 2009), parthenocarpy in varieties of *Malus domestic*a (Yao *et al*. 2001) and an observed dwarf phenotype in *Oryza sativa* (Chen *et al.* 2017) have been linked to insertional mutagenesis causes by class I elements.

3.3.2.2 Impact on splicing

Besides causing loss of function mutations, retroTE-insertions can also lead to changes in splicing. In case of three different alleles of the *waxy* gene involved in amylose biosynthesis in maize it was shown that insertions of class I elements into introns can cause alternative splicing (Varagona *et al.* 1992). A recent study in oil palm further illustrated that also dynamic epigenetic modifications of inserted retroTEs can result in drastic phenotypic changes. Hence, it was reported that alternative splicing caused by hypo-methylation of a LINE element inserted in the intron of the homeotic gene *DEFICIENS* underlies the mantled somaclonal variant of oil palm (Ong-Abdullah *et al.* 2015).

3.3.2.3 Expressional changes in retrotransposon-flanking regions

Functional *copia* and *gypsy* elements are flanked by identical LTRs with the same orientation. Hence, transcription originating at the TSS in the 3`LTR (Fig. 1) may also drive expression of flanking genomic regions (Chuong *et al.* 2017; Grandbastien 2015). Depending on the *cis* regulatory elements of the retroTE, this so-called "read-through" transcription can therefore also be stress-dependent. In blood oranges, this kind of stress-dependent gene activation mediated by the 3`LTR of a retroTE has been shown to account for the accumulation of anthocyanins following cold-stress. Mechanistically, the cold responsive promoter in the 3`LTR drives the expression of a MYB transcriptional activator of anthocyanin production (Butelli *et al.* 2012). It was also reported that following a recombination event in some orange cultivars, the remaining soloLTR of the same retroTE is sufficient to cause the same cold-dependent phenotype. Striking evidence for the importance of various TE-super-families including retroTEs as mediators of abiotic stress responses was found in maize. It was revealed that as much as 33 % of genes that are only expressed in response to abiotic stresses are associated with upstream inserted TEs (Makarevitch *et al.* 2015).

Besides read-through transcription, the insertion of a retroTE close to a gene can also give rise to the recruitment and spreading of repressive epigenetic marks resulting in slicing of flanking regions (Sigman and Slotkin 2016). A famous example for the alteration of the epigenetic landscape flanking a class I element (SINE-element) has been reported for the *FLOWERING WAGENINGEN (FWA)*-locus in *Arabidopsis*

*(*Kinoshit*a et al.* 2007*)*. In this case, the closely inserted retroTE mediates establishment of DNA-methylation and hence forms the basis for imprinting of the *FWA*-locus (Fujimoto *et al.* 2008). TEs can thus bring endogenous genes under epigenetic control.

3.3.2.4 Gene movement

The regulation and expression of a gene can be fundamentally changed if it is relocated into a different genomic region. There is increasing evidence that TEs can contribute to such major events in plants (Lisch 2013). A striking example for a retroTE-mediated gene movement was found in *Solanum lycopersicum* where a 24.7 kB gene duplication event resulted in the evolution of the *SUN* locus that is underlying the oval shape of Roma tomatoes (van der Knaap *et al.* 2004; Xiao *et al.* 2008).

3.4 Retrotransposons as a tool for crop improvement and breeding

As described above and reviewed by (Vitte *et al.* 2014) ongoing retrotransposition events substantially contribute to phenotypic diversity of various crops. With growing knowledge of the causes and the consequences of class I element mobility in crops, they are increasingly seen as a valuable resource to advance plant breeding (Mirouze and Vitte 2014; Paszkowski 2015). However, due to sophisticated TE-silencing mechanism described above, there are relatively few cases where the frequency of retrotransposition was efficiently increased in plants. Exceptions are found in cases where plants are exposed to major stresses such as growth in tissue culture (Hirochika *et al.* 1996; Masuta *et al.* 2017) or when strong genomic stresses, like during inter-specific hybridization, occur (Wang *et al.* 2009).

3.4.1 Detection of retrotransposition events

A prerequisite for harnessing class I elements for plant breeding is their efficient detection during and after transposition (Vitte *et al.* 2014). Advances in sequencing technologies and downstream data processing significantly contributed to the recent progress in the field of TE-research. Especially the availability of longer sequencing reads overcomes detection limitations of TE-associated structural variants caused by

the repetitive nature of TEs (Debladis *et al.* 2017). Recently developed methods to specifically sequence the active “mobilome” (mobile genetic elements in a cell) will further enhance the use of TEs in plant breeding (Lanciano *et al.* 2017).

3.4.2 Genetic engineering

Fundamental discoveries elucidating TE-silencing pathways in plants were achieved by studying knock-out mutants of the *Arabidopsis* model plant (Ito *et al.* 2011; Mirouze *et al.* 2009; Miura *et al.* 2001). In principle, the approach of using mutants defective in TE-silencing could also be promising to induce retrotransposition in crops. Indeed, there are examples for the analogical mobilization of retroTEs in crops. In rice for instance, the mobilization of the *copia*-like retroTE Tos17 was enhanced by knocking out a H3K9 methyltransferase (Ding *et al.* 2007). However, the limited availability of mutant alleles (Paszkowski 2015) that are in addition possibly accompanied by severe side-effects like observed in rice (Hu *et al.* 2014) or maize (Li *et al.* 2014), currently restricts the use of retroTEs in crop breeding. Yet, recent major advances in the field of genetic engineering and DNA-sequencing opens up completely new possibilities (Springer and Schmitz 2017). Besides targeted mutagenesis of TE-silencers by means of zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) or CRISPR-Cas9 systems, especially specific “epigenome editing” using for example a nuclease deficient Cas9 protein (dCAS9) fused to a methyltransferase (Park *et al.* 2016) are catalyzing the progress in the field.

3.4.3 Transient inhibition of retrotransposon silencing

An alternative approach to increase retroTE mobility in plants is to transiently inhibit enzymes involved in epigenetic silencing (Pecinka and Liu 2014). By targeting highly conserved silencing-pathways, the drug-mediated retroTE-activation does neither require prior knowledge of the DNA-sequence nor elaborate genetic engineering. Originally developed as agents in cancer therapy (Lyko and Brown 2005), mainly inhibitors of DNMtases such as the cytidine-analogues 5-azacytidine (AZA) and zebularine (Z) are now also commonly used to induce epigenetic changes in plants (Griffin *et al.* 2016). During cell replication, both cytidine-analogues are incorporated into the DNA. In bacterial model systems, it was shown that the presence of these

cytidine-analogues leads to a very stable (Z) (Champion *et al.* 2010) or even covalent (AZA) (Santi *et al.* 1983) binding of DNMtases to the DNA resulting in a depletion of active DNMtases and subsequent loss of DNA methylation. In plants, transient inhibition of DNMtases leads to a reduction of methylation levels in all sequence contexts (Griffin *et al.* 2016). In case of Z, that is considered to be more stable compared to AZA, the observed reduced DNA-demethylation resulted in the transcriptional release of the non-LTR retroTE LINE1-4 in *Arabidopsis* (Baubec *et al.* 2009).

3.5 Policy related aspects of bridging science and innovation

Due to the limitations described above, the approach of harnessing epigenetics and endogenous mobile genetic elements for plant breeding has so far not been actively implemented. At the same time, recent technological achievements are catalyzing research and contribute to the rediscovery of TEs as important drivers of evolution (McClintock 1950; Mirouze and Vitte 2014; Paszkowski 2015). Hence, with the aim to develop a method to allow for the stress-dependent mobilization of retroTEs for plant breeding, this project was situated at the interface of basic research and the implementation of a new breeding technology. At this stage of development it was therefore- besides providing a proof of concept in the laboratory- also important to set the right course for futures steps in the process of bringing the method developed in the frame of this thesis to the market. Thus, important policy-related aspects linked to the project can be summarized as follows:

- Assess the patentability of the method;
- Disseminate scientific results to inform different stakeholders involved;
- Initiate a stakeholder dialog on the acceptance of the new method.

3.6 Aims of the thesis

The versatility of TEs as main components of eukaryotic genomes is currently attracting the attention of scientists from various fields of research. As natural drivers of plant evolution, TEs are particularly interesting as a so far widely underutilized genetic resource for crop breeding. Even though new methodologies are catalyzing the overall gain of knowledge, there are still major scientific questions that remain to be answered. Besides the elucidation of silencing mechanisms that safeguard genome stability, research in the field of TE-biology also aims to understand how transposable elements shape genomes and their epigenetic landscapes. These findings are, in turn, used to assess the role of TEs during development, in creating phenotypic diversity or in mediating adaptation

The aims of my dissertation are to elucidate the fundamental mechanisms involved in retroTE-silencing in the model plant *Arabidopsis* and to apply these findings to induce retrotransposition in crops such as soybean and rice. In doing so, I will follow the approach of mobilizing retroTEs using inhibitor molecules that specifically target highly conserved pathways in plants. More specifically I will try to address the presumably so far underestimated role of Pol II as a key-regulator of retroTEs mobility in plants and more generally in eukaryotes. Due to the fact that retrotransposition can entail a broad panel of (epi)genetic changes, one objective of this work is also to track the phenotypes linked to novel retroTE-insertions in the successive generations.

As the approach of using specific inhibitors does not comprise genetic engineering, it could in principle also open up new avenues to accelerate plant breeding for the organic sector. This work also aims at providing a basis that will allow for a fruitful discourse between different stakeholders in the field of conventional and organic breeding. Hence, the implementation of different measures at the science and policy interface are also key elements of this thesis.

4 RNA polymerase II regulates retrotransposon mobility in *Arabidopsis*

A modified version of this chapter was published in Thieme, M. *et al.*, (2017). Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. Genome Biology 18, 134.

4.1 Abstract

To ensure genome integrity several silencing mechanisms have evolved to repress retrotransposon mobility in plants. Even though retroTEs fully depend on transcriptional activity of the host RNA polymerase II (Pol II) for their mobility, it was so far unclear whether Pol II is directly involved in repressing their activity. Here I show that plants defective in Pol II activity lose DNA methylation at repeat sequences and produce more extrachromosomal retroTE DNA of the *ONSEN*-retroTE upon heat stress in *Arabidopsis*. I demonstrate that Pol II acts at the root of transposon silencing presumably in a DCL-dependent and RDR-independent manner. Preliminary data further suggest that pairing of Pol II-derived sense and antisense transcripts could generally initiate RDR-independent retroTE-silencing in plants.

4.2 Introduction

Like retroviruses, LTR-retroTEs (class I elements), which represent the most abundant class of transposable elements (TEs) in eukaryotes, transpose via a copy and paste mechanism. This process requires the conversion of a full length RNA-polymerase II (Pol II) transcript into extrachromosomal complementary DNA (ecDNA) by reverse transcription (Schulman 2013). In their life cycle LTR-retroTEs can produce extrachromosomal circular DNA (eccDNA) which is an indicator for their ongoing activity (Lanciano *et al.* 2017). As transposition of class I elements can jeopardize genome integrity, plants have evolved several regulatory pathways to retain control over the activity of these potentially harmful mobile genetic elements. Cytosine methylation (^{m}C) plays a central role in TE-silencing in plants (Miura *et al.* 2001). In addition, plants have evolved two Pol II-related RNA-polymerases Pol IV and Pol V that are essential to provide specific silencing signals leading to RNA-directed DNA

methylation (RdDM) at retroTEs (Matzke *et al.* 2015) thereby limiting their mobility (Ito *et al.* 2011; Mirouze *et al.* 2009; Tsukahara *et al.* 2009). More recently, various additional non-canonical Pol IV-independent RdDM pathways have been described (Cuerda-Gil and Slotkin 2016). Notably it was found that Pol II itself also plays an important role in RdDM (Gao *et al.* 2010; Zheng *et al.* 2009) by feeding template RNAs into downstream factors such as RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) resulting in a dicer-dependent or independent initiation and establishment of TE-specific DNA methylation (Cuerda-Gil and Slotkin 2016). Beyond that, recent work suggests a new “non-canonical” branch of RdDM that specialized in targeting transcriptionally active full-length TEs (Panda *et al.* 2016). This pathway functions independently of RDRs via Pol II transcripts that are directly processed into siRNAs by DCL3.

Here, we wanted to investigate if Pol II could play a direct role in repressing retroTE mobility in plants. For that purpose we chose the well-characterized heat-responsive *copia*-like *ONSEN* retroTE (Cavrak *et al.* 2014; Ito *et al.* 2011) of *Arabidopsis* and took advantage of the hypomorphic *nrpb2-3* mutant allele that accumulates reduced NRPB2 (the second-largest component of Pol II) protein levels (Zheng *et al.* 2009).

4.3 Results

4.3.1 RNA polymerase II represses the heat-dependent activation of *ONSEN*

Using quantitative real-time PCR (qPCR), we determined that challenging *nrpb2-3* seedlings by heat stress (HS) led to a mild increase in total *ONSEN* copy number (sum of extrachromosomal complementary DNA (ecDNA), extrachromosomal circular DNA (eccDNA) and new genomic insertions) relative to control stress (CS) and compared to the wild-type (WT, Fig. 4a). This result was supported by the observed dose responsive increase in *ONSEN* copy number after HS and pharmacological inactivation of Pol II with α-amanitin (A), a potent Pol II inhibitor (Lindell *et al.* 1970) that does not affect Pol IV or Pol V (Haag *et al.* 2012) (Fig. 4a and b). In order to test the interaction between Pol II-mediated repression of retroTE activation and DNA methylation we grew WT and *nrpb2-3* plants on media supplemented with zebularine (Z), an inhibitor of DNA methyltransferases active in plants (Baubec *et al.* 2009) and submitted them to HS.

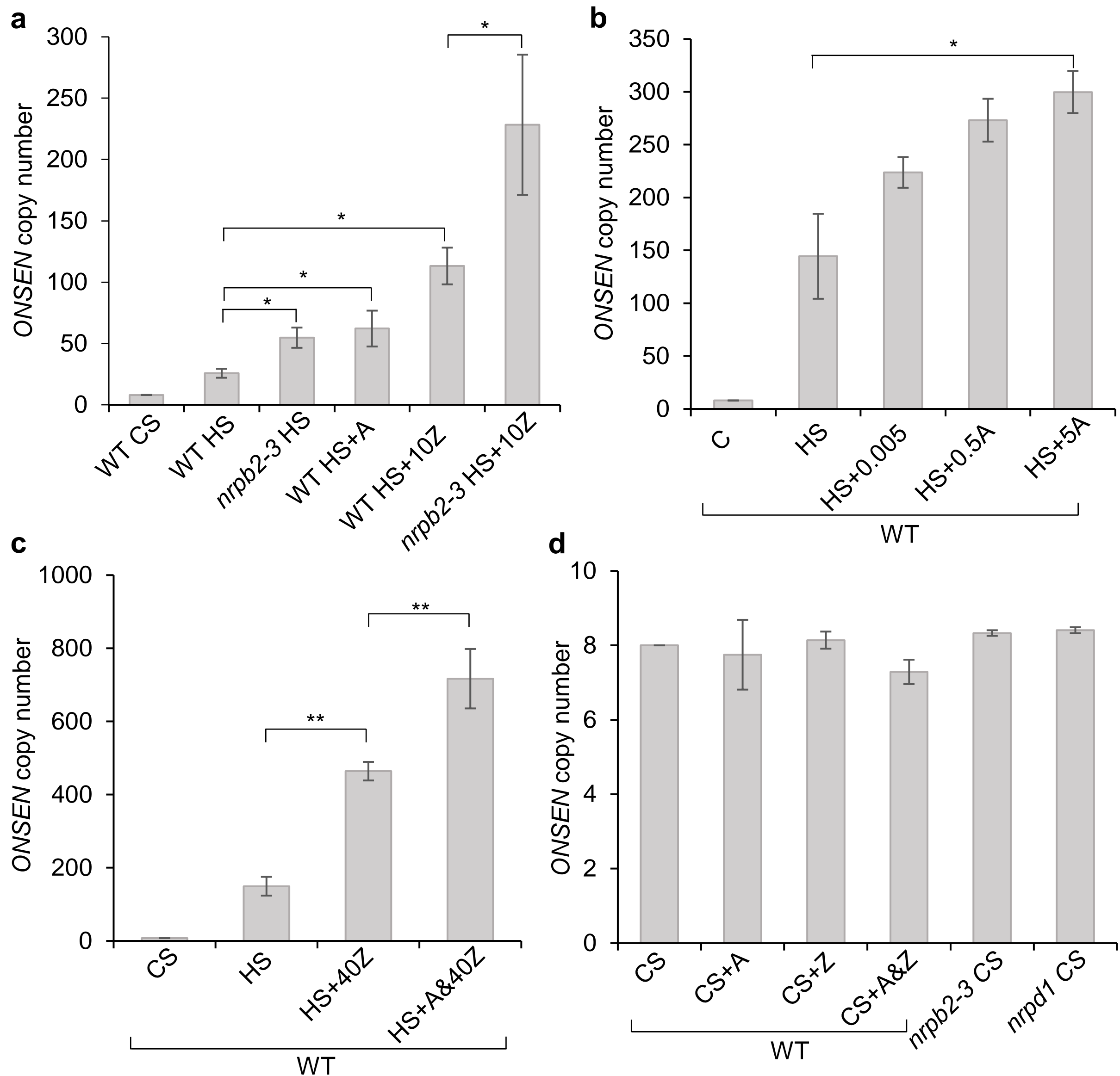


Figure 4 **Pol II represses the heat stress dependent mobility of the *ONSEN* retrotransposon in *Arabidopsis***. *ONSEN* copy number in *Arabidopsis* seedlings measured by qPCR directly after CS and HS-treatments. **a** In the WT and the *nrpb2-3* mutant and after HS plus treatments with α-amanitin (A, 5 µg/ml) or zebularine (Z, 10 µM) (mean ± s.e.m., n= 6 biological repetitions). **b** In the WT and after HS plus treatment with A at different concentrations (µg/ml) as specified in the figure labels (mean ± s.e.m., n= 4 biological repetitions). **c** In the WT and after HS plus treatment with Z (40 µM) or a combination of A (5 µg/ml) and Z (A&40Z) (mean ± s.e.m., n= 3 biological repetitions). **d** In the WT after chemical treatment with A (5 µg/ml), Z (40 µM) the combination of A and Z (A&Z) or in the *nrpb2-3* and *nrpd1* backgrounds following CS. (mean ± s.e.m., n= 3 biological repetitions). All values are relative to *ACTIN2*. * = P<0.05; **=P<0.01.

To ensure the viability of the *nrpb2-3* seedlings we choose a moderate amount of Z (10 µM). The presence of Z in the medium during HS generally enhanced the production of *ONSEN* copies. Importantly, this induced increase in *ONSEN* copy number was more distinct in the *nrpb2-3* background (Fig. 4a). This indicated that both, DNA methylation and Pol II transcriptional activity contribute to the repression of *ONSEN* ecDNA production. Because both DNA methylation and Pol II can be inhibited by the addition of specific drugs we wanted to test if treating WT plants with both A and Z at the same time could strongly activate and even mobilize *ONSEN* after a HS treatment. We grew WT seedlings on MS medium supplemented with Z (40 µM) (Baubec *et al.* 2009) individually or combined with A (5 µg/ml, A&Z). In conformity with the strong activation of *ONSEN* in HS and Z-treated *nrpb2-3*-seedlings, the combined treatment (A&Z) of the WT gave rise to a high (Fig. 4c) HS-dependent (Fig. 4d) increase of *ONSEN* copy number, comparable to *nrpd1* (Fig. 9).

We noted that the overall amplitude of HS-dependent *ONSEN* activation could vary between different waves of stress applications in terms of copy number (Fig. 4a-c, 5 and 9). Yet, the observed enhancing effect of an inhibition of Pol II and DNA-methyltransferases with A and Z on *ONSEN* activation was consistent in independent experiments (Fig. 4a-c, 5 and 9).

Additional data further indicate that the observed additive effect of the combined treatment of A and Z also occurs when the two inhibitors are used in different ratios such as 10 µM zebularine combined with 20 µg/ml α-amanitin (Fig. 5).

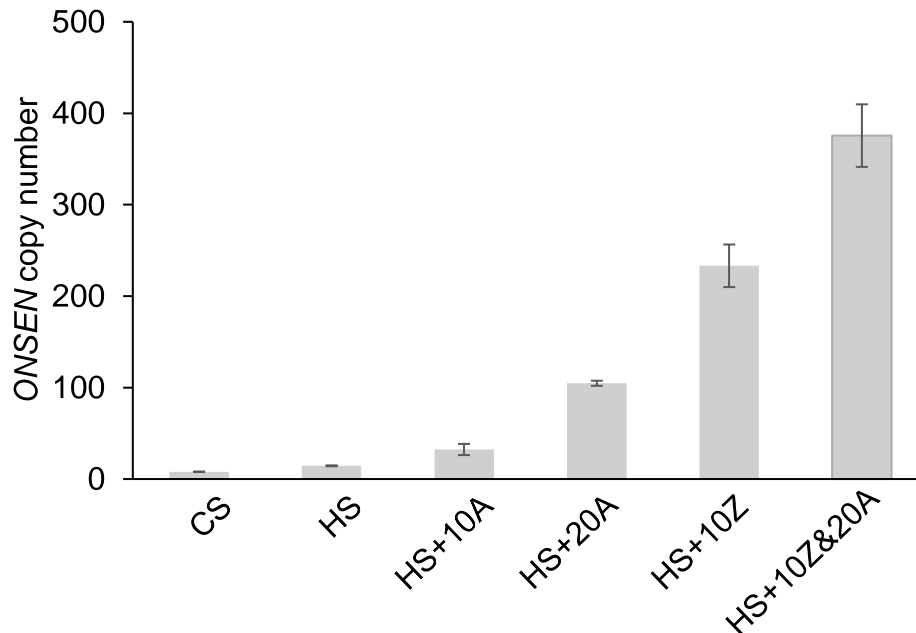


Figure 5 **Different concentrations of A&Z can be combined to induce the heat-dependent *ONSEN*-activation in the WT of *Arabidopsis*.** *ONSEN* copy number in seedlings measured by qPCR directly after CS and HS-treatments and after HS plus treatments with Z (10 µM), A (10-20 µg/ml) or a combination of A (20 µg/ml) and Z (mean ± s.e.m., n= 3 biological repetitions, all values are relative to *ACTIN2*).

To detect activated retroTEs at the genome-wide level we took advantage of the production of extrachromosomal circular DNA (eccDNA) by active retroTEs. EccDNA is a byproduct of the LTR-retroTE life cycle (Flavell and Ish-Horowicz 1981). Using mobilome sequencing, that comprises a specific amplification step of circular DNA followed by high-throughput sequencing to identify eccDNA derived from active LTR-retroTEs (Lanciano *et al.* 2017), we found that only *ONSEN* was activated by HS in combination with A&Z (Fig. 6). Confirming our qPCR data, more *ONSEN*-specific reads were detected in the presence of A and Z in the medium.

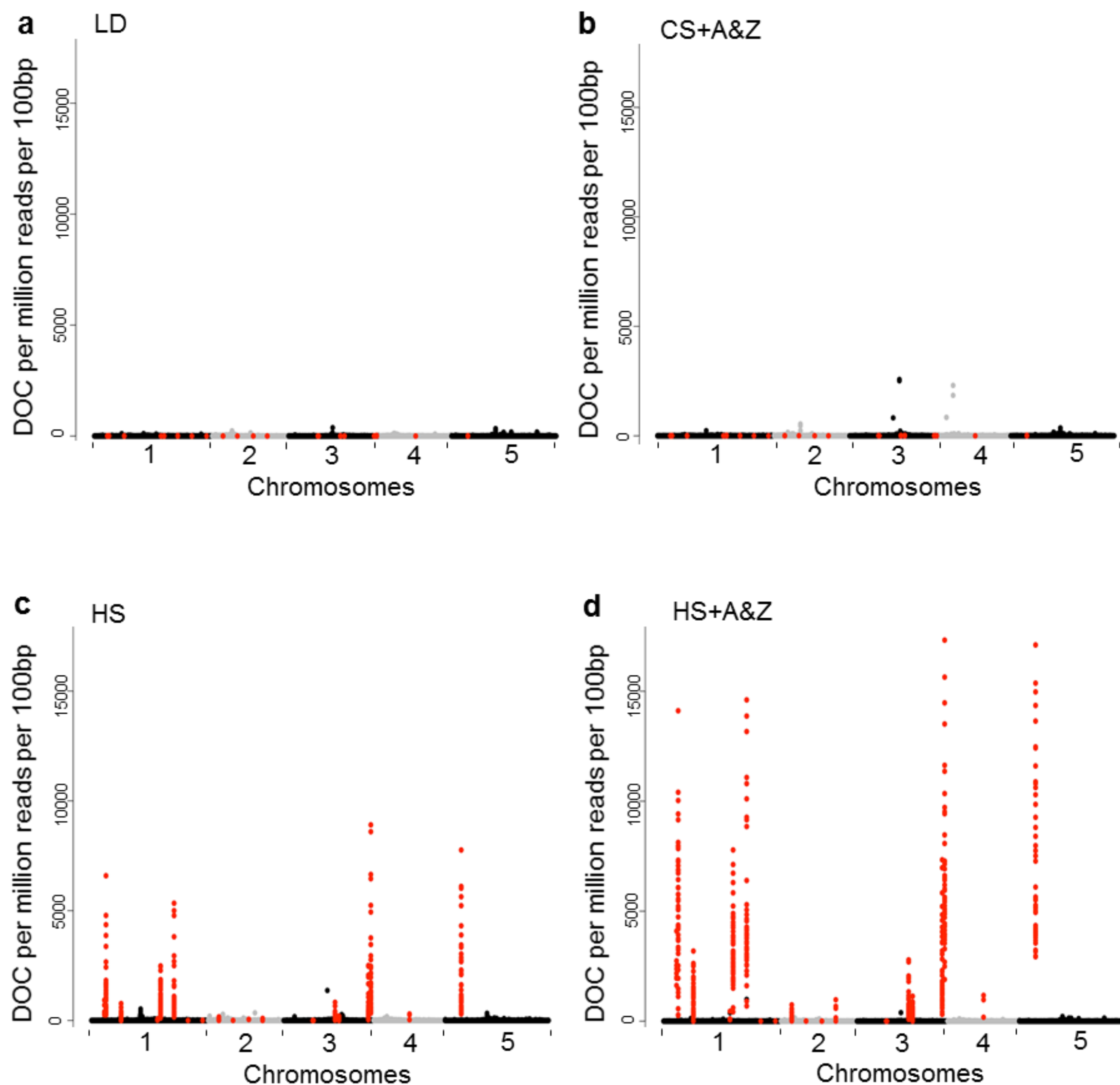


Figure 6 **Combined treatment of A&Z induces the heat-dependent detection of reads originating from *eccONSEN* in a mobilome of the *Arabidopsis* WT.** Abundance of reads from the mobilome-sequencing libraries of WT *Arabidopsis* plants mapping at TE-annotated loci from seedlings after: **a** growth under long day conditions (LD), **b** CS plus treatment with A (5 µg/ml) and Z (40 µM) (A&Z), **c** HS and **d** HS plus treatment with A&Z. Each dot represents the normalized coverage per million mapped reads per all TE-containing 100bp windows obtained after aligning the sequenced reads on the five chromosomes (black and grey circles). Red dots indicate the position of 100bp windows corresponding to *ONSEN* loci.

4.3.2 Inhibition of Pol II reduces global DNA-methylation

To better understand the mechanisms by which the α-amanitin and zebularine-treatments enhanced the activation of *ONSEN* after HS at the DNA level, we assessed how they influenced genome-wide DNA methylation using whole-genome bisulfite sequencing after CS. Overall, we found that all drug treatments affected global DNA methylation levels. While the treatment with Z affected all sequence contexts, we observed that inhibition of Pol II primarily affected cytosine methylation in the CHG and CHH sequence contexts (where H is an A, T or a G). The combined A&Z treatment had a slightly additive de-methylating effect in the CHG and CHH contexts compared to A or Z alone (Fig. 7a, b and d). DNA methylation levels at one *ONSEN* locus (AT1TE12295) is depicted in Fig. 7c. Treatment with A lead to a slight decrease in DNA methylation, which was more apparent in Z and A&Z treated plants. We further confirmed these results by a chopPCR at the long terminal repeat (LTR) of a selected *ONSEN* endogenous locus (*AT1TE12295*) (Fig. 7d). Treating plants with A or Z individually resulted in reduced DNA methylation levels in the CHH context at the *ONSEN* LTR after CS (Fig. 7d). The simultaneous inhibition of Pol II and DNMTases led to a loss of DNA methylation at a level comparable to the RdDM-deficient *nrpd1* mutant.

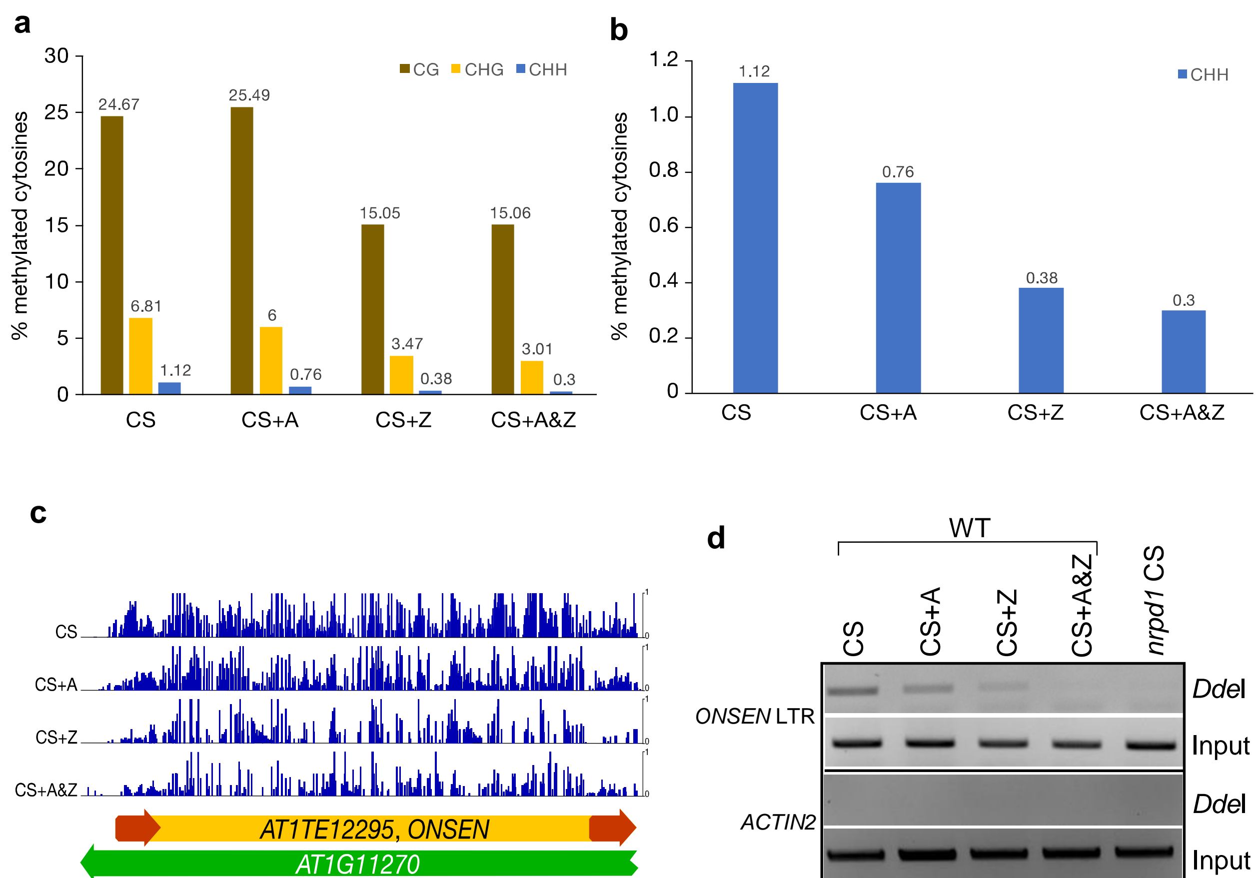


Figure 7 **Simultaneous inhibition of DNA methyltransferases and Pol II reduces global CHH methylation and mimics TE-silencing deficiency of *nrpd1*. a** Genome-wide DNA methylation levels in the WT after CS and CS plus treatments with A (5 µg/ml), Z (40 µM) or a combination of A and Z (A&Z) shown for the three sequence contexts (brown for CG, yellow for CHG and blue for CHH). **b** Same as (**a**) but only depicting the CHH context data for clarity. **c** Methylome-data of treated and untreated plants at an *ONSEN* locus located on Chromosome 1 (*ONSEN* is indicated in yellow, its LTRs in red). **d** DNA methylation analysis of the *ONSEN* LTR in untreated and A (5 µg/ml), Z (40 µM) or A&Z-treated seedlings of the WT and the *nrpd1* mutant after CS. Gels depict PCR products obtained from genomic DNA that was either undigested (input) or digested with the methylation sensitive restriction enzyme *DdeI* (reporting on CHH methylation here). *ACTIN2* is included as a control for complete *Dde*I digestion.

4.3.3 Pol II transcripts feed into DCL3-RdDM

We then wanted to test how the drug-induced loss of DNA methylation could affect *ONSEN* transcript accumulation after a heat shock (HS). For that purpose a northern blot was performed directly after HS in WT plants in the presence or absence of the aforementioned drugs. We found that treatment with Z alone resulted in the highest *ONSEN*-transcript level after HS (Fig. 8). Considering the data obtained on *ONSEN* ecDNA (Fig. 4c), we concluded that a substantial proportion of these Z-induced transcripts were not suitable templates for *ONSEN* ecDNA synthesis.

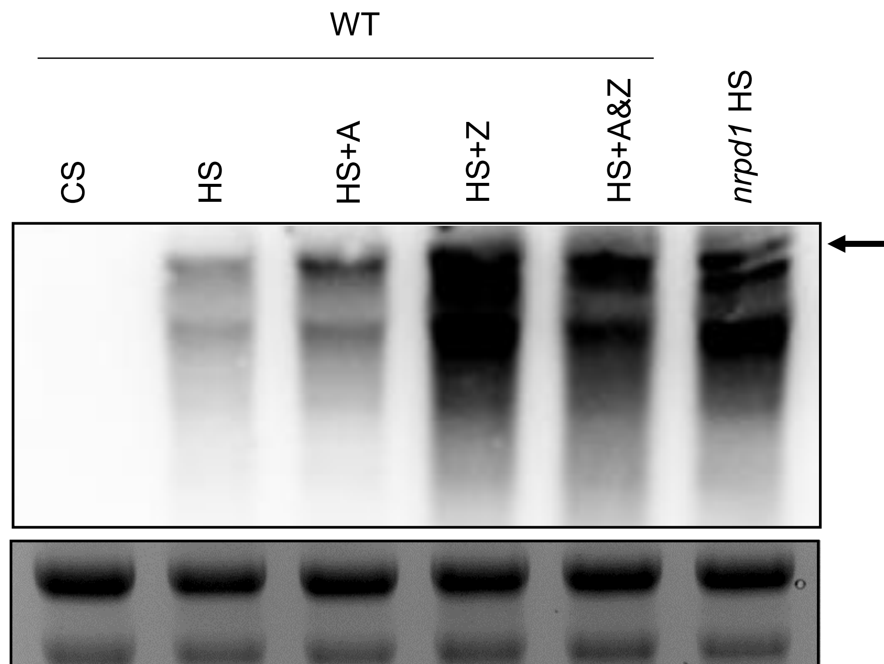


Figure 8 **Treatment with the inhibitors A and Z and HS specifically alters the heat-dependent *ONSEN*-transcript accumulation in the WT.** Northern blot detecting *ONSEN*-transcripts directly after CS, HS and HS plus treatment with A (5 µg/ml), Z (40 µM) or a combination of A and Z (A&Z) in the WT and after HS in *nrpd1*. Black arrow indicates the *ONSEN* full length transcript. Below, a Midori-stained agarose-gel is shown as a loading control.

In *Drosophila*, it has been shown that Pol II-mediated antisense transcription results in the production of TE-derived siRNAs in a Dicer-2 dependent manner (Russo *et al.* 2016). Supporting this notion for *Arabidopsis*, a recent publication pointed out the importance of DCL3 in regulating *ONSEN* in the *ddm1* background (Panda *et al.* 2016).

To elucidate whether the effect of the inhibition of Pol II was also dicer-dependent, we grew both an *rdr6* and a *dcl2/3/4*-triple mutant of *Arabidopsis* on A, applied HS and measured *ONSEN* ecDNA levels. Strikingly, we found that A was still enhancing ecDNA accumulation in *rdr6*, whereas inhibition of Pol II had no additional effect in the *dcl2/3/4* triple-mutant (Fig. 9).

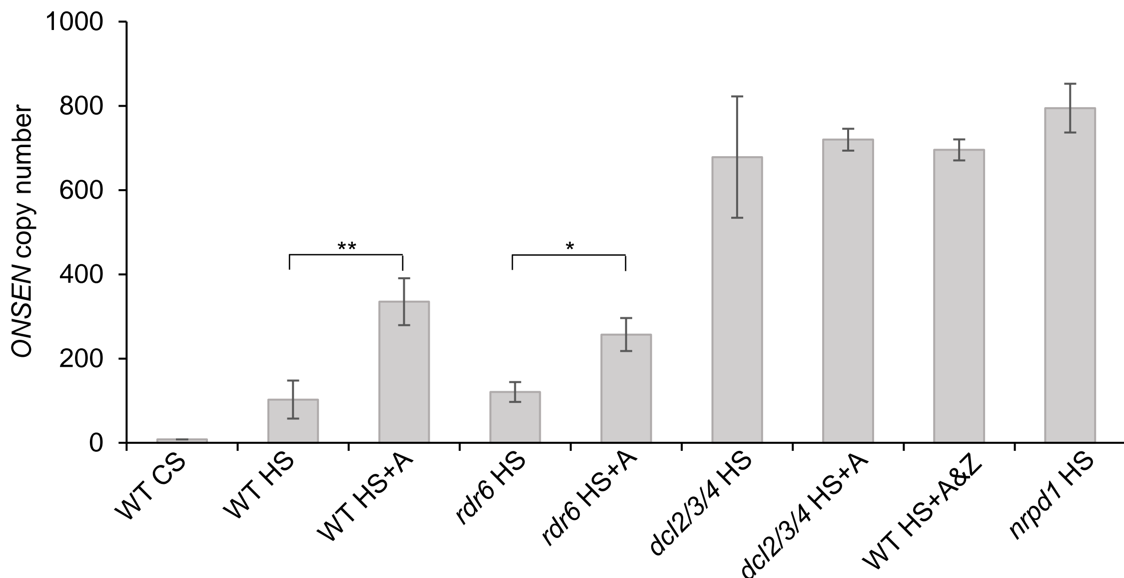


Figure 9 **Inhibition of Pol II leads to a dicer-dependent activation of *ONSEN* following HS**. *ONSEN* copy number measured by qPCR directly after CS and HS-treatments in seedlings of WT, *rdr6*, *dcl2/3/4* and *nrpd1* plants directly after CS, HS and HS plus treatment with A (5 µg/ml), Z (40 µM) or a combination of A&Z (*n*= 3 biological repetitions, values relative to *ACTIN2*; * = $P<0.05$; ** = $P<0.01$).

To further examine whether the inhibition of Pol II would impair RDR6-independent RdDM pathways, we analyzed the recently described DCL3-RdDM target *At*SINE4 (*At*3TE40740) (Panda *et al.* 2016) and found a strong decline of DNA-methylation in all sequence contexts after inhibition of Pol II. Strikingly, the combined treatment with A and Z resulted in an almost complete loss of DNA-methylation in the CHH-context at this locus (Fig. 10).

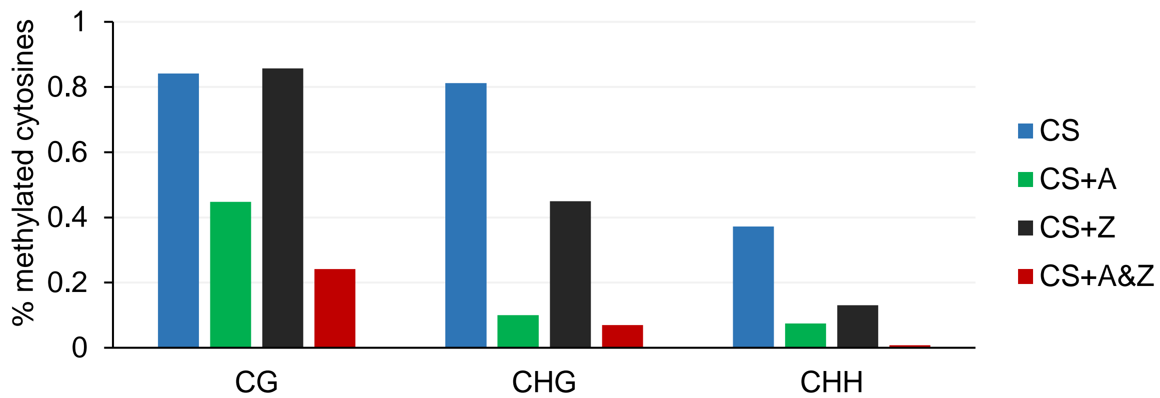


Figure 10 **Inhibition of Pol II reduces methylation at the DCL3-RdDM locus *At*SINE4.** Methylome-data of the WT after CS and CS plus treatment with A (5 µg/ml) and Z (40 µM) or a combination of A and Z (A&Z) at the DCL3-RdDM target *At*SINE4 (*At*3TE40740). The percentage of methylated cytosines in the CG, CHG and CHH sequence contexts (where H is an A,T or a G) are depicted. See legend for color code.

Using strand-specific northern probes, we wanted to test whether a switch in *ONSEN*-antisense transcription following the inhibition of Pol II could be responsible for the observed general reduction in *ONSEN*-repression. Using three different probes specific for the HRE, the GAG and the RVT (Fig. 11, *preliminary data*) we were able to detect *ONSEN* antisense transcripts both in the WT and the mutant backgrounds. Interestingly, we already detected antisense-transcripts after CS in the WT. Most importantly, we found that especially the double treatment (HS+A&Z) significantly changed the observed size-patterns, resulting in an accumulation of shorter transcripts in all three blots. This effect was most obvious for the RVT-probe where also a slight reduction in the longest band compared to Col HS was observed.

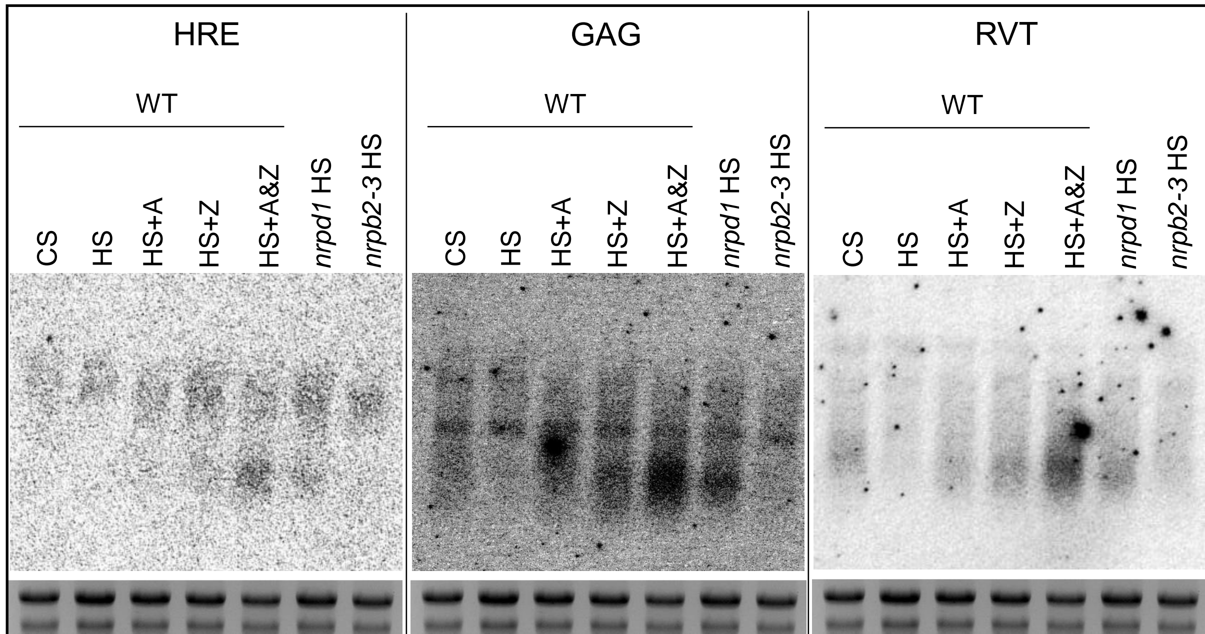


Figure 11 **Inhibition of Pol II and DNMTases alters the accumulation of *ONSEN* antisense transcripts.** Detection of *ONSEN* antisense transcripts in *Arabidopsis* seedlings directly after CS and HS-treatments in WT, *nrpb2-3* and *nrpd1* plants and in the WT after HS plus treatments with α-amanitin (A, 5 µg/ml) or zebularine (Z, 40 µM) or the combination of both inhibitors (A&Z). Region-specific probes for *ONSEN* were used for hybridization. From the left: HRE, GAG and RVT. The membrane was stripped twice prior to hybridizations with new probes. The loading control is shown under the blot (three times the same image). As the size marker is not visible, only qualitative comparisons can be drawn. *Preliminary data.*

4.4 Discussion

In this study, we show the importance of Pol II in the repression of TE mobility in plants. By choosing the well-characterized heat inducible *ONSEN*-retroTE, we were able to specifically address the role of Pol II in silencing transcriptionally active endogenous retroTEs in wild-type plants. Recent studies propose Pol II as the primary source for the production of retroTE-silencing signals that can then feed into the RNA silencing and DNA methylation pathways (Cuerda-Gil and Slotkin 2016). Our data strongly support these findings at two levels: First, we found that inhibition of Pol II activity reduced the degree of DNA methylation at *ONSEN* demonstrating its distinct role in this process and that Pol II also contributes to reinforcing silencing at the genome-wide level primarily in the CHH but also in the CHG contexts. Second, our finding that DCL-enzymes are sufficient to process the silencing-signal produced by Pol II suggest that Pol II acts at very early steps in the retroTE silencing pathway by providing substrates to these enzymes. The observation that inhibition of Pol II in the *rdr6* background still further enhanced *ONSEN* accumulation after HS supports the notion that Pol II plays a central role in the previously proposed expression-dependent RdDM pathway (Panda *et al.* 2016).

The strong de-methylation of the DCL3-RdDM-traget *At*SINE4 (*AT3TE40740*) after the A-treatment supports the model in which Pol II-derived transcripts are processed in a DCL3-dependent manner. The fact that the degree of de-methylation following the A-treatment at this specific locus was even stronger than after a general inhibition of DNMtases with zebularine underlines the importance of Pol II as an epigenetic regulator of retroTEs. Furthermore this is a clear indicator for the locus-specificity of different variants of silencing pathways that regulate retroTE-activity in plants. Our findings may indicate that Pol II is primarily involved in silencing young, recently active retroTEs and perhaps to a lesser extent other tightly silenced TEs. Indeed, there are indications for very recent transposition events for *ONSEN* in natural populations of *Arabidopsis* (Quadrana *et al.* 2016).

Key to the understanding of the mechanism behind the epigenetic regulation of transcriptionally active TEs may not only lie in the quantity but rather the quality of Pol II derived transcripts. Yet the fact that Pol II- derived transcripts are partially translated or used as a template for reverse transcription opens up an additional level for potential regulation of retroTE activity. Indeed, it was found that the *BARE1*-retrotransposon in

barley is transcribed into three different types of Pol II-derived transcripts with distinct functions needed during *BARE1* retrotransposition (Chang *et al.* 2013). To account for such differences in transcript qualities, we favored to perform a northern blot over using RT-PCRs to measure *ONSEN*-expression. The observed diversity of size classes on the northern blot clearly indicated heat-stress dependent production of subgenomic *ONSEN*- transcripts. The induction of unfavorable relative amounts of different types of transcript pools could be an explanation for the observed discrepancy of total *ONSEN* transcript accumulation and measured ecDNA in Z-treated seedlings. This substantiates the notion that both quantity and quality of retroTE-transcripts affect regulation, reverse transcription and successful integration of class I elements in plants.

In *S. cerevisiae and D. melanogaster* Pol II-mediated intra-element antisense transcription is known to regulate TE-activity (Berretta *et al.* 2008; Russo *et al.* 2016). Besides a potential drug-induced shift in the ratio of different classes of subgenomic sense-transcripts, we hypnotize that altered antisense-transcription of *ONSEN* could account for its observed drug-induced activation following HS. In a preliminary experiment, we were indeed able to detect various *ONSEN* antisense-transcripts even in the absence of HS. This finding suggests that Pol II-mediated transcription originating from bidirectional TSSs in the LTRs of plant retroTEs plays a key-role in initiating and maintaining retroTE silencing. Although we observed clear shifts in the patterns of *ONSEN*-antisense transcripts following the drug-treatments, the consequences observed in this preliminary experiment remain elusive. Still, the fact that sense-antisense pairing of Pol II-transcripts, can result in the RDR-independent initial formation of TE dsRNAs substantiated the importance of these preliminary data. Once processed by DCL-enzymes, these primary dsRNAs can be amplified under the involvement of RDR6 in RNAi-loops ultimately resulting in PTGS and robust epigenetic silencing (Cuerda-Gil and Slotkin 2016). Due to such RDR-dependent amplifications of siRNAs, already small alterations of the initial dsRNA production caused by minor drug-induced shifts of antisense-transcripts can have a major impact on retroTE-silencing (Lisch and Bennetzen 2011). The fact that both DNA-methylation and the observed changes in *ONSEN*-antisense-transcript levels already occurred in the absence of HS, could furthermore indicate that Pol II actively surveils potentially mobile genetic elements already prior to their stress-induced activation.

Therefore, as a next step it will be of great interest to systematically test if Pol II-dependent (subgenomic) antisense transcription of TEs and subsequent dicer-dependent processing prior to or during stress may be the key to solve “the chicken and the egg problem” of *de novo*-silencing functional retroTEs in eukaryotes (see general discussion).

Overall, these findings lead to the question as to when plants do lower their guards: under which condition could Pol II be less effective in silencing TEs? Certain stresses that affect the cell cycle have been reported to lead to the inactivation of Pol II (Oelgeschläger 2002; Palancade and Bensaude 2003). That would provide a window of opportunity for TEs to be mobilized. Therefore, combined stresses that affect the cell cycle and activate TEs may lead to actual TE bursts under natural growth conditions. Interestingly, it has been reported that retroTE-derived short interspersed element (SINE) transcripts can inhibit Pol II activity (Pai *et al.* 2014). This strongly suggests the presence of an ongoing arms-race between retroTEs and Pol II. Considering that almost all organisms analyzed so far have TEs (Huang *et al.* 2012) and RNA polymerases (Lazcano *et al.* 1988) and the reliance of TEs on host RNA polymerases, it may - from an evolutionary point of view - not come as a surprise that Pol II also has a function as an important regulator of retroTE activity.

Without challenging the importance of Pol IV and V as key regulators of retroTEs including *ONSEN* (Ito *et al.* 2011) our data indicate that the role of Pol II at the root of TE-silencing has for a long time been underestimated. The approach of specifically inhibiting Pol II with α-amanitin circumvents the adverse effects that arise by mutating this vital enzyme and opens up new possibilities to elucidate processes at the origin of TE-silencing in eukaryotes. As a next step this will for the first time allow to assess the phenotypic consequences caused by potential heritable (epi)genetic changes triggered by induced retroTE-busts in wild-type plants.

4.5 Material and Methods

Experiments were conducted as described in (Thieme *et al.* 2017).

Additional experiments were performed as follows:

4.5.1 ChopPCR for CHH-methylation at the *ONSEN*-LTR

20 ng of total genomic DNA was isolated from the aerial part of at least ten *Arabidopsis* seedlings and digested with the methylation sensitive restriction enzyme *DdeI* (NEB) at 37 °C over night. Following heat inactivation at 60°C for 20 min, the digested DNA was used as a template for the chopPCR (primers listed in table 1). *ACTIN2* served as a control for the digest. Undigested DNA was used as a loading control. PCR products were separated on a 1% agarose gel stained with Midori Green.

Table 1 **Primer names and sequence used for chopPCRs.**

Name	Sequence 5`->3`
286 OnsenBis F1	GGTTGAAGGGTYAAAGAGTAAAT
287 OnsenBis R1	CCTCCAAACTACAAAATATCTAAAA
835 Chop PCR ACT2 F	TGTAGTGTCGTACGTTGAACAGAAAGC
836 Chop PCR ACT2 R	TTGGCACAGTGTGAGACACACCA

4.5.2 Detection of *ONSEN*-derived antisense transcripts

Total RNA from the aerial part of at least ten *Arabidopsis* seedlings was isolated using the TRI Reagent (Sigma) according to the manufacturer’s recommendations. RNA concentration was measured (Qubit RNA HS Assay Kit, Thermo Fisher) and 12 µg of RNA separated on a denaturing (formaldehyde) 1.5% agarose gel, blotted on a Hybond-N+ (GE Healthcare) membrane and hybridized with radioactive [γP^{32}]-ATP labelled (T4 Polynucleotide Kinase, NEB) and purified (illustra, MicroSpin G-25 Columns, GE Healthcare) probes (table 2) over night at 42°C. Prior to re-hybridization membranes were stripped by adding boiling solutions of 0.5xSSC+0.5%SDS and 0.1xSSC+0.5% SDS.

Table 2 **Probes used for the detection of *ONSEN* antisense transcripts.**

Name	Sequence 5`->3`
ONSEN 5`HRE probe for antisense	AAGTTCTAGAGTTTTCTCTAGAAATATCATCAT TTCCACCTCCTTAAAAG
ONSEN GAG RT fw	GTCGTTGAAGCTACGTCGGCG
ONSEN REV RT fw	GGCAATTGGCGTGAAGTGGGT

5 Evolutionary consequences of an induced burst of *ONSEN* in *Arabidopsis*

A modified version of this chapter was published in Thieme, M. *et al.*, (2017). Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. Genome Biol 18, 134.

5.1 Abstract

Retrotransposons play a central role in plant evolution and could be a powerful endogenous resource to induce genetic and epigenetic variability for crop breeding. In the previous chapter I have shown that inhibition of RNA-polymerase II and DNA-methyltransferases leads to a very strong HS-dependent activation of the *ONSEN*-retroTE in *Arabidopsis* (chapter 4). Here, I demonstrate that the combined inhibition of both DNA methylation and restriction of Pol II activity lead to a strong stress-dependent mobilization of the heat responsive *ONSEN* retroTE in *Arabidopsis* seedlings. The progenies of treated WT-plants contained up to 75 new *ONSEN* insertions in their genome. These novel *ONSEN* insertions are stably inherited over three generations of selfing. Repeated application of HS in progeny plants containing increased numbers of *ONSEN* copies did not result in an increased activation of this transposon compared to control lines. Supporting the role of retroTEs as important drivers of plant evolution, progenies with additional *ONSEN* copies showed a broad panel of environment-dependent phenotypic diversity. This suggests that Pol II can regulate the speed of plant evolution by fine-tuning the amplitude of transposon mobility in plants. These findings represent a proof of concept for futures application of inhibitors to induce controlled retrotransposition for plant breeding.

5.2 Introduction

In plants, TEs are increasingly seen as a source for genetic and epigenetic variability and thus important drivers of evolution (Belyayev 2014; Huang *et al.* 2012; Lisch 2013; Paszkowski 2015) (see general introduction). Induced mobilization of endogenous TEs in plants has so far been very inefficient, thus limiting their use in basic research and plant breeding (Paszkowski 2015).

Although the activity of retroTEs depends on transcription by the host RNA polymerase II, we have previously shown that a functional loss of Pol II leads to an increased accumulation of *ONSEN* ecDNA following HS (chapter 4). The additional treatment with zebularine, a general inhibitor of DNMTases, resulted in a very strong *ONSEN* ecDNA-accumulation in the WT. In fact, the level of ecDNA in treated WT-seedlings was comparable to that in the *nrpd1*-mutant. It was previously shown that the strong *ONSEN*-accumulation in heat-stressed *nrpd1*-plants can result in the acquisition of novel *ONSEN* insertions that can be tracked in the following generations (Ito *et al.* 2011). Although the endogenous *ONSEN* copy number is known to vary between different accessions (Quadrana *et al.* 2016) its mobility in WT plants has so far not been observed in real time.

Here, we wanted to investigate whether HS and simultaneous inhibition of TE-defense with A and Z would lead to a mobilization of *ONSEN* resulting in novel insertions of this retroTE in *Arabidopsis*. As TE-bursts lead to a strong increase in (epi)genetic diversity, we also addressed the question whether we could detect phenotypes emerging in plants containing an induced increase of *ONSEN* copy numbers.

5.3 Results

5.3.1 Pol II-deficiency and inhibition of DNA-methylation leads to retrotransposition of *ONSEN*

To complete their lifecycle, the reverse transcribed ecDNA of activated retroTEs has to integrate back into the genome (Schulman 2013). Given that we observed a strong increase in *ONSEN* copy number after HS and treatment with moderate amounts of Z in the *nrpb2-3* background (chapter 4), we wanted to address the inheritance of additional *ONSEN* copies to the offspring. For this, we collected all seeds of individual heat stressed and Z-treated WT and *nrpb2-3* plants and grew several S1 (selfed 1st generation) -descendants per parental plant under controlled conditions on soil. We then extracted and pooled DNA from n>7 S1-individuals from the same parent and compared the average *ONSEN* copy number by qPCR. We observed a distinct increase of the overall *ONSEN* copy number exclusively in the *nrpb2-3* background (Fig. 12).

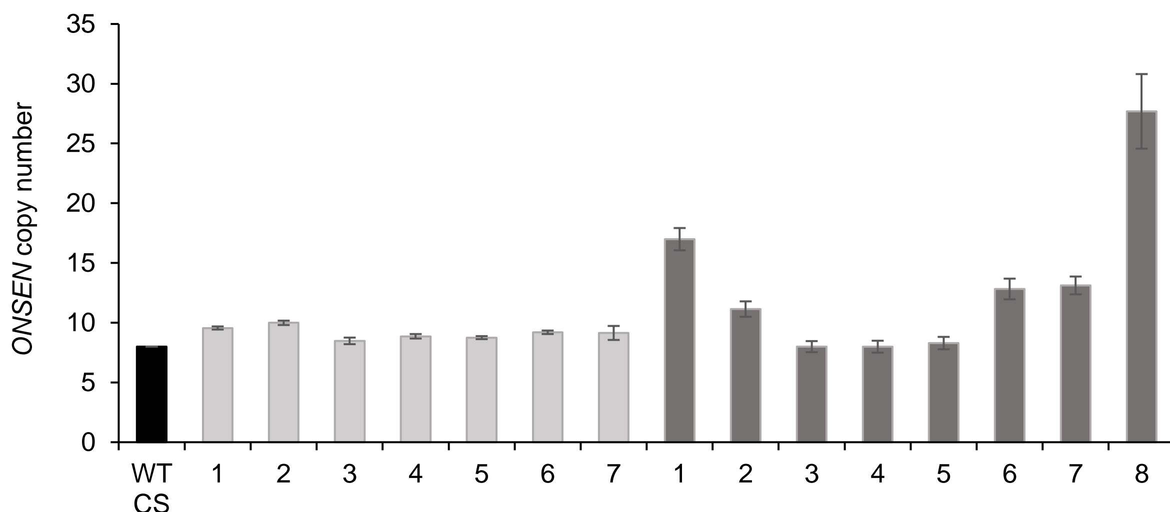


Figure 12 **Detection of increased *ONSEN* copy numbers in S1 pools of heat-stressed and Z-treated *nrpb2-3* plants**. *ONSEN* copy number measured by qPCR in pooled seedlings of the S1-generation of heat stressed and zebularine-treated (10 µM) WT (light grey bars) and *nrpb2-3* plants (dark grey bars) that were grown under control conditions on soil relative to a control stressed WT-plant (black bar) (mean ± s.e.m., *n*=3 technical repetitions, all values relative to *ACTIN2).*

Due to this observation and the fact that the combined inhibition of DNMtases and Pol II resulted in a very high accumulation of *ONSEN* copy numbers after HS - essentially mimicking plants defective in NRPD1 (chapter 4) – we wanted to test whether the

induced activation of *ONSEN* would also lead to novel insertions in the WT. In a similar approach (Fig. 13) like for the Z-treated and heat stressed *nrpb2-3* plants, we first screened by qPCR if, and at which frequencies, new *ONSEN* copies could be detected in the progeny of A&Z-treated and heat stressed plants (Fig. 14).

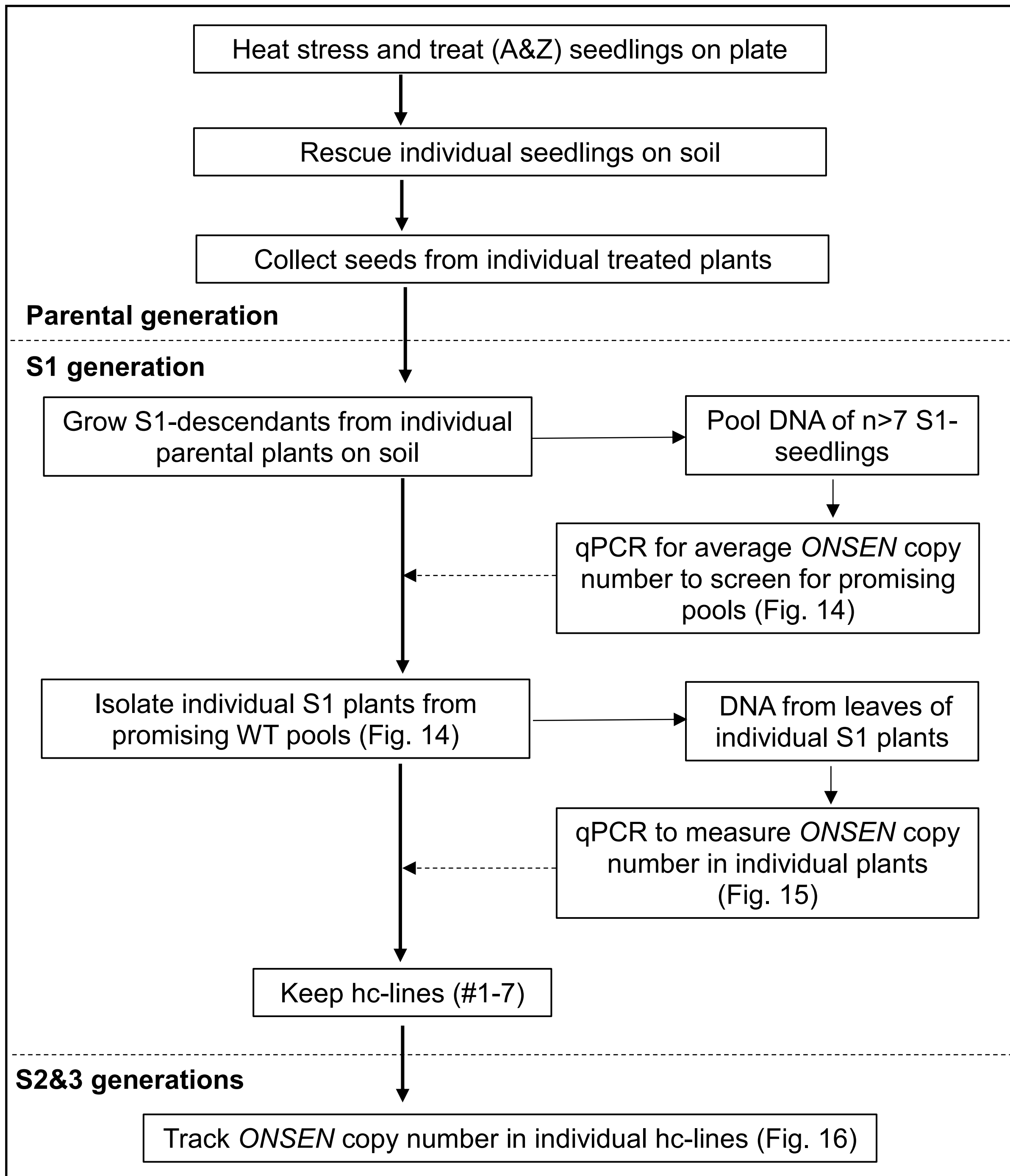


Figure 13 **Schematic representation of the experimental setup used to select for *ONSEN* hc-lines following the HS and A&Z-treatment in the parental-generation.** Dataobtained by qPCRs on pooled or individual progeny plants are depicted in Fig. 14-16.

In fact, we found new *ONSEN* insertions in 29.4% of the tested (n=51) S1 pools (DNA extracted from n>7 S1-individuals descending from the same parent) with measured *ONSEN* copy-numbers reaching up to 52 insertions (Fig. 14).

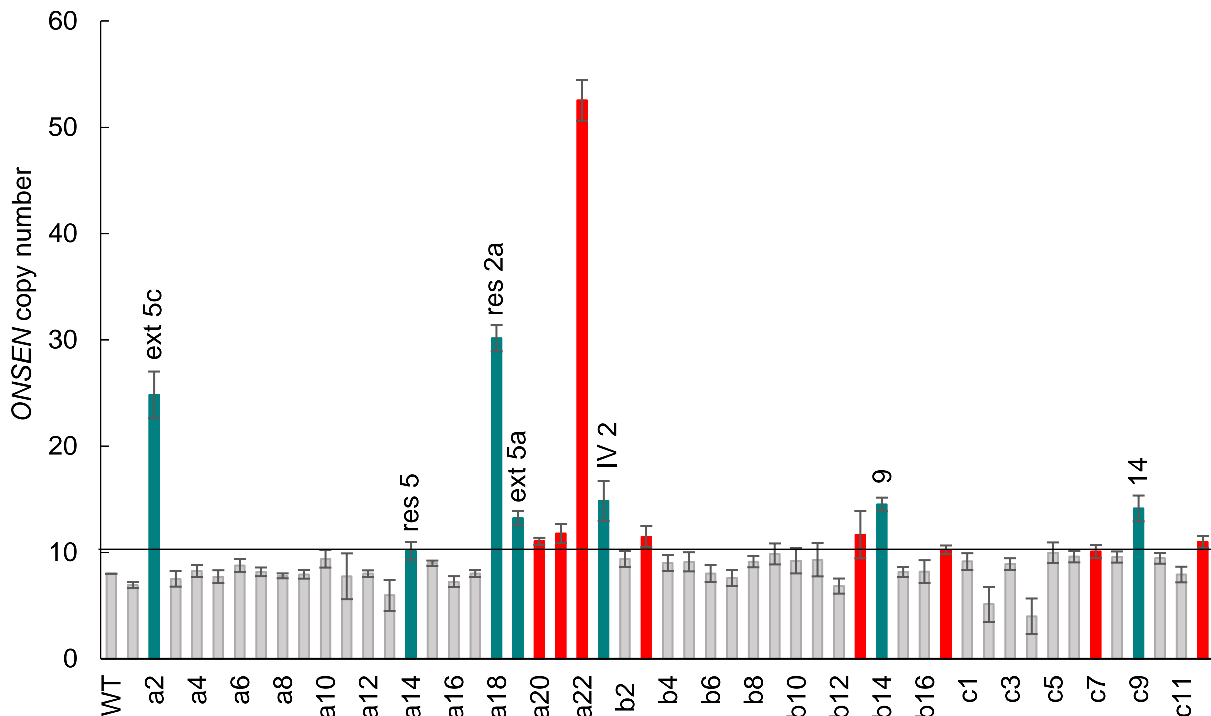


Figure 14 **Detection of increased *ONSEN* copy numbers in S1 pools of heat-stressed and A&Z-treated WT plants.** Parental plants were heat stressed and treated in independent experiments (characters a-c) with a combination of A (5 µg/ml) and Z (40 µM). Pooled (n>7) S1-descendants of individual treated plants were analyzed here. Pools with clearly increased *ONSEN*-copy numbers (>10) are marked in red and turquoise. Pools marked in turquoise were selected for an analysis on an individual basis. Names of selected pools are specified above the respective bars (see Fig. 15a-g). *ONSEN*-copy number measured by qPCR (mean ± s.e.m., n=3 technical repetitions, values relative to *ACTIN2*).

To confirm new *ONSEN* insertions in the progeny of treated plants, we separated seedlings (n= 4- 7, named with individual numbers or letters as supplements the actual names of the pool) from a subset of positive S1-pools depicted in Fig. 14 grew them under controlled long day conditions and used qPCRs to measure total *ONSEN*-copy numbers on an individual basis (Fig. 15). The analysis of individual lines from selected positive S1-pools with an average *ONSEN*-copy number >10 clearly indicated *ONSEN*-transposition following HS and A&Z-treatment in the parental plant. The

individual *ONSEN*-copy number differed between S1-individuals originating from the same parent and ranged up to 83 copies (Fig. 15f). In the majority of cases, we observed that siblings that were selected from the same pool resembled each other in total *ONSEN* copy numbers (Fig. 15a-d and f) showing either a similar gain in copies or the original number of eight endogenous copies initially present in the Columbia accession of *Arabidopsis*. Based on this observation we choose to keep one *ONSEN* high-copy line (hc-line) per pool and confirmed the presence of novel *ONSEN* copies in this subset of independent individual high copy plants in the S2 generation by qPCR (Fig. 16a).

To ascertain the integration of novel *ONSEN* copies into the genome, we ligated previously digested DNA that was extracted from hc-lines and three control lines to adaptors and amplified the ligation product in a PCR using primers specific to *ONSEN* and to the adaptor (transposon display). The separation of the obtained PCR-products clearly showed the presence of additional bands compared to the control lines indicating the presence of novel inserted *ONSEN* copies in the genome of all seven selected hc-lines (Fig. 16b). Tracking *ONSEN* copy numbers in these hc-lines over three generations of selfing and growth under controlled long-day conditions indicated that the new insertions were stably inherited (Fig. 16a). Notably, one of the selected hc-lines (#6) produced only very few seeds and became extinct after the S2 generation (Fig. 16a).

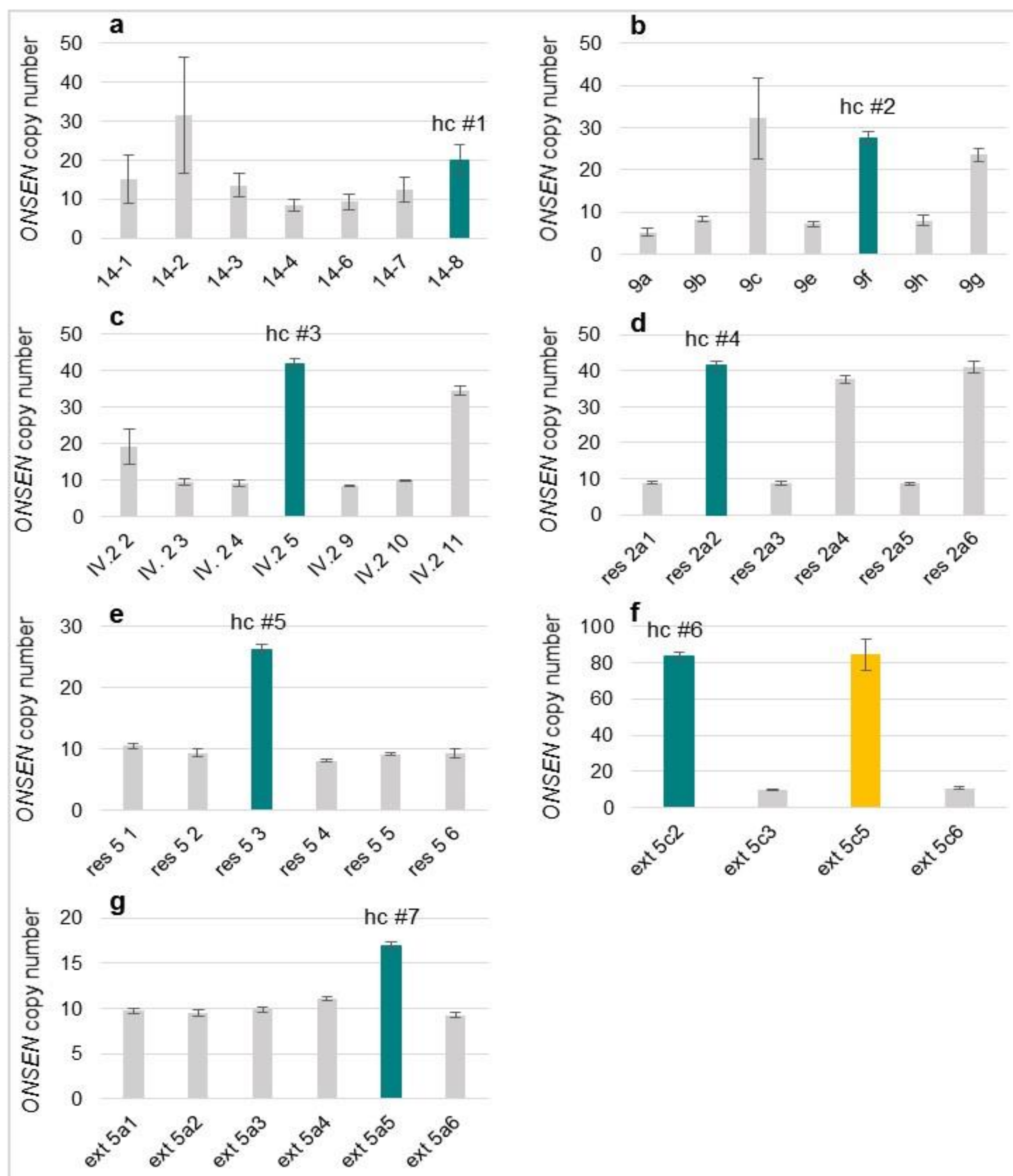


Figure 15 **Detection of distinct patterns of *ONSEN* copy numbers among S1-siblings originating from a heat stressed and A&Z-treated WT plant.** *ONSEN* copy numbers in individual plants (named with the pool name and an individual name supplement) isolated from selected S1-pools of heat-stressed and A (5 µg/ml) and Z (40 µM)-treated WT plants (Fig. 14) are depicted. Plants that were kept for analysis of successive generations (S2 and 3) are highlighted in turquoise. Selected *ONSEN* high-copy lines (hc-lines) were renamed with numbers (#1-7) above the turquoise bars in (**a**-**g**). The progeny-pool (S2 generation) of line ext5c5 (orange bar) showed segregation for an albino phenotype (Fig. 20). *ONSEN*-copy number measured by qPCR (mean ± s.e.m., n=3 technical repetitions, values relative to *ACTIN2*).

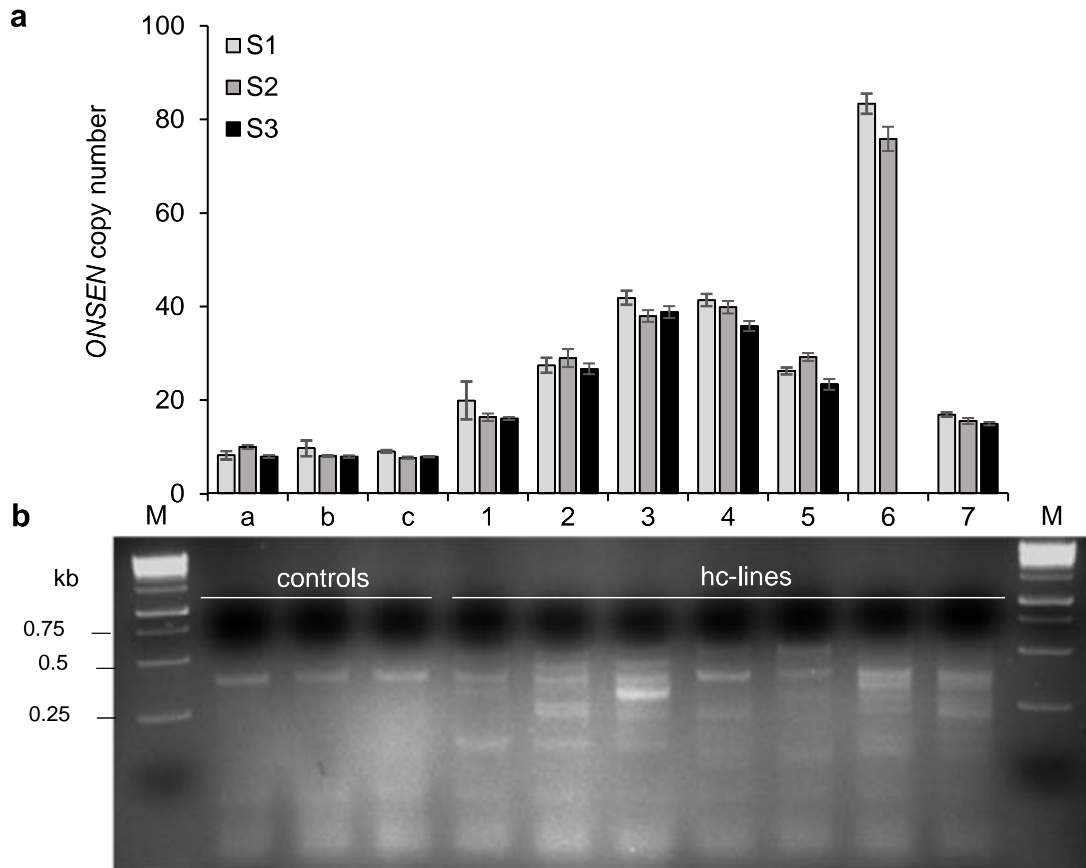
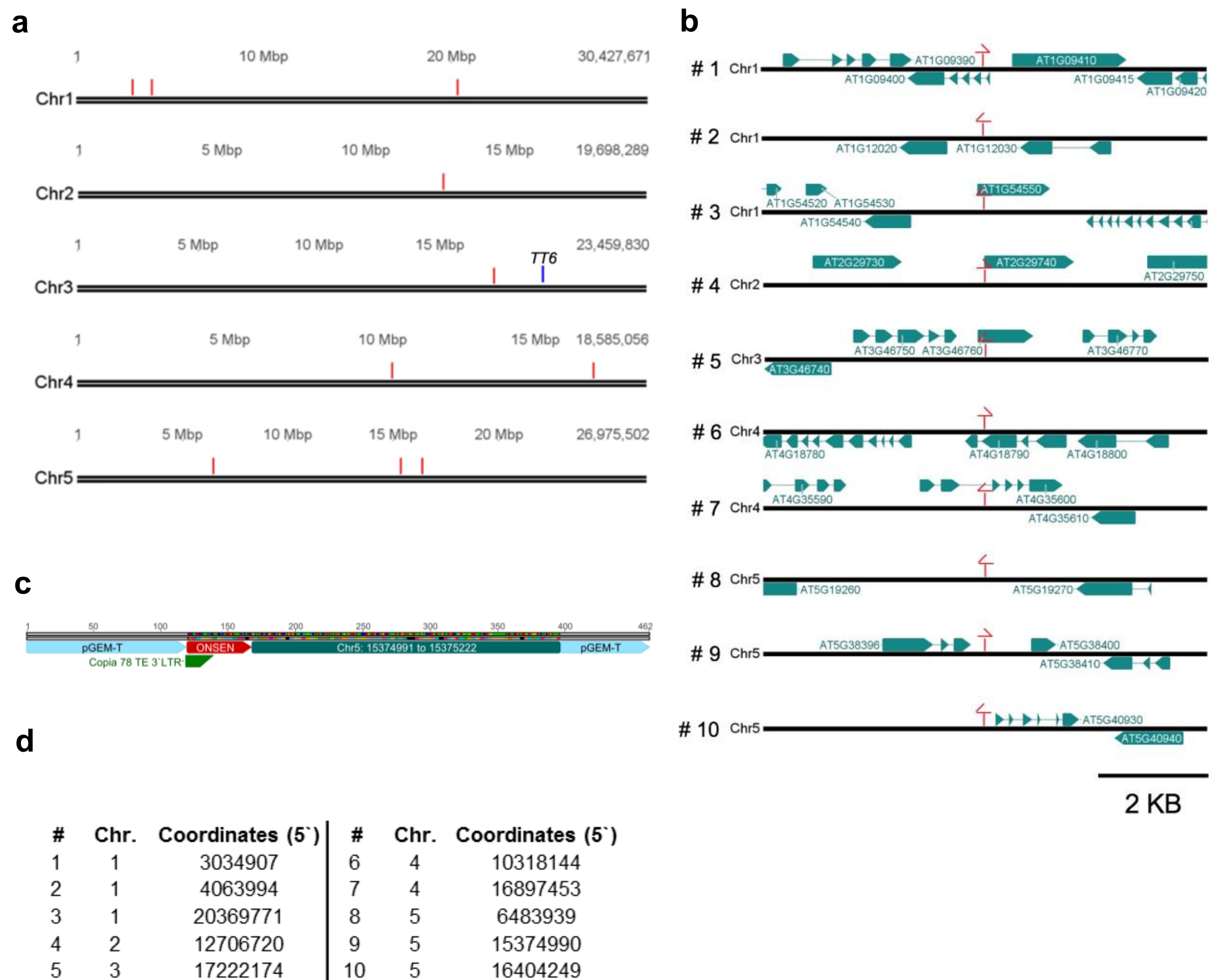


Figure 16 **Drug-induced mobilization of *ONSEN* in wild-type *Arabidopsis* plants**. **a** *ONSEN* copy number in the S1, S2 and S3 generation of HS (a-c) and HS with A (5 µg/ml) and Z (40 µM) treated WT plants (hc-lines 1-7) measured by qPCR (n= 3 technical replicates, values relative to *ACTIN2*). qPCR-Data for the S3-generation of line 6 in (**a**) are missing due to severe infertility and extinction of this line. **b** Transposon display testing seedlings in the S2 generation for novel *ONSEN* insertions. A GeneRuler 1 kb DNA Ladder (Thermo Scientific) was used a size marker (M).

By cloning and sequencing of PCR products obtained from the transposon display (Fig. 16b) followed by genotyping we were able to track down 11 novel *ONSEN* insertions in hc-line 3. We found *ONSEN* to be inserted in exons, introns and between genes on all five chromosomes of *Arabidopsis* (Fig. 17 and 19).



#	Chr.	Coordinates (5`)	#	Chr.	Coordinates (5`)
1	1	3034907	6	4	10318144
2	1	4063994	7	4	16897453
3	1	20369771	8	5	6483939
4	2	12706720	9	5	15374990
5	3	17222174	10	5	16404249

Figure 17 **Summary of confirmed novel *ONSEN* insertions in hc-line 3. a** Overview of insertion sites shown in (**b**) (red bar) and the location of the *ONSEN* insertion in the *TT6*-gene depicted in close-up in Fig. 19 (blue bar) **b** Close-up of regions with new *ONSEN* insertions (red bar) in the S2 generation of a selected heat stressed and A (5 µg/ml) and Z (40 µM) treated WT plant (hc-line 3). Orientation of novel *ONSEN* insertions is indicated with red arrows. **c** A scheme to exemplify the annotation of sequences that lead to the identification of novel *ONSEN* insertion sites depicted in (**b**) shown for insertion # 9. Colors correspond to the pGEM-T vector (light blue) used for cloning, the *ONSEN*- 3`LTR (red), the Copia 78 TE 3`LTR primer (dark green) that was used for the preceding TE-Display PCR and the genomic region (turquoise) flanking the 3`LTR of the new *ONSEN* insertion. **d** Summary of coordinates (base 5` of insertion) of new *ONSEN* insertions shown in (**a**) and (**b**). Numbering corresponds to (**b**). Sequences of primers used to confirm new *ONSEN* insertions are given with the numbering corresponding to (**b**) in table S1 in (Thieme *et al.* 2017).

5.3.2 *ONSEN* high-copy lines show phenotypic diversity

TE insertions can interrupt genes or alter their expression by recruiting epigenetic marks or by stress-dependent readout transcription from the 3'LTR into flanking regions (Lisch 2013). To test this, we grew the S2-plants of the selected hc-lines under long and short-day conditions. Interestingly, we observed that many hc-lines showed clear and homogenous phenotypes in response to the different growth conditions (plant size, chlorophyll content and flowering time (Fig. 18a and b)).

To demonstrate that *ONSEN* insertions could directly influence such developmental phenotypes, we closely investigated hc-line 3 that produced white seeds (Fig. 19a). Using a candidate gene approach, we found that an *ONSEN* insertion in *transparent testa 6* (*TT6*, *AT3G51240*, Flavanone 3-hydroxylase) (Fig. 19b) was responsible for the recessive white seed phenotype (Appelhagen *et al.* 2014; Rosso *et al.* 2003). This was confirmed by the segregation analysis of the F2 generation of a cross between the WT and hc-line 3 (Fig. 19a) followed by genotyping (Fig. 19c).

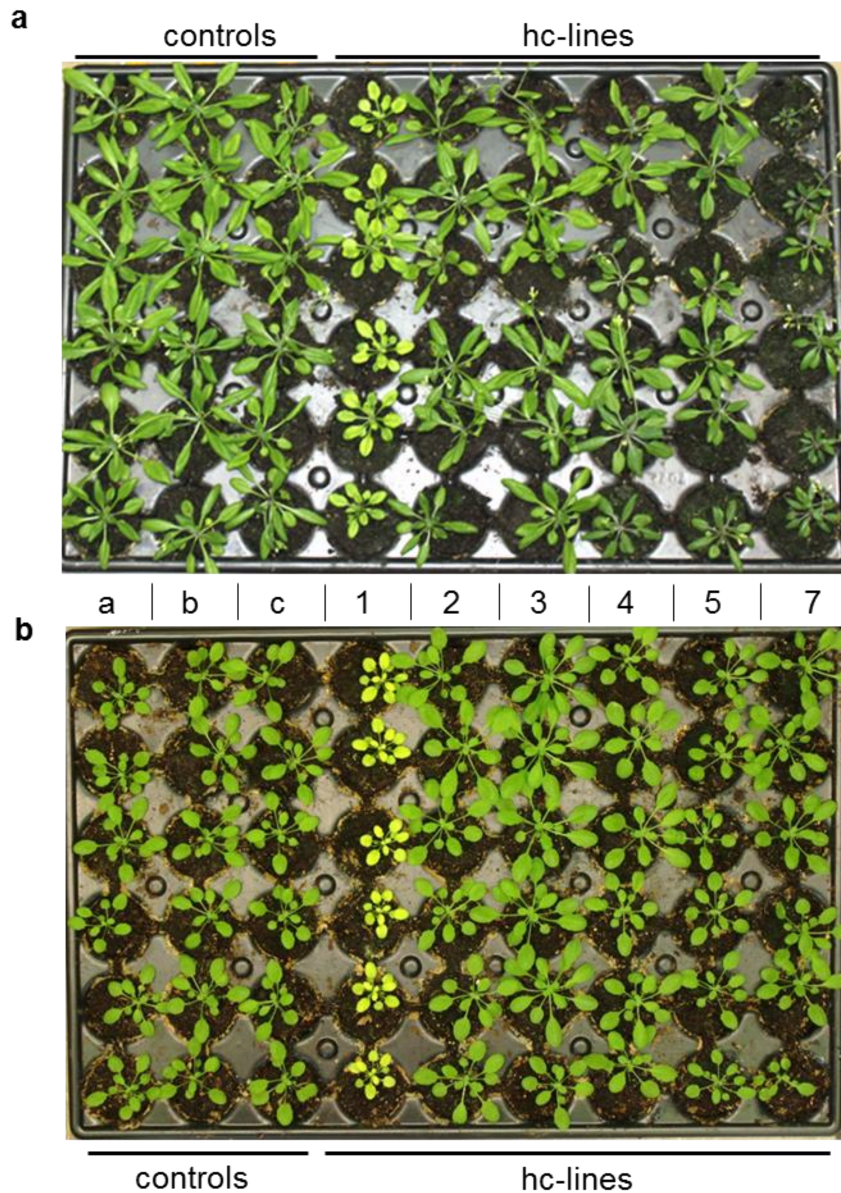


Figure 18 ***ONSEN* high-copy lines show a broad panel of phenotypes**. Photographs of selected hc-lines in the S2-generation originating from HS (a-c) or HS and A (5 µg/ml) and Z (40 µM) treated WT *Arabidopsis* plants showing both homogeneous and environment-dependent phenotypic variability induced by the *ONSEN* mobilization when grown under long (**a**) and short day conditions (**b**). Pictures of hc-line 6 are missing due to severe infertility and extinction of this line.

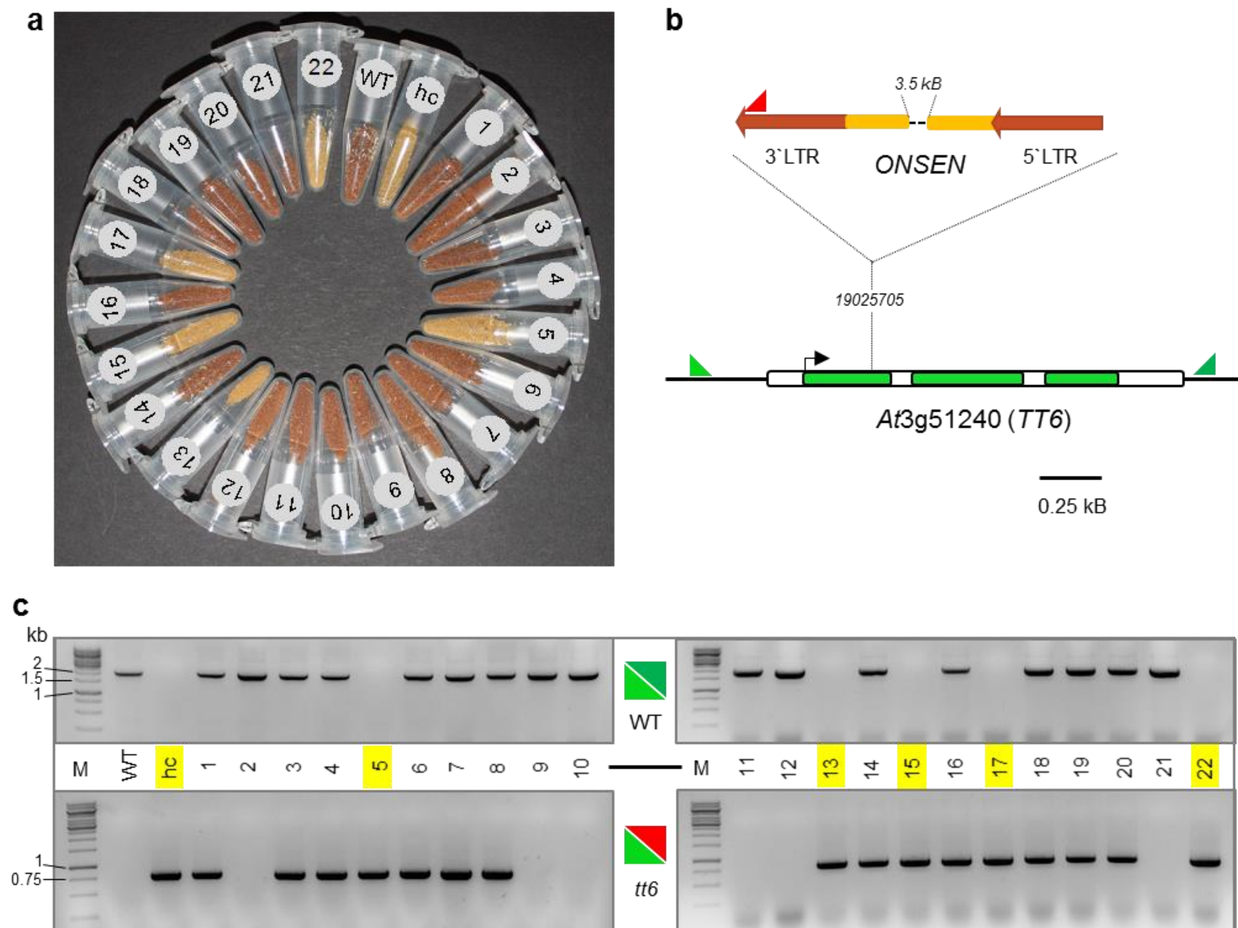


Figure 19 **Transparent testa phenotype of hc-line 3 co-segregates with an *ONSEN* insertion in *TT6*.** Seed-phenotypes (**a**) and corresponding genotypes (**c**) of a segregating F2 population (1-22) obtained from a cross between the WT and hc-line 3 (hc) are shown. **b** Schematic representation of primers used for genotyping of the *ONSEN* insertion. For the WT-PCR depicted in the upper part of (**c**) the light (*tt6* fw) and dark (*tt6* rev) green primers flanking the *TT6* locus (AT3G51240) were used. The *ONSEN* insertion in *TT6* was detected by a combination of the light green primer with the red primer specific to the *ONSEN* LTR (*Copia* 78 3`LTR, red arrow). Primer sequences are given in table S1 of (Thieme *et al.* 2017).

When growing a pool of plants in the S2-progeny of a sibling of hc-line 6 (name: ext 5c5 see Fig. 15f) we made the striking and reproducible observation that a certain percentage (~15-20%) of individuals in this pool showed a lethal albino phenotype at the seedling stage (Fig. 20). This indicated the segregation of a causative genetic mutation in this line.

Figure 20 **Segregation of an albino phenotype within the progeny of an hc-line**. Seedlings in the progeny of the S1- generation of plant ext 5c5 (see Fig. 15f) showed segregation of a mutation causing an albino phenotype when grown under long day conditions. Albino plants are highlighted with white arrows.

5.3.3 Genetic stability of *ONSEN* high-copy lines under HS

To address the question whether re-application of heat-stress would lead to a burst of *ONSEN*-transposition in hc-lines, we applied heat-stress with and without A&Z-treatment to the S3-generation of two selected hc-lines (#3 & 4) and measured *ONSEN* copy number by qPCR. We found that the re-application of HS and drugs in the S3 generation of these hc-lines did not lead to a stronger accumulation of *ONSEN* copies compared to control lines and rather observed a stronger silencing in lines with more integrated *ONSEN* copies (Fig. 21).

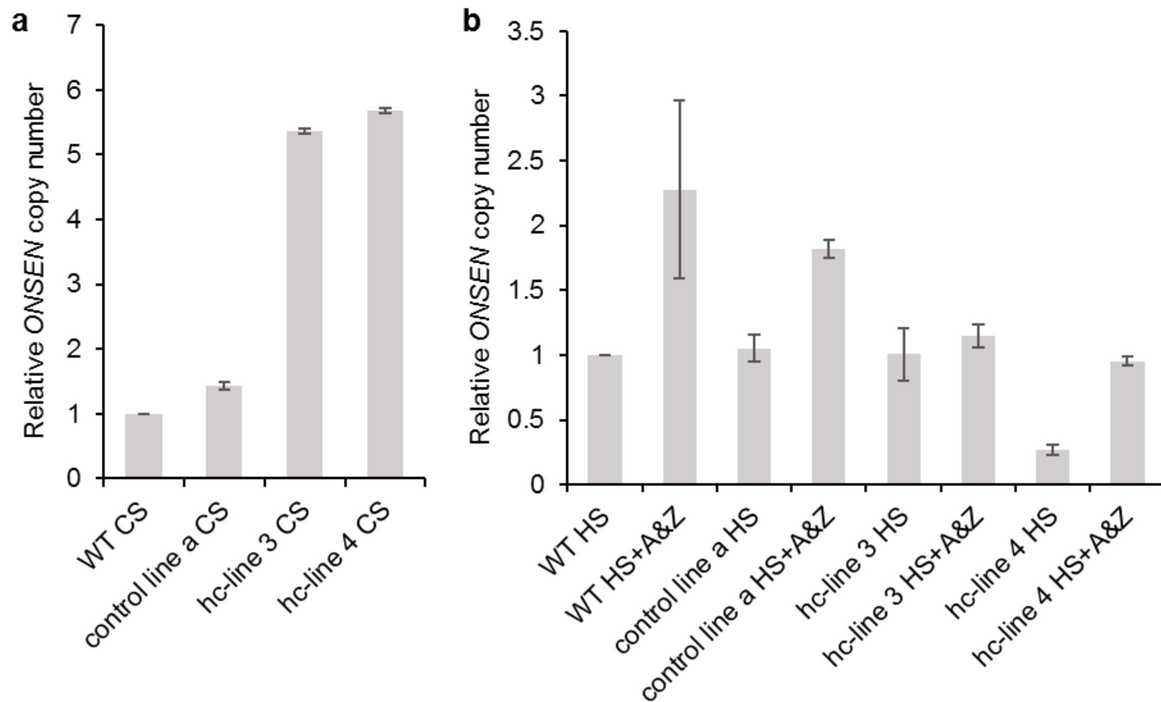


Figure 21 **Heat-stress-induced activation of *ONSEN* in the S3 generation after initial HS-treatment.** *ONSEN* copy number measured by qPCR directly after CS (**a**) and HS or HS plus treatments with α-amanitin (A, 5 µg/ml) and zebularine (Z, 40 µM) (**b**) in seedlings of the WT, the control line a and the hc-lines 3 and 4. *ONSEN* copy number is shown relative to the WT CS (**a**) or HS (**b**) (mean ± s.e.m., n= 3 biological repetitions, all values relative to *ACTIN2*).

5.4 Discussion

We previously observed that both a hypomorphic mutation in the second largest subunit of Pol II and its inhibition with α-amanitin in combination with a chemical DNA de-methylation with Z can trigger a very strong heat-dependent accumulation of *ONSEN* in *Arabidopsis* (chapter 4). This led us to investigate whether some of these *ONSEN* copies would stably integrate into the genome and could be detected in the progeny of heat-stressed seedlings. As transposition is known to be a stochastic process (Ito *et al.* 2011; Matsunaga *et al.* 2015) we first choose to measure the average *ONSEN*-copy number of several pooled S1 seedlings that originated from the same parent. The observation that only pools in the progeny of heat-stressed and Z-treated *nrpb2-3* plants showed a significant increase in *ONSEN* copy number supports our previous findings on the importance of Pol II in TE-silencing (chapter 4). Based on this result, we used the same approach to screen for additional *ONSEN* copies in the progeny of heat-stressed and A&Z-treated WT plants grown under control conditions and found an increase of *ONSEN*-copy number in 29.4% of the tested S1 pools. By first demonstrating an increase in copy number by qPCR on an individual basis, followed by sequencing and genotyping of true novel *ONSEN* copies in a selected hc-plant, we substantiated the approach of testing pooled S1-seedlings. Thus, the possibility that increased *ONSEN*-copy numbers detected in soil-grown S1 pools could have been caused by the induction of ecDNA-production by other stresses such as bacterial infection known to activate *ONSEN*-transcription (Yu *et al.* 2013) is unlikely. However, further evidence gained through additional experiments like transposon display or sequencing would be needed to certify novel *ONSEN*- insertions in the *nrpb2-3* background.

Interestingly, we observed that *ONSEN* copy numbers in individual S1-progeny plants often depicted a binary pattern meaning that hc-lines within one pool showed a very similar number of *ONSEN* copies whereas the others did not gain any copies. This could suggests that *ONSEN* transposed at an early developmental stage in the germline-progenitors of the parent resulting in mosaic-plants producing self-pollinating flowers with little diversity for *ONSEN*-copy numbers. The fact that we also observed homozygous insertions already in the S1-generation of treated plants supports this conclusion. Furthermore, this clearly indicated that *ONSEN* at least partially already transposed before the development of female and male gametes. However, this

question should be systematically addressed by testing individual siliques of treated parental plants as reported previously in the *nrpd1*-mutant background (Matsunaga *et al.* 2015). In doing so, also the exact degree of homozygosity of novel insertions in the S1-generation as well as in the somatic tissue of the paternal plant should be taken into account.

When tracking *ONSEN*- copy numbers over three generations we generally observed a high degree of stability within each of the tested hc-lines. The observed fluctuations of copy numbers can be explained by segregation of heterozygous insertions. The observed occurrence of around 15-20 % albino seedlings among siblings in the S2 generation of an hc-line is likely to be yet another illustrative indicator for the segregation of *ONSEN* insertions. Moreover, we also detected heterozygous insertions by genotyping novel *ONSEN* copies identified in hc-line 3.

On the other hand, we also found homozygous insertions of *ONSEN* that resulted in very stable phenotypes such as the *transparent testa* phenotype in hc-line 3. Interestingly, we also noted several environment-dependent alterations of complex traits such as early flowering and biomass production that could be caused by various (epi)genetic changes caused by novel *ONSEN* insertions (Lisch 2013). Indeed, our observations confirm the previous finding that *ONSEN* tends to integrate close to or into gene-rich region in *Arabidopsis* (Ito *et al.* 2011; Underwood *et al.* 2017). In line with this, it was recently reported that *ONSEN*-transposition into euchromatic regions of a heat stressed *nrpd1*-plant was the reason for an abscisic acid-insensitivity phenotype that was observed in its progeny (Ito *et al.* 2016). Even though we were able to clearly link the observed *transparent testa* phenotype to the homozygous insertion of *ONSEN* in the *TT6* locus (*AT3G51240)* it remains unclear whether transposition of other TEs or possible drug-induced epigenetic or genotoxic effects could underlie some of the observed phenotypes. For instance it was shown that a Z-treatment can induce DNA damage during strand synthesis in DNA replication (Liu *et al.* 2015) and lead to the induction of class II *CACTA*-elements in *Arabidopsis* (Griffin *et al.* 2016). Nonetheless, our findings will allow future studies on the potential beneficial role TEs play in the adaption to stresses in wild-type plants. Indeed, two recent studies point out the adaptive potential of retroTEs and -more specifically- of *ONSEN* copy number variation in natural accessions (Quadrana *et al.* 2016) and in the RdDM-mutant backgrounds of *Arabidopsis* (Ito *et al.* 2016). Upon mobilization, the

heat-responsive elements in the LTRs of *ONSEN* (Pietzenuk *et al.* 2016) can create new gene regulatory networks responding to HS (Ito *et al.* 2011). Therefore, it will now be of great interest to test if the *ONSEN* hc-lines obtained in this study are better adapted to HS. This will open up the possibility to test if retroTE-induced genetic and epigenetic changes more rapidly create beneficial alleles than if it would occur by random mutagenesis. Whole-genome (bisulfite) sequencing in combination with repeated backcrossing, phenotyping and assessment of stress-dependent transcriptional changes in hc-lines will allow for a differentiation of possible mechanisms underlying the overserved phenotypic diversity.

Finally, the observation that HS did not lead to a stronger activation of *ONSEN* in hc-lines compared to WT plants suggests that genome stability is not compromised in these lines. This result can be explained by at least two possible mechanisms: (i) The occurrence of insertions of inverted duplications of *ONSEN*, such as has been observed for the Mu killer locus in maize (Slotkin *et al.* 2005). Such insertions will lead to the production of double-stranded RNAs that feed into gene silencing and thereby limit the activity of that TE. (ii) Another possibility is the balance of retroTE-activity and integrated copy number as it has been described for *EVADE* in *Arabidopsis* (Mari-Ordonez *et al.* 2013). In this case when a certain TE copy number threshold is reached robust transcriptional gene silencing takes over, thereby limiting retroTE mobility and ensuring genome stability. Although it is hard to draw general conclusions from this two case observations, it can be speculated that the observed balance of copy number and activity of retroTEs is one of the fundamental mechanisms that explains their success in co-evolution with their host. In addition, the stability of new retroTE insertions is an important aspect in light of the future use of TEs in crop breeding and trait stability.

TEs are important contributors to genome evolution. The ability to mobilize them in plants and possibly in other eukaryotes in a controlled manner with the straightforward drug application shown here opens up the possibility to study their importance in inducing genetic and epigenetic changes resulting from external stimuli. Because the induced transposition of *ONSEN* can efficiently produce TE-tagged (epi)genetic diversity resulting in developmental changes in *Arabidopsis*, it will be very interesting to test if specific stress-induced TE activation can be used for directed crop breeding for improved stress tolerance in the near future.

5.5 Material and Methods

Experiments were performed according to (Thieme *et al.* 2017).

6 Induced activation of retrotransposons in crops

A modified version of this chapter was published in Thieme, M. *et al.*, (2017). Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. Genome Biol 18, 134.

6.1 Abstract

Retrotransposons are increasingly seen as a valuable endogenous genetic resource that could be harnessed for plant breeding. However, as they are under strict epigenetic regulation their controlled mobilization has so far been limited. Here I show that the combined chemical inhibition of Pol II and DNMTases that has previously been reported to mobilize a stress responsive retroTE in *Arabidopsis* (chapter 5) also leads to the activation of a class I element in rice (*Oryza sativa*). Based on this observation, I conclude that RNA polymerase II is a highly conserved player of TE-silencing in plants that can now be targeted by a straightforward drug application. As rice and *Arabidopsis* differ significantly in their epigenetic landscape but still show a response to the combined drug-treatment, I further conclude that this new approach could be used to mobilize retroTEs in virtually any plant. In a first attempt to broaden the reported applicability of the two drugs to induce retrotransposition, I thus present and discuss a preliminary approach to activate class I elements in soybean (*Glycine max*).

6.2 Introduction

We have previously shown that the combined inhibition of DNMtases and RNA polymerase II leads to heat-stress- dependent activation (chapter 4) and mobilization (chapter 5) of the *ONSEN* retroTE in *Arabidopsis*. Plants with additional *ONSEN* copies showed a striking phenotypic diversity (chapter 5) that was partially dependent on the exposure to different environmental conditions. As retroTEs are important drivers of genome evolution (Lisch 2013) and could therefore be an interesting genetic resource for plant breeding (Paszkowski 2015) these observations in the model plant *Arabidopsis* lead us to test whether the same approach could also trigger ecDNA production, an intermediary step in the life cycle of retrotransposons, in crops. With the aim to generalize our findings from *Arabidopsis* we choose to apply the same treatment to rice (*Oryza sativa*) a genetically well-characterized (Kawahara *et al.* 2013) monocotyledonous crop.

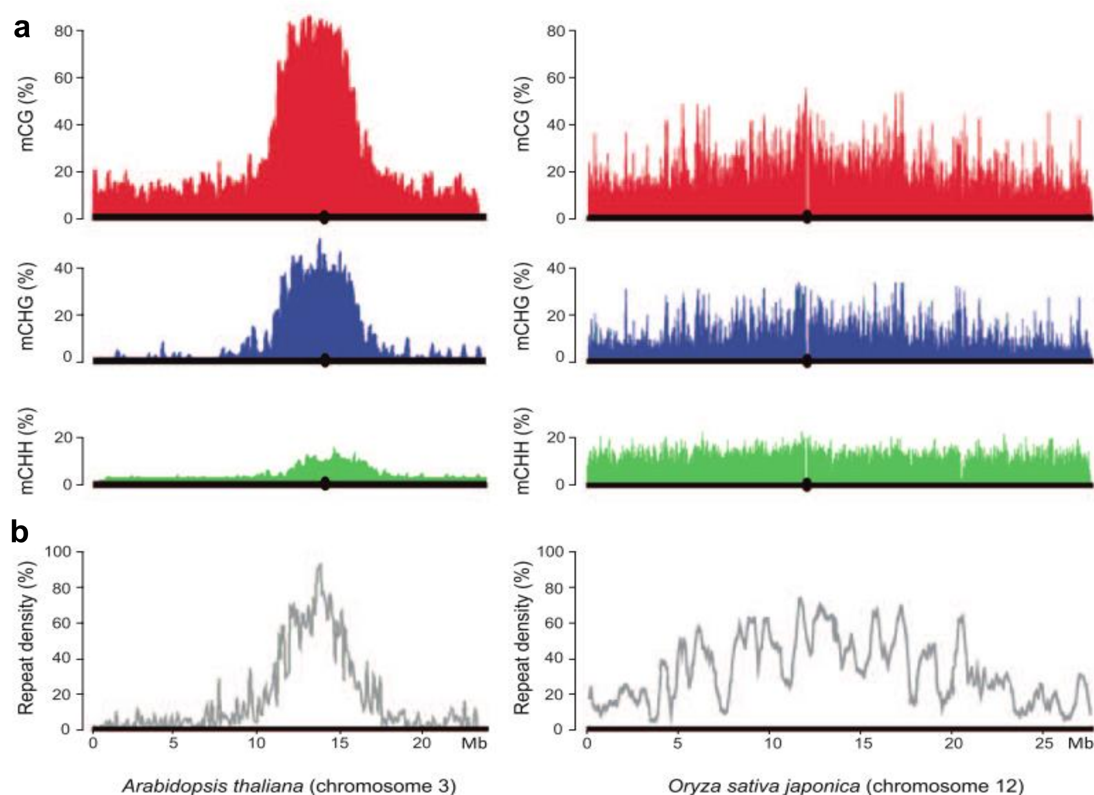


Figure 22 **Differences in the epigenetic and TE-landscapes of *Arabidopsis* and rice. a** Chromosome-wide distribution of DNA methylation (red: mCG, blue: mCHG, green: mCHH) in sliding windows of 100kb or 50kb in *Arabidopsis* and rice, respectively. **b** Chromosome-wide distribution of repeat density (in %) in sliding windows of 100kb. Chromosome coordinates are given in Mb. Figure adapted from (Mirouze and Vitte 2014).

Importantly, both the epigenetic and the TE landscape of rice differ significantly from *Arabidopsis* (Liu *et al.* 2017; Mirouze and Vitte 2014). Thus, TEs that are generally more abundant in the rice genome (see general introduction) are also more equally distributed across the rice chromosomes. As a consequence, the degree of DNA-methylation, which is known to correlate with the presence of TEs in the genome, also differs drastically between *Arabidopsis* and rice (Fig. 22).

As a second genetically well-characterized crop (Fang *et al.* 2017) to test the efficiency of the two inhibitors A&Z we choose soybean (*Glycine max*) which is currently one of the most important crops worldwide (Wilson 2008). Similar to rice, soybean significantly differs from *Arabidopsis* both at the DNA-methylation level (Mirouze and Vitte 2014) and the content and distribution of TE annotations in its genome (Fig. 2). As retroTEs are known to often depend on (a)biotic stress for their activation (Negi *et al.* 2016) and we have previously shown that inhibition of TE-silencing in combination with the occurrence of an abiotic stress can lead to efficient retrotransposition in *Arabidopsis* (chapters 4 & 5) we also wanted to address the question whether we could mobilize stress-responsive retroTEs in soybean. In a preliminary experiment, we chose to apply cold-stress as an inducer of retroTEs in soybean. In soybean, cold-tolerance is one of the main breeding objectives to reduce yield losses in Switzerland. Low temperatures can for instance trigger flower abscission (Schori and Charles 2004). The identification and potential mobilization of a cold-stress responsive retroTE in soybean could therefore significantly accelerate soybean-breeding for Switzerland and other areas with suboptimal climate.

6.3 Results

6.3.1 Simultaneous inhibition of Pol II and DNMtases mobilizes a *copia*-like retrotransposon in rice

To capture drug-induced mobilized retroTEs, we followed the same approach as for *Arabidopsis* (chapter 4) and characterized the active mobilome in rice seedlings that were grown on MS medium supplemented either with no drugs, A (5 µg/ml) only, Z (40 µM) only or the combination of A and Z. We generally noted that both drugs and especially Z present in the medium negatively affected the growth of rice seedlings (Fig. 23) *in vitro*.

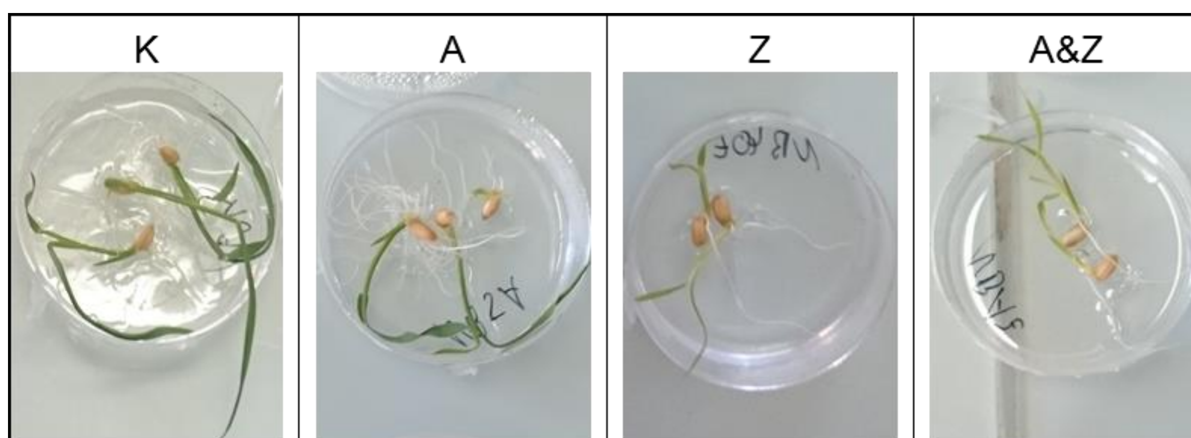


Figure 23 **Representative phenotypes of seedlings of *Oryza sativa japonica* after growth on medium containing A, Z or a combination of both inhibitors.** Seedlings were germinated and grown for 10 days on medium containing no drugs (K), α-amanitin (5 µg/ml) (A), zebularine (40 µM) (Z) or a combination of both inhibitors (A&Z). The aerial parts of three treated seedlings were used for the mobilome analysis.

After the extraction and processing of DNA from the aerial parts of treated rice seedlings, we performed mobilome-sequencing (as explained in chapter 4) in order to detect activated retroTEs. We found that *Houba*, a *copia*-like retroTE (Panaud *et al.* 2002), was highly activated only when plants were treated with the combination of A&Z (Fig. 24). *Bona fide* activity of *Houba* was supported by the detection of mobilome-sequencing reads originating from eccDNA containing LTR-LTR junctions (Fig. 25). The presence of closed *Houba*-circles and thus the mobility (see chapter 4) of *Houba* was further confirmed by an inverse PCR using a pair of primers designed to only amplify closed *Houba* eccDNA (Fig. 24b). As a control for equal amounts of template in the inverse PCR we used primers specific to chloroplast DNA (Fig. 24d).

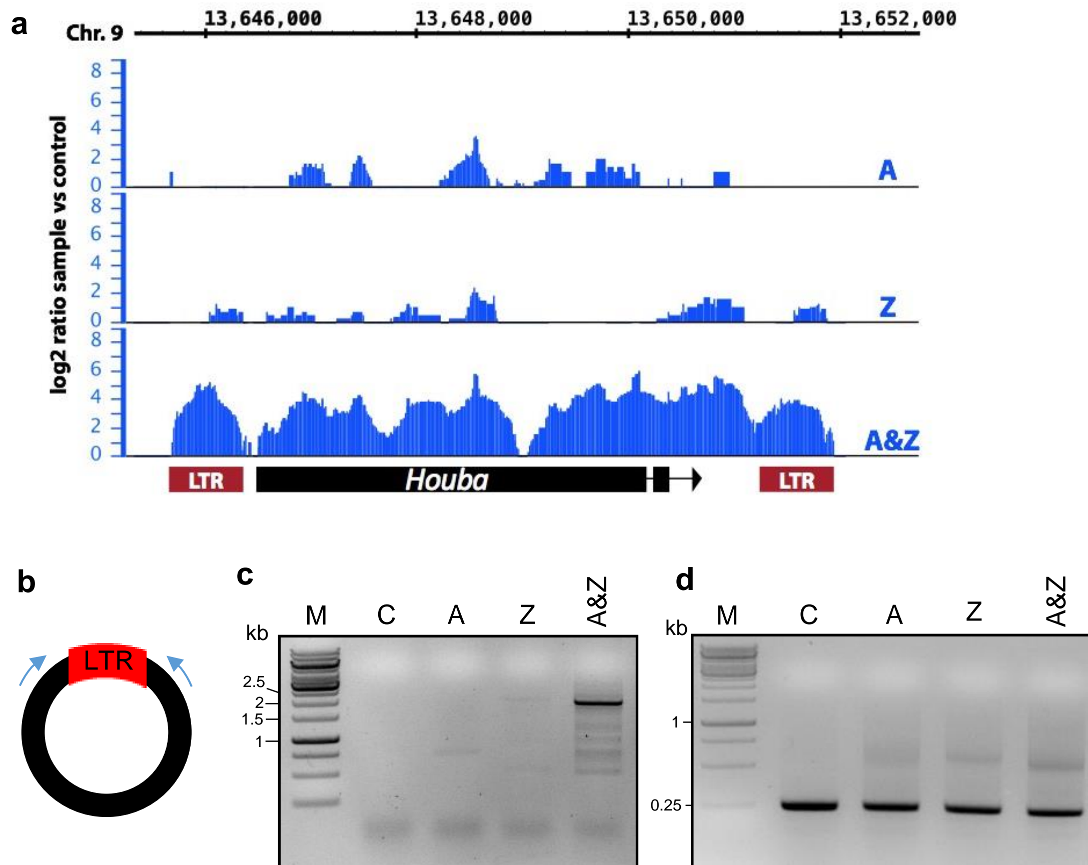


Figure 24 **Drug-induced activation of the *Houba* retrotransposon in *Oryza sativa***. Mobilome analysis of DNA extracted from seedlings after growth on control conditions (C), A (5 µg/ml), Z (40 µM), and the combination of A&Z. **a** Logarithmic ratio of detail of the depth of coverage obtained after aligning the sequenced reads on one *Houba* element. **b** Scheme of primers localization (*black bar*: *Houba* element, *arrows*: PCR primers, *red box*: LTR). **c** circular forms of *Houba* are specifically detected in plants treated with both A&Z using inverse PCR with primers shown in (**b**). **d** Specific PCR on chloroplast DNA is shown as a loading control. Total DNA subjected to a rolling circle amplification was used as a template.

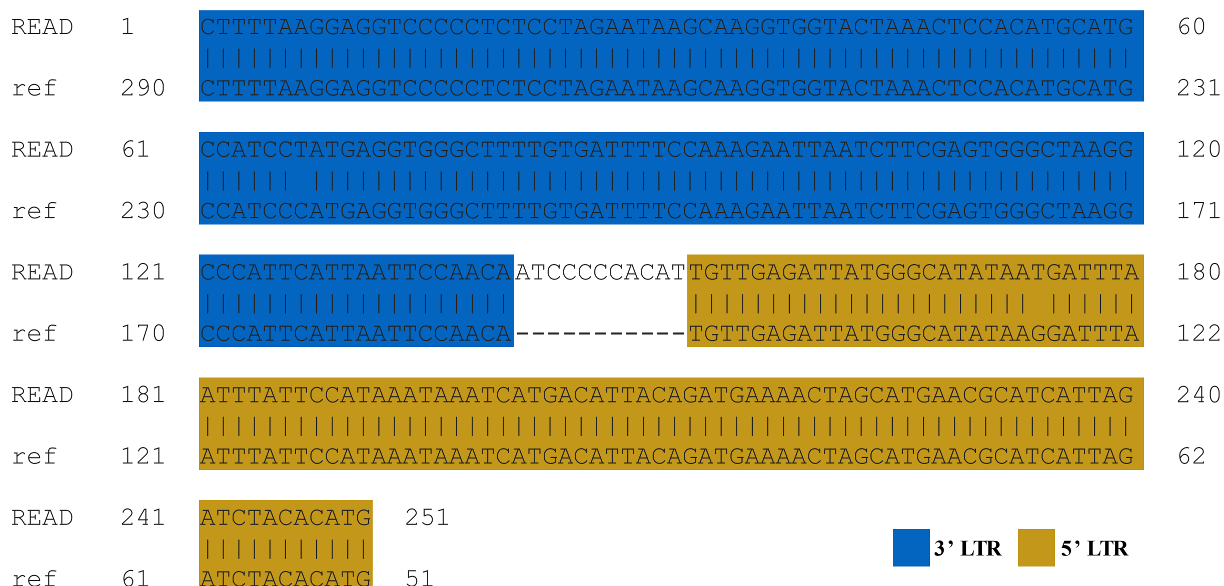


Figure 25 ***Houba* forms LTR-LTR junction eccDNAs when treated with the combination of A&Z.** Alignment between a sequencing read resulting from the mobilome sequencing of A (5 µg/ml) and Z (40 µM) -treated plants (top) and an artificial junction corresponding to the 3' part of the 3' LTR (blue box) fused to the 5' part of the 5' LTR (yellow box).

In a preliminary experiment, we also applied HS (37°C, 24h) to rice seedlings grown on control and A&Z-containing medium. However, we could not detect a distinct activation of a heat-responsive retroTE in the mobilomes of these plants. However, in line with the previous observation that the A&Z-treatment specifically led to an accumulation of eccDNA of *Houba*, we were also able to show by inverse PCR that production of *Houba* eccDNA was induced independently of the applied HS (Fig. 26, *preliminary data*).

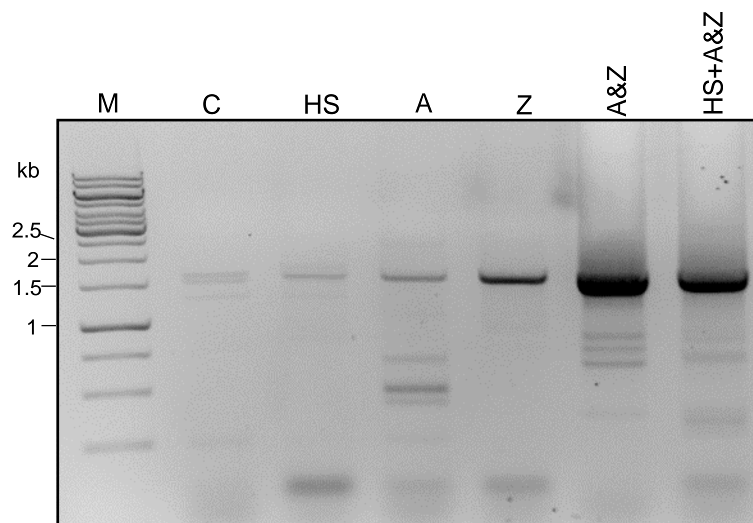


Figure 26 **Drug-induced activation of the *Houba* retroTE in *Oryza sativa.*** The agarose-separated products of an inverse PCR (as explained in Fig. 24b) on total DNA extracted from seedlings after growth on control conditions (C), heat stress (HS), A (5 µg/ml), Z (40 µM) the combination of A&Z and HS+A&Z that was subjected to a rolling circle amplification is depicted. Circular forms of *Houba* are predominantly detected in plants treated with both A&Z. *Preliminary data*.

In the case of the heat responsive retroTE *ONSEN,* we have previously linked the detection of mobilome-reads originating from eccDNA to the presence of novel *ONSEN* copies in individuals of the successive generation (chapter 5). To address the question whether extrachromosomal *Houba* copies induced by the A&Z-treatment could also be detected in the progeny of a A&Z-treated rice plants, we performed an IRAP (inter retrotransposon amplified polymorphism)-analysis (Yuzbasioglu *et al.* 2016). In a preliminary approach, we choose to compare individuals originating from four different panicles of a control plant that was only heat-stressed and in total twelve individuals from four panicles of one A&Z-treated and heat-stressed rice plant (Fig. 27, *preliminary data*). We noted distinct band patterns in each of the samples indicating a high degree of diversity of *Houba*-insertions between individual plants and also within one panicle (Fig.27, *preliminary data*). Notably, distinct band patterns were also observed between individuals originating from the tested control plant. However, due to a limited resolution, we observed a high degree of overlap between bands of different sizes. Furthermore, the PCR-efficiency between samples was highly variable, resulting in systematic and possibly unspecific differences between band intensities.

We could therefore not determine whether actual transposition events occurred as a result of our treatments.

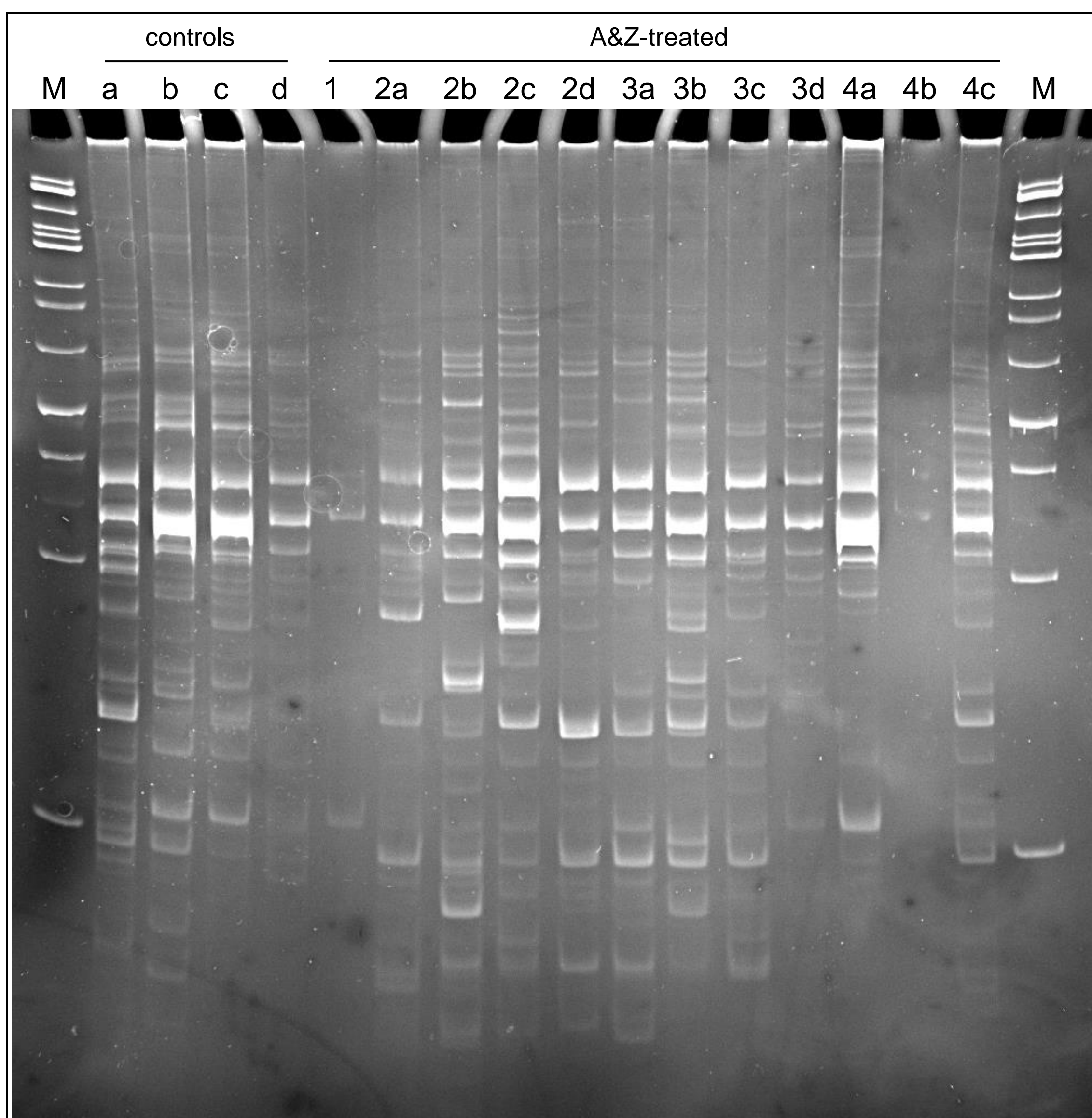


Figure 27 **IRAP-analysis for the detection of *Houba*-copy number variation in rice**. IRAP-analysis of the *Houba*-retroTE in the S1-generation of heat-stressed (a-d) and heat-stress and A (5 µg/ml) and Z (40 µM) treated *O. sativa* seedlings (1-4c). Control plants a-d originate from seeds of four independent panicles of one heat stressed rice seedling. Same numbers in labels of A&Z-treated plants indicate that they origin from the same panicle of the tested parental plant. The PCR products of the IRAP-analysis were separated on a Midori-stained acrylamide gel, a GeneRuler 1 kb DNA Ladder (Thermo Scientific) was used as a size marker (M). *Preliminary data.*

6.3.2 Testing A&Z-treatments to induce retrotransposition in soybean

Due to our previous observation that the combination of 5 µg/ml α-amanitin and 40 µM zebularine led to the detection of extrachromosomal DNA of the retroTEs *Houba* in rice and the heat-responsive *ONSEN* element in *Arabidopsis* (chapter 4) we wanted to test whether the same treatment would also activate class I elements in the dicotyledonous crop soybean. For that we germinated seeds of soybean on control medium or medium containing 5 µg/ml α-amanitin and 40 µM zebularine. As already previously observed in rice (Fig. 23), we noted that the drug-treatment caused a distinct negative effect on the root and shoot growth of soybean seedlings (Fig. 28).

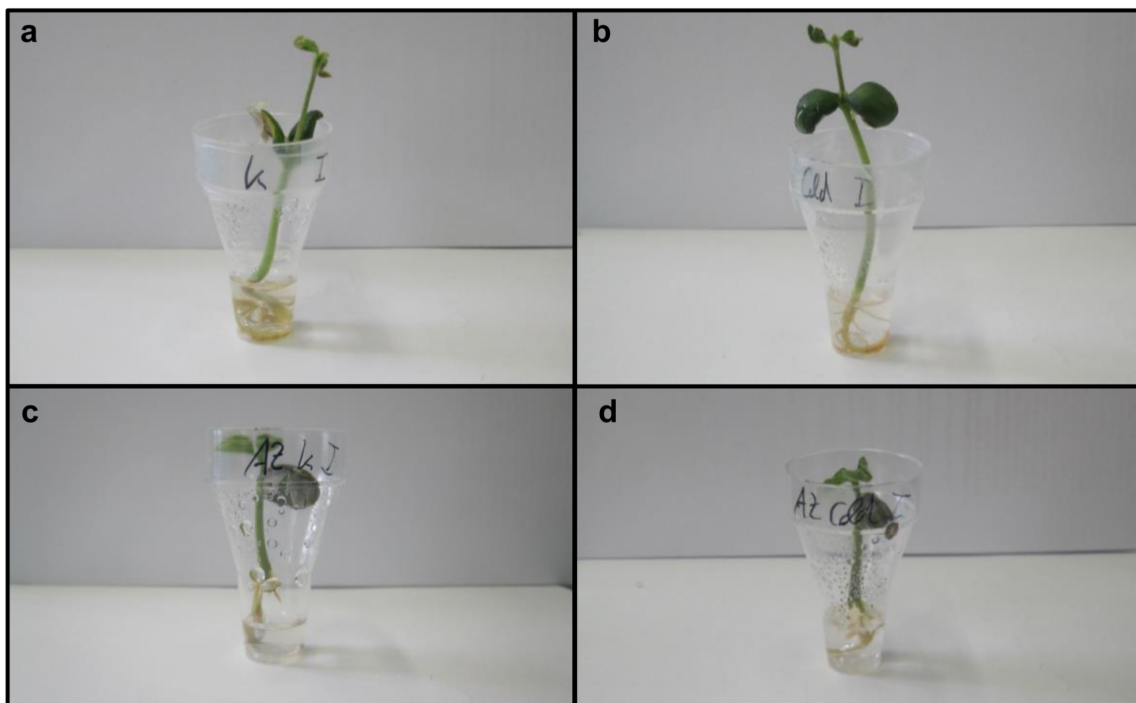


Figure 28 **Representative phenotypes of seedlings of *Glycine max (Wm82)* after growth on medium containing a combination of the inhibitors A&Z.** Seedlings were germinated and grown for 16 days on medium containing no drugs (**a** and **b**) or a combination of α-amanitin (5 µg/ml) and (A), zebularine (40 µM) (Z) (**c** and **d**). Plants were either grown under control conditions (**a** and **c**) or cold stressed at 4°C for 48 h (**b** and **d**).

Despite the negative effects of A&Z on the growth performance of soybean *in vitro*, we observed a high degree of recovery and survival when seedlings were subsequently rescued on soil and grown until seed maturity under controlled conditions in the greenhouse (Fig. 29).

Figure 29 **Representative example for a cold and A&Z-treated seedling of soybean (Wm82) that was rescued from A&Z-containing medium and grown for ten days on soil.**

We further attempted to analyze the mobilomes of soybean in DNA extracted from the first emerging true leaves following growth on control medium or the A&Z-treatment in combination with or without cold-stress. However, by mapping all reads of the individual mobilomes to a concatenation of all known LTR-TEs of soybean (Du *et al.* 2010), we could not detect major condition-dependent distinct patterns of coverage that would indicate the activation of a specific LTR-retroTE in this preliminary experiment (Fig. 30).

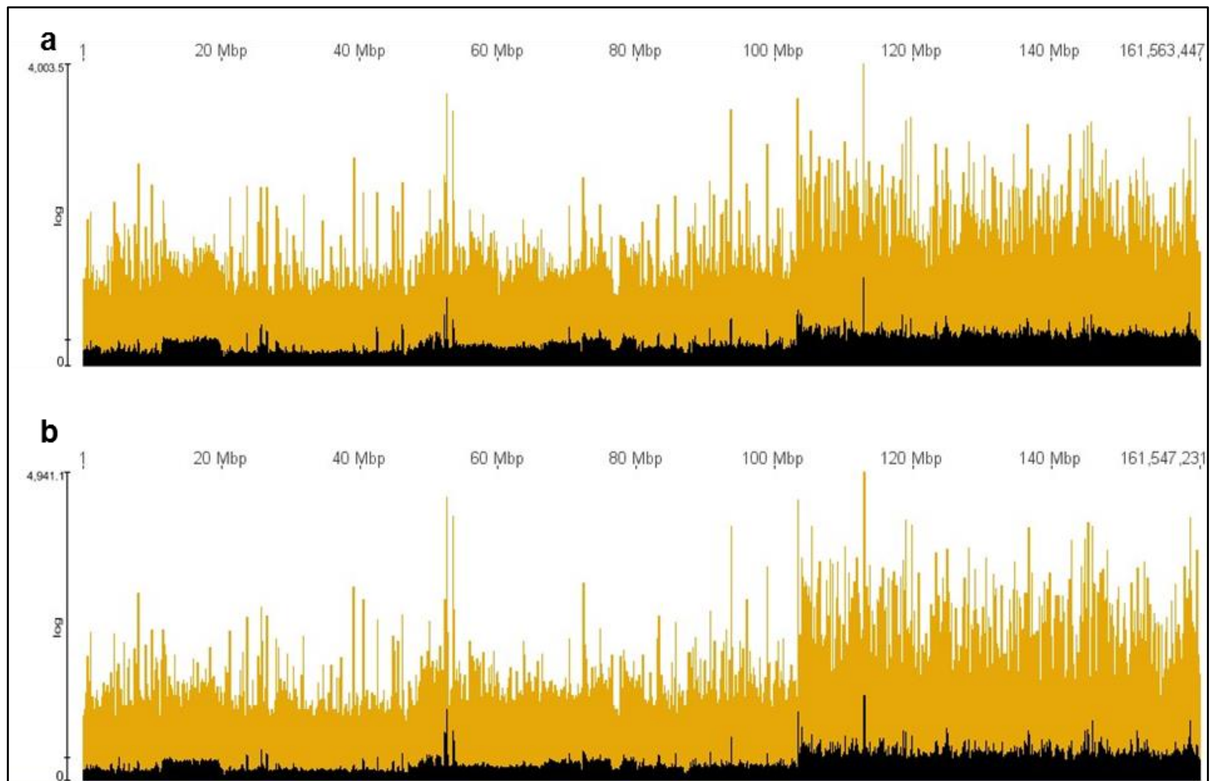


Figure 30 **Coverage of mobilome-reads mapped to the concatenation of all known LTR-retroTEs of soybean.** Mobilome-reads from DNA extracted from the control sample (**a**) and the sample obtained after the A&Z-treatment and cold stress (**b**) are depicted. The average (black bars) and maximum (yellow bars) coverage is depicted in a log scale.

6.4 Discussion

As described above, experiments on TEs or epigenetics in the model plant *Arabidopsis* are not necessarily valid in crops such as rice or soybean (Mirouze and Vitte 2014). However, from a mechanistic point of view, the high degree of enzymatic homology especially of RNA-polymerase II between different plant species (Nawrath *et al.* 1990) and eukaryotes in general (Ream *et al.* 2009; Sweetser *et al.* 1987), strongly suggested that A&Z would also interfere with TE-silencing in plants other than *Arabidopsis.* Indeed, we found that the double-treatment specifically activated the *Houba* retroTE in rice confirming this assumption. By testing several *Arabidopsis* mutants impaired in different branches of the RdDM-pathway we have previously found strong evidence that the recently discovered DCL3-RdDM (Panda *et al.* 2016) is a main target responsible for the strong activation of *ONSEN* after HS and the A&Z-treatment (chapter 4). As discussed above and also suggested by (Panda *et al.* 2016) the DCL3-dependent route of siRNA biogenesis is assumed to be especially important to silence evolutionary young and transcriptionally active retroTEs in plants. Indeed, there are indications that *Houba* (Vitte et al. 2007) just like *ONSEN* (Quadrana *et al.* 2016) showed very recent natural transposition events in the rice and the *Arabidopsis* genomes, respectively. *Houba* is the most abundant retroTE of the *copia* family in rice and has been active in the last 500’000 years (Wicker and Keller 2007). Based on these two cases including distantly related species such as *Arabidopsis* and rice it can be speculated that similar expression dependent silencing mechanisms that directly depend on a silencing signal produced by Pol II generally exist in plants. Moreover, it would be interesting to determine whether there are structure-dependent subtle subclass-specific differences in TE-silencing, potentially caused by subgenomic transcription (chapter 4). This could explain the observation that the A&Z-treatment only induced two very distantly related members of the *copia*-subfamily in rice and *Arabidopsis*. The analysis of genome-wide methylation patterns after inhibition of Pol II and/or DNMtases, possibly in combination with reverse genetic approaches (Stroud *et al.* 2013) could further elucidate the role of Pol II in surveilling genome integrity in rice.

In contrast to the strict heat-stress dependence of *ONSEN* (chapter 4 and (Cavrak *et al.* 2014)) we noted that the production of eccDNA of *Houba* in rice was already triggered after germination and growth on ½ MS medium containing the combination

of A&Z in the absence of additional stresses. It is known that drastic events such as tissue culture can activate retroTEs in rice (Hirochika *et al.* 1996). Keeping this in mind, it would be interesting to systematically test to what extend different parameters such as pH, UV-light, humidity or sucrose present in the medium could influence the level of *Houba* activation in rice seedlings grown on sterile culture. Indeed, we observed that the vitality of rice seedlings was negatively affected by the A&Z- treatment, suggesting suboptimal growth conditions. In *Arabidopsis*, several key-findings such as the important role of Pol IV in TE-silencing (Ito *et al.* 2011) have been substantiated by using *ONSEN* as a by now well-characterized stress responsive model-TE (Cavrak *et al.* 2014). Further studies elucidating the exact mechanisms underlying the activation of *Houba* could allow its application as an endogenous model-retroTE for basic research in the important rice crop.

A recent publication revealed, that in some rice cultivars, transposition of *Houba* can cause a high degree of polymorphic insertions in somatic tissue extracted from roots and leaves (Yuzbasioglu *et al.* 2016). By using the same technical approach to detect inter retrotransposon amplified polymorphisms (IRAPs) we attempted to demonstrate the correlation of *Houba* eccDNA accumulation following A&Z-treatment and the detection of novel insertions that could be tracked in the successive generation of rice cultivar *Nipponbare*. Although we also observed a high degree of polymorphisms between tested lines, we could not exclude the possibility that either somatic transposition of *Houba* or random fluctuations of the efficiency of the IRAP-analysis have led to the observed band patterns. Therefore, the transgenerational transposition of *Houba* still needs to be confirmed and may be hampered by the already very high *Houba* copy number present in the rice genome (Vitte and Panaud 2005). However, novel powerful tools such as highly sensitive digital droplet PCRs to determine copy number variations (White *et al.* 2014) or newly developed sequencing technologies such as Oxford Nanopore Sequencing (Debladis *et al.* 2017) might overcome current limitations in the detection of transposition events of high-copy TEs such as *Houba* in rice. Thus, the technological advance that allows for the generation of significantly longer sequencing reads compared to Illumina sequencing will facilitate a more reliable detection of novel TE-copies also in complex genomes (Debladis *et al.* 2017). In that case, the preliminary observation that heat-stress did not affect the A&Z-dependent induction of *Houba*, opens up the possibility to determine whether a potential insertion-

site preference of *Houba* would be affected by the concurrent presence of heat-stress. In addition, the successful identification and knowledge of *Houba* as being highly active after the A&Z-treatment might also allow for forward genetic approaches using only endogenous genetic material as previously shown for *ONSEN* in *Arabidopsis* (chapter 5).

Due to the promising results from rice, we also attempted to activate and mobilize retroTEs by inhibiting Pol II and DNMTases in soybean. In a first trial we used the same concentrations of the two inhibitors α-amanitin and zebularine that we demonstrated to be efficient in both *Arabidopsis* and rice. Like in rice, we observed a distinct negative phenotypic effect of the two inhibitors on the seedling growth of soybean suggesting the efficient uptake of at least one of the two drugs. Due to this observation, we choose to analyze the active mobilome of the first true leaves of soybean. However, unlike the activation of *Houba* in rice and *ONSEN* in *Arabidopsis*, we failed to detect significant levels of treatment-specific eccDNA that would indicate the distinct activity of class I elements of soybean. There are several different possible explanations for this observation. Although we saw an obvious impact of the two chemicals on the growth performance of soybean seedlings it is possible that the effect of the two drugs in the first true leaves was buffered by the considerably bigger seeds and higher biomass of soybean compared to *Arabidopsis* and rice. As the Z-mediated reduction of DNA-methylation depends on the inhibition of maintenance methylation and therefore on repeated cell replication (Baubec *et al.* 2009), it is likely that its effect was buffered by the significantly larger soybean cotyledons compared to *Arabidopsis.* A tissue-specific release of repeat-repression following Z-treatment and RdDM-dependent restoration of silencing in true leaves has also been reported for *Arabidopsis* (Baubec *et al.* 2014). It is known that siRNAs produced upon active de-methylation of DNA in companion cells ensure proper TE-silencing in gametes of *Arabidopsis* (Ibarra *et al.* 2012). Based on such observations it has been speculated that mobile siRNAs originating from cotyledons could be involved in ensuring TE-silencing in early vegetative tissue in *Arabidopsis* (Baubec *et al.* 2014). Whether such effects could also counteract or buffer the chemical inhibition of silencing pathways in soybean remains to be determined. More importantly, similar to rice, it was reported that the epigenetic landscape of soybean differs dramatically from *Arabidopsis* (Schmitz *et al.* 2013). Thus it is known that global ^{m}C-content in all three sequence contexts of soybean is considerably higher

compared to both *Arabidopsis* and rice (Mirouze and Vitte 2014). The strong activation of retroTEs following the A&Z- treatment is based on the reduction of ^{m}C-levels in the genome. It is possible that the used concentrations of the two inhibitors was not high enough to shift silenced retroTEs into an active state (see general discussion). To counteract the generally increased levels of DNA-methylation in soybean it might therefore be necessary to test increased concentrations and different ratios of the two inhibitors A&Z. As a control, it could be useful to monitor the efficiency and impact of A&Z on the level of DNA-methylation both in cotyledons and true leaves. The high vitality of cold-stressed soybean-plants on soil rescued from medium containing A (5 µg/ml) and Z (40 µM) supports this notion. Therefore, soybean but also rice could also be treated with significantly higher amounts of the two inhibitors. Indeed, we have previously shown a dose-response for the induced activation of *ONSEN* in *Arabidopsis*. Once the optimal balance between survival rate and efficient DNA-demethylation is found, the mobilome analysis could also be repeated with variations in or combinations of different stress treatments such as cold or heat.

As TE-activity can differ between various varieties of the same crop as it was observed for *Houba* in rice (Yuzbasioglu *et al.* 2016), it would be very interesting to investigate and compare cultivar-specific mobilomes in response to stress. The tremendous impact of retrotransposition on gene-expression and the emergence of phenotypic diversity in crops has been demonstrated in a multitude of examples (see general introduction). After a potential future detection of mobile class I elements it would therefore be very interesting to address the question whether the A&Z-treatment would also result in striking phenotypic diversity and potentially increased stress-tolerance in the F1-generation of soybean or rice. After the proof of concept in the model plant *Arabidopsis* (chapter 5) the generation of retroTE-mediated phenotypic diversity in crops will be the next coherent step required to harness transposable elements for plant breeding.

6.5 Materials and methods

Experiments were performed as described below and according to (Thieme *et al.* 2017).

6.5.1 Heat-stress treatment of rice

Rice seedlings were heat stressed after nine days of growth under control conditions (12 h at 28 °C (day) and 27 °C (night)) on ½ MS plates for 24h at 37°C in a Sanyo MLR-350 growth chamber.

6.5.2 IRAP-analysis to detect *Houba*-copy number variation

Total DNA from the leaves of rice plants that were grown under controlled conditions (12 h at 28 °C (day) and 27 °C (night)) on soil in a Sanyo MLR-350 growth chamber was extracted with the DNeasy Plant Mini Kit (Qiagen) following the manufacturer's recommendations. The IRAP-analysis in the F1 of control or treated individuals was performed according to (Yuzbasioglu *et al.* 2016). After separation of the PCR-products, the acrylamide gel was stained with Midori Green.

6.5.3 Seed material and cultivation of soybean

Seeds material of the *Glycine max* cultivar Williams 82 (Wm 82) (Strain: PI 518671; Seed source: 13U-9281; Subcollection: Modern) was obtained from the USDA Soybean Germplasm Collection, Urbana, Illinois 61801, United States. Prior to growth on soil, soybean was inoculated with the HiStick® Soybean Inoculant (Becker Underwood) according to manufacturer's recommendations. Soybean was grown under controlled conditions (14 h at 25 °C (day) and 22 °C (night)) in a Sanyo MLR-350 growth chamber (for the mobilome-analysis).

6.5.4 *In vitro* culture of soybean

Prior to growth on autoclaved ½ MS-medium pH 5.8, 1% sucrose, 0.5 % Phytagel, soybean seeds were incubated in 30 % bleach (Javel) under stirring and rinsed twice in autoclaved ddH_2O.

6.5.5 Cold treatment of soybean

After twelve days of growth under controlled conditions (14 h at 25 °C (day) and 22 °C (night)) in a Sanyo MLR-350 growth chamber on control or A&Z-containing ½ MS-medium, seedlings of soybean were cold stressed for 48h at 4°C in a Sanyo MLR-350 growth chamber.

6.5.6 Mobilome-analysis of soybean

For the mobilome-analysis DNA from the first true leaves of 14 days old seedlings of soybean (Fig. 28), grown under control conditions or after cold stress and medium with or without A&Z was extracted using a DNeasy DNA extraction kit (Quiagen). For the mobilome-analysis DNA was processed as described in (Lanciano *et al.* 2017; Thieme *et al.* 2017). The obtained reads were mapped to a concatenation of all LTR-retroTEs of the the Wm 82 reference genome downloaded from the SoyTEdb (Du *et al.* 2010) using the “map to reference” tool of Geneious (version 8.1.9). Multiple mappings of reads were allowed.

7 Policy-related aspects of introducing a novel breeding technology to the market

7.1 Patent

Due to the novelty (I), the inventive step (II), and the potential industrial application (III) all three requirements for a patent application to protect the discovery made in my dissertation were met. After a positive report of a patent research at the Swiss Federal Institute of Intellectual Property (Bern, CH) we decided to apply together with the technology transfer organization of the University of Basel (Unitectra) for a patent to protect the process of using the two inhibitors A and Z to induce the "Mobilization of Transposable Elements to Enhance Genetic and Epigenetic Variability in a Population" (Bucher* and Thieme* 2017) (appendix II). The patent has been granted under the number WO2017093317 of the European Patent Office. An exclusive use license of this patent for the application in plants has been granted by the University of Basel to the start-up company epibreed AG (Basel, CH).

7.2 Dissemination

7.2.1 Article for the Newsletter of the PSC

With the aim to disseminate my scientific findings and the idea of using retroTEs for plant breeding to the scientific community (Zürich-Basel Plant Science Center) and to interested readers of the PSC-Newsletter (fall edition 2017, 500 printed copies) I wrote an article with the title "Putting plants in school: on the potential of epigenetic memory for crop breeding" (Thieme 2017b) (english, appendix III). Related to the idea of training plants to transfer acquired knowledge of how to resist a certain stress to successive generations, I pictured retroTEs as highly informative "cheat sheets" that can be copied and inserted somewhere else into the "genetic book" of the plant.

7.2.2 Article pflanzenforschung.de

In cooperation with the PSC I further contacted the editorial office of the online platform "pflanzenforschung.de" which is founded by the German Federal Ministry of Education and Research. They thereupon published the article "Hitze lässt *ONSEN* hüpfen-Retrotransposonen kontrolliert zum Springen bringen" (german, appendix IX) explaining the scientific backgrounds and potential applications to improve crop breeding in lay language to the interested public.

7.2.3 Open access publication

We choose to publish our scientific paper in the open access online journal "Genome Biology" (appendix I). The open access publication is another important element for the successful dissemination but also transparency for possible clients, interested consumes and other stakeholders with restricted access to other scientific journals.

7.3 Stakeholder dialog on the acceptance of new breeding technologies

7.3.1 Identification and characterization of stakeholders

Prior to a successful dialog on the acceptance of new breeding technologies such as the induced amplification of retroTEs, it was important to identify and correlate the most important stakeholders linked to the topic. Therefore, I created a stakeholder-map (Fig. 31).

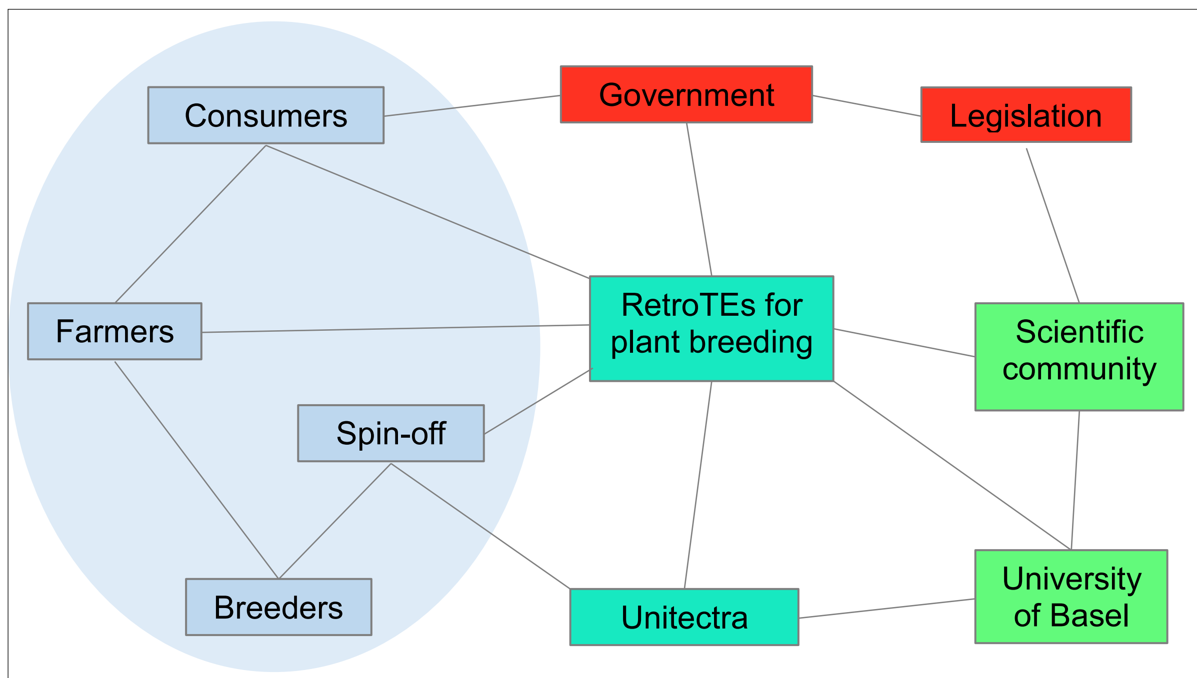


Figure 31 **Stakeholder map depicting the most important stakeholders and their relation in the context of the current stage of my research project “retroTEs for plant breeding”.** Individual stakeholders were arranged according to their role in the stakeholder dialog to be related to economy (blue), government/legislation (red), science (green) or to be situated at the interface of science and policy (turquoise).

In a next step, I attempted to characterize the identified stakeholders concerning their role, interests, contributions and power relative to my own position in the stakeholder dialog regarding the acceptance of our new breeding approach (table 3) according to published tools (http://wageningenportals.nl and according to the course script “Stakeholder Engagement”, offered by the Zürich-Basel Plant Science Center).

Table 3 **Characterization of stakeholders (SHs) concerning their general role, interests, contributions and power in the stakeholder dialog regarding the acceptance of inducing retroTEs for plant breeding.**

SH	Role	Interests	Contributions	Power
Government/EU	Funder Beneficiary Regulator	Clear and evidence based regulations of new breeding technologies; Follow the public opinion to survive next elections; Ensure economic growth to keep people satisfied; Fund research to develop new breeding tech.	Provide funding for research; Employ experts; Organize public events; Initiate a referendum; Decide about regulations of labelling; Fund platforms;	High
Scientific Community	Partner Beneficiary Knowledge provider Influencer Informer	Raise funding; Stay up to date in latest research; Collaborate to get data published;	Provide information; Increase pressure to generate new data; “Market” for new publications; Add a “value” to and review generated data;	High
University of Basel	Regulator Beneficiary	Train the next generation of researchers; Raise funding;Climb up in ranking;	Provide facilities; Infrastructure; Salary etc.	High
Unitectra	Beneficiary	Improve portfolio of University; Convert data of employees into revenue;	Financial support and expertise for patent application;	Medium
Breeding companies	Contractor	Generate innovate cultivars; Profit from progress in research; Be the first to have innovative product; Partially: search for ways to increase farmers/consumers dependency and to become more efficient and compatitive in breeding;	Early adopters, decide if new technology will be implemented; Form the market for innovation and patents; Apply and develop new breeding technologies further; Push government for regulation to obtain level playing field in the international context;	High
Consumers	Beneficiary	Obtain healthy products that are produced devoid of ethical concerns; Food security;	Push government towards desired regulations; Decide which products to buy (if labelled)	Medium
Economy / Spin-off	Beneficiary Partner	Convert research data into money;	Enhance research by adding a monetary value to research data;	Medium

SH	Role	Interests	Contributions	Power
Farmers	Beneficiary	Take choice which cultivars they want to grow, of those offered by seed companies; Profit from progress in research; Have stable yield/income; Retain independence;	Build the market for breeders,	Low
NGOs	Interest group	Concerned about safety issues;	Initiate and participate in public discussions, consumer awareness, policy recommendations;	Medium

The awareness of various stakeholders and their characteristics was very helpful for the further planning and organization of the stakeholder dialog.

7.3.2 Fachtagung Dialog Grün 2016

The symposium "Fachtagung Dialog Grün 2016, Neue Verfahren in der Pflanzenforschung – eine Alternative zu Pflanzenschutzmitteln?" was launched to discuss with experts from science, economy, politics and administration whether current developments in plant research could contribute to a reduction in pesticide use in Switzerland. The aim of my presentation (20 min) was to explain our scientific findings in lay langue (german). This was a first attempt to disseminate and discuss our new breeding approach with various stakeholders from both the conventional and organic sectors. My presentation was further published in a publically available conference transcript (Thieme 2017a) (200 printed copies, appendix IV) that was posted on the PSC-website, on the platform "naturwissenschaften.ch" and included into the Research Collection of the ETH Zürich.

7.3.3 Presentation and World-Café at FiBL

7.3.3.1 Aims an planning

The target audience was composed of different stakeholders including breeders, researchers, consumers and employees (FiBL) working in the organic sector (Fig. 31).

By using the format of a World-Café at the FiBL in Frick, CH we aimed to:

(I) Disseminate and explain the scientific background of using retroTEs in plant breeding to generate a proper basis for discussion;

(II) Collect and discuss critical aspects linked to the possible application of this innovative breeding approach in the organic sector.

For (I) the event was initiated with a 30 min presentation (german, appendix VIII, pictures were removed) in lay language easy to understand for an audience without specific knowledge in (epi)genetics or breeding followed by a Q&A-session (30 min).

Subsequently, the degree of understanding of the scientific background in the audience was assessed in a multiple choice test (appendix V). To achieve a maximum learning effect, the correct answers were handed out prior to the actual discussions in the World-Café. Questions with multiple possible answers covering the following parts of the presentation were posed:

- 1: Mechanism of retroTE movement.
- 2: Effects of stresses on retroTEs and flanking regions.
- 3: Effect of DNA-methylation on the activity of retroTEs.
- 4: Properties of the two inhibitors A&Z.
- 5: RetroTE-induced genetic diversity as the basis for breeding.
- 6: Genetic and trait stability of high-copy lines.
- 7: Applicability of the inhibitors A&Z in crops.

The World-Café was planned and carried out according to the open source guidelines “Café to go” offered by (http://theworldcafe.com). The audience was split into three small groups (size three to four people) and allocated to tables where three different topics were discussed (Fig. 32).

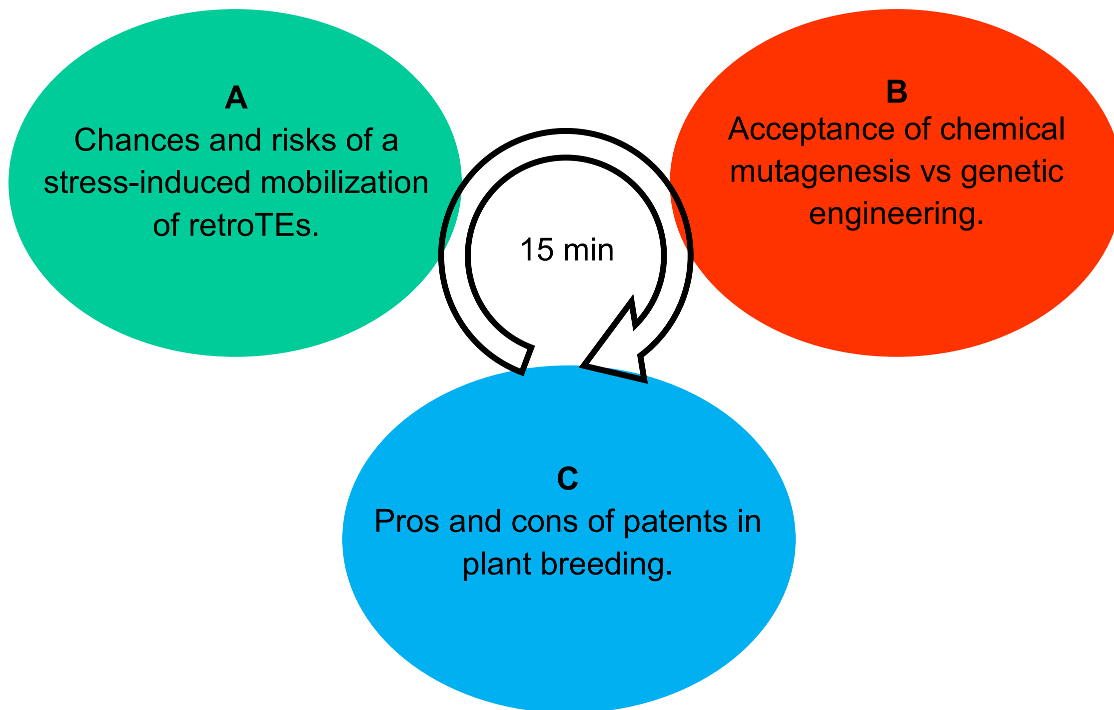


Figure 32 **Topics (A, B and C) discussed in the World-Café**. After a discussion for 15 min that was facilitated by a table host who remained on the table, groups of three to four persons moved to the next table.

After 15 min of discussion, groups moved to the next table. The discussions were facilitated by table hosts (Dr. Messmer, Dr. Bucher and me) that were leading the discussions according to a set of guiding questions that were prepared in advance (german, appendix VII). Following three rounds of discussion, a synopsis from each table was presented by the table hosts in the plenum (3 x 7min).

By choosing the format of a World-Café we intended to collect various issues linked to the use of retroTEs in organic agriculture. In addition to the discussion on the three tables we also designed a questionnaire (german, appendix VI) that would reflect the diversity of opinions in the audience. The following questions (translated from german) were asked:

- Question 1: Should varieties that were bred using retroTEs be used in the organic sector? (*Rate 1-5; 1: not at all, 5 yes, necessarily*).
- Question 2: What are potential risks linked to the use of retroTEs in plant breeding? (*Open field*).
- Question 3: Should natural substances that activate retroTEs be used for plant

breeding in the organic sector? (*Rate 1-5; 1: not at all, 5 yes, necessarily*).

- Question 4: How high is the potential of using retroTEs in plant breeding for the organic sector? (*Rate 1-3; 1: not any, 3: high*).
- Question 5: Process patents are the basis for innovation in agriculture (*Rate 1-5; 1 fully disagree, 5: fully agree*).

7.3.3.2 Results and discussion

7.3.3.2.1 Multiple choice test to assess the understanding of scientific background

A prerequisite for a fruitful dialogue and discussion about a novel technology is a successful dissemination of the scientific background and a reasonable knowledge transfer to the audience or the public. By using a multiple choice test, we wanted to generally assess to which extend the audience was able to understand the scientific backgrounds of using the two inhibitors A&Z to induce retrotransposition for plant breeding (aim (I), see above).

Overall, 78% of given answers were correct. However, we observed major differences between the tested subject areas. Questions addressing the influence of DNA-methylation on transposition (question 3) and concerning a potential use of A&Z to induce transposition in crops (question 7) were answered mostly correct (92%) while difficulties were observed concerning genetic and trait stability of high-copy lines (61% correct answers) (Fig. 33).

As the correct answers to the multiple choice test were handed out and if required also debated during the World-Café it can be assumed that further discussions were based on a very high degree of understanding of the scientific backgrounds. Despite the limited number of twelve participants, this result could be used to adapt future communication strategies about this novel breeding technology. Accompanying multiple choice tests could be used to further validate successful communication strategies.

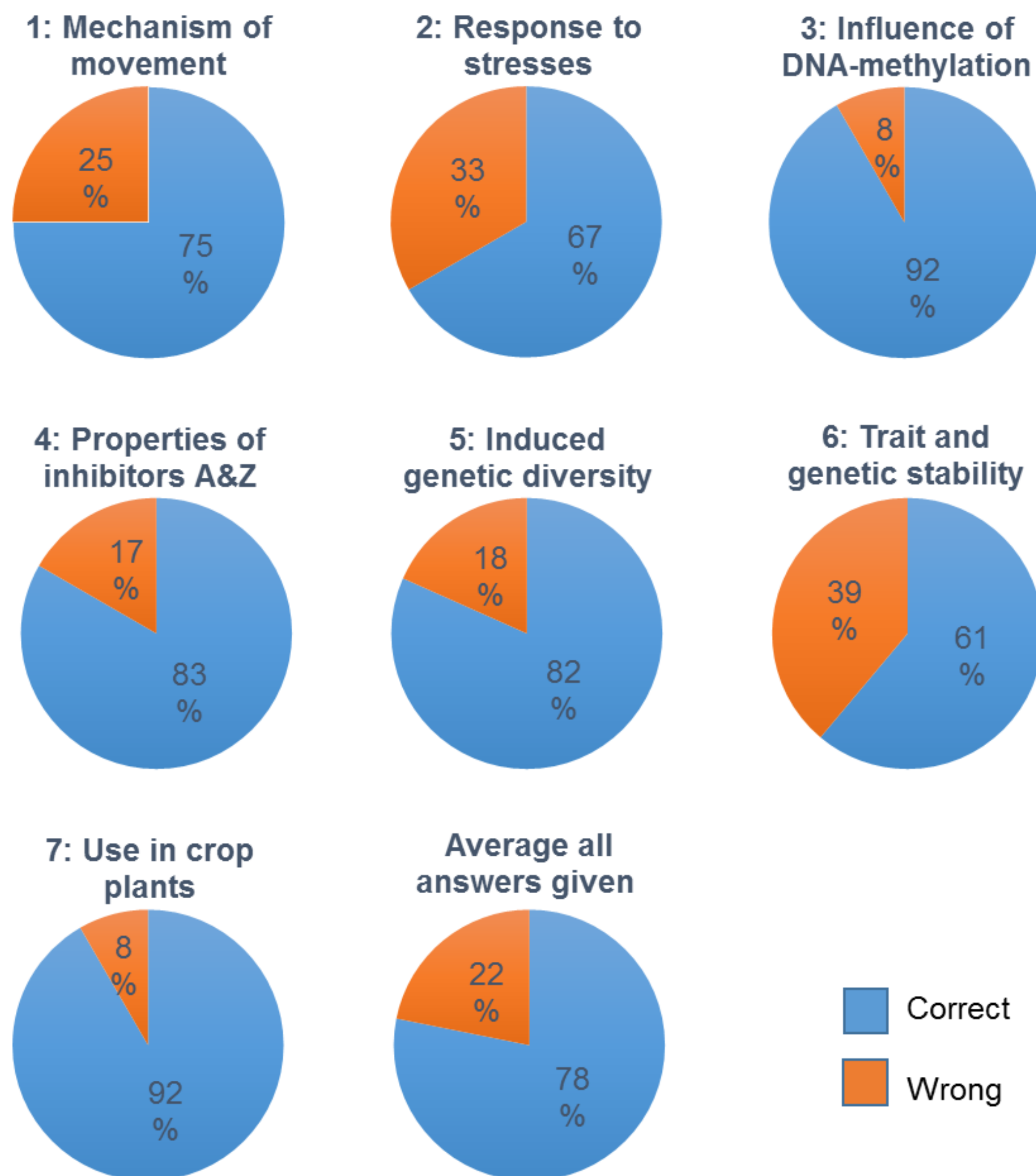


Figure 33 **Percentage of correct (blue) and wrong (orange) answers to individual questions (1-7) and on average in the multiple choice test (german, appendix V).** The summary of answers given by twelve participants is shown. The total number of given answers in the multiple choice test was n=238.

7.3.3.3 Synopsis and key findings from the World-Café and open questions from the questionnaire

Based on the overall successful communication of scientific results (78% of correct answers in the audience, Fig. 33) we were able to proceed with an open discussion aiming to collect a broad range of aspects and ideas linked to the use of retroTEs in plant breeding (aim (II), see above). In doing so, we were explicitly not aiming to agree on a common position concerning the use of retroTEs in organic agriculture.

When discussing the different topics (A, B and C, Fig. 32) in the plenum, it turned out that many aspects discussed at the individual tables were overlapping. Hence, for the validation of the World-Café and the questionnaire we choose to sort the emerging arguments to eight subcategories. A list of discussed ideas/topics and issues can be summarized as follows:

Risks & limits

Crop/ Economic

- Induction of side and off target effects (mutations and altered gene regulation, less resistance to other stresses);
- Forcing plants to release retroTEs could generate a too high diversity;
- Use of A&Z might lead to uncontrollable effects (gene-transcription, transpositions and in general);
- Interference with self-regulation of the plant- plant might experience a shock;
- Selection of new traits might be hampered due to too many mutations;
- High-copy lines might behave differently under other/combined stresses;
- Crossing might still be needed in very poor genetic backgrounds (no emergence of novel traits out of nothing);
- Method might not be accepted by breeders and consumers in the organic sector (see below);
- Too early to assess benefits and risks and thus to predict the acceptance;
- Regulation of breeding technologies is likely to change, mutation breeding could also be considered to be genetic engineering (see discussion below);

Environment

- Regulatory balance and self-regulation of the plant is affected;
- Long-term effects and altered performance in the field (multiple stresses) hard to determine;
- Outcrossing meaning crossing of high-copy lines with wild relatives possible;
- Accelerated evolution of super-weeds possible;

Health

- Zebularine (synthetic nucleotide) is incorporated into the DNA and could be inherited to the next generation, accumulate in food chain and become active at very low concentrations;
- Handling and degradation of the toxic inhibitors A&Z;
- Induction of random gain of function mutations could lead to allergies;
- Impairing the general "order and regulating entity of an organism" and extreme acceleration of evolution could affect food quality;

Social and ethical

- Patents on life are not accepted in the organic sector while high-copy lines and retroTEs in novel varieties might be easy to patent;
- Process patents are partially accepted, however under discussion (Fig. 34);
- Restriction of freedom of choice for farmers and consumers;
- Data/knowledge generated with public resources leads to profit in the private sector;

Chances

- Massive increase in (epi)genetic diversity within one generation without crossing;
- So far “hidden phenotypes/characteristics” might become more obvious through the amplification of class I elements;
- Method might be directed, evolution according to Lamarck`s theory;
- Stimulation and amplification of inherent characteristics like “educating” a plant;
- Gain-of-function phenotypes possible and likely;
- Very fast evolution (already next generation has homozygous insertions);
- No need for genetic sequence information (which is a prerequisite for e.g. TALEN or CRIPSR-Cas9);
- No need for genetic engineering/transformation of the crop:

Acceptance

- Product would generally not be a GMO and should rather be regulated like products obtained from EMS /irradiation mutagenesis or colchicine-treatment;
- Incorporation of zebularine into DNA could still be considered as “genetic engineering”;
- Stimulation and amplification of inherent characteristics is a popular idea in the organic sector;
- The organic sector would prefer to use breeding material that was not treated with colchicine, EMS or irradiation but there are currently few alternatives;
- New varieties would need to perform much better to gain acceptance in organic sector;
- Breeding should be done by farmers not by chemists;
- Whether a compound used for the treatments is synthetic or natural does change much for the acceptance;
- Self-regulation is generally still possible when seed/plant is treated (not *in vitro*) which could positively influence acceptance in the organic sector;
- The concept of “back to the roots” could limit acceptance of a new breeding approach in the organic sector;

- Food as "a shelter" from global changes and a transmitter of a certain "order" should not be impaired;
- "1000 years of evolution within only one year" could be frightening to consumers
- Whenever possible to use, other new techniques such as CRISPR-Cas9 might be more precise and controllable;

Recommendations

- More research/safety assessment before product is grown in the field;
- More funding for research for example to find alternatives to the use of zebularine;
- The ultimate goal for using retroTEs in the organic sector would be to develop a method that does not use A&Z;

As mentioned before, we also determined the acceptance in the audience by using an questionnaire (german, appendix VI) that allowed for a rating of four different positions or statements (see above). Although also limited by the small number of participants (n=12), the evaluation clearly reflected the full range of different opinions in the audience (Fig. 34). This was in line with the high diversity of opinions and ideas that came up during the discussion in the World-Café. Overall, we observed a trend towards a moderate rating of all four questions asked (Fig. 34).

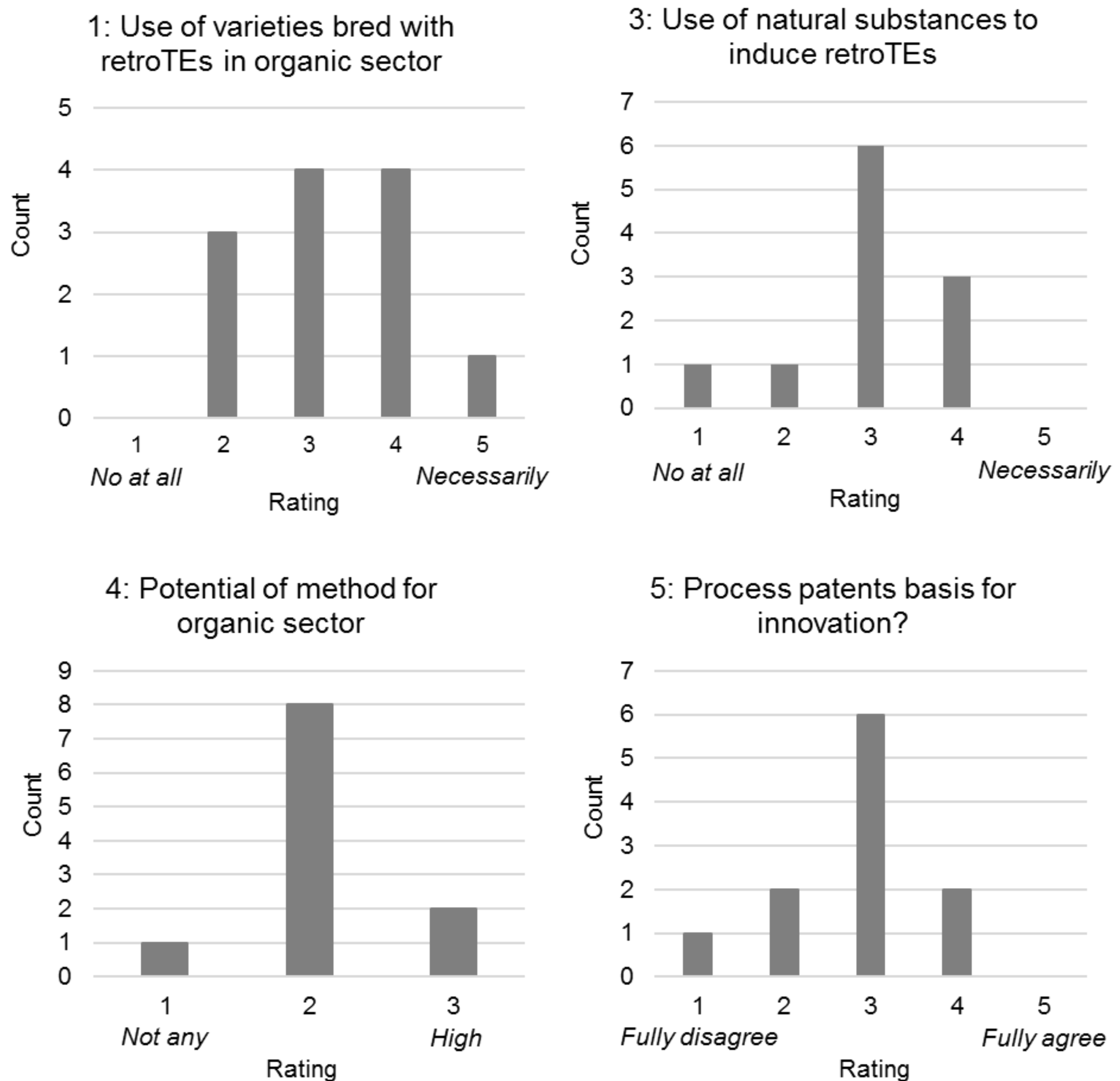


Figure 34 **Distribution of frequency of ratings for individual questions in the questionnaire (german, appendix VI).** The summary of answers given by eleven participants is shown.

7.3.3.4 Summary and outlook

In summary, the two objectives to disseminate and explain the scientific background and to collect and discuss critical aspects linked to the possible application of retroTEs in plant breeding for the organic sector were met. Although these findings are based on a small sample size and therefore not generally valid, they could still be used as a first indicator about major issues arising with the development and possible application of this new breeding approach in the organic sector. As a consequence, insights gained by using the World-Café tool with different stakeholders from the organic sectors have the potential to guide and shape futures discussion and developments not only in the organic sector. Thanks to the early stage of development of this novel breeding approach, important necessary adjustments both in the context of directions of research but also with regard to communication or marketing strategies can still be made. One important issue that should be addressed in futures experiments is the assessment of the biosafety of plants that have been bred using mobilized endogenous retroTEs. A potential emergence of allergens in products as well as unforeseen side effects including an interference with gene expression and the effects of outcrossing should be tracked and evaluated over several generations. These findings will in turn have major implications for the acceptance and regulation of this newly developed breeding approach. Indeed, current EU-level discussions on the regulation of breeding approaches including mutation breeding (EuGH 2017) could be indicative for futures amendments of legislation. The potential obligation for a declaration of products obtained from plants that were bred using induced retrotransposition will heavily influence the commercialization of the method developed in this work.

8 General discussion and outlook

Currently, especially breeders in the organic sector are facing a challenging situation. They have to develop new crop varieties with the potential to build the basis for food security of a growing global population under rapidly changing environmental conditions (Fischer *et al.* 2014). Besides knowledge gained through recent significant advances in genomics (Abberton *et al.* 2016) especially new insights into the field of epigenetics and epigenetic memory has become a focus of attention of plant breeders (Gallusci *et al.* 2017). Situated at the interface of genetics and epigenetics, also the awareness of retroTEs as a valuable genetic resource that could be used to cope with future challenges in agriculture has increased significantly (Paszkowski 2015). However, as described above, it was so far not possible to access this "(epi)genetic treasure" (Mirouze and Vitte 2014) without using genetic engineering. Key to harnessing these elements for plant breeding is to gain further knowledge about the mechanisms of their regulation.

Although it was presumed that Pol II could play a central role in repressing retroTE-transcription in plants (Zheng *et al.* 2009) the observation that inhibition of Pol II with α-amanitin led to an increased accumulation of extrachromosomal DNA is remarkable. As most class I elements depend on the host Pol II for their transcription, the detected accumulation of retroTE-DNA after the A-treatment and in the background of the hypomorphic mutation of Pol II was counterintuitive. This finding provides striking evidence that this highly conserved and vital enzyme also holds a mechanism to actively repress retrotransposition in plants. From an evolutionary point of view this may not come as a surprise. Although plants have evolved the two additional RNA-polymerases IV & V to ensure genome stability (Matzke *et al.* 2015) other organisms including humans still get along without these two Pol II-related RNA-polymerases. RetroTEs are known to be of central importance for stress-tolerance, adaptation and evolution of plants (Makarevitch *et al.* 2015; Quadrana *et al.* 2016; Zhang and Gao 2017). As sessile organisms, plants indeed have to face and tolerate a multitude of (a)biotic stresses. Considering this, it can be speculated whether the evolution of the two plant-specific RNA-polymerases Pol IV and V as "fine-tuners" and later key-regulators of retroTE activity was a prerequisite for the optimal utilization of these potential hazardous genetic elements to cope with various suboptimal growth conditions. Indeed, the current understanding of TE-silencing in plants is likely to be

biased by the lethality of Pol II-knockout mutants (Cuerda-Gil and Slotkin 2016; Zheng *et al.* 2009). The approach of using the highly specific (Haag *et al.* 2012) inhibitor α-amanitin and, in case of well-established model-organisms such as *Arabidopsis*, also new gene-editing techniques will certainly advance the understanding of the exact mechanisms underlying the dual-role of Pol II in controlling class I elements. On the enzymatic-level, especially the C-terminal domain (CTD) of the large subunit 1 of Pol II (NRPB1) is in the focus of attention as a key-player of epigenetic regulation (Eick and Geyer 2013; Matzke *et al.* 2015; Palancade and Bensaude 2003). Notably, it was found that the Pol II-CTD is crucial to repress transcriptional activation and mobility of *Ty1* elements in *S. cerevisiae* (Aristizabal *et al.* 2015). Further, it is known that iterating GW/WG-motifs in the CTD of Pol V mediate the interaction of Pol V and AGO4 in RdDM (El-Shami *et al.* 2007; Li *et al.* 2006). Interestingly, Zheng and colleagues also found evidence for the presence of four GW/WG-motifs in the second largest subunit of Pol II (Zheng *et al.* 2009). Presumably, this forms the basis of the interaction of Pol II and AGO4 that was observed despite the structural differences of the CTDs of Pol II and V (Matzke *et al.* 2015; Zheng *et al.* 2009). Hence, this clearly indicates that further functional studies aiming to investigate the role of Pol II as an epigenetic regulator should comprise all elements of this multisubunit enzyme (Ream *et al.* 2009).

A so far neglected mechanism of retroTE-regulation in plants is the Pol II-depended production of antisense-transcripts originating from bi-directional transcriptional start sites located in both LTRs of class I elements, as previously reported from *Drosophila* (Russo *et al.* 2016). Indeed, in a preliminary experiment we were able to detect antisense transcripts originating from the *ONSEN*-TE in *Arabidopsis*. The potential ability of Pol II to produce antisense-transcripts originating from the TSS in the 3`LTR of the retroTEs could be the yet unknown fundamental step for the initiation of retroTE-silencing in plants. Given that functional *copia* and *gypsy*-elements depend on the complete identity of their two LTRs including their TSSs (Schulman 2013), a potential regulation via sense/antisense pairing of TE-transcripts would be a highly specific and robust mechanism to specifically control transcriptionally competent LTR-retroTEs. To what extend also sub-genomic (anti)sense transcription influences TE-expression and genome evolution currently remains elusive. Based on our findings and considering recent advances in the field (Cuerda-Gil and Slotkin 2016; Panda *et al.* 2016) we updated the current model of the expression-dependent retroTE-regulation by

substantiating the role of the RDR-independent route of siRNA production via the presumed dicer-dependent processing of paired sense and antisense retroTE-transcripts (Fig. 35).

In line with previous studies in *ddm1* (Panda *et al.* 2016) we also conclude that the expressional state of a retroTE predetermines downstream silencing pathways in plants. According to our updated model, the Z-induced reduction of DNA-methylation shifts retroTEs into a “transcriptional active state” resulting in an increased Pol II-dependent production of sense and antisense transcripts that would normally antagonize retroTE-activity via the DCL3-dependent production of siRNAs. Based on our preliminary detection of *ONSEN*-antisense transcripts, we conclude that the inhibition of Pol II with A stabilizes the “active state” by generally reducing the levels of antisense transcripts and consequently depleting the primary siRNAs that would normally be massively amplified in the RNAi-loop during PTGS (Cuerda-Gil and Slotkin 2016). As a consequence of this concerted action of A and Z, we observed the massive accumulation of ecDNA and in case of *ONSEN* the insertion of novel retroTE-copies into the genome of treated plants (Fig. 35).

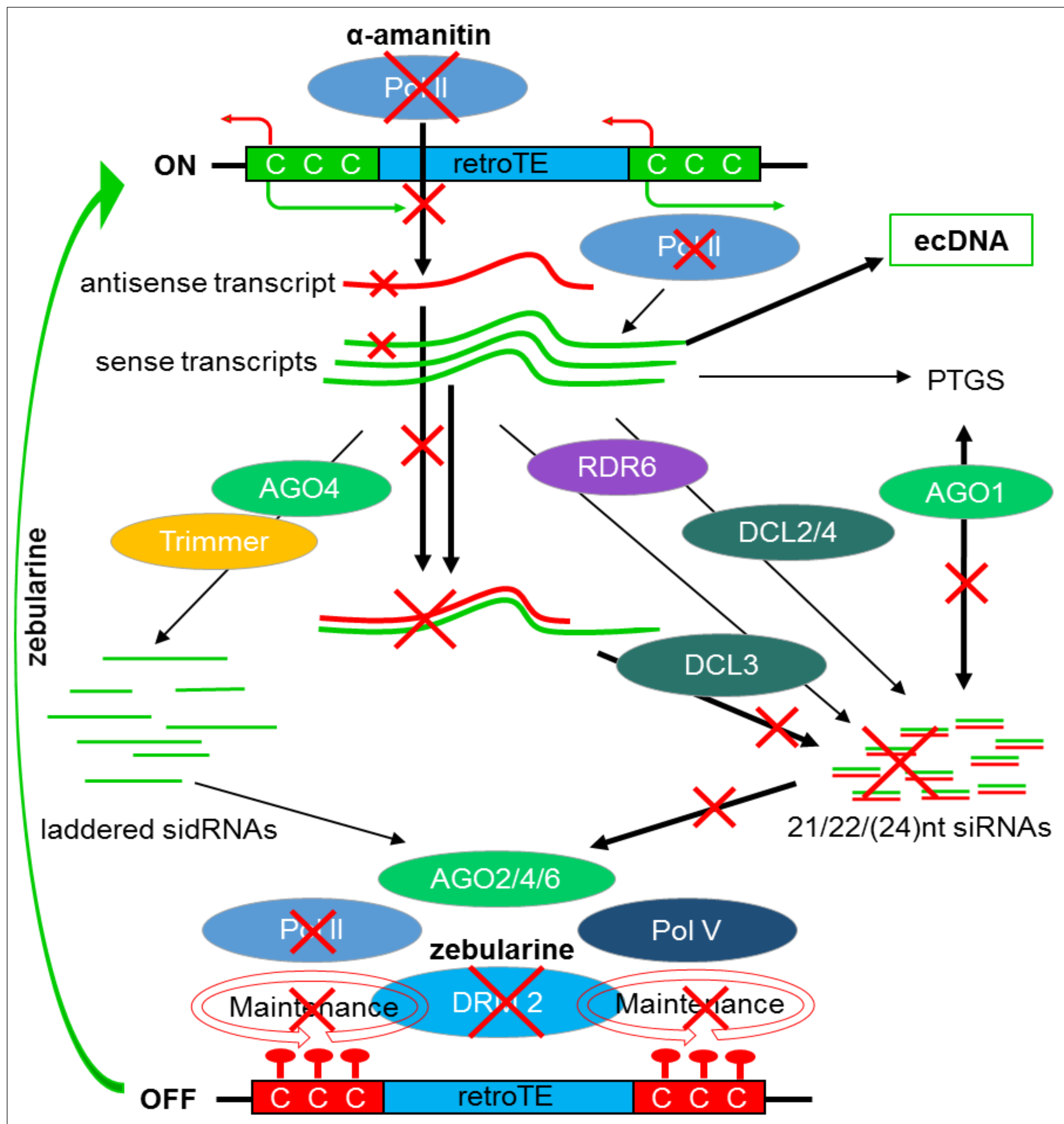


Figure 35 **Updated model for the presumed mechanism of the "non-canonical" expression-dependent RdDM leading to retroTE silencing in plants.** The initial silencing of retroTE is presumably triggered by the Pol II-dependent production of antisense-transcripts. As previously suggested by (Panda *et al.* 2016) paired sense-antisense transcripts can be processed in a RDR6-independent manner resulting in the DCL3-dependent production of specific siRNAs. Inhibition of DNMtases (DRM2) with Z results in a loss of DNA-methylation of retroTE converting them into a transcriptional active (ON) state. As previously reported (Panda *et al.* 2016) this would lead to an increased (bold arrows) DCL3-dependent production of primary siRNAs. The additional general inhibition of Pol II with A reduces both sense and antisense-transcription, leading to less sense-antisense pairing of TE-transcripts, a reduced accumulation of primary siRNAs, a drop in PTGS and an increased production of retroTE-ecDNA. The specific targets of A and Z and resulting impacts on retroTE-silencing are marked with red crosses.

As described above, our fundamental findings from the model plant *Arabidopsis* that are supported by the activation of the *Houba* retroTE in rice open up new avenues to enhance plant breeding to face futures agronomic challenges. The planned testing of (heat) stress tolerance and a large-scale phenotyping of hc-lines generated in this study will be the next step of elucidating the evolutionary potential of class I elements in plants. Although the correlation of *ONSEN* copy number in various accessions of *Arabidopsis* was recently reported to be linked to the annual climate range of their natural habitat (Quadrana *et al.* 2016) there is to our knowledge so far no direct evidence showing an obvious link of copy number variations and adaptation to different environmental conditions in plants. Notably, also the observed gain of heat-responsiveness of genes flanking a novel *ONSEN* insertion in the F1 of heat stressed *Arabidopsis* plants was so far only observed in the *nrpd1*-mutant background (Ito *et al.* 2011). These plants are defective in siRNA-production and thus also impeded in potential spreading of silencing to flanking regions (Sigman and Slotkin 2016). Hence, analogous experiments addressing the influence of novel TE-insertions on gene regulatory networks also considering potential alterations in chromatin marks triggered by the canonical RdDM (Holoch and Moazed 2015; Matzke and Mosher 2014) should now be repeated in the WT-background.

Besides the detected induced environment-dependent phenotypic diversity of hc-lines, especially the observed high degree of their genetic stability upon heat-stress can be discussed in a more general context of evolution. Considering that each novel *ONSEN* copy is in principal capable to again generate additional copies of itself, it could be concluded that there is a stable balance between TE-copy number and TE-activity that restricts the uncontrolled proliferation of retroTEs in plants (Mari-Ordonez *et al.* 2013). This presumed “self-regulation” of retroTE-activity is a strong indicator that the former characterization of retroTEs as parasitic or selfish elements was a common misconception. Indeed, the inversion of argumentation coming to the conclusion that TE-copy numbers in eukaryotic genomes increase because of and not despite of epigenetic silencing has already been suggested previously (Fedoroff 2012). The above discussed presumed coevolution of the symmetric structure of LTR class I elements that could allow for the Pol II-mediated simultaneous production of sense and antisense transcripts further substantiates this notion.

In regard of using retroTEs for plant breeding it would be very interesting to test whether detected insertion biases of various retroTE-families (Bennetzen and Wang 2014) can be reproduced in real time by using the inhibitors A and Z. Possible insertion site preferences could have relevant implications for the actual use of retroTEs in the breeding processes. In the context of targeted insertions of retroTEs, the approach of using CRISPR-Cas9 to introduce double-strand breaks that have in yeast been shown to attract retroTEs (Moore and Haber 1996) could prove promising. Likewise, the question whether previously reported potential Z-induced DNA-damages (Liu *et al.* 2015) could facilitate or influence retroTE-insertions should be addressed in further experiments. Taking such potential side effects of the treatment itself into account, it could in principle also be highly interesting to screen a high number of hc-lines for preferential retroTE insertion sites. In doing so, it could be assessed whether the growth condition during or subsequent to an induced ecDNA-production in the parent influences the insertion site-preference of retroTEs. As discussed above, the *Houba* retroTE in rice that showed a similar induction of ecDNA production under control and heat-stress conditions (*preliminary data*) following the A&Z-treatment, could be a promising model.

Assuming that the frequency of the integration of ecDNA correlates with the actual amount of ecDNA produced, it could, based on the observed dose-response of the *ONSEN* ecDNA-production generally be possible to control the rate of transposition by adjusting the concentrations of inhibitors used. From the observation that concentrations that have been demonstrated to be effective in rice and *Arabidopsis* did not lead to the activation of retroTEs in soybean, it can further be concluded that inhibitor concentrations need to be optimized case by case. In addition, the treatment could also limit the re-silencing of active retroTEs or attenuate so far unknown TE-defense mechanisms downstream of the production of ecDNA. Hence, also the duration of the treatment following the stress application could heavily influence transposition rates. Besides species and stress-dependent necessary optimizations of the protocol, especially the potential induction of class I elements to generate (epi)genetic diversity in vegetatively propagated crops would also require adjustments in the application of the inhibitors.

As described above and illustrated in Fig. 35, the strong activation of retroTEs is based on a massive decrease of global DNA-methylation in all three sequence contexts following the double-treatment. Stable inheritance of changed DNA methylation patterns as previously observed in a population of so called epigenetic recombinant inbred lines (epiRILs) generated from a cross between the WT and a homozygous *met1*-mutant (Reinders *et al.* 2009) can contribute to phenotypic diversity (Latzel *et al.* 2013). The stable inheritance of epialleles could therefore provide valuable contributions for plant breeding (Gallusci *et al.* 2017; Springer and Schmitz 2017). We have also observed a nearly complete loss of asymmetric DNA-methylation at the tested DCL3-RdDM locus following the A&Z-treatment in the parental generation of *Arabidopsis*. Hence, it would be of great interest to test whether changed methylation patterns could be transferred to successive generations resulting in the emergence of drug-induced stable epialleles in the offspring of treated plants. In the context of breeding it would in a next step also be exciting to test whether hybrid vigour as previously reported to occur in the F1 of a cross between the WT and individual lines from the aforementioned epiRIL population (Dapp *et al.* 2015) could also be induced by crossing drug-treated with untreated WT-plants. In this context, it should also be tested to what extend a potential heterosis effect in the F1 of a cross between the WT and an hc-line could be explained by the difference in retroTE-copy numbers in the genomes of the parents. Further information gained through whole genome-sequencing of selfed or crossed high-copy lines will also provide insight into the influence of an induced burst of retroTEs in on the genome architecture and evolution of WT plants in real time.

In summary, the fundamental findings suggesting the Pol II-dependent production of retroTE-antisense transcripts that lead to the dicer-dependent formation of primary siRNAs significantly contribute to the solution of the “chicken and the egg problem” of TE-silencing (Nuthikattu *et al.* 2013; Thieme *et al.* 2017). This opens up completely new possibilities for basic research and to actively harness retroTEs for plant breeding (Fig. 36).

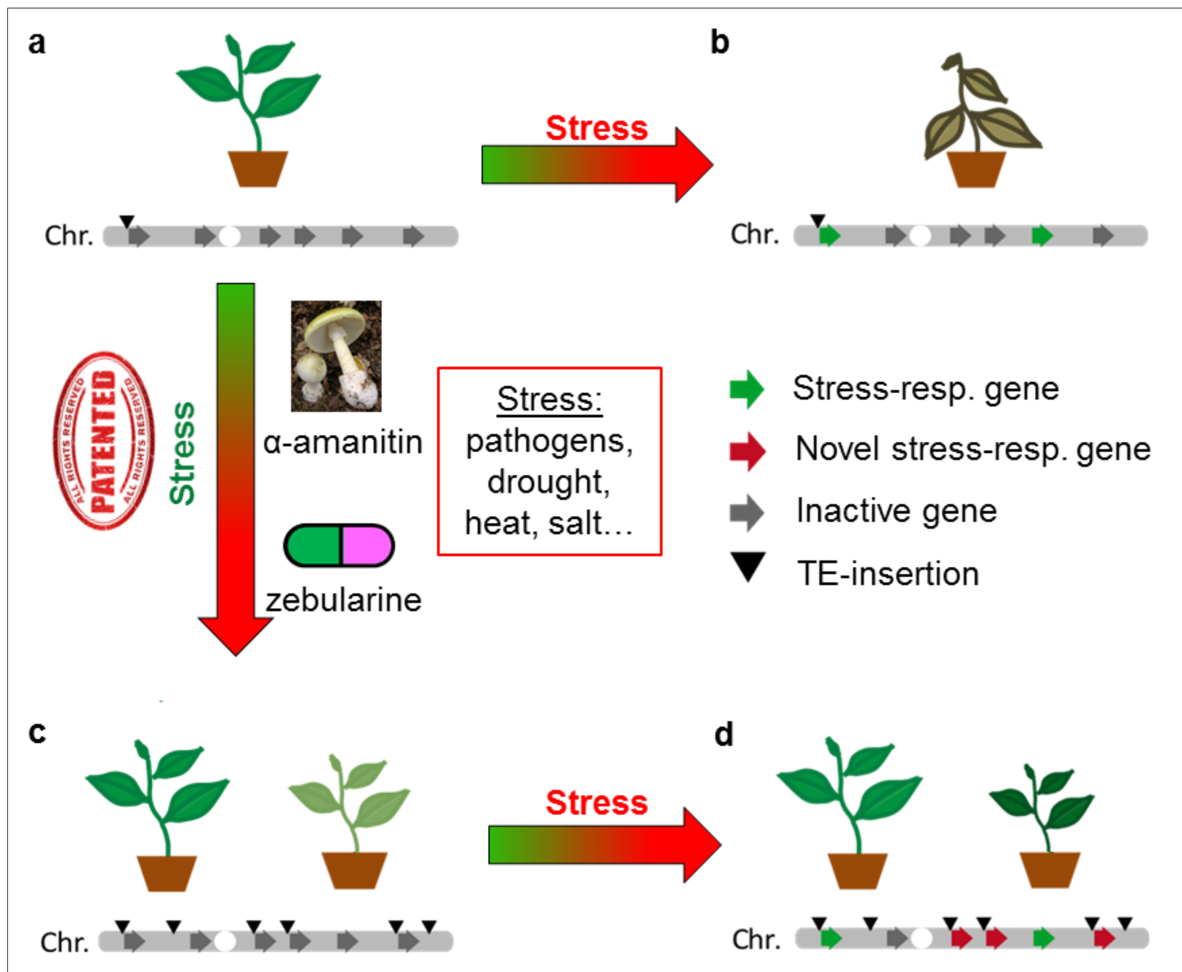


Figure 36 **Stress-dependent induced amplification of retroTEs for plant breeding.** The occurrence of a stress (biotic or abiotic) leads to retroTE-mediated gene-activation in the original cultivar (**a** and **b**). However, due to a limited number of activated genes needed for stress tolerance the plant is heavily affected by the stress (**b**). The patented stress-induced activation of stress-responsive retroTEs using the two inhibitors α-amanitin and zebularine leads to the stable insertion of novel retroTE-copies into the genome, resulting in increased (epi)genetic diversity and a broad panel of phenotypes in the offspring (**c** and **d**). The occurrence of the same stress can lead to the retroTE-mediated stress-dependent activation of novel genes. This can result in heritable stress-inducible phenotypes and/or increased stress-tolerance of hc-lines (**d**).

Importantly, it should be noted that due to high degree of conservation of targeted mechanisms, these findings from *Arabidopsis* also have far-reaching implications for other eukaryotes including human cells (DeNicola *et al.* 2015). Indeed, we have preliminary evidence (appendix II, *unpublished data*) suggesting that the chemical inhibition of Pol II and DNMtases leads to a very strong de-methylation of LINE-1 retroTEs in human cells. DNA-methylation and activity of LINE1 elements is known to play a central role in various developmental processes including neuronal development (Erwin *et al.* 2014). In addition, the dysregulation and transposition of LINE1-elemets which can also lead the transduction of non-repetitive DNA fragments (Tubio *et al.* 2014) has been linked to the emergence of cancer (Xiao-Jie *et al.* 2016).

Considering the aforementioned versatile possible directions of research and applications entailed by this work it will be very important to continuously take measures at the science and policy interface. This will form the basis for an objective and less biased social debate on using endogenous retroTEs to cope with imminent challenges in agriculture.

9 References

Abberton, M., Batley, J., Bentley, A., Bryant, J., Cai, H., Cockram, J., et al. (2016) Global agricultural intensification during climate change: a role for genomics. *Plant Biotechnol J* 14: 1095-1098.

Anca, I.A., Fromentin, J., Bui, Q.T., Mhiri, C., Grandbastien, M.A. and Simon-Plas, F. (2014) Different tobacco retrotransposons are specifically modulated by the elicitor cryptogein and reactive oxygen species. *J Plant Physiol* 171: 1533-1540.

Appelhagen, I., Thiedig, K., Nordholt, N., Schmidt, N., Huep, G., Sagasser, M., et al. (2014) Update on transparent testa mutants from Arabidopsis thaliana: characterisation of new alleles from an isogenic collection. *Planta* 240: 955-970.

Aristizabal, M.J., Negri, G.L. and Kobor, M.S. (2015) The RNAPII-CTD Maintains Genome Integrity through Inhibition of Retrotransposon Gene Expression and Transposition. *PLoS Genet* 11: e1005608.

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.

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

Baucom, R.S., Estill, J.C., Chaparro, C., Upshaw, N., Jogi, A., Deragon, J.M., et al. (2009) Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. *PLoS Genet* 5: e1000732.

Belyayev, A. (2014) Bursts of transposable elements as an evolutionary driving force. *J Evol Biol* 27: 2573-2584.

Bennett, M.D. (1972) Nuclear DNA Content and Minimum Generation Time in Herbaceous Plants. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 181: 109-135.

Bennetzen, J.L. and Kellogg, E.A. (1997) Do Plants Have a One-Way Ticket to Genomic Obesity? *The Plant Cell* 9: 1509-1514.

Bennetzen, J.L., Ma, J. and Devos, K.M. (2005) Mechanisms of recent genome size variation in flowering plants. *Annals of botany* 95: 127-132.

Bennetzen, J.L. and Wang, H. (2014) The Contributions of Transposable Elements to the Structure, Function, and Evolution of Plant Genomes. *Annual Review of Plant Biology* 65: 505-530.

Bernatavichute, Y.V., Zhang, X., Cokus, S., Pellegrini, M. and Jacobsen, S.E. (2008) Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. *PLoS One* 3: e3156.

Berretta, J., Pinskaya, M. and Morillon, A. (2008) A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. *Genes Dev* 22: 615-626.

Böhmdorfer, G., Rowley, M.J., Kuciński, J., Zhu, Y., Amies, I. and Wierzbicki, A.T. (2014) RNA-directed DNA methylation requires stepwise binding of silencing factors to long non-coding RNA. *The Plant Journal* 79: 181-191.

Bucher*, E. and Thieme*, M. (2017) Mobilization of Transposable Elements to Enhance Genetic and Epigenetic Variability in a Population. *Patent* WO2017/093317A1.

Buisine, N., Quesneville, H. and Colot, V. (2008) Improved detection and annotation of transposable elements in sequenced genomes using multiple reference sequence sets. *Genomics* 91: 467-475.

Burns, K.H. (2017) Transposable elements in cancer. *Nat Rev Cancer* 17: 415-424.

Butelli, E., Licciardello, C., Zhang, Y., Liu, J., Mackay, S., Bailey, P., et al. (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* 24: 1242-1255.

Cao, X. and Jacobsen, S.E. (2002) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A* 99 Suppl 4: 16491-16498.

Casacuberta, J.M. and Grandbastien, M.A. (1993) Characterisation of LTR sequences involved in the protoplast specific expression of the tobacco Tnt1 retrotransposon. *Nucleic Acids Research* 21: 2087-2093.

Castrillo, G., Sanchez-Bermejo, E., de Lorenzo, L., Crevillen, P., Fraile-Escanciano, A., Tc, M., et al. (2013) WRKY6 transcription factor restricts arsenate uptake and transposon activation in Arabidopsis. *Plant Cell* 25: 2944-2957.

Cavrak, V.V., Lettner, N., Jamge, S., Kosarewicz, A., Bayer, L.M. and Mittelsten Scheid, O. (2014) How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genet* 10: e1004115.

Champion, C., Guianvarc'h, D., Senamaud-Beaufort, C., Jurkowska, R.Z., Jeltsch, A., Ponger, L., et al. (2010) Mechanistic insights on the inhibition of c5 DNA methyltransferases by zebularine. *PLoS One* 5: e12388.

Chang, W., Jaaskelainen, M., Li, S.P. and Schulman, A.H. (2013) BARE retrotransposons are translated and replicated via distinct RNA pools. *PLoS One* 8: e72270.

Chen, J., Greenblatt, I.M. and Dellaporta, S.L. (1987) Transposition of Ac from the P locus of Maize into Unreplicated Chromosomal Sites *Genetics* 117: 109-116.

Chen, J., Zhao, H., Zheng, X., Liang, K., Guo, Y. and Sun, X. (2017) Recent amplification of Osr4 LTR-retrotransposon caused rice D1 gene mutation and dwarf phenotype. *Plant Diversity* 39: 73-79.

Chung, H., Bogwitz, M.R., McCart, C., Andrianopoulos, A., ffrench-Constant, R.H., Batterham, P., et al. (2007) Cis-Regulatory Elements in the Accord Retrotransposon Result in Tissue-Specific Expression of the Drosophila melanogaster Insecticide Resistance Gene Cyp6g1. *Genetics* 175: 1071-1077.

Chuong, E.B., Elde, N.C. and Feschotte, C. (2017) Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet* 18: 71-86.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452: 215-219.

Creasey, K.M., Zhai, J., Borges, F., Van Ex, F., Regulski, M., Meyers, B.C., et al. (2014) miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis. *Nature* 508: 411-415.

Cuerda-Gil, D. and Slotkin, R.K. (2016) Non-canonical RNA-directed DNA methylation. *Nat Plants* 2: 16163.

Daccord, N., Celton, J.M., Linsmith, G., Becker, C., Choisne, N., Schijlen, E., et al. (2017) High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat Genet* 49: 1099-1106.

Dapp, M., Reinders, J., Bediee, A., Balsera, C., Bucher, E., Theiler, G., et al. (2015) Heterosis and inbreeding depression of epigenetic Arabidopsis hybrids. *Nat Plants* 1: 15092.

de la Chaux, N., Tsuchimatsu, T., Shimizu, K.K. and Wagner, A. (2012) The predominantly selfing plant Arabidopsis thaliana experienced a recent reduction in transposable element abundance compared to its outcrossing relative Arabidopsis lyrata. *Mobile DNA* 3: 2-2.

Debladis, E., Llauro, C., Carpentier, M.C., Mirouze, M. and Panaud, O. (2017) Detection of active transposable elements in Arabidopsis thaliana using Oxford Nanopore Sequencing technology. *BMC Genomics* 18: 537.

DeNicola, G.M., Karreth, F.A., Adams, D.J. and Wong, C.C. (2015) The utility of transposon mutagenesis for cancer studies in the era of genome editing. *Genome Biol* 16: 229.

Devos, K.M., Brown, J.K.M. and Bennetzen, J.L. (2002) Genome Size Reduction through Illegitimate Recombination Counteracts Genome Expansion in Arabidopsis. *Genome Research* 12: 1075-1079.

Diez, C.M., Gaut, B.S., Meca, E., Scheinvar, E., Montes-Hernandez, S., Eguiarte, L.E., et al. (2013) Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytol* 199: 264-276.

Ding, Y., Wang, X., Su, L., Zhai, J., Cao, S., Zhang, D., et al. (2007) SDG714, a histone H3K9 methyltransferase, is involved in Tos17 DNA methylation and transposition in rice. *Plant Cell* 19: 9-22.

Du, J., Grant, D., Tian, Z., Nelson, R.T., Zhu, L., Shoemaker, R.C., et al. (2010) SoyTEdb: a comprehensive database of transposable elements in the soybean genome. *BMC Genomics* 11: 113.

Duan, K., Ding, X., Zhang, Q., Zhu, H., Pan, A. and Huang, J. (2008) AtCopeg1, the unique gene originated from AtCopia95 retrotransposon family, is sensitive to external hormones and abiotic stresses. *Plant Cell Reports* 27: 1065-1073.

Ebbs, M.L. and Bender, J. (2006) Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. *Plant Cell* 18: 1166-1176.

Eick, D. and Geyer, M. (2013) The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem Rev* 113: 8456-8490.

El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., et al. (2007) Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. *Genes Dev* 21: 2539-2544.

Elrouby, N. and Bureau, T.E. (2010) Bs1, a new chimeric gene formed by retrotransposon-mediated exon shuffling in maize. *Plant Physiol* 153: 1413-1424.

Erwin, J.A., Marchetto, M.C. and Gage, F.H. (2014) Mobile DNA elements in the generation of diversity and complexity in the brain. *Nat Rev Neurosci* 15: 497-506.

EuGH (2017) Vorlagefragen zur Einordnung von neuen gentechnischen Verfahren-Rechtssache C 528/16 *Amtsblatt der Europäischen Union* C 014/29: 23-24.

Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., et al. (2017) Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol* 18: 161.

Fedoroff, N.V. (2012) Transposable Elements, Epigenetics, and Genome Evolution. *Science* 338: 758-767.

Feschotte, C., Swamy, L. and Wessler, S.R. (2003) Genome-wide analysis of mariner-like transposable elements in rice reveals complex relationships with stowaway miniature inverted repeat transposable elements (MITEs). *Genetics* 163: 747-758.

Finnegan, E.J. and Dennis, E.S. (1993) Isolation and identification by sequence homology of a putative cytosine methyltransferase from Arabidopsis thaliana. *Nucleic Acids Res* 21: 2383–2388.

Finnegan, E.J., Peacock, W.J. and Dennis, E.S. (1996) Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development. *Proc Natl Acad Sci U S A* 93.

Fischer, R.A., Byerlee, D. and Edmeades, G.O. (2014) Crop yields and global food security: will yield increase continue to feed the world? Australian Centre for International Agricultural Research, Canberra.

Flavell, A.J. and Ish-Horowicz, D. (1981) Extrachromosomal circular copies of the eukaryotic transposable element copia in cultured Drosophila cells. *Nature* 292: 591-595.

Fuchs, J., Demidov, D., Houben, A. and Schubert, I. (2006) Chromosomal histone modification patterns--from conservation to diversity. *Trends Plant Sci* 11: 199-208.

Fujimoto, R., Kinoshita, Y., Kawabe, A., Kinoshita, T., Takashima, K., Nordborg, M., et al. (2008) Evolution and control of imprinted FWA genes in the genus Arabidopsis. *PLoS Genet* 4: e1000048.

Gallusci, P., Dai, Z., Genard, M., Gauffretau, A., Leblanc-Fournier, N., Richard-Molard, C., et al. (2017) Epigenetics for Plant Improvement: Current Knowledge and Modeling Avenues. *Trends Plant Sci* 22: 610-623.

Gao, D., Chen, J., Chen, M., Meyers, B.C. and Jackson, S. (2012) A highly conserved, small LTR retrotransposon that preferentially targets genes in grass genomes. *PLoS One* 7: e32010.

Gao, D., Jiang, N., Wing, R.A., Jiang, J. and Jackson, S.A. (2015) Transposons play an important role in the evolution and diversification of centromeres among closely related species. *Front Plant Sci* 6: 216.

Gao, Z., Liu, H.L., Daxinger, L., Pontes, O., He, X., Qian, W., et al. (2010) An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. *Nature* 465: 106-109.

Gasciolli, V., Mallory, A.C., Bartel, D.P. and Vaucheret, H. (2005) Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr Biol* 15: 1494-1500.

Gendrel, A.-V., Lippman, Z., Yordan, C., Colot, V. and Martienssen, R.A. (2002) Dependence of Heterochromatic Histone H3 Methylation Patterns on the Arabidopsis Gene DDM1. *Science* 297: 1871-1873.

Gouil, Q. and Baulcombe, D.C. (2016) DNA Methylation Signatures of the Plant Chromomethyltransferases. *PLoS Genet* 12: e1006526.

Grandbastien, M.A. (2015) LTR retrotransposons, handy hitchhikers of plant regulation and stress response. *Biochim Biophys Acta* 1849: 403-416.

Griffin, P.T., Niederhuth, C.E. and Schmitz, R.J. (2016) A Comparative Analysis of 5-Azacytidine- and Zebularine-Induced DNA Demethylation. *G3 (Bethesda)* 6: 2773-2780.

Haag, J.R., Ream, T.S., Marasco, M., Nicora, C.D., Norbeck, A.D., Pasa-Tolic, L., et al. (2012) In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol Cell* 48: 811-818.

Hawkins, J.S., Proulx, S.R., Rapp, R.A. and Wendel, J.F. (2009) Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants. *Proc Natl Acad Sci U S A* 106: 17811-17816.

Henderson, I.R. and Jacobsen, S.E. (2007) Epigenetic inheritance in plants. *Nature* 447: 418-424.

Herr, A.J., Jensen, M.B., Dalmay, T. and Baulcombe, D.C. (2005) RNA Polymerase IV Directs Silencing of Endogenous DNA. *Science* 308: 118-120.

Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H. and Kanda, M. (1996) Retrotransposons of rice involved in mutations induced by tissue culture. *Proceedings of the National Academy of Sciences of the United States of America* 93: 7783-7788.

Holoch, D. and Moazed, D. (2015) RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* 16: 71-84.

Hu, L., Li, N., Xu, C., Zhong, S., Lin, X., Yang, J., et al. (2014) Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proceedings of the National Academy of Sciences* 111: 10642-10647.

Hu, T.T., Pattyn, P., Bakker, E.G., Cao, J., Cheng, J.F., Clark, R.M., et al. (2011) The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. *Nat Genet* 43: 476-481.

Huang, C.R., Burns, K.H. and Boeke, J.D. (2012) Active transposition in genomes. *Annu Rev Genet* 46: 651-675.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.-F., Uzawa, R., Rodrigues, J.A., et al. (2012) Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes. *Science* 337: 1360-1364.

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I. and Paszkowski, J. (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* 472: 115-119.

Ito, H., Kim, J.M., Matsunaga, W., Saze, H., Matsui, A., Endo, T.A., et al. (2016) A Stress-Activated Transposon in Arabidopsis Induces Transgenerational Abscisic Acid Insensitivity. *Sci Rep* 6: 23181.

Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., et al. (2004) Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. *Chromosoma* 112: 308-315.

Jackson, J.P., Lindroth, A.M., Cao, X. and Jacobsen, S.E. (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416: 556-560.

Jacob, Y., Stroud, H., Leblanc, C., Feng, S., Zhuo, L., Caro, E., et al. (2010) Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. *Nature* 466: 987-991.

Jameson, N., Georgelis, N., Fouladbash, E., Martens, S., Hannah, L.C. and Lal, S. (2008) Helitron mediated amplification of cytochrome P450 monooxygenase gene in maize. *Plant molecular biology* 67: 295-304.

Jeddeloh, J.A., Stokes, T.L. and Richards, E.J. (1999) Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat Genet* 22: 94-97.

Jiang, N., Bao, Z., Zhang, X., Eddy, S.R. and Wessler, S.R. (2004) Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 431: 569-573.

Joly-Lopez, Z. and Bureau, T.E. (2014) Diversity and evolution of transposable elements in Arabidopsis. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 22: 203-216.

Kalendar, R., Vicient, C.M., Peleg, O., Anamthawat-Jonsson, K., Bolshoy, A. and Schulman, A.H. (2004) Large Retrotransposon Derivatives: Abundant, Conserved but Nonautonomous Retroelements of Barley and Related Genomes. *Genetics* 166: 1437–1450.

Kanazawa, A., Liu, B., Kong, F., Arase, S. and Abe, J. (2009) Adaptive Evolution Involving Gene Duplication and Insertion of a Novel Ty1/copia-Like Retrotransposon in Soybean. *Journal of Molecular Evolution* 69: 164-175.

Kankel, M.W., Ramsey, D.E., Stokes, T.L., Flowers, S.K., Haag, J.R., Jeddeloh, J.A., et al. (2003) Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics* 163.

Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., et al. (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* 37: 761-765.

Kapitonov, V.V. and Jurka, J. (2001) Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A* 98: 8714-8719.

Kawahara, Y., de la Bastide, M., Hamilton, J.P., Kanamori, H., McCombie, W.R., Ouyang, S., et al. (2013) Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6: 4.

Kimura, Y., Tosa, Y., Shimada, S., Sogo, R., Kusaba, M., Sunaga, T., et al. (2001) OARE-1, a Ty1-copia Retrotransposon in Oat Activated by Abiotic and Biotic Stresses. *Plant and Cell Physiology* 42: 1345-1354.

Kinoshita, Y., Saze, H., Kinoshita, T., Miura, A., Soppe, W.J., Koornneef, M., et al. (2007) Control of FWA gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. *The Plant journal : for cell and molecular biology* 49: 38-45.

Lanciano, S., Carpentier, M.C., Llauro, C., Jobet, E., Robakowska-Hyzorek, D., Lasserre, E., et al. (2017) Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. *PLoS Genet* 13: e1006630.

Latzel, V., Allan, E., Bortolini Silveira, A., Colot, V., Fischer, M. and Bossdorf, O. (2013) Epigenetic diversity increases the productivity and stability of plant populations. *Nat Commun* 4: 2875.

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

Lazcano, A., Fastag, J., Gariglio, P., Ramírez, C. and Oró, J. (1988) On the early evolution of RNA polymerase. *Journal of Molecular Evolution* 27: 365-376.

Lee, S.I. and Kim, N.S. (2014) Transposable elements and genome size variations in plants. *Genomics Inform* 12: 87-97.

Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Chan, S.W., et al. (2006) An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. *Cell* 126: 93-106.

Li, Q., Eichten, S.R., Hermanson, P.J., Zaunbrecher, V.M., Song, J., Wendt, J., et al. (2014) Genetic perturbation of the maize methylome. *Plant Cell* 26: 4602-4616.

Lindell, T.J., Weinberg, F., Morris, P.W., Roeder, R.G. and Rutter, W.J. (1970) Specific Inhibition of Nuclear RNA Polymerase II by α-Amanitin. *Science* 170: 447-449.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., et al. (2001) Requirement of the CHROMOMETHYLASE3 for Maintenance of CpXpG Methylation. *Science* 292: 2077-2080.

Lippman, Z., Gendrel, A.-V., Black, M., Vaughn, M.W., Dedhia, N., Richard McCombie, W., et al. (2004) Role of transposable elements in heterochromatin and epigenetic control. *Nature* 430: 471-476.

Lippman, Z., May, B., Yordan, C., Singer, T. and Martienssen, R. (2003) Distinct Mechanisms Determine Transposon Inheritance and Methylation via Small Interfering RNA and Histone Modification. *PLOS Biology* 1: e67.

Lisch, D. (2013) How important are transposons for plant evolution? *Nat Rev Genet* 14: 49-61.

Lisch, D. and Bennetzen, J.L. (2011) Transposable element origins of epigenetic gene regulation. *Curr Opin Plant Biol* 14: 156-161.

Liu, C., Cheng, Y.J., Wang, J.W. and Weigel, D. (2017) Prominent topologically associated domains differentiate global chromatin packing in rice from Arabidopsis. *Nat Plants* 3: 742–748.

Liu, C.H., Finke, A., Diaz, M., Rozhon, W., Poppenberger, B., Baubec, T., et al. (2015) Repair of DNA Damage Induced by the Cytidine Analog Zebularine Requires ATR and ATM in Arabidopsis. *Plant Cell* 27: 1788-1800.

Liu, X., Yu, C.W., Duan, J., Luo, M., Wang, K., Tian, G., et al. (2012) HDA6 directly interacts with DNA methyltransferase MET1 and maintains transposable element silencing in Arabidopsis. *Plant Physiol* 158: 119-129.

Lu, H.-F., Dong, H.-T., Sun, C.-B., Qing, D.-J., Li, N., Wu, Z.-K., et al. (2011) The panorama of physiological responses and gene expression of whole plant of maize inbred line YQ7-96 at the three-leaf stage under water deficit and re-watering. *Theoretical and Applied Genetics* 123: 943–958.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8[thinsp]A resolution. *Nature* 389: 251-260.

Lyko, F. and Brown, R. (2005) DNA methyltransferase inhibitors and the development of epigenetic cancer therapies. *J Natl Cancer Inst* 97: 1498-1506.

Ma, J., Devos, K.M. and Bennetzen, J.L. (2004) Analyses of LTR-Retrotransposon Structures Reveal Recent and Rapid Genomic DNA Loss in Rice. *Genome Research* 14: 860-869.

Makarevitch, I., Waters, A.J., West, P.T., Stitzer, M., Hirsch, C.N., Ross-Ibarra, J., et al. (2015) Transposable elements contribute to activation of maize genes in response to abiotic stress. *PLoS Genet* 11: e1004915.

Mari-Ordonez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V. and Voinnet, O. (2013) Reconstructing de novo silencing of an active plant retrotransposon. *Nat Genet* 45: 1029-1039.

Masuta, Y., Nozawa, K., Takagi, H., Yaegashi, H., Tanaka, K., Ito, T., et al. (2017) Inducible Transposition of a Heat-Activated Retrotransposon in Tissue Culture. *Plant Cell Physiol* 58: 375-384.

Matsunaga, W., Ohama, N., Tanabe, N., Masuta, Y., Masuda, S., Mitani, N., et al. (2015) A small RNA mediated regulation of a stress-activated retrotransposon and the tissue specific transposition during the reproductive period in Arabidopsis. *Front Plant Sci* 6: 48.

Matzke, M.A., Kanno, T. and Matzke, A.J. (2015) RNA-Directed DNA Methylation: The Evolution of a Complex Epigenetic Pathway in Flowering Plants. *Annu Rev Plant Biol* 66: 243-267.

Matzke, M.A. and Mosher, R.A. (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* 15: 394-408.

McClintock, B. (1950) The Origin and Behavior of Mutable Loci in Maize. *Proceedings of the National Academy of Sciences of the United States of America* 36: 344-355.

McClintock, B. (1956) Intranuclear systems controlling gene action and mutation. *Brookhaven Symp Biol.* 8: 58-74.

McClintock, B. (1984) The significance of responses of the genome to challenge. *Science* 226: 792-801.

McCue, A.D., Panda, K., Nuthikattu, S., Choudury, S.G., Thomas, E.N. and Slotkin, R.K. (2015) ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *EMBO J* 34: 20-35.

Mirouze, M., Reinders, J., Bucher, E., Nishimura, T., Schneeberger, K., Ossowski, S., et al. (2009) Selective epigenetic control of retrotransposition in Arabidopsis. *Nature* 461: 427-430.

Mirouze, M. and Vitte, C. (2014) Transposable elements, a treasure trove to decipher epigenetic variation: insights from Arabidopsis and crop epigenomes. *J Exp Bot* 65: 2801-2812.

Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H. and Kakutani, T. (2001) Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. *Nature* 411: 212-214.

Montanini, B., Chen, P.-Y., Morselli, M., Jaroszewicz, A., Lopez, D., Martin, F., et al. (2014) Non-exhaustive DNA methylation-mediated transposon silencing in the black truffle genome, a complex fungal genome with massive repeat element content. *Genome Biology* 15: 411.

Moore, J.K. and Haber, J.E. (1996) Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. *Nature* 383: 644-646.

Naito, K., Cho, E., Yang, G., Campbell, M.A., Yano, K., Okumoto, Y., et al. (2006) Dramatic amplification of a rice transposable element during recent domestication. *Proc Natl Acad Sci U S A* 103: 17620-17625.

Nawrath, C., Schell, J. and Koncz, C. (1990) Homologous domains of the largest subunit of eucaryotic RNA polymerase II are conserved in plants. *Molecular and General Genetics MGG* 223: 65-75.

Negi, P., Rai, A.N. and Suprasanna, P. (2016) Moving through the Stressed Genome: Emerging Regulatory Roles for Transposons in Plant Stress Response. *Front Plant Sci* 7: 1448.

Neumann, P., Požárková, D. and Macas, J. (2003) Highly abundant pea LTR retrotransposon Ogre is constitutively transcribed and partially spliced. *Plant molecular biology* 53: 399-410.

Nobuta, K., Venu, R.C., Lu, C., Belo, A., Vemaraju, K., Kulkarni, K., et al. (2007) An expression atlas of rice mRNAs and small RNAs. *Nat Biotechnol* 25: 473-477.

Nuthikattu, S., McCue, A.D., Panda, K., Fultz, D., DeFraia, C., Thomas, E.N., et al. (2013) The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. *Plant Physiol* 162: 116-131.

Oelgeschläger, T. (2002) Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. *Journal of Cellular Physiology* 190: 160-169.

Ohno, S. (1972) So much "junk" DNA in our genome. *Brookhaven Symp Biol.* 23: 366-370.

Ong-Abdullah, M., Ordway, J.M., Jiang, N., Ooi, S.E., Kok, S.Y., Sarpan, N., et al. (2015) Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature* 525: 533-537.

Onodera, Y., Haag, J.R., Ream, T., Costa Nunes, P., Pontes, O. and Pikaard, C.S. (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 120: 613-622.

Pai, D.A., Kaplan, C.D., Kweon, H.K., Murakami, K., Andrews, P.C. and Engelke, D.R. (2014) RNAs nonspecifically inhibit RNA polymerase II by preventing binding to the DNA template. *RNA* 20: 644-655.

Palancade, B. and Bensaude, O. (2003) Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. *European Journal of Biochemistry* 270: 3859-3870.

Panaud, O., Vitte, C., Hivert, J., Muzlak, S., Talag, J., Brar, D., et al. (2002) Characterization of transposable elements in the genome of rice (Oryza sativa L.) using Representational Difference Analysis (RDA). *Mol Genet Genomics* 268: 113-121.

Panda, K., Ji, L., Neumann, D.A., Daron, J., Schmitz, R.J. and Slotkin, R.K. (2016) Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biol* 17: 170.

Park, M., Keung, A.J. and Khalil, A.S. (2016) The epigenome: the next substrate for engineering. *Genome Biol* 17: 183.

Paszkowski, J. (2015) Controlled activation of retrotransposition for plant breeding. *Current opinion in biotechnology* 32C: 200-206.

Pecinka, A., Dinh, H.Q., Baubec, T., Rosa, M., Lettner, N. and Mittelsten Scheid, O. (2010) Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis. *Plant Cell* 22: 3118-3129.

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

Piegu, B., Bire, S., Arensburger, P. and Bigot, Y. (2015) A survey of transposable element classification systems--a call for a fundamental update to meet the challenge of their diversity and complexity. *Mol Phylogenet Evol* 86: 90-109.

Piegu, B., Guyot, R., Picault, N., Roulin, A., Sanyal, A., Kim, H., et al. (2006) Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in Oryza australiensis, a wild relative of rice. *Genome Res* 16: 1262-1269.

Pietzenuk, B., Markus, C., Gaubert, H., Bagwan, N., Merotto, A., Bucher, E., et al. (2016) Recurrent evolution of heat-responsiveness in Brassicaceae COPIA elements. *Genome Biol* 17: 209.

Pontier, D., Picart, C., Roudier, F., Garcia, D., Lahmy, S., Azevedo, J., et al. (2012) NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in Arabidopsis. *Mol Cell* 48: 121-132.

Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.-A., et al. (2005) Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. *Genes & Development* 19: 2030-2040.

Probst, A.V., Fagard, M., Proux, F., Mourrain, P., Boutet, S., Earley, K., et al. (2004) Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. *Plant Cell* 16: 1021-1034.

Quadrana, L., Bortolini Silveira, A., Mayhew, G.F., LeBlanc, C., Martienssen, R.A., Jeddeloh, J.A., et al. (2016) The Arabidopsis thaliana mobilome and its impact at the species level. *Elife* 5: e15716.

Ramallo, E., Kalendar, R., Schulman, A.H. and Martinez-Izquierdo, J.A. (2008) Reme1, a Copia retrotransposon in melon, is transcriptionally induced by UV light. *Plant molecular biology* 66: 137-150.

Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.K., et al. (2009) Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol Cell* 33: 192-203.

Reinders, J., Wulff, B.B., Mirouze, M., Mari-Ordonez, A., Dapp, M., Rozhon, W., et al. (2009) Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes Dev* 23: 939-950.

Rocheta, M., Carvalho, L., Viegas, W. and Morais-Cecílio, L. (2012) Corky, a gypsy-like retrotransposon is differentially transcribed in Quercus suber tissues. *BMC Research Notes* 5: 432-432.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B. (2003) An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant molecular biology* 53: 247-259.

Russo, J., Harrington, A.W. and Steiniger, M. (2016) Antisense Transcription of Retrotransposons in Drosophila: An Origin of Endogenous Small Interfering RNA Precursors. *Genetics* 202: 107-121.

Santi, D.V., Garrett, C.E. and Barr, P.J. (1983) On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. *Cell* 33: 9-10.

Schmidt, T. (1999) LINEs, SINEs and repetitive DNA: non-LTR retrotransposons in plant genomes. *Plant molecular biology* 40: 903-910.

Schmitz, R.J., He, Y., Valdes-Lopez, O., Khan, S.M., Joshi, T., Urich, M.A., et al. (2013) Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome Res* 23: 1663-1674.

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., et al. (2009) The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* 326: 1112-1115.

Schori, A. and Charles, R. (2004) Sojabohne: Züchtung, Agronomie und Produktion in der Schweiz. *Agrarforschung* 10: 1-8.

Schubert, I. and Oud, J.L. (1997) There Is an Upper Limit of Chromosome Size for Normal Development of an Organism. *Cell* 88: 515-520.

Schulman, A.H. (2013) Retrotransposon replication in plants. *Curr Opin Virol* 3: 604-614.

Sharma, A., Wolfgruber, T.K. and Presting, G.G. (2013) Tandem repeats derived from centromeric retrotransposons. *BMC Genomics* 14: 142.

Shirasu, K., Schulman, A.H., Lahaye, T. and Schulze-Lefert, P. (2000) A Contiguous 66-kb Barley DNA Sequence Provides Evidence for Reversible Genome Expansion. *Genome Research* 10: 908-915.

Sigman, M.J. and Slotkin, R.K. (2016) The First Rule of Plant Transposable Element Silencing: Location, Location, Location. *Plant Cell* 28: 304-313.

Slotkin, R.K., Freeling, M. and Lisch, D. (2005) Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. *Nat Genet* 37: 641-644.

Springer, N.M. and Schmitz, R.J. (2017) Exploiting induced and natural epigenetic variation for crop improvement. *Nat Rev Genet* 18: 563–575.

Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V. and Jacobsen, S.E. (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* 152: 352-364.

Stroud, H., Otero, S., Desvoyes, B., Ramírez-Parra, E., Jacobsen, S.E. and Gutierrez, C. (2012) Genome-wide analysis of histone H3.1 and H3.3 variants in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* 109: 5370-5375.

Sultana, T., Zamborlini, A., Cristofari, G. and Lesage, P. (2017) Integration site selection by retroviruses and transposable elements in eukaryotes. *Nat Rev Genet* 18: 292-308.

Sun, F.J., Fleurdepine, S., Bousquet-Antonelli, C., Caetano-Anolles, G. and Deragon, J.M. (2007) Common evolutionary trends for SINE RNA structures. *Trends Genet* 23: 26-33.

Sweetser, D., Nonet, M. and Young, R.A. (1987) Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. *Proceedings of the National Academy of Sciences of the United States of America* 84: 1192-1196.

Takeda, S., Sugimoto, K., Otsuki, H. and Hirochika, H. (1998) Transcriptional activation of the tobacco retrotransposon Tto1 by wounding and methyl jasmonate. *Plant molecular biology* 36: 365-376.

Tanskanen, J.A., Sabot, F., Vicient, C. and Schulman, A.H. (2007) Life without GAG: the BARE-2 retrotransposon as a parasite's parasite. *Gene* 390: 166-174.

Tapia, G., Verdugo, I., Yanez, M., Ahumada, I., Theoduloz, C., Cordero, C., et al. (2005) Involvement of ethylene in stress-induced expression of the TLC1.1 retrotransposon from Lycopersicon chilense Dun. *Plant Physiol* 138: 2075-2086.

Tenaillon, M.I., Hollister, J.D. and Gaut, B.S. (2010) A triptych of the evolution of plant transposable elements. *Trends Plant Sci* 15: 471-478.

The_Arabidopsis_Genome_Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408: 796-815.

Thieme, M. (2017a) Der mobile Teil des Pflanzengenoms als Ressource für den zukünftigen Pflanzenschutz. *Tagungsband Fachtagung Dialog Grün 2016*: 36-39.

Thieme, M. (2017b) Putting plants in shool: On the potential of epigenetic memory in crop breeding. *Plant Science News* 32: 4-5.

Thieme, M., Lanciano, S., Balzergue, S., Daccord, N., Mirouze, M. and Bucher, E. (2017) Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. *Genome Biol* 18: 134.

Tittel-Elmer, M., Bucher, E., Broger, L., Mathieu, O., Paszkowski, J. and Vaillant, I. (2010) Stress-induced activation of heterochromatic transcription. *PLoS Genet* 6: e1001175.

To, T.K., Kim, J.M., Matsui, A., Kurihara, Y., Morosawa, T., Ishida, J., et al. (2011) Arabidopsis HDA6 regulates locus-directed heterochromatin silencing in cooperation with MET1. *PLoS Genet* 7: e1002055.

Tsukahara, S., Kobayashi, A., Kawabe, A., Mathieu, O., Miura, A. and Kakutani, T. (2009) Bursts of retrotransposition reproduced in Arabidopsis. *Nature* 461: 423-426.

Tubio, J.M.C., Li, Y., Ju, Y.S., Martincorena, I., Cooke, S.L., Tojo, M., et al. (2014) Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* 345: 1251343.

Underwood, C.J., Henderson, I.R. and Martienssen, R.A. (2017) Genetic and epigenetic variation of transposable elements in Arabidopsis. *Curr Opin Plant Biol* 36: 135-141.

van der Knaap, E., Sanyal, A., Jackson, S.A. and Tanksley, S.D. (2004) High-resolution fine mapping and fluorescence in situ hybridization analysis of sun, a locus controlling tomato fruit shape, reveals a region of the tomato genome prone to DNA rearrangements. *Genetics* 168: 2127-2140.

Varagona, M.J., Purugganan, M. and Wessler, S.R. (1992) Alternative splicing induced by insertion of retrotransposons into the maize waxy gene. *The Plant Cell* 4: 811-820.

Vitte, C., Fustier, M.A., Alix, K. and Tenaillon, M.I. (2014) The bright side of transposons in crop evolution. *Brief Funct Genomics* 13: 276-295.

Vitte, C. and Panaud, O. (2005) LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenet Genome Res* 110: 91-107.

Vitte, C., Panaud, O. and Quesneville, H. (2007) LTR retrotransposons in rice (Oryza sativa, L.): recent burst amplifications followed by rapid DNA loss. *BMC Genomics* 8: 218.

Vongs, A., Kakutani, T., Martienssen, R.A. and Richards, E.J. (1993) *Arabidopsis thaliana* DNA methylation mutants. *Science* 260: 1926-1928.

Voronova, A., Belevich, V., Jansons, A. and Rungis, D. (2014) Stress-induced transcriptional activation of retrotransposon-like sequences in the Scots pine (Pinus sylvestris L.) genome. *Tree Genetics & Genomes* 10: 937-951.

Wang, H., Chai, Y., Chu, X., Zhao, Y., Wu, Y., Zhao, J., et al. (2009) Molecular characterization of a rice mutator-phenotype derived from an incompatible cross-pollination reveals transgenerational mobilization of multiple transposable elements and extensive epigenetic instability. *BMC Plant Biol* 9: 63.

Wassenegger, M., Heimes, S., Riedel, L. and Sänger, H.L. (1994) RNA-directed de novo methylation of genomic sequences in plants. *Cell* 76: 567-576.

Wendel, J.F., Jackson, S.A., Meyers, B.C. and Wing, R.A. (2016) Evolution of plant genome architecture. *Genome Biol* 17: 37.

Wendte, J.M. and Pikaard, C.S. (2017) The RNAs of RNA-directed DNA methylation. *Biochim Biophys Acta* 1860: 140-148.

Wenke, T., Dobel, T., Sorensen, T.R., Junghans, H., Weisshaar, B. and Schmidt, T. (2011) Targeted identification of short interspersed nuclear element families shows their widespread existence and extreme heterogeneity in plant genomes. *Plant Cell* 23: 3117-3128.

Wheeler, B.S. (2013) Small RNAs, big impact: small RNA pathways in transposon control and their effect on the host stress response. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 21: 587-600.

White, T.B., McCoy, A.M., Streva, V.A., Fenrich, J. and Deininger, P.L. (2014) A droplet digital PCR detection method for rare L1 insertions in tumors. *Mob DNA* 5: 30.

Wicker, T. and Keller, B. (2007) Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and Arabidopsis reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. *Genome Res* 17: 1072-1081.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., et al. (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8: 973-982.

Wierzbicki, A.T., Ream, T.S., Haag, J.R. and Pikaard, C.S. (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet* 41: 630-634.

Wilson, R.F. (2008) Soybean: Market Driven Research Needs. In *Genetics and Genomics of Soybean*. Edited by Stacey, G. pp. 3-15. Springer New York, New York, NY.

Witte, C.-P., Le, Q.H., Bureau, T. and Kumar, A. (2001) Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. *Proceedings of the National Academy of Sciences of the United States of America* 98: 13778-13783.

Wolfgruber, T.K., Sharma, A., Schneider, K.L., Albert, P.S., Koo, D.H., Shi, J., et al. (2009) Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic Loci shaped primarily by retrotransposons. *PLoS Genet* 5: e1000743.

Xiao-Jie, L., Hui-Ying, X., Qi, X., Jiang, X. and Shi-Jie, M. (2016) LINE-1 in cancer: multifaceted functions and potential clinical implications. *Genet Med* 18: 431-439.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J. and van der Knaap, E. (2008) A Retrotransposon-Mediated Gene Duplication Underlies Morphological Variation of Tomato Fruit. *Science* 319: 1527-1530.

Yang, Z., Ebright, Y.W., Yu, B. and Chen, X. (2006) HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res* 34: 667-675.

Yao, J., Dong, Y. and Morris, B.A. (2001) Parthenocarpic apple fruit production conferred by transposon insertion mutations in a MADS-box transcription factor. *Proc Natl Acad Sci U S A* 98: 1306-1311.

Ye, R., Chen, Z., Lian, B., Rowley, M.J., Xia, N., Chai, J., et al. (2016) A Dicer-Independent Route for Biogenesis of siRNAs that Direct DNA Methylation in Arabidopsis. *Mol Cell* 61: 222-235.

Yelagandula, R., Stroud, H., Holec, S., Zhou, K., Feng, S., Zhong, X., et al. (2014) The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell* 158: 98-109.

Yu, A., Lepere, G., Jay, F., Wang, J., Bapaume, L., Wang, Y., et al. (2013) Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. *Proc Natl Acad Sci U S A* 110: 2389-2394.

Yuzbasioglu, G., Yilmaz, S., Marakli, S. and Gozukirmizi, N. (2016) Analysis ofHopi/Osr27andHouba/Tos5/Osr13retrotransposons in rice. *Biotechnology & Biotechnological Equipment* 30: 213-218.

Zedek, F., Šmerda, J., Šmarda, P. and Bureš, P. (2010) Correlated evolution of LTR retrotransposons and genome size in the genus eleocharis. *BMC Plant Biology* 10: 265.

Zeller, G., Henz, S.R., Widmer, C.K., Sachsenberg, T., Ratsch, G., Weigel, D., et al. (2009) Stress-induced changes in the Arabidopsis thaliana transcriptome analyzed using whole-genome tiling arrays. *The Plant journal : for cell and molecular biology* 58: 1068-1082.

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

Zhang, J., Zuo, T., Wang, D. and Peterson, T. (2014) Transposition-mediated DNA re-replication in maize. *Elife* 3: e03724.

Zhang, J.J., Zhou, Z.S., Song, J.B., Liu, Z.P. and Yang, H. (2012) Molecular dissection of atrazine-responsive transcriptome and gene networks in rice by high-throughput sequencing. *J Hazard Mater* 219-220: 57-68.

Zhang, Q.J. and Gao, L.Z. (2017) Rapid and Recent Evolution of LTR Retrotransposons Drives Rice Genome Evolution During the Speciation of AA-Genome Oryza Species. *G3 (Bethesda)* 7: 1875-1885.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* 126: 1189-1201.

Zhao, D., Ferguson, A.A. and Jiang, N. (2016) What makes up plant genomes: The vanishing line between transposable elements and genes. *Biochim Biophys Acta* 1859: 366-380.

Zheng, B., Wang, Z., Li, S., Yu, B., Liu, J.Y. and Chen, X. (2009) Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. *Genes Dev* 23: 2850-2860.

Zheng, X., Zhu, J., Kapoor, A. and Zhu, J.-K. (2007) Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *The EMBO Journal* 26: 1691-1701.

Zilberman, D., Cao, X. and Jacobsen, S.E. (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* 299: 716-719.

10 Acknowledgements

This project has fundamentally changed my personal perception on evolution and sparked my interest to further study how environmental cues are captured, stored and transmitted to successive generations.

Against this background, I would first of all like to thank the European Union and the Freie Akademische Gesellschaft for the generous funding and Prof. Thomas Boller and Dr. Etienne Bucher for offering me the great opportunity to start a PhD in the field of epigenetics at the University of Basel.

Especially the open-minded and thought-provoking discussions at the beginning of the project followed by frequent support in many skype-meetings with Etienne provided the basis for the successful outcome of this work. I would also like to explicitly thank Monika Messmer, Manuela Dahinden and Luisa Last for their fantastic support at the Science & Policy- Interface. In addition I would like to thank Peter Eckard from Unitectra Basel for the support with the patent application. I would also like to thank Laurin and Emilija for the nice teamwork and help at the beginning of my work in the Bucher Lab and Arthur for his great assistance and work during the last year of my project.

Furthermore, the successful outcome of this project was made possible by the very successful collaborations with Marie, Sophie, Sandrine, Nicolas and facilitated by the administrative and technical support by Maura, Olivier, Markus and Giacomo.

Moreover, I`m very thankful for all the interesting, helpful and entertaining discussions with other colleagues from the Botanical Institute Basel and the INRA, Angers and more particular with Silvia, Lukas, Emilija, Arthur, Olivier, Jürg, Ines, Kay and Jean-Marc.

Finally, I would like to specially highlight and acknowledge the great support I received from my family and my friends that helped me to successfully finish this dissertation.

11 List of appendices

Appendix I Scientific publication: Thieme, M., Lanciano, S., Balzergue, S., Daccord, N., Mirouze, M. and Bucher, E. (2017) Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. *Genome Biol* 18: 134. (17 pages)

Appendix II Patent: Bucher*, E. and Thieme*, M. (2017) Mobilization of Transposable Elements to Enhance Genetic and Epigenetic Variability in a Population. *Patent* WO2017/093317A1 (50 pages)

Appendix III Newsletter article: Thieme, M. (2017) Putting plants in shool: On the potential of epigentic memory in crop breeding. *Plant Science News* 32: 4-5. (2 pages)

Appendix IV Transcript of presentation for “Fachtagung Dialog Grün 2016”: Thieme, M. (2017) Der mobile Teil des Pflanzengenoms als Ressource für den zukünftigen Pflanzenschutz. *Neue Technologien in der Pflanzenforschung-eine Alternative zu Pflanzenschutzmitteln?*: 36-39 (4 pages)

Appendix V Multiple choice test (World-Café), correct answers in red (2 pages)

Appendix VI Questionnaire (World-Café) (1 page)

Appendix VII Questions for guiding the discussion of World-Café (1 page)

Appendix VIII Presentation given prior to the discussion in the World-Café at FiBL (pictures were removed) (5 pages)

Appendix IX Article published by “pflanzenforschung.de” “Hitze lässt *ONSEN* hüpfen-Retrotransposonen kontrolliert zum Springen bringen” (2 pages)

11.1 Appendix I: Scientific publication: Thieme, M., Lanciano, S., Balzergue, S., Daccord, N., Mirouze, M. and Bucher, E. (2017) Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. *Genome Biol* 18: 134. (17 pages)

Thieme et al. Genome Biology (2017) 18:134
DOI 10.1186/s13059-017-1265-4

Genome Biology

RESEARCH Open Access

Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding

Michael Thieme[1], Sophie Lanciano[2,3], Sandrine Balzergue[4], Nicolas Daccord[4], Marie Mirouze[2,3] and Etienne Bucher[4*]

Abstract

Background: Retrotransposons play a central role in plant evolution and could be a powerful endogenous source of genetic and epigenetic variability for crop breeding. To ensure genome integrity several silencing mechanisms have evolved to repress retrotransposon mobility. Even though retrotransposons fully depend on transcriptional activity of the host RNA polymerase II (Pol II) for their mobility, it was so far unclear whether Pol II is directly involved in repressing their activity.

Results: Here we show that plants defective in Pol II activity lose DNA methylation at repeat sequences and produce more extrachromosomal retrotransposon DNA upon stress in *Arabidopsis* and rice. We demonstrate that combined inhibition of both DNA methylation and Pol II activity leads to a strong stress-dependent mobilization of the heat responsive *ONSEN* retrotransposon in *Arabidopsis* seedlings. The progenies of these treated plants contain up to 75 new *ONSEN* insertions in their genome which are stably inherited over three generations of selfing. Repeated application of heat stress in progeny plants containing increased numbers of *ONSEN* copies does not result in increased activation of this transposon compared to control lines. Progenies with additional *ONSEN* copies show a broad panel of environment-dependent phenotypic diversity.

Conclusions: We demonstrate that Pol II acts at the root of transposon silencing. This is important because it suggests that Pol II can regulate the speed of plant evolution by fine-tuning the amplitude of transposon mobility. Our findings show that it is now possible to study induced transposon bursts in plants and unlock their use to induce epigenetic and genetic diversity for crop breeding.

Keywords: Epigenetics, DNA methylation, Genome integrity, Evolution, *Oryza sativa*, *Arabidopsis thaliana*

Background

Like retroviruses, long terminal repeat (LTR) retrotransposons (class I elements), which represent the most abundant class of transposable elements (TEs) in eukaryotes, transpose via a copy and paste mechanism. This process requires the conversion of a full length RNA polymerase II (Pol II) transcript into extrachromosomal complementary DNA (ecDNA) by reverse transcription [1]. In their life cycle LTR retrotransposons can produce extrachromosomal circular DNA (eccDNA), which is an indicator for their ongoing activity [2]. In plants, TEs are increasingly seen as a source of genetic and epigenetic variability and thus important drivers of evolution [3–6]. However, plants have evolved several regulatory pathways to retain control over the activity of these potentially harmful mobile genetic elements. Cytosine methylation (^{m}C) plays a central role in TE silencing in plants [7]. In addition, plants have evolved two Pol II-related RNA polymerases, Pol IV and Pol V, that are essential to provide specific silencing signals leading to RNA-directed DNA methylation (RdDM) at TEs [8], thereby limiting their mobility [9–11]. More recently, various additional non-canonical Pol IV-independent RdDM pathways have been described [12]. Notably it was found that Pol II

* Correspondence: etienne.bucher@inra.fr
[4]IRHS, Université d'Angers, INRA, AGROCAMPUS-Ouest, SFR4207 QUASAV, Université Bretagne Loire, 49045 Angers, France
Full list of author information is available at the end of the article

© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

itself also plays an important role in RdDM [13, 14] by feeding template RNAs into downstream factors such as RNA-DEPENDENT RNA POLYMERASE 6 (RDR6), resulting in dicer-dependent or -independent initiation and establishment of TE-specific DNA methylation [15]. Beyond that, recent work suggests a new "non-canonical" branch of RdDM that specializes in targeting transcriptionally active full-length TEs [16]. This pathway functions independently of RDRs via Pol II transcripts that are directly processed by DCL3 into small interfering RNAs (siRNAs).

Results

Here, we wanted to investigate if Pol II could play a direct role in repressing TE mobility in plants. For this purpose we chose the well-characterized heat-responsive *copia*-like *ONSEN* retrotransposon [11] of *Arabidopsis* and took advantage of the hypomorphic *nrpb2-3* mutant allele that causes reduced NRPB2 (the second-largest component of Pol II) protein levels [14]. Using quantitative real-time PCR (qPCR), we determined that challenging *nrpb2-3* seedlings by heat stress (HS) led to a mild increase in total *ONSEN* copy number (sum of ecDNA, eccDNA and new genomic insertions) relative to control stress (CS) and compared to the wild type (WT) (Fig. 1a). This result is supported by the observed dose-responsive increase in *ONSEN* copy number after HS and pharmacological inactivation of Pol II with α-amanitin (A), a potent Pol II inhibitor [17] that does not affect Pol IV or Pol V [18] (Fig. 1b). In order to test the interaction between Pol II-mediated repression of TE activation and DNA methylation, we grew WT and *nrpb2-3* plants on media supplemented with zebularine (Z), an inhibitor of DNA methyltransferases active in plants [19], and subjected them to HS. To ensure the viability of the *nrpb2-3* seedlings we choose a moderate amount of Z (10 μM). The presence of Z in the medium during HS generally enhanced the production of *ONSEN* copies. Importantly, this induced increase in *ONSEN* copy number was more distinct in the *nrpb2-3* background (Fig. 1a). This indicated that both DNA methylation and Pol II transcriptional activity contribute to the repression of *ONSEN* ecDNA production. To complete their lifecycle, the reverse transcribed ecDNA of activated retrotransposons

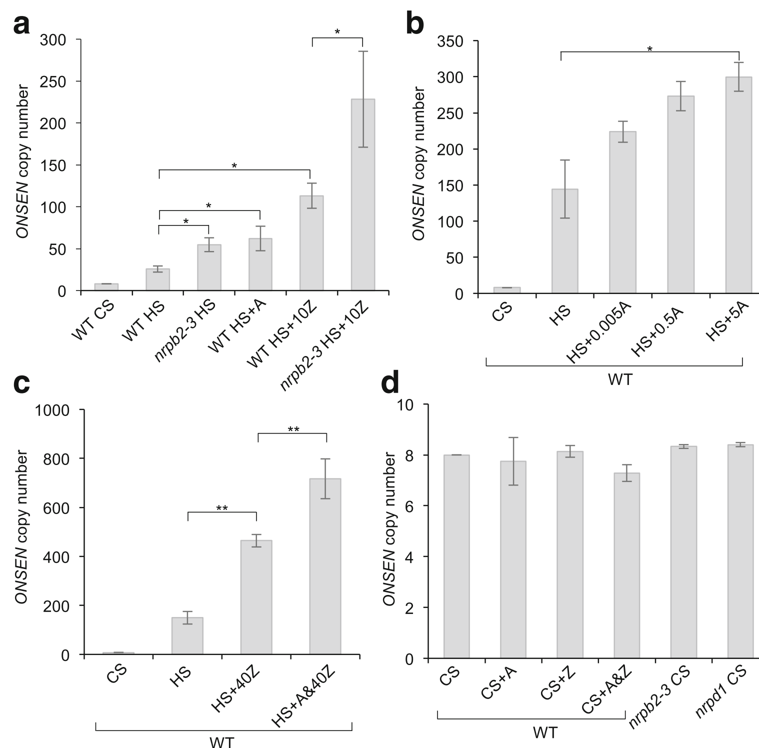


Fig. 1 Pol II represses the HS-dependent mobility of the *ONSEN* retrotransposon in *Arabidopsis*. *ONSEN* copy number in *Arabidopsis* seedlings measured by qPCR directly after CS and HS treatments. **a** In the WT and the *nrpb2-3* mutant and after HS plus treatments with α-amanitin (*A*; 5 μg/ml) or zebularine (*Z*; 10 μM) (mean ± standard error of the mean (s.e.m.), *n* = 6 biological repetitions). **b** In the WT and after HS plus treatment with A at different concentrations (μg/ml) as specified on the x-axis (mean ± s.e.m., *n* = 4 biological repetitions). **c** In the WT and after HS plus treatment with Z (40 μM) or a combination of A (5 μg/ml) and Z (*A&40Z*) (mean ± s.e.m., *n* = 3 biological repetitions). **d** In the WT after chemical treatment with A (5 μg/ml), Z (40 μM), a combination of A and Z (*A&Z*) or in the *nrpb2-3* and *nrpd1* backgrounds following CS (mean ± s.e.m., *n* = 3 biological repetitions). All values are relative to *ACTIN2*. *$P < 0.05$, **$P < 0.01$

has to integrate back into the genome [1]. Given that we observed a strong increase in *ONSEN* copy number after HS and treatment with moderate amounts of Z in the *nrpb2-3* background, we wanted to address the inheritance of additional *ONSEN* copies by the offspring. For this we compared the average *ONSEN* copy number of pooled S1 seedlings obtained from Z-treated and heat-stressed WT and *nrpb2-3* plants grown under controlled conditions on soil by qPCR. We observed a distinct increase in the overall *ONSEN* copy number exclusively in the *nrpb2-3* background (Additional file 1: Figure S1).

Because both DNA methylation and Pol II can be inhibited by the addition of specific drugs, we wanted to test if treating WT plants with both A and Z at the same time could strongly activate and even mobilize *ONSEN* after a HS treatment. We grew WT seedlings on MS medium supplemented with Z (40 μM) [19] individually or combined with A (5 μg/ml, A&Z). Consistent with the strong activation of *ONSEN* in HS and Z-treated *nrpb2-3* seedlings, the combined treatment (A&Z) of the WT gave rise to a very high (Fig. 1c) HS-dependent (Fig. 1d) increase in *ONSEN* copy number, comparable to that in the *nrpd1* background (Fig. 2e). We noted that the overall amplitude of HS-dependent *ONSEN* activation could vary between different waves of stress applications in terms of copy number (Fig. 1a, b). Yet, the observed enhancing effect of Pol II and DNA methyltransferase inhibition with A and Z on *ONSEN* activation was consistent in

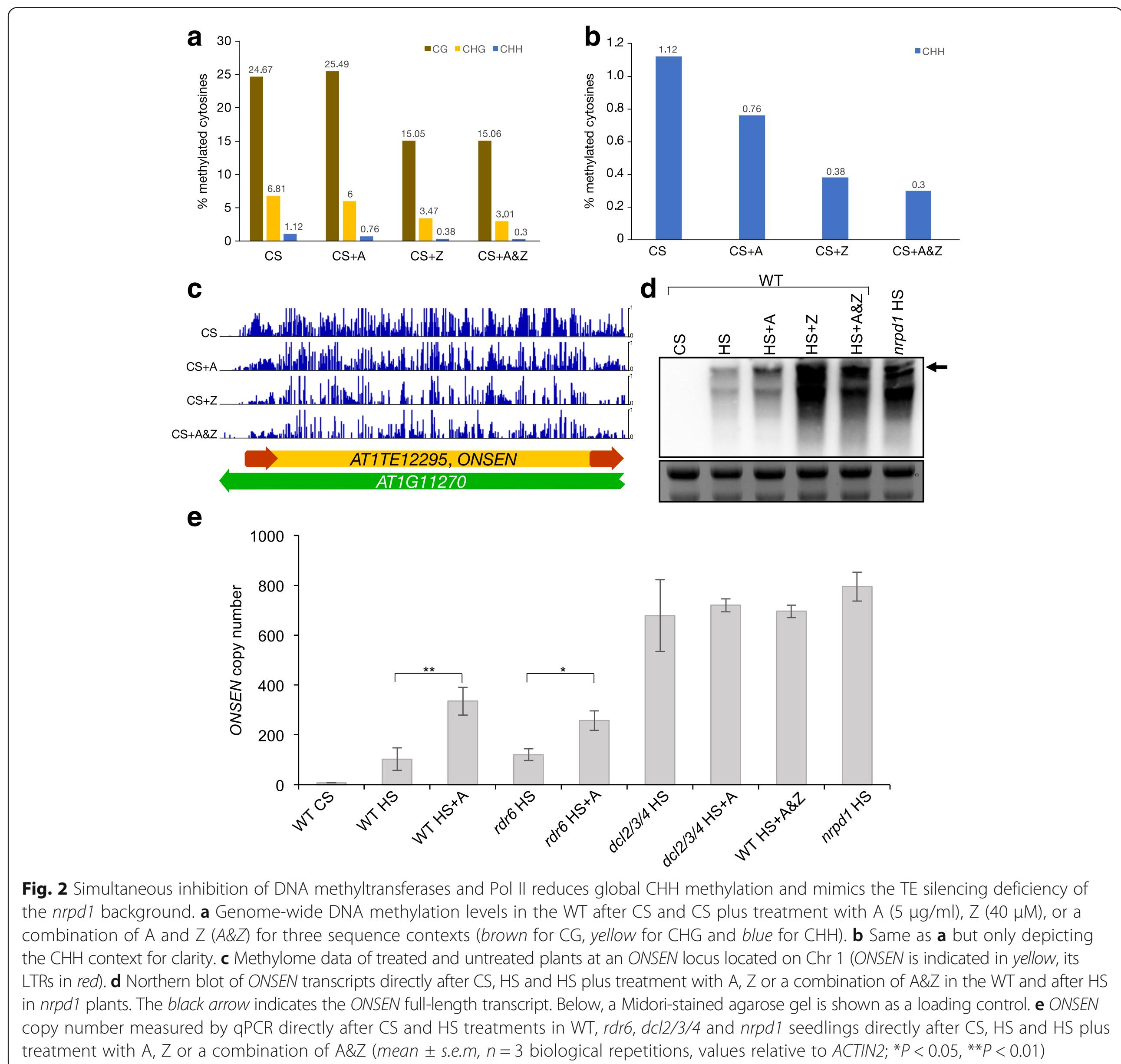


Fig. 2 Simultaneous inhibition of DNA methyltransferases and Pol II reduces global CHH methylation and mimics the TE silencing deficiency of the *nrpd1* background. **a** Genome-wide DNA methylation levels in the WT after CS and CS plus treatment with A (5 μg/ml), Z (40 μM), or a combination of A and Z (*A&Z*) for three sequence contexts (*brown* for CG, *yellow* for CHG and *blue* for CHH). **b** Same as **a** but only depicting the CHH context for clarity. **c** Methylome data of treated and untreated plants at an *ONSEN* locus located on Chr 1 (*ONSEN* is indicated in *yellow*, its LTRs in *red*). **d** Northern blot of *ONSEN* transcripts directly after CS, HS and HS plus treatment with A, Z or a combination of A&Z in the WT and after HS in *nrpd1* plants. The *black arrow* indicates the *ONSEN* full-length transcript. Below, a Midori-stained agarose gel is shown as a loading control. **e** *ONSEN* copy number measured by qPCR directly after CS and HS treatments in WT, *rdr6*, *dcl2/3/4* and *nrpd1* seedlings directly after CS, HS and HS plus treatment with A, Z or a combination of A&Z (*mean* ± *s.e.m*, $n = 3$ biological repetitions, values relative to *ACTIN2*; $^{*}P < 0.05$, $^{**}P < 0.01$)

independent experiments (Figs. 1a–c and 2e). To detect activated TEs at the genome-wide level we took advantage of the production of eccDNA by active retrotransposons. eccDNA is a byproduct of the LTR retrotransposon life cycle [20]. Using mobilome sequencing, which comprises a specific amplification step of circular DNA followed by high-throughput sequencing to identify eccDNA derived from active LTR retrotransposons [2], we found that only *ONSEN* was activated by HS in combination with A&Z (Additional file 1: Figure S2). Confirming our qPCR data, more *ONSEN*-specific reads were detected in the presence of A and Z in the medium.

To better understand the mechanisms by which the drugs enhanced the activation of *ONSEN* after HS at the DNA level, we assessed how they influenced DNA methylation at the genome-wide level using whole-genome bisulfite sequencing (WGBS) after CS. Overall, we found that all drug treatments affected global DNA methylation levels. While the treatment with Z affected all sequence contexts, we observed that inhibition of Pol II primarily affected cytosine methylation in the CHG and CHH sequence contexts (where H is an A, T or G). The combined A&Z treatment had a slight additive demethylating effect in the CHG and CHH contexts compared to A or Z alone (Fig. 2a, b). DNA methylation levels at one *ONSEN* locus (*AT1TE12295*) is depicted in Fig 2c. Treatment with A led to a slight decrease in DNA methylation, which was more apparent in Z- and A&Z-treated plants. We then checked by northern blot whether the degree of reduction in DNA methylation would coincide with increased *ONSEN* transcript levels directly after HS. We found that treatment with Z alone resulted in the highest *ONSEN* transcript level after HS (Fig. 2d). Considering the data obtained on *ONSEN* ecDNA (Fig. 1c), we concluded that a substantial proportion of these Z-induced transcripts were not suitable templates for *ONSEN* ecDNA synthesis.

In *Drosophila*, it has been shown that Pol II-mediated antisense transcription results in the production of TE-derived siRNAs in a Dicer-2-dependent manner [21]. In support of this in *Arabidopsis*, a recent publication pointed out the importance of DCL3 in regulating *ONSEN* in the *ddm1* background [16]. To elucidate whether the effect of Pol II inhibition was also dicer-dependent, we grew both *rdr6* and *dcl2/3/4* triple mutant plants on A, applied HS and measured *ONSEN* ecDNA levels. Strikingly, we found that A still enhanced ecDNA accumulation in *rdr6* plants, whereas inhibition of Pol II had no additional effect in the *dcl2/3/4* triple mutant (Fig. 2e).

Induced mobilization of endogenous TEs in plants has so far been very inefficient, thus limiting their use in basic research and plant breeding [3]. In the case of *Arabidopsis*, transposition of *ONSEN* in HS-treated WT plants has not been observed [11, 22]. Because the A&Z drug treatment resulted in high accumulation of *ONSEN* copy numbers—essentially mimicking plants defective in NRPD1 (Fig. 2e)—we wanted to test if the combined drug treatment could lead to efficient *ONSEN* mobilization in WT plants. First, we assessed by qPCR if, and at what frequencies, new *ONSEN* copies could be detected in the progeny of A&Z-treated and heat stressed plants. In fact, we found new *ONSEN* insertions in 29.4% of the tested S1 (selfed first generation) pools (n = 51), with pools having up to 52 insertions (Additional file 1: Figure S3). We then confirmed stable novel *ONSEN* insertions in a subset of independent individual high copy plants by transposon display (Fig. 3a), qPCR (Fig. 3b) and sequencing of 11 insertions in a selected high-copy line (hc line 3; Fig. 4; Additional file 1: Figure S4). Tracking *ONSEN* copy numbers over three generations of selfing indicated that the new insertions were stably inherited (Fig. 3b). Furthermore, the re-application of heat stress and drugs in the S3 generation of two hc lines did not lead to greater accumulation of *ONSEN* copies compared to control lines, but we instead observed stronger silencing in lines with more *ONSEN* copies (Additional file 1: Figure S5).

TE insertions can interrupt genes or alter their expression by recruiting epigenetic marks or by stress-dependent readout transcription from the 3′ LTR into flanking regions [6]. To test this, we grew the S2 generation of the selected hc lines under long- and short-day conditions. Interestingly, we observed that many hc lines showed clear and homogenous phenotypes in response to the different growth conditions (plant size, chlorophyll content and flowering time; Fig. 3c, d).

To demonstrate that *ONSEN* insertions could directly influence such developmental phenotypes, we closely investigated hc line 3, which produced white seeds (Fig. 4a). Using a candidate gene approach, we found that an *ONSEN* insertion in *transparent testa 6* (*TT6*, *AT3G51240*; Fig. 4b) was responsible for the recessive white seed phenotype [23, 24]. This was confirmed by segregation analysis of the F2 generation of a cross between WT and hc line 3 (Fig. 4a) followed by genotyping (Fig. 4c).

Next, we wanted to test if Pol II plays a more general role in repressing TEs in plants. Due to its significantly different epigenetic and TE landscape compared to *Arabidopsis*, we wanted to test if we could mobilize TEs in rice (*Oryza sativa*) [25], a genetically well-characterized monocotyledonous crop. To capture drug-induced mobilized TEs, we characterized the active mobilome in *O. sativa* seedlings that were grown on MS medium supplemented with no drugs, A only, Z only or a combination of A and Z, using the same approach as we used for *Arabidopsis*. We identified *Houba*, a copia-like retrotransposon [26], as highly activated only when plants were treated with A&Z (Fig. 5a). Bona fide activity of

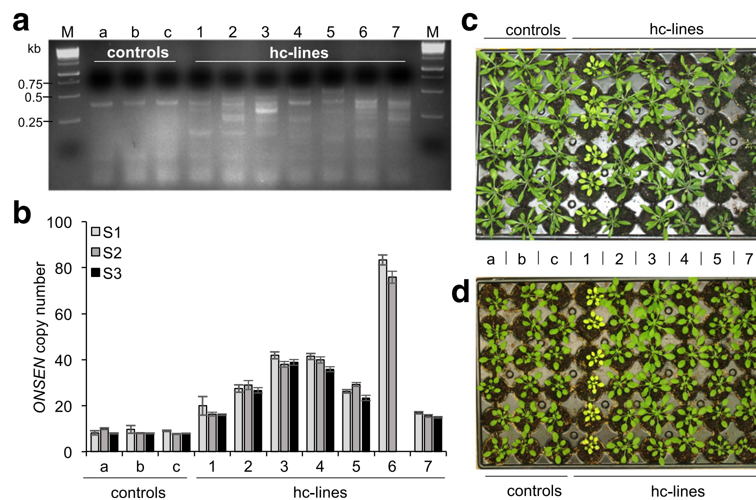


Fig. 3 Drug-induced mobilization of *ONSEN* in WT *Arabidopsis* plants. **a** Transposon display testing seedlings in the S2 generation of WT plants for novel *ONSEN* insertions: lanes a to c show HS-treated plants; lanes 1 to 7 show hc lines 1–7 treated with HS and A (5 μg/ml) and Z (40 μM), M indicates the size marker. **b** *ONSEN* copy number in the S1, S2 and S3 generations measured by qPCR (mean ± s.e.m, $n = 3$ technical replicates, values relative to *ACTIN2*). **c**, **d** Photographs of S2 plants showing both homogeneous and environment-dependent phenotypic variability induced by the *ONSEN* mobilization when grown under long (**c**) and short day (**d**) conditions. qPCR data for the S3 generation of line 6 in **b** as well as pictures of phenotypes in **c** and **d** are missing due to severe infertility and extinction of this line

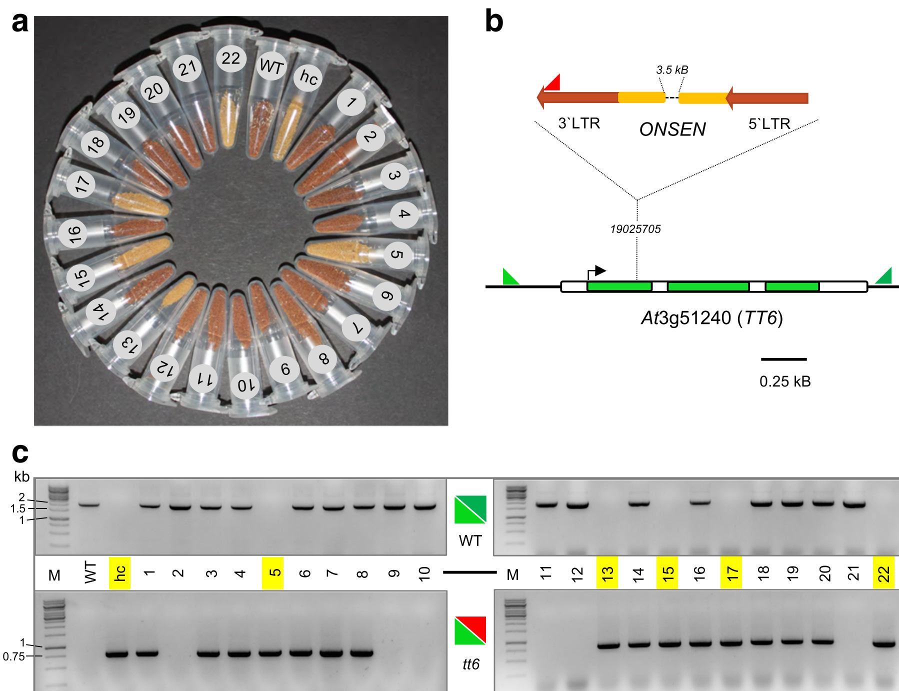


Fig. 4 *Transparent testa* phenotype of hc line 3 co-segregates with an *ONSEN* insertion in *TT6*. Seed phenotypes (**a**) and corresponding genotypes (**c**) of a segregating F2 population (lanes 1–22) obtained from a cross between the WT and hc line 3 (*hc*) are shown. **b** Primers used for genotyping of the *ONSEN* insertion. For the WT-PCR depicted in the upper part of **c** the light (*tt6 fw*) and dark (*tt6 rev*) green primers flanking the *TT6* locus (*AT3G51240*) were used. The *ONSEN* insertion in *TT6* was detected by a combination of the light green primer with the red primer specific to the *ONSEN* LTR (*Copia* 78 3′ LTR, *red arrow*). M indicates the size marker. Primer sequences are given in Additional file 1: Table S1

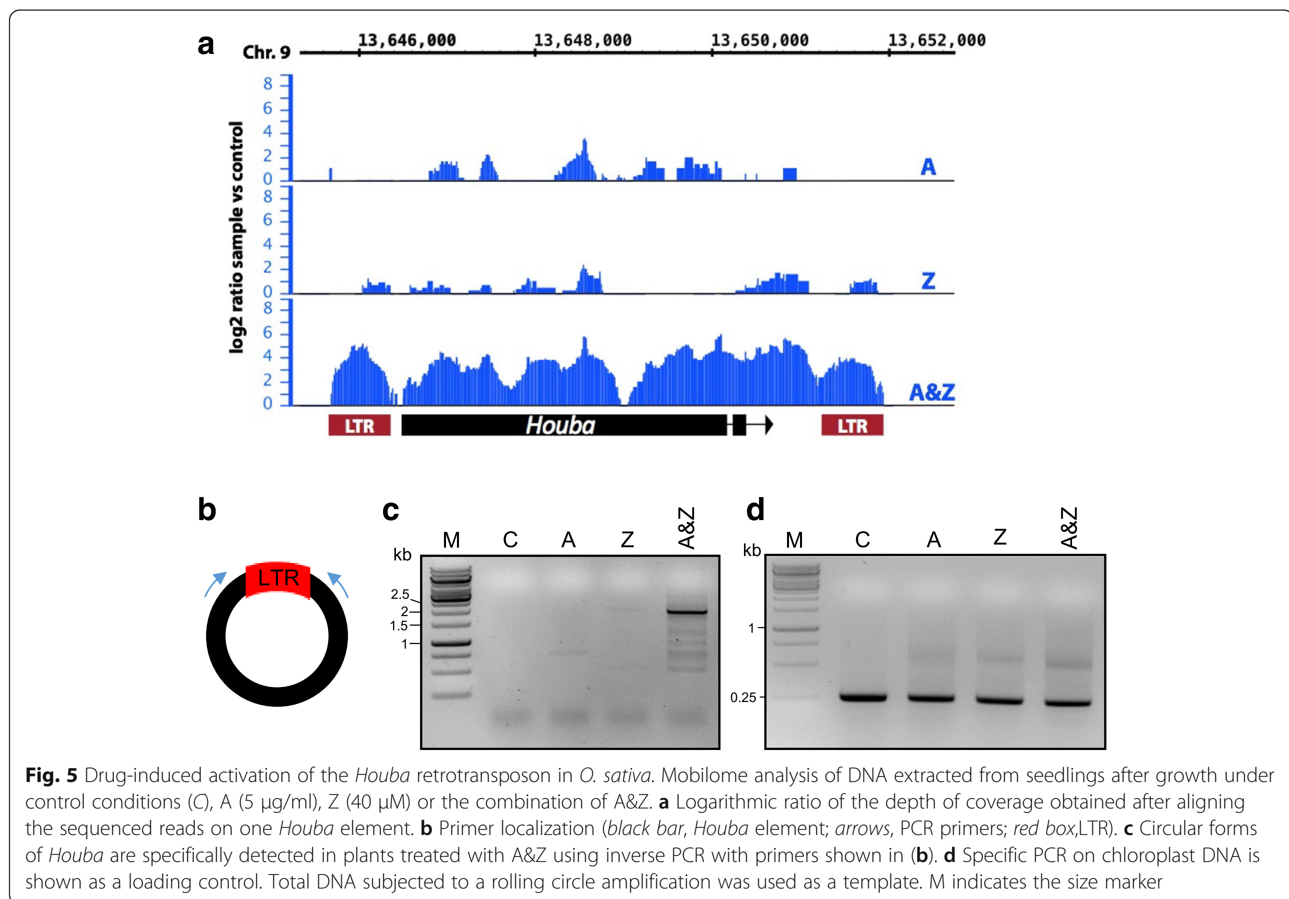


Fig. 5 Drug-induced activation of the *Houba* retrotransposon in *O. sativa*. Mobilome analysis of DNA extracted from seedlings after growth under control conditions (*C*), A (5 μg/ml), Z (40 μM) or the combination of A&Z. **a** Logarithmic ratio of the depth of coverage obtained after aligning the sequenced reads on one *Houba* element. **b** Primer localization (*black bar*, *Houba* element; *arrows*, PCR primers; *red box*,LTR). **c** Circular forms of *Houba* are specifically detected in plants treated with A&Z using inverse PCR with primers shown in (**b**). **d** Specific PCR on chloroplast DNA is shown as a loading control. Total DNA subjected to a rolling circle amplification was used as a template. M indicates the size marker

Houba was supported by the detection of eccDNA containing LTR–LTR junctions (Additional file 1: Figure S6). The activation of *Houba* was further confirmed by eccDNA-specific PCR on the *Houba* circles (Fig. 5b–d).

Discussion

In this study, we show the importance of Pol II in the repression of TE mobility in plants. By choosing the well-characterized heat inducible *ONSEN* retrotransposon, we were able to specifically address the role of Pol II in silencing transcriptionally active endogenous TEs in WT plants. Recent studies propose Pol II as the primary source for the production of TE-silencing signals that can then feed into the RNA silencing and DNA methylation pathways [15]. Our data strongly support these findings at two levels. First, we found that inhibition of Pol II activity reduced the degree of DNA methylation at *ONSEN*, demonstrating its distinct role in this process, and that Pol II also contributes to reinforcing silencing at the genome-wide level, primarily in the CHH but also in the CHG context. Second, our finding that DCL enzymes are sufficient to process the silencing signal produced by Pol II suggest that Pol II acts at very early steps in the TE silencing pathway by providing substrates to these enzymes. The observation that inhibition of Pol II in the *rdr6* background still further enhanced *ONSEN* accumulation after HS supports the notion that Pol II plays a central role in the previously proposed expression-dependent RdDM pathway [16].

Using mobilome sequencing we confirmed previous findings [2] that this approach is a powerful diagnostic tool to detect mobile retrotransposons: we detected highest levels of eccDNA of *ONSEN* in HS and drug-treated *Arabidopsis* seedlings and found new insertions in successive generations of these plants. Using the same approach on rice we were able to detect production of *Houba* eccDNA after drug treatments, suggesting that the progeny will then contain novel *Houba* insertions. This is still to be confirmed and may be hampered by the already very high *Houba* copy number present in the genome [27].

Our findings may indicate that Pol II is primarily involved in silencing young, recently active retrotransposons and perhaps to a lesser extent other tightly silenced TEs. Indeed, there are indications of very recent natural transposition events for *ONSEN* [28] and *Houba* [29] in the *Arabidopsis* and rice genomes, respectively. For instance, the annual temperature range has and may still contribute to contrasting *ONSEN* mobilization events in different *Arabidopsis* accessions [28]. *Houba* is the most

abundant TE of the *copia* family in rice and has been active in the last 500,000 years [30].

Overall, our findings lead to the question of when plants lower their guard: under what conditions could Pol II be less effective in silencing TEs? Certain stresses that affect the cell cycle have been reported to lead to the inactivation of Pol II [31, 32]; this would provide a window of opportunity for TEs to be mobilized. Therefore, combined stresses that affect the cell cycle and activate TEs may lead to actual TE bursts under natural growth conditions. Interestingly, it has been reported that retrotransposon-derived short interspersed element (SINE) transcripts can inhibit Pol II activity [33]. This strongly suggests the presence of an ongoing arms race between retrotransposons and Pol II. Considering that almost all organisms analyzed so far have TEs [4] and RNA polymerases [34] and the reliance of TEs on host RNA polymerases, it may—from an evolutionary point of view—not come as a surprise that Pol II also has a function as an important regulator of retrotransposon activity. Strikingly, it has been shown in both *Saccharomyces cerevisiae* and *Drosophila melanogaster* that Pol II-dependent intra-element antisense transcription plays an important role in TE silencing [21, 35]. In addition, we observed a discrepancy in *ONSEN* transcript accumulation and measured ecDNA after HS in seedlings that were treated with zebularine only. This substantiates the notion that both the quantity and quality of transcripts affect regulation, reverse transcription and successful integration of retrotransposons. This is well in line with previous observations demonstrating that different TE-derived transcripts have distinct functions in the regulation of TE activity [36]. As a next step it will be of great interest to investigate if Pol II-dependent antisense transcription of TEs and subsequent dicer-dependent processing may be the key to solve “the chicken and the egg problem” of *de novo* silencing functional retrotransposons in eukaryotes.

Finally, our findings will allow future studies on the potential beneficial role TEs play in adaptation to stresses. Indeed, two recent studies point out the adaptive potential of retrotraonsposon and, more specifically, *ONSEN* copy number variation in natural accessions [28] and RdDM mutant backgrounds of *Arabidopsis* [37]. Upon mobilization, the heat-response elements in the LTRs of *ONSEN* [38] can create new gene regulatory networks responding to heat stress [11]. Therefore, it will now be of great interest to test if the *ONSEN* hc lines obtained in this study are better adapted to heat stress. This will allow us to test if retrotransposon-induced genetic and epigenetic changes more rapidly create beneficial alleles than would occur by random mutagenesis. Furthermore, the observation that HS did not lead to a stronger activation of *ONSEN* in hc lines compared to WT plants suggests that genome stability is not compromised in these lines. This result can be explained by at least two possible mechanisms: (i) the occurrence of insertions of inverted duplications of *ONSEN*, such as has been observed for the *Mu killer* locus in maize [39]—such insertions will lead to the production of double-stranded RNA feeding into gene silencing and thereby limit the activity of that TE; and (ii) balancing of TE activity and integrated copy number as has been described for *EVADE* in *Arabidopsis* [40]. In this case, when a certain TE copy number threshold is reached robust transcriptional gene silencing takes over, thereby limiting TE mobility and ensuring genome stability. The stability of new TE insertions is an important aspect in light of the future use of TEs in crop breeding and trait stability.

Conclusions

TEs are important contributors to genome evolution. The ability to mobilize them in plants and possibly in other eukaryotes in a controlled manner with straightforward drug application, as shown here, opens the possibility to study their importance in inducing genetic and epigenetic changes resulting from external stimuli. Because the induced transposition of *ONSEN* can efficiently produce developmental changes in *Arabidopsis*, it will be very interesting to test if specific stress-induced TE activation can be used for directed crop breeding for better stress tolerance in the near future.

Methods

Plant material

All *Arabidopsis* mutants used in this study (*nrpb2-3* [14], *nrpd1-3* [41], *rdr6* [42], *dcl2/3/4* triple mutant [43]) are in the Col-0 background. For *O. sativa japonica*, the cultivar Nipponbare was used.

Growth conditions

Prior to germination, *Arabidopsis* seeds were stratified for 2 days at 4 °C. Before and during stress treatments plants were grown under controlled conditions in a Sanyo MLR-350 growth chamber on solid ½ MS medium (1% sucrose, 0.5% Phytagel (Sigma), pH 5.8) under long day conditions (16 h light) at 24 °C (day) and 22 °C (night) (*Arabidopsis*) and 12 h at 28 °C (day) and 27 °C (night) (*O. sativa*).

To analyze successive generations, seedlings were transferred to soil and grown under long day conditions (16 h light) at 24 °C (day) and 22 °C (night) (*Arabidopsis)* in a Sanyo MLR-350 growth chamber until seed maturity.

For phenotyping, *Arabidopsis* plants were grown under long day conditions (16 h light) at 24 °C (day) and 22 °C (night) and short day conditions (10 h light) at 21 °C (day) and 18 °C (night).

Stress and chemical treatments

Surface sterilized seeds of *Arabidopsis* and *O. sativa* were germinated and grown on solid ½ MS medium that was supplemented with sterile filtered zebularine (Sigma; stock, 5 mg/ml in DMSO), α-amanitin (Sigma; stock, 1 mg/ml in water) or a combination of both chemicals. Control stresses (6 °C for 24 h followed by control conditions for 24 h, CS) and heat stresses (6 °C for 24 h followed by 37 °C for 24 h, HS) of *Arabidopsis* seedlings were conducted as described previously [11].

DNA analysis

For qPCR and prior to digestions, total DNA from *Arabidopsis* plants was extracted with the DNeasy Plant Mini Kit (Qiagen) following the manufacturer's recommendations. For the qPCRs to measure the *ONSEN* copy number following HS and chemical treatments the aerial parts of at least ten *Arabidopsis* plants per replicate were pooled prior to DNA extraction. To track *ONSEN* copy numbers in the S1–3 generations of controls (only HS) and hc lines (HS + A&Z treatment) DNA from true leaves was extracted. For the estimation of the *ONSEN* transposition frequency, total DNA of pools consisting of at least eight seedlings of the progeny of HS + A&Z-treated plants was isolated. The DNA concentration was measured with a Qubit Fluorometer (Thermo Fisher Scientific). The copy numbers of *ONSEN* were determined with qPCRs on total DNA using a TaqMan master mix (Life Technologies) in a final volume of 10 μl in the Light-Cycler 480 (Roche). *ACTIN2* (*AT3G18780*) was used to normalize DNA levels. Primer sequences are given in Additional file 1: Table S1.

For the mobilome-seq analysis total DNA from the pooled aerial parts of three 10-day-old *O. sativa* seedlings was extracted as previously reported [44]. Genomic DNA (5 μg) for each sample was purified using a Geneclean kit (MPBio, USA) according to the manufacturer's instructions. ecDNA was isolated from the GeneClean product using PlasmidSafe DNase (Epicentre, USA) according to the manufacturer's instructions, except that the 37 °C incubation was performed for 17 h. DNA samples were precipitated by adding 0.1 volume of 3 M sodium acetate (pH 5.2), 2.5 volumes of ethanol and 1 μl of glycogen (Fisher, USA) and incubating overnight at −20 °C. The precipitated circular DNA was amplified by random rolling circle amplification using the Illustra TempliPhi kit (GE Healthcare, USA) according to the manufacturer's instructions except that the incubation was performed for 65 h at 28 °C. The DNA concentration was determined using the DNA PicoGreen kit (Invitrogen, USA) using a LightCycler480 (Roche, USA). One nanogram of amplified ecDNA from each sample was used to prepare the libraries using the Nextera XT library kit (Illumina, USA) according to the manufacturer's instructions. DNA quality and concentration were determined using a high sensitivity DNA Bioanalyzer chip (Agilent Technologies, USA). Samples were pooled and loaded onto a MiSeq platform (Illumina, USA) and 2 × 250-nucleotide paired-end sequencing was performed. Quality control of FASTQ files was done using the FastQC tool (version 0.10.1). To remove any read originating from organelle circular genomes, reads were mapped against the mitochondria and chloroplast genomes using the program Bowtie2 version 2.2.2 71 with –sensitive local mapping. Unmapped reads were mapped against the reference genome IRGSP1.0 (http://rgp.dna.affrc.go.jp/E/IRGSP/Build5/build5.html) using the following parameters: –sensitive local, -k 1. DNA from both mitochondria and chloroplast genomes integrated in nuclear genomes was masked (1,697,400 bp). The TE-containing regions cover 194,224,800 bp in *O. sativa*. Finally, the bam alignment files were normalized and compared using deeptools [45] and visualized with the Integrative Genomics Viewer (IGV) software (https://www.broadinstitute.org/igv/). Data from the mobilome analysis were submitted to GEO (accession number GSE90484).

The presence of circular *Houba* copies was tested by an inverse PCR on 7 ng of the rolling-circle amplified template that was also used for sequencing. A PCR specific to chloroplast DNA served as a loading control. PCR products were separated on a 1% agarose gel that was stained with a Midori Green Nucleic Acid Staining Solution (Nippon Genetics Europe). Primer sequences are given in Additional file 1: Table S1.

Transposon display

The integration of additional copies of *ONSEN* into the genome of heat stressed and treated plants was ascertained by a simplified transposon display based on the GenomeWalker Universal kit (Clontech Laboratories), as previously described [11] with the following modifications: 300 ng of total DNA from adult plants in the S2 generation of heat stressed and A&Z-treated plants was extracted with a DNeasy Plant Mini Kit (QIAGEN) and digested with blunt cutter restriction enzyme *Dra*I (NEB). After purification with a High Pure PCR Product Purification Kit (Roche) digested DNA was ligated to the annealed GenWalkAdapters 1&2. The PCR was performed with the adaptor-specific primer AP1 and the *ONSEN*-specific primer Copia78 3′ LTR. The PCR products were separated on a 2% agarose gel that was stained with Midori Green. For primer sequence information, see Additional file 1: Table S1.

Cloning, sequencing and genotyping of new insertions

To identify the genomic region of new *ONSEN* insertions, the PCR product of the transposon display was purified using a High Pure PCR Product Purification Kit

(Roche), ligated into a pGEM-T vector (Promega) and transformed into *Escherichia coli*. After a blue white selection, positive clones were used for the insert amplification and sequencing (StarSEQ). The obtained sequences were analyzed with Geneious 8.2.1 and blasted against the *Arabidopsis* reference genome. The standard genotyping PCRs to prove novel *ONSEN* insertions were performed with combinations of the *ONSEN*-specific primer Copia78 3′ LTR and primers listed in Additional file 1: Table S1.

RNA analysis and northern blotting

Total RNA from the aerial part of at least ten *Arabidopsis* seedlings was isolated using the TRI Reagent (Sigma) according to the manufacturer's recommendations. RNA concentration was measured (Qubit RNA HS Assay Kit, Thermo Fisher) and 15 μg of RNA was separated on a denaturing 1.5% agarose gel, blotted on a Hybond-N^+ (GE Healthcare) membrane and hybridized with 25 ng of a gel-purified and P^{32}-labelled probe (Megaprime DNA Labelling System, GE Healthcare) specific to the full length *ONSEN* transcript (see Additional file 1: Table S1 for primer sequences). Northern blots were repeated in three independent experiments with the same results.

Whole-genome DNA methylation analysis

Whole-genome bisulfite sequencing library preparation and DNA conversion were performed as previously reported [46]. Bisulphite read mapping and methylation value extraction were done on the *Arabidopsis* TAIR10 genome sequence using BSMAP v2.89 [47]. Following mapping of the reads the fold coverages of the genome for CS, CS + A, CS + Z and CS + A&Z were 13.4, 13.2, 18.4 and 16.3, respectively. Data from the bisulphite sequencing analysis have been submitted to GEO (accession number GSE99396).

Statistics

Statistical analyses were performed with SigmaPlot (v. 11.0). Depending on the normality of the data, either an H-test or a one-way ANOVA was performed. The Student-Newman-Keuls method was used for multiple comparisons.

Additional file

Additional file 1: Table S1. Table of all primers used in this study. **Figure S1.** Increase in *ONSEN* copy numbers in S1 pools of heat-stressed and Z-treated *nrpb2-3* plants. **Figure S2.** Detection of eccDNAs originating from ONSEN loci following heat stress and chemical treatments in *Arabidopsis*. **Figure S3.** Increase in ONSEN copy numbers in S1 pools of heat-stressed and A&Z-treated WT plants. **Figure S4.** Summary of confirmed novel *ONSEN* insertions in hc line 3. **Figure S5.** Stress-induced activation of *ONSEN* in the S3 generation after initial HS treatment. **Figure S6.** *Houba* forms LTR–LTR junction eccDNAs after combined A&Z treatment. (PDF 1660 kb)

Acknowledgements
We wish to thank Emilija Hristova for her support at the beginning of this project. We thank Christel Llauro for technical support on the production of the mobilomes and Todd Blevins for providing the *dcl*2/3/4 triple mutant line. We thank the IMAC platform from the Structure Fédérative de Recherche 'Qualité et Santé du Végétal' (SFR QUASAV) for their technical support (Illumina sequencing).

Funding
This work was supported by grants provided by the European Commission (PITN-GA-2013-608422–IDP BRIDGES to MT and ERC grant 725701 BUNGEE to EB) and the region of Pays de la Loire (ConnecTalent EPICENTER project awarded to EB).

Availability of data and materials
The mobilome sequencing data and whole-genome DNA methylation analysis data are available in the GEO (accession numbers GSE90484 and GSE99396, respectively).

Authors' contributions
MT and EB conceived the study. MT, SL, SB and MM performed experiments. ND performed methylome analyses. MT and EB wrote the paper with contributions from SL and MM. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Competing interests
MT and EB declare that a patent application based on the presented discoveries has been submitted to the European Patent Office (PCT/EP2016/079276). EB is CEO of epibreed Ltd, a company that has an exclusive use license for this patent.

Author details
[1]Botanical Institute, Zürich-Basel Plant Science Center, University of Basel, Hebelstrasse 1, 4056 Basel, Switzerland. [2]Institut de Recherche pour le Développement, UMR232 DIADE Diversité Adaptation et Développement des Plantes, Université Montpellier 2, Montpellier, France. [3]University of Perpignan, Laboratory of Plant Genome and Development, 58 Avenue Paul Alduy, 66860 Perpignan, France. [4]IRHS, Université d'Angers, INRA, AGROCAMPUS-Ouest, SFR4207 QUASAV, Université Bretagne Loire, 49045 Angers, France.

Received: 13 February 2017 Accepted: 27 June 2017
Published online: 07 July 2017

References

1. Schulman AH. Retrotransposon replication in plants. Curr Opin Virol. 2013;3:604–14.
2. Lanciano S, Carpentier M-C, Llauro C, Jobet E, Robakowska-Hyzorek D, Lasserre E, et al. Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. PLoS Genet. 2017;13:e1006630.
3. Paszkowski J. Controlled activation of retrotransposition for plant breeding. Curr Opin Biotechnol. 2015;32C:200–6.
4. Huang CRL, Burns KH, Boeke JD. Active transposition in genomes. Annu Rev Genet. 2012;46:651–75.
5. Belyayev A. Bursts of transposable elements as an evolutionary driving force. J Evol Biol. 2014;27:2573–84.
6. Lisch D. How important are transposons for plant evolution? Nat Rev Genet. 2013;14:49–61.
7. Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T. Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. Nature. 2001;411:212–4.
8. Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet. 2014;15:394–408.
9. Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. Bursts of retrotransposition reproduced in Arabidopsis. Nature. 2009;461:423–U125.
10. Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, et al. Selective epigenetic control of retrotransposition in Arabidopsis. Nature. 2009;461:427–30.

11. Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature. 2011;472:115–9.
12. Matzke MA, Kanno T, Matzke AJM. RNA-directed DNA methylation: the evolution of a complex epigenetic pathway in flowering plants. Annu Rev Plant Biol. 2015;66:243–67.
13. Gao Z, Liu H-L, Daxinger L, Pontes O, He X, Qian W, et al. An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. Nature. 2010;465:106–9.
14. Zheng B, Wang Z, Li S, Yu B, Liu J-Y, Chen X. Intergenic transcription by RNA Polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. Genes Dev. 2009;23:2850–60.
15. Cuerda-Gil D, Slotkin RK. Non-canonical RNA-directed DNA methylation. Nature Plants. 2016;2:16163.
16. Panda K, Ji L, Neumann DA, Daron J, Schmitz RJ, Slotkin RK. Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. Genome Biol. 2016;17:1–19.
17. Lindell TJ, Weinberg F, Morris PW, Roeder RG, Rutter WJ. Specific inhibition of nuclear RNA polymerase II by alpha-Amanitin. Science. 1970;170:447–9.
18. Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolić L, et al. In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. Mol Cell. 2012;48:811–8.
19. Baubec T, Pecinka A, Rozhon W, Mittelsten SO. Effective, homogeneous and transient interference with cytosine methylation in plant genomic DNA by zebularine. Plant J. 2009;57:542–54.
20. Flavell AJ, Ish-Horowicz D. Extrachromosomal circular copies of the eukaryotic transposable element Copia in cultured Drosophila cells. Nature. 1981;292:591–5.
21. Russo J, Harrington AW, Steiniger M. Antisense transcription of retrotransposons in Drosophila: an origin of endogenous small interfering RNA precursors. Genetics. 2016;202:107–21.
22. Matsunaga W, Ohama N, Tanabe N, Masuta Y, Masuda S, Mitani N, et al. A small RNA mediated regulation of a stress-activated retrotransposon and the tissue specific transposition during the reproductive period in Arabidopsis. Front Plant Sci. 2015;6:48.
23. Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B. An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol Biol. 2003;53:247–59.
24. Appelhagen I, Thiedig K, Nordholt N, Schmidt N, Huep G, Sagasser M, et al. Update on transparent testa mutants from Arabidopsis thaliana: characterisation of new alleles from an isogenic collection. Planta. 2014;240:955–70.
25. Kawahara Y, la Bastide de M, Hamilton JP, Kanamori H, Mccombie WR, Ouyang S. Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. Rice. 2013;6:4–10.
26. Panaud O, Vitte C, Hivert J, Muziak S, Talag J, Brar D, et al. Characterization of transposable elements in the genome of rice (Oryza sativa L.) using Representational Difference Analysis (RDA). Mol Genet Genomics. 2002;268:113–21.
27. Vitte C, Panaud O. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. Cytogenet Genome Res. 2005;110:91–107.
28. Quadrana L, Silveira AB, Mayhew GF, LeBlanc C, Martienssen RA, Jeddeloh JA, et al. The Arabidopsis thaliana mobilome and its impact at the species level. elife. 2016;5:e15716.
29. Vitte C, Panaud O, Quesneville H. LTR retrotransposons in rice (Oryza sativa, L.): recent burst amplifications followed by rapid DNA loss. BMC Genomics. 2007;8:218–15.
30. Wicker T, Keller B. Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and Arabidopsis reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. Genome Res. 2007;17:1072–81.
31. Oelgeschläger T. Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. J Cell Physiol. 2001;190:160–9.
32. Palancade B, Bensaude O. Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. Eur J Biochem. 2003;270:3859–70.
33. Pai DA, Kaplan CD, Kweon HK, Murakami K, Andrews PC, Engelke DR. RNAs nonspecifically inhibit RNA polymerase II by preventing binding to the DNA template. RNA. 2014;20:644–55.
34. Lazcano A, Fastag J, Gariglio P, Ramirez C, Oro J. On the early evolution of RNA-polymerase. J Mol Evol. 1988;27:365–76.
35. Berretta J, Pinskaya M, Morillon A. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. Genes Dev. 2008;22:615–26.
36. Chang W, Jääskeläinen M, Li S-P, Schulman AH. BARE retrotransposons are translated and replicated via distinct RNA pools. PLoS One. 2013;8:e72270–12.
37. Ito H, Kim J-M, Matsunaga W, Saze H, Matsui A, Endo TA, et al. A Stress-activated transposon in Arabidopsis induces transgenerational abscisic acid insensitivity. Sci Rep. 2016;6:23181.
38. Pietzenuk B, Markus C, Gaubert H, Bagwan N, Merotto A, Bucher E, et al. Recurrent evolution of heat-responsiveness in Brassicaceae COPIA elements. Genome Biol. 2016;17:209.
39. Slotkin R, Freeling M, Lisch D. Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. Nat Genet. 2005;37:641–4.
40. Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A. Reconstructing de novo silencing of an active plant retrotransposon. Nat Genet. 2013;45:1029–39.
41. Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. Science. 2005;308:118–20.
42. Peragine A, Yoshikawa M, Wu G, Albrecht H, Poethig R. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes Dev. 2004;18:2368–79.
43. Blevins T, Pontes O, Pikaard CS, Meins F. Heterochromatic siRNAs and DDM1 independently silence aberrant 5S rDNA transcripts in Arabidopsis. PLoS One. 2009;4:e5932–2.
44. Mette MF, van der Winden J, Matzke MA, Matzke AJ. Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters in trans. EMBO J. 1999;18:241–8.
45. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 2014;42:W187–91.
46. Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, et al. Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. Nature. 2011;480:245–9.
47. Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics. 2009;10:232.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at www.biomedcentral.com/submit

Supplemental Table 1. Names purpose and sequences of primers used in this study.

Name	Sequence 5`->3`	Experiment	
GenWalkAdaptator1	GTAATACGACTCACTATAGGGCACGCGT GGTCGACGGCCCGGGCTGGT	Transposon Display	
GenWalkAdaptator2	(PHOS) ACCAGCCC (AMINO)		
AP1	GTAATACGACTCACTATAGGGC		
Copia78 3`LTR	AACACTTAAACACTTTCTCCA		
284 COPIA78-4219F_RT	CCACAAGAGGAACCAACGAA	qPCR	
285 COPIA78-4219R_RT	TTCGATCATGGAAGACCGG		
ONSEN probe	(FAM) AAG TCG GCA ATA GCT TTG GCG AAG A (BHQ1)		
ACT2_QT_F	TGCCAATCTACGAGGGTTTC		
ACT2_QT_R	TTACAATTTCCCGCTCTGCT		
ACT2_QT_probe	(JOE) TCCGTCTTGACCTTGCTGGACG (BHQ-1)		
OnsenFull_F	AAGTGGTATCAGAGCTTGAAGATCC	Northern blot	
OnsenFull_R	CAACACCCCCTCTTAAACTTGATTTTGC		
M13F	CGCCAGGGTTTTCCCAGTCACGAC	Cloning and sequencing	
M13R	TCACACAGGAAACAGCTATGAC		
houba_F2	ATCCTGGGAAGAACAAACCATTAA	PCR on circular rice TE and the chloroplast control	
houba_R2	GAGTTCGAGTACCTTAGCCATGGT		
Chloroplast cyc F	ACAACCACTGATGAAGGATT		
Chloroplast cyc R	AGAAAGAAAAGCAACGACTG		
Prove TED 2_20 R	ACCTAGCTCTGAGTGATGAA	# 1	Genotyping of novel *ONSEN* insertions
Prove TED4_27 F	TGGATATACACATTGGTTGCA	# 2	
Prove TED 2_19 F	GGAGAAAGCTGAAAACTTGG	# 3	
Prove TED4_30_rev	CTAGGTTGGTGACTGATGAG	# 4	
Prove TED 2_17 F	AAGAATGGGAGCAGCATTAA	# 5	
Prove3_2R	GCAGTACTATAACCGGGACT	# 6	
prove TED3_1 Fw	GAACTTTCCGTTGTTACCGG	# 7	
Prove TED3 F	ATGAGACAGGGAGCTTATCT	# 8	
Prove TED1 R	GGTGTGAACCGAACCTAAAT	# 9	
Prove TED 4_25 F	AAACACCAGAAATCTTTCGC	# 10	
tt6 fw	CACAGACCACAAGCATTTTT	TT6 gene	
tt6 rev	TGTCGATTTTCTTGGTGCTA		

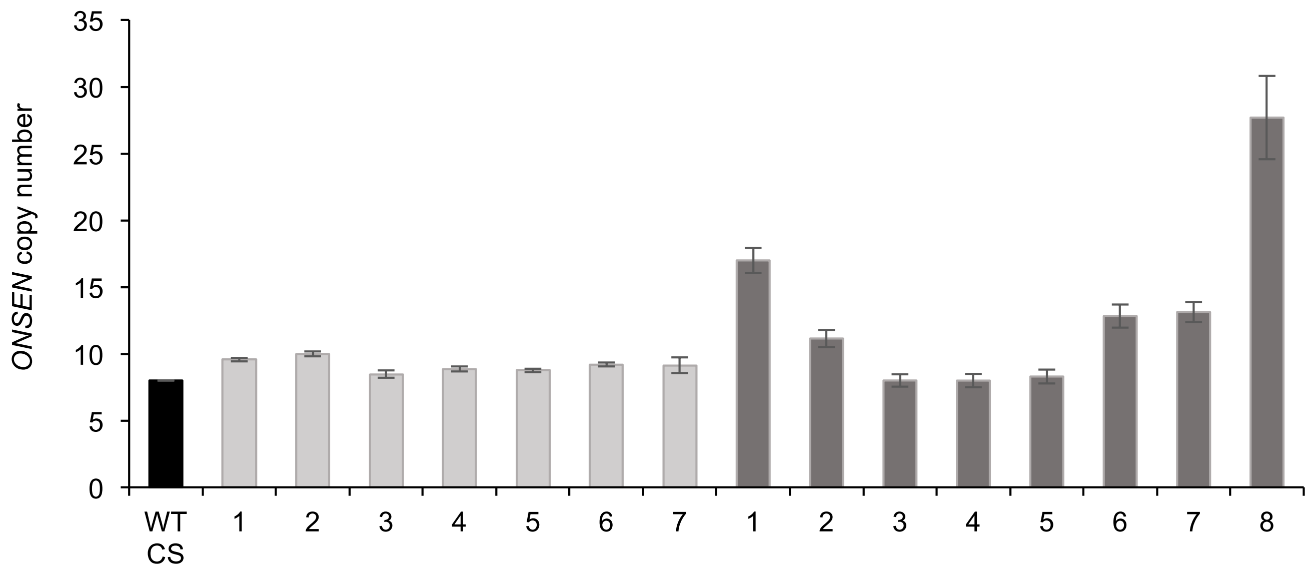


Fig. S1. Increase in *ONSEN* copy numbers in S1 pools of heat-stressed and Z-treated *nrpb2-3* plants. *ONSEN* copy number measured by qPCR in pooled seedlings of the S1-generation of heat stressed and zebularine-treated (10 µM) WT (light grey bars) and *nrpb2-3* plants (dark grey bars) that were grown under control conditions on soil relative to a control stressed WT-plant (black bar) (mean ± s.e.m., *n*=3 technical repetitions, all values relative to *ACTIN2).*

Fig. S1.

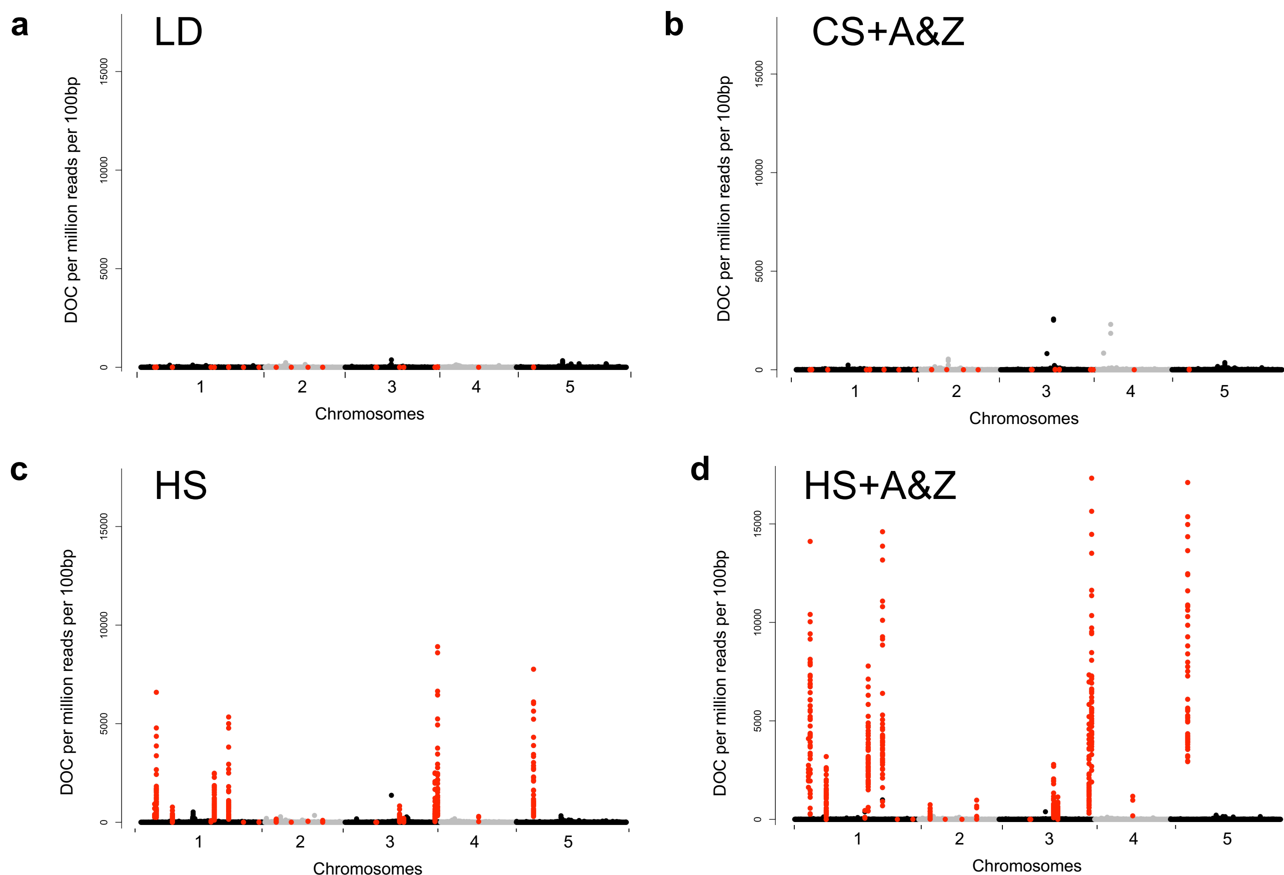


Fig. S2. Detection of eccDNAs originating from *ONSEN* loci following heat stress and chemical treatments in *Arabidopsis*. Abundance of reads from the mobilome-seq libraries of WT *Arabidopsis* plants mapping at TE-annotated loci from seedlings after: **a** growth under long day conditions (LD), **b** CS plus treatment with A (5 µg/ml) and Z (40 µM) (A&Z), **c** HS and **d** HS plus treatment with A&Z. Each dot represents the normalized coverage per million mapped reads per all TE-containing 100bp windows obtained after aligning the sequenced reads on the five chromosomes (black and grey circles). Red dots indicate the position of 100bp windows corresponding to *ONSEN* loci.

Fig. S2.

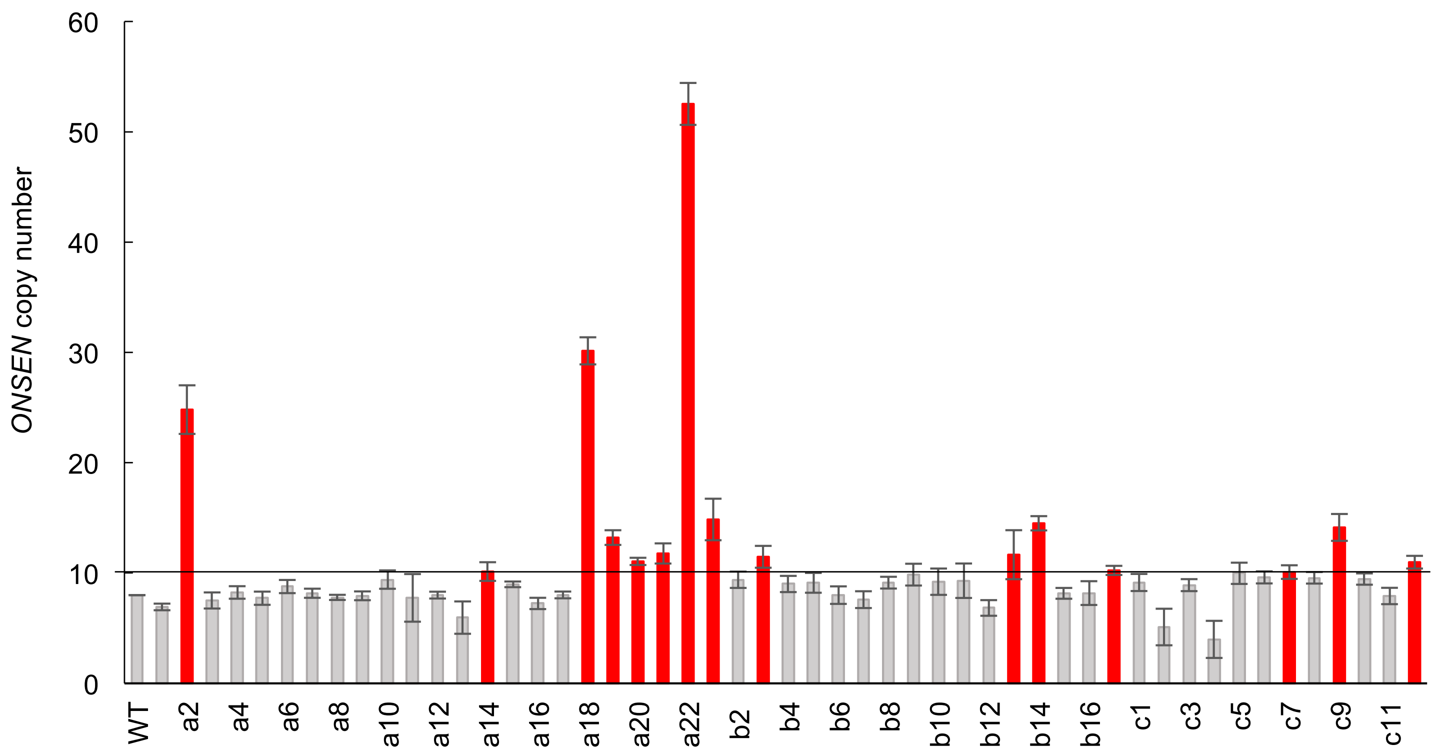
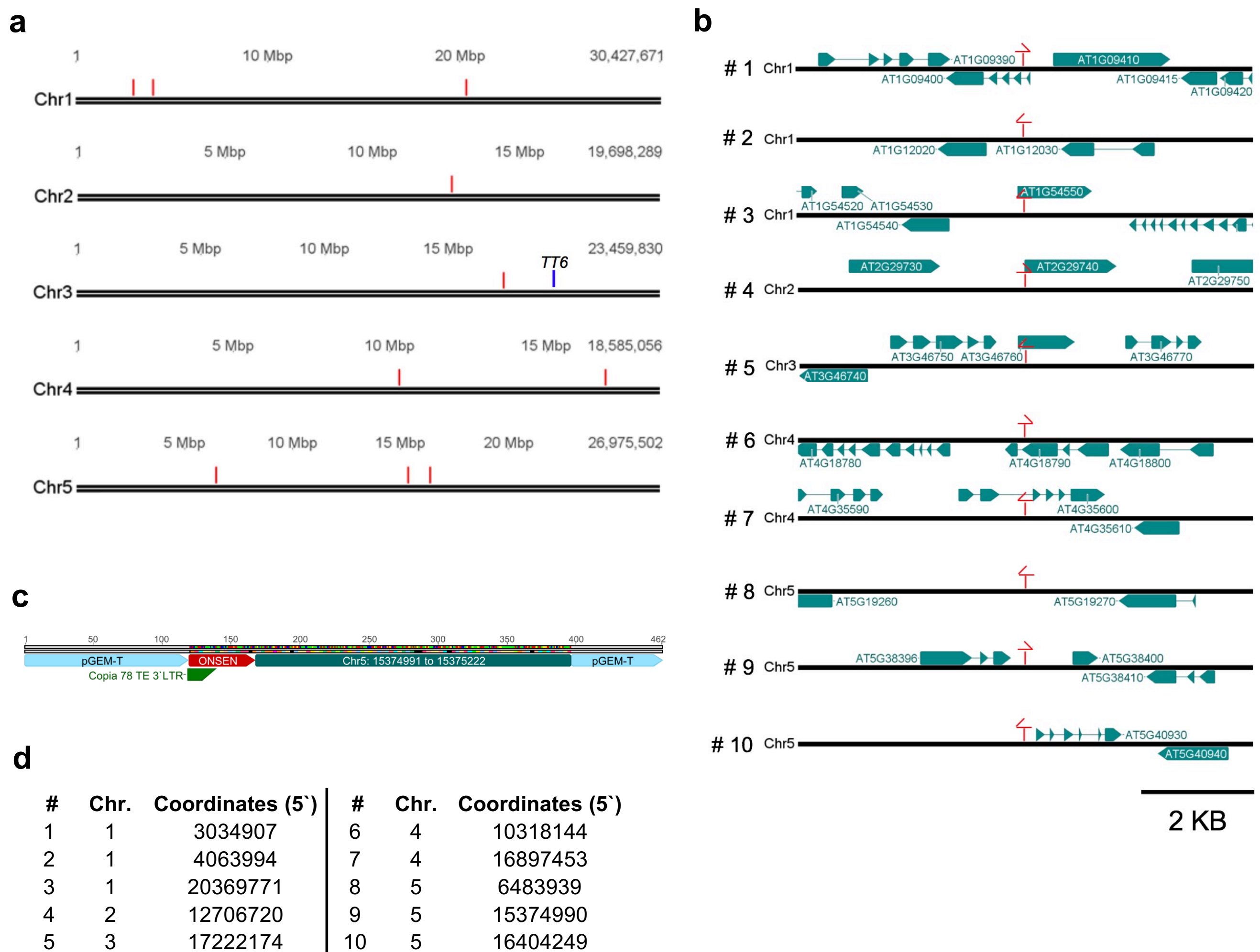


Fig. S3. Increase in *ONSEN* copy numbers in S1 pools of heat-stressed and A&Z-treated WT plants. Parental plants were heat stressed and treated in independent experiments (characters a-c) with a combination of A (5 µg/ml) and Z (40 µM). Pools with clearly increased *ONSEN*-copy numbers (>10) are marked in red. *ONSEN*-copy number measured by qPCR (mean ± s.e.m., n=3 technical repetitions, values relative to *ACTIN2*).

Fig. S3.



#	Chr.	Coordinates (5`)	#	Chr.	Coordinates (5`)
1	1	3034907	6	4	10318144
2	1	4063994	7	4	16897453
3	1	20369771	8	5	6483939
4	2	12706720	9	5	15374990
5	3	17222174	10	5	16404249

Fig. S4. Summary of confirmed novel *ONSEN* insertions in hc-line 3. a Overview of insertion sites shown in (**b**) (red bar) and the location of the *ONSEN* insertion in the *TT6*-gene depicted in Fig. 4 (blue bar) **b** Close-up of regions with new *ONSEN* insertions (red bar) in the S2 generation of a selected heat stressed and A (5 µg/ml) and Z (40 µM) treated WT plant (hc-line 3). Orientation of novel *ONSEN* insertions is indicated with red arrows. **c** A scheme to exemplify the annotation of sequences that lead to the identification of novel *ONSEN* insertion sites depicted in (**b**) shown for insertion # 9. Colors correspond to the pGEM-T vector (light blue) used for cloning, the *ONSEN*- 3`LTR (red), the Copia 78 TE 3`LTR primer (dark green) that was used for the preceding TE-Display PCR and the genomic region (turquoise) flanking the 3`LTR of the new *ONSEN* insertion. **d** Summary of coordinates (base 5` of insertion) of new *ONSEN* insertions shown in (**a**) and (**b**). Numbering corresponds to (**b**). Sequences of primers used to confirm new *ONSEN* insertions are given with the numbering corresponding to (**b**) in in Additional file 1: Table S1.

Fig. S4.

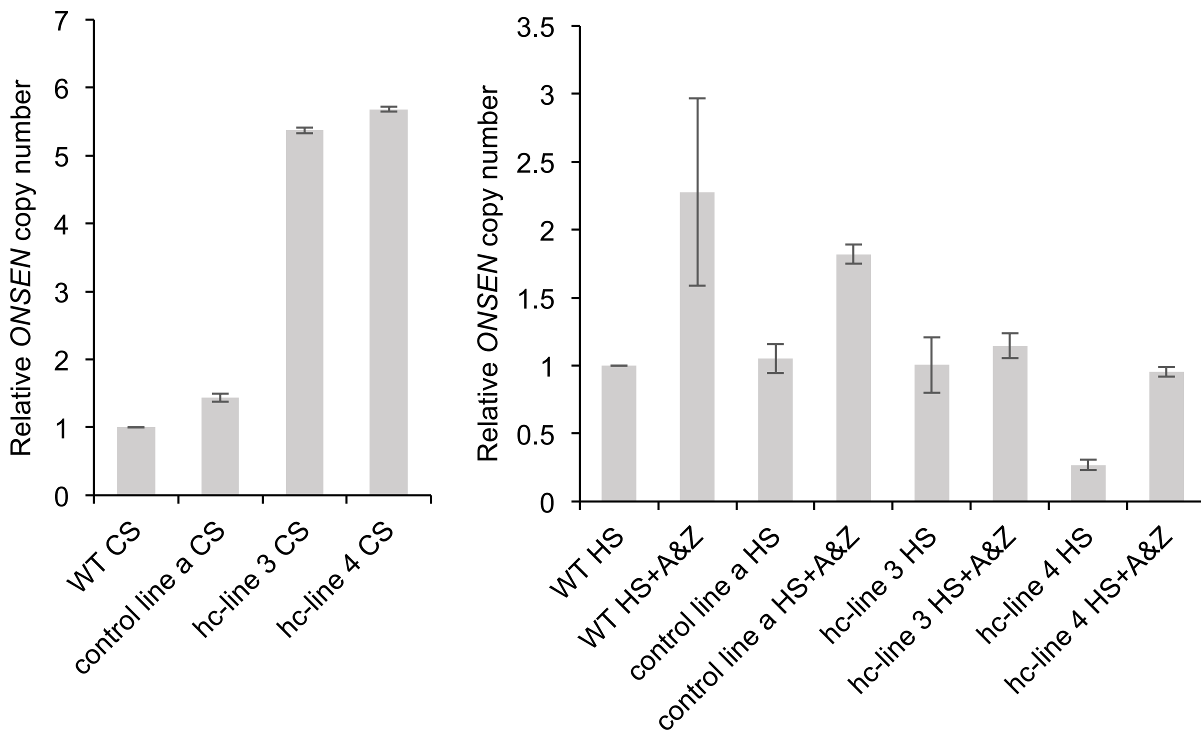


Fig. S5. Stress-induced activation of *ONSEN* in the S3 generation after initial HS-treatment. *ONSEN* copy number measured by qPCR directly after HS and HS plus treatments with α-amanitin (A, 5 µg/ml) and zebularine (Z, 40 µM) in seedlings of the WT, the control line a and the hc-lines 3 and 4. *ONSEN* copy number is shown relative to the WT HS (mean ± s.e.m., *n*= 3 biological repetitions, all values relative to *ACTIN2).*

Fig. S5.

```
READ    1    CTTTTAAGGAGGTCCCCCTCTCCTAGAATAAGCAAGGTGGTACTAAACTCCACATGCATG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ref     290  CTTTTAAGGAGGTCCCCCTCTCCTAGAATAAGCAAGGTGGTACTAAACTCCACATGCATG  231

READ    61   CCATCCTATGAGGTGGGCTTTTGTGATTTTCCAAAGAATTAATCTTCGAGTGGGCTAAGG  120
             |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
ref     230  CCATCCCATGAGGTGGGCTTTTGTGATTTTCCAAAGAATTAATCTTCGAGTGGGCTAAGG  171

READ    121  CCCATTCATTAATTCCAACAATCCCCCACATTGTTGAGATTATGGGCATATAATGATTTA  180
             ||||||||||||||||||||           |||||||||||||||||||||| ||||||
ref     170  CCCATTCATTAATTCCAACA-----------TGTTGAGATTATGGGCATATAAGGATTTA  122

READ    181  ATTTATTCCATAAATAAATCATGACATTACAGATGAAAACTAGCATGAACGCATCATTAG  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ref     121  ATTTATTCCATAAATAAATCATGACATTACAGATGAAAACTAGCATGAACGCATCATTAG  62

READ    241  ATCTACACATG   251
             |||||||||||
ref     61   ATCTACACATG   51
```

3' LTR   5' LTR

Fig. S6. ***Houba*** **forms LTR-LTR junction eccDNAs when treated with the combination of A&Z.** Alignment between a sequencing read resulting from the mobilome sequencing of A (5 µg/ml) and Z (40 µM) -treated plants (top) and an artificial junction corresponding to the 3' part of the 3' LTR (blue box) fused to the 5' part of the 5' LTR (yellow box).

Fig. S6.

11.2 Appendix II: Patent: Bucher*, E. and Thieme*, M. (2017) Mobilization of Transposable Elements to Enhance Genetic and Epigenetic Variability in a Population. *Patent* WO2017/093317A1 (50 pages)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(10) International Publication Number
WO 2017/093317 A1

(43) International Publication Date
8 June 2017 (08.06.2017)

(51) International Patent Classification:
C12N 15/01 (2006.01) *A61L 31/16* (2006.01)
A61K 31/7068 (2006.01) *A01H 1/06* (2006.01)
C07H 19/06 (2006.01) *A01H 3/04* (2006.01)

(21) International Application Number:
PCT/EP2016/079276

(22) International Filing Date:
30 November 2016 (30.11.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
15197663.6 2 December 2015 (02.12.2015) EP

(71) Applicant: **UNIVERSITÄT BASEL** [—/CH]; Petersgraben 35, Vizerektorat Forschung, 4003 Basel (CH).

(72) Inventors: **BUCHER, Etienne**; 47 Rue des Ormes, 49070 Beaucouzé (CH). **THIEME, Michael**; Schönaustr. 48, 4058 Basel (CH).

(74) Agent: **SCHULZ JUNGHANS PATENTANWÄLTE PARTGMBB**; Großbeerenstraße 71, 10963 Berlin (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2017/093317 A1

(54) **Title**: MOBILISATION OF TRANSPOSABLE ELEMENTS TO ENHANCE GENETIC AND EPIGENETIC VARIABILITY IN A POPULATION

(57) **Abstract**: A method for the mobilization of a transposable element is provided. The method comprises the steps of a) providing an inhibitor of DNA methylation, and/or an inhibitor of transcription, and b) contacting the inhibitor(s) with a cell comprising inactivated transposable elements, yielding a cell with mobilized transposable elements. In a second aspect of the invention a method for increasing the genetic and/or epigenetic variation in a plurality of eukaryotic organisms is provided. The method comprises the steps of i. providing an inhibitor of DNA methylation and/or an inhibitor of transcription, ii. contacting the organism with the inhibitor(s) and iii. propagating the organism.

Mobilisation of Transposable Elements to Enhance Genetic and Epigenetic Variability in a Population

Field of the invention

The present invention relates to the mobilisation of transposable elements and related uses
5 thereof.

Background of the invention

Transposable elements (TEs) were initially discovered in the early 1950s by Barbara McClintock due to their mutagenic activity that could influence kernel pigmentation (variegation) in maize (McClintock, PNAS, 1950 36(6):344-55). Since their initial discovery
10 numerous functions have been attributed to TEs. Indeed, TEs now tend to be viewed as natural molecular tools that can reshape the genome (Bire et al., Methods Mol Biol, 2012; 859:1-28). TEs have been identified in playing important (if not major) roles in structuring host genomes; especially centromeric regions are rich in TEs. The copy number of long terminal repeats (LTR) retrotransposons has been found to strongly correlate with host
15 genome size and mobilization of TEs can have an impact on genome organization by inducing chromosome breakage and by influencing homologous recombination. At the gene level, TEs can have multiple effects: Cause mutations by directly inserting into genes, move genes within the genome, duplicate and/or create novel genes, regulate gene expression, create novel regulatory pathways and bring genes under epigenetic control. Currently, TEs
20 are considered as a mutagen that can accentuate the positive outcome of the mutagenesis to the host (Bennetzen et al, Annu Rev Plant Biol, 2014, 65:505-30).

TEs have proven to be very useful genetic tools and have been broadly exploited for gene disruption and transgenesis in a wide variety of organisms. However, because TEs naturally very rarely get activated under normal growth conditions only few active TEs are currently
25 known. Thus only a very limited number of TEs are actively being used for genetic modification. Some examples include P elements (Drosophila), PiggyBac (insects, human cell lines), L1 LINE elements (mouse), Mariner (vertebrates), SleepingBeauty (animals). In order to create genetic diversity, these TEs are introduced into the organisms of interest via transgenesis. However, this limits use of organisms modified in such a manner because they
30 are considered as genetically modified organisms (GMOs) by current legislation.

In plants, it has been demonstrated that the mobility of transposable elements is limited by DNA methylation and certain histone marks (Miura et al., Nature, 2001, 411(6834):212-4; Mirouze et al., Nature, 2009, 461(7262):427-30). Suppression of DNA methylation in genetic

mutants can therefore result in the mobilization of transposable elements. It was also shown that drugs that reduce DNA methylation (e.g. 5-aza-2'-deoxycytidine) can mobilize certain DNA TEs (Scortecci et al., Plant Cell Physiol, 1997, 38(3):336-43). Furthermore, it has been reported that stresses on plants defective in RNA-directed DNA methylation (RdDM) activate transposable elements (Ito, H. et al., Nature, 2011, 472:115–19). However, the requirement of genetic mutants in components involved in the defense against TEs is limiting the possibility to activate TEs in non-model organisms or organisms that are difficult to transform. Therefore, the exploitation of endogenous TEs to obtain genetic and epigenetic diversity is currently very limited.

Under normal growth conditions, TEs are very rarely mobilized and different treatments to activate TEs have so far been very inefficient in eukaryotes. Treatments with drugs that reduce genomic DNA methylation levels have been shown to allow mild activation of TEs (Baubec, T. et al., Plant J, 2009 57:542–54), but without resulting in novel insertions of those TEs. It has been shown in plants that mutations in factors involved in the RNA-directed DNA methylation pathway could mobilize TEs at a high frequency. An important limitation in these approaches is that they are either inefficient (aforementioned drug treatment) or they require genetic mutations that are difficult to obtain, especially in non-model organisms. These technical problems therefore limit the use and the study of transposable elements in most organisms.

The problem underlying the present invention is to provide the means for efficient mobilization of transposable elements. This problem is solved by the subject-matter of the independent claims.

Specific description of the invention

The inventors provide herein a drug-based treatment that can mobilize transposable elements in eukaryotes. Additionally, the combination of this treatment with specific stresses leads to the mobilization of specific TEs that respond to this particular stress. The treatment leads to a high accumulation of extrachromosomal DNA of the activated TEs in the treated organism. Furthermore, the progeny of the treated organism shows stable integration of a high number of TE copies in the genome and increased resistance to the stress that is part of the treatment. Therefore, the method of the invention overcomes the necessity of genetic mutations to inactivate TE defense, thus allowing transposable elements to be efficiently activated in virtually any eukaryote. This invention enables the induction of TE mediated changes in genome size and structure, modulation of endogenous gene expression, gene transduplication, heterosis, homologous recombination and stress adaptation. Furthermore, this invention allows the identification of novel functional TEs.

According to a first aspect of the invention a method for the mobilization of a transposable element, particularly within the genome, of a eukaryotic cell is provided. The method comprises:

a) providing a eukaryotic cell comprising one or several dormant, i.e. inactive, transposable elements, and
b) contacting the cell(s) with an inhibitor of transcription,

and optionally, contacting the cell addtionally with an inhibitor of DNA-methylation,

thereby yielding a eukaryotic cell with one or several mobilized transposable elements.

In the context of the present specification the terms *transposable element* or *transposons* are used in their meaning known in the art of molecular genetics; they refer to DNA sequences in the genome of an organism that are able to change their position within the genome (cut and paste mechanism) or being able to produce novel copies of themselves that integrate into the genome (copy and paste mechanism). Transposition can result in multiplication of the element thereby influencing the size of the genome. There are two classes of transposons, class 1 transposons also referred to as retrotransposons and class 2 transposons also referred to as DNA transposons. Retrotransposons are first transcribed into RNA by the molecular apparatus provided by the host cell, and are then reverse transcribed into a double stranded DNA copy of the RNA, termed complementary DNA (cDNA) before they are inserted at a new position into the genome. They share some characteristics such as the dependency on a reverse transcriptase with retroviruses. DNA transposons do not have a RNA intermediate and are transferred to their new position in the genome by a transposase. The majority of transposons in the genome are inactive and will not duplicate or change position. The activation of transposons is therefore also referred to as mobilization of transposons. Examples of transposons that are responsive to certain stresses are provided in Table 1. These transposons are activated by the indicated stress up to a certain degree. However, use of the method of the invention mobilizes these transposons to a much larger extent as can be seen in the examples provided.

In certain embodiments of any aspect of the invention, a class 1 transposon is mobilized.

In certain embodiments of any aspect of the invention, a class 2 transposon is mobilized.

In the context of the present specification the term *DNA methylation* is used in its meaning known in the art of molecular biology and molecular genetics; it refers to the addition of methyl groups to the DNA, which in eukaryotes occurs mainly on cytosines. Methylation of DNA is catalyzed by DNA methyltransferases (DNMT) and can be divided into maintenance methylation, which is necessary to transfer methylation patterns on newly synthesized DNA

strands, and de novo methylation. DNA-methylation is associated with the inactivation of gene expression and the silencing of transposons. DNA methylation can be passed on to following generations and therefore represents a common form of epigenetic modification.

Table 1: Examples of transposons

Transposable element	Activating stress	Organism	Reference
ONSEN	heat, flagellin	*Arabidopsis thaliana*	Ito et al., 2011, Nature; Yu et al., 2012, PNAS
TLC1.1	salicylic acid, abscisic acid, methyl jasmonate, hydrogen peroxide and the synthetic auxin 2,4-D.	*Solanum chilense*	Salazar et al., 2007, Plant Cell
Tnt1A	wounding, biotic elicitors and pathogen attacks of fungal extracts	*Nicotiana tabacum*	Melayah et al., 2001, Plant Journal
Erika1 **Sabrina**	heat, drought and wounding cell culture	*Hordeum vulgare*	Alzohairy, et al., 2012; Life Science Journal
Tcs1	cold	*Citrus sinensis*	Butelli et al., 2012; Plant Cell

5 In certain embodiments, the inhibitor of DNA methylation is an exogenous compound.

In certain embodiments, the inhibitor of transcription is an exogenous compound.

In certain embodiments, the exogenous compound is a small molecule compound having a molecular mass of ≤1000u, particularly ≤920u.

In the context of the present specification the term *exogenous compound* refers to molecules
10 that are not present in the cell under physiological conditions unless added technically.

In certain embodiments, the inhibitor of DNA methylation might be present in at least some of the cells under at least some particular physiological conditions in trace amounts, but is added in the method of the invention at much higher concentrations to exert a significant impact on cell physiology. To achieve this, the compound is present in the cell’s medium at a

concentration being selected to be at least 10 times higher than the concentration of the inhibitor of DNA methylation found in the interior of the cell.

In certain embodiments, the inhibitor of transcription is present in the cell under physiological conditions and present in a medium at a concentration being selected to be at least 10 times, 100 times, 1000 times, or even 10.000 times higher than the concentration of the inhibitor of transcription found in the interior of the cell.

In certain embodiments, the method of the invention as specified in any aspect or embodiment disclosed herein additionally comprises a step c):

c) exposing the cell to an abiotic stress, biotic stress or chemical stress.

In the context of the present specification the term *abiotic stress* refers to the negative impact of non-living factors on a living organism in a specific environment. The non-living variable influences the environment beyond its normal range of variation. Non-limiting examples of abiotic stress are heat, cold, drought, submergence/ water excess, wind, UV-radiation, nuclear radiation, salinity, heavy metals, soil pH, tissue culture cultivation and starvation of phosphorous, nitrogen, light, CO_2 etc.. In contrast the term *biotic stress* refers to the negative impact of fungi, bacteria, viruses, insects, wounding by herbivores and biological competition etc..

The term *chemical stress* refers to the negative impact of chemical substances (“stressors”) on a living organism. These substances may also comprise substances that are stress-mimicking substances that mimic an abiotic or biotic stress. Non-limiting examples of chemical stressors are herbicides, herbicide safener, insecticides, fungicides, plant secondary metabolites, synthetic or natural compounds that induce plant defense.

The term *herbicide safener* refers to a compound that selectively protects monocotyledonous plants from herbicide damage whereas dicotyledonous plants are still affected by the herbicides. The common crop plants such as rice, wheat, maize etc but also forage grass, sugar cane and bamboo are monocotyledonous plants whereas most weed species are dicotyledonous plants. Herbicide safeners can be applied as a dressing for the seeds before sowing, to prepare the soil of agricultures or be applied to the foils of grown plants. In the two latter cases herbicide safeners can be applied together with the herbicides. Examples of common herbicide safeners are: Benoxacor (CAS 98730-04-2), Cloquintocet-mexyl (CAS 99607-70-2), Cyometrinil (CAS 63278-33-1), Dichlormid (CAS 37764-25-3), Fenchlorazole-ethyl (CAS 103112-35-2), Fenclorim (CAS 3740-92-9), Flurazole (CAS 72850-64-7), Fluxofenim (CAS 88485-37-4), Furilazole (CAS 121776-33-8), Mefenpyr-diethyl (CAS 135590-91-9), MG 191 (CAS 96420-72-3), Naphthalic anhydride (CAS 81-84-5), MON-13900 (CAS 121776-33-8), LAB 145138 (CAS 79260-71-2) and Oxabetrinil (CAS 74782-23-3).

In certain embodiments, the transposable element is a retrotransposon.

In certain embodiment, the cell is part of a multicellular organism. In certain embodiments, the eukaryotic cell is part of a non-human organism.

In certain embodiments, the eukaryotic cell is a plant cell. In certain embodiments, the plant
5 cell is a cell from *Arabidopsis*, particularly *Arabidopsis thaliana.*

In certain embodiments, the plant cell is part of crop plants, particularly of the family of *Poaceae* that comprises plants such as rice, sugar cane, maize, wheat, rye, barley, oat or millet. In certain embodiments, the method comprises a subsequent step of isolating said cell and determining whether a phenotype of the cell has been changed.

10 In certain embodiments, the eukaryotic cell is part of a multicellular organism, particularly a plant, and wherein subsequent to exposure of the cell to step c), the cell is cultivated to render a multicellular organism, and the phenotype of the multicellular organism is determined.

In certain embodiments, the phenotype of the organism comprises determining resistance to
15 the stressor, wherein the stressor causes the stress applied in step c).

In certain embodiments, the resistance to the stressor that causes the stress applied in step c) is increased after application of the method of the invention.

According to a second aspect of the invention a method for increasing genetic and/or epigenetic variation in a population of eukaryotic organisms is provided. The method
20 comprises:

i. providing an eukaryotic organism,
ii. contacting the eukaryotic organism with
    - an inhibitor of DNA methylation, and/or
    - an inhibitor of transcription,
25 iii. propagating the eukaryotic organism, yielding the eukaryotic population with increased genetic and/or epigenetic variation.

The method mobilizes dormant, i.e. inactive, not currently transcribed or reverse transcribed, transposable elements within the eukaryotic organism. Since to the knowledge of the inventors, all eukaryotic organisms comprise dormant transposable elements within their
30 genome, the element “eukaryotic organism” is synonymous with “eukaryotic organism comprising a dormant transposable element”.
In certain embodiments, the method is employed on a eukaryotic organism comprising any one of the specific transposable elements recited in the current specification.

In certain embodiments of the second aspect of the invention, the method additionally comprises a step ii.a, which is following step ii.:

ii.a exposing the eukaryotic organism to an abiotic stress, biotic stress or chemical stress.

In certain embodiments, the inhibitor of DNA methylation and/or the inhibitor of transcription are provided as a solution in a polar solvent, in particular a polar aprotic solvent, more particularly Dimethyl sulfoxide (DMSO).

In certain embodiments, the inhibitor of DNA methylation and/or the inhibitor of transcription are provided as a solution in a polar solvent, in particular water.

In certain embodiments, the method comprises the subsequent step iv. comprising:

a. Determining any genetic and/or epigenetic changes or
b. Determining any changes in the phenotype, particularly the resistance to any stressors applied in step ii.a

wherein these changes are determined in the individual constituent eukaryotic organisms or for a representative sample of the population of eukaryotic organisms, or for all of the constituent eukaryotic organisms of the population.

In certain embodiments of the first and the second aspect of the invention, the abiotic stress is selected from heat, cold, drought, submergence/ water excess, wind, UV-radiation, nuclear radiation, salinity, heavy metals, soil pH, tissue culture cultivation and starvation (phosphorous, nitrogen, light, CO_2 etc.).

In certain embodiments of the first and the second aspect of the invention, the biotic stress is selected from the negative impact of fungi, bacteria, viruses, insects, wounding by herbivores and biological competition. Non-limiting examples of fungi having a negative impact would be *Phytophthora infestans* (potato blight) and *Magnaporthe grisea* (rice blast). Non-limiting examples for bacteria having a negative impact are *Botrytis cinerea* (gray mold), *Xylella fastidiosa* (Olive Quick Decline Syndrome) and *Puccinia spp*. (wheat rust). Non-limiting examples of viruses having a negative impact are Tobacco mosaic virus and Tomato spotted wilt virus. Non-limiting examples for insects having a negative impact are *Mamestra brassicae* (Cabbage moth), *Helicoverpa zea* (corn earworm) and *Ostrinia nubilalis* (European corn borer). Non limiting examples of other organisms that can have a negative impact due to biological competition are *Orobanche* (broomrape) and *Ambrosia trifida* (giant ragweed).

In certain embodiments of the first and second aspect of the invention, the chemical stress is selected from herbicides, herbicide safener, insecticides, fungicides, plant secondary metabolites, synthetic or natural compounds that induce plant defense. Non-limiting examples of compounds that induce plant defense are flagellin (natural compound, bacterial elicitor; Felix et al., 1999, Plant J.), a 22-amino acid sequence of the conserved N-terminal

part of flagellin (flg22), salicylic acid and analogues e.g. Bion® (natural compound with synthetic analogues; (Vlot et al., 2009, Annu. Rev. Phytopathol.;Friedrich et al., 1996, Plant J.)), jasmonic acid and jasmonic methyl ester (natural compounds; Cohen et al., 1993, Phytopathology), ethylene (natural compound; van Loon et al., 2006, Trends Plant Sci.), 5 abscisic acid (natural compound; Mauch-Mani and Mauch, 2005, Curr. Opin. Plant Biol.) and volatiles such as terpenes and green leaf volatiles (natural compounds; reviewed by Unsicker et al., 2009, Curr Opin Plant Biol).

In certain embodiments of the first and the second aspect of the invention, the DNA-methylation inhibitor is a nucleoside analogue.

10 In certain embodiments of the first and the second aspect of the invention, the DNA-methylation inhibitor is selected from 5-azacytidine, 5-aza-2'-deoxycytidine, 5-fluoro-2'-deoxycytidine, 5,6-dihydro-5-azacytidine and zebularine.

In certain embodiments of the first and the second aspect of the invention, the inhibitor of transcription is a RNA polymerase inhibitor, in particular a RNA polymerase II inhibitor, a 15 RNA polymerase IV inhibitor or a RNA polymerase V inhibitor, more particularly a RNA polymerase II inhibitor.

In certain embodiments of the first and the second aspect of the invention, the RNA polymerase II inhibitor is selected from

- amatoxins, in particular alpha-amanitin (CAS 23109-05-9),
- 20 derivatives of amatoxins, in particular alpha-amanitin oleate,
- nucleoside analogues, in particular 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB; CAS 53-85-0),
- actinomycin D (CAS 50-76-0),
- flavopiridol (CAS 146426-40-6),
- 25 triptolide (CAS 38748-32-2).

In certain embodiments of any aspect of the invention disclosed herein, the amatoxin, in particular alpha-amanitin is used with a concentration of 0.0005µg/ml to 50µg/ml, in particular 0.001µg/ml to 25µg/ml, more particular 0.005µg/ml to 20µg/ml, even more particular 0.005 µg/ml to 5 µg/ml.

30 In certain embodiments of any aspect of the invention disclosed herein, the inhibitor of DNA methylation, in particular zebularine, is used at a concentration of 5 µM to 100 µM, in particular 10 µM to 80 µM, more particular 10 µM to 40 µM, even more particular 20 µM to 40 µM.

In certain embodiments of the second aspect of the invention, the increased genetic and/or 35 epigenetic variation in a plurality of eukaryotic organisms results in increased resistance of

the organisms to the abiotic or biotic stress the organisms have been exposed to. In other words the increase in genetic and/or epigenetic variation is not random as for example would be expected from a chemical mutagen. The increase is directed toward resistance against the stress used in the method. For example using the abiotic stress heat would preferentially result in heat-resistant organisms. Without wishing to be bound by theory the inventors assume that transposons are preferentially integrated into the genome in the vicinity of genes thereby creating novel gene regulatory pathways that are able to respond to the previously applied stress. This may lead to genetic variety in genes activated by the respective stress and thereby confers increased resistance to the respective stress.

According to a third aspect of the invention, the use of a composition in a method according to the first and second aspect of the invention is provided. The composition comprises an inhibitor of DNA-methylation and an inhibitor of transcription.

In certain embodiments, the DNA-methylation inhibitor is a nucleoside analogue.

In certain embodiments, the DNA-methylation inhibitor is selected from 5-azacytidine (CAS 320-67-2), 5-aza-2'-deoxycytidine (CAS 2353-33-5), 5-fluoro-2'-deoxycytidine (CAS 10356-76-0), 5,6-dihydro-5-azacytidine (CAS 62488-57-7) and zebularine (CAS 3690-10-6).

In certain embodiments, the inhibitor of transcription is a RNA polymerase inhibitor, in particular a RNA polymerase II inhibitor, a RNA polymerase IV inhibitor or a RNA polymerase V inhibitor, more particularly a RNA polymerase II inhibitor.

In certain embodiments, the RNA polymerase II inhibitor is selected from

- amatoxins, in particular alpha-amanitin (CAS 23109-05-9),
- derivatives of amatoxins, in particular alpha-amanitin oleate,
- nucleoside analogues, in particular 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB; CAS 53-85-0),
- actinomycin D (CAS 50-76-0),
- flavopiridol (CAS 146426-40-6),
- triptolide (CAS 38748-32-2).

In certain embodiments, the amatoxin, in particular alpha-amanitin is used with a concentration of 0.5 nM to 55 µM, in particular 1 nM to 27.5 µM, more particular 5 nM to 20 µM, even more particular 5 nM to 5 µM.

In certain embodiments of any of the aspects of the invention disclosed herein, the ratio of the molar concentrations of the inhibitor of transcription, in particular alpha-amanitin, to the inhibitor of DNA-methylation, in particular zebularine, is 0.000005 to 11, more particular 0.000125 to 2, even more particular 0.000125 to 0.125.

In certain embodiments, the ratio of the molar concentration depends on the concentrations a and b, which are as follows:

a) the inhibitor of DNA-methylation, in particular zebularine, is used at a concentration of 5 µM to 100 µM, in particular 10 µM to 80 µM, more particular 10 µM to 40 µM, even more particular 20 µM to 40 µM
b) amatoxin, in particular alpha-amanitin is used at a concentration of 0.0005 µg/ml to 50 µg/ml, in particular 0.001 µg/ml to 25 µg/ml, more particular 0.005 µg/ml to 20 µg/ml, even more particular 0.005 µg/ml to 5 µg/ml.

A fourth aspect of the invention provides a kit of parts for use in the method according to the first and second aspect of the invention. The kit of parts comprises an inhibitor of DNA-methylation and an inhibitor of transcription.

In certain embodiments, the DNA-methylation inhibitor is a nucleoside analogue.

In certain embodiments, the DNA-methylation inhibitor is selected from 5-azacytidine (CAS 320-67-2), 5-aza-2'-deoxycytidine (CAS 2353-33-5), 5-fluoro-2'-deoxycytidine (CAS 10356-76-0), 5,6-dihydro-5-azacytidine (CAS 62488-57-7) and zebularine (CAS 3690-10-6).

In certain embodiments, the inhibitor of transcription is a RNA polymerase inhibitor, in particular a RNA polymerase II inhibitor, a RNA polymerase IV inhibitor or a RNA polymerase V inhibitor, more particularly a RNA polymerase II inhibitor.

In certain embodiments the RNA polymerase II inhibitor is selected from

- amatoxins, in particular alpha-amanitin (CAS 23109-05-9),
- derivatives of amatoxins, in particular alpha-amanitin oleate,
- nucleoside analogues, in particular 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB; CAS 53-85-0),
- actinomycin D (CAS 50-76-0),
- flavopiridol (CAS 146426-40-6),
- triptolide (CAS 38748-32-2).

Wherever alternatives for single separable features such as, for example, a type of inhibitor or organism are laid out herein as "embodiments", it is to be understood that such alternatives may be combined freely to form discrete embodiments of the invention disclosed herein.

The invention is further illustrated by the following examples and figures, from which further embodiments and advantages can be drawn. These examples are meant to illustrate the invention but not to limit its scope.

<u>Short description of the figures</u>

Fig. 1 shows accumulation of *ONSEN* extrachromosomal DNA upon pharmacological treatment and heat stress. (a) *ONSEN* DNA accumulation measured by qPCR directly after control stress (CS) heat stress (HS)-treatment in wild-type (WT) and *nrpb2-3* plants and treatments with alpha-amanitin (A, 5 µg/ml) or zebularine (Z, 10 µM) (mean ± s.e.m., *n*=6
5 biological repetitions, values relative to *ACTIN2*). (b) *ONSEN* copy number measured by quantitative PCR (qPCR) in seedlings of Columbia (Col) WT directly after control stress (CS; 24h 6°C), heat stress (HS; 24h 6°C and 24h 37°C) and a treatment with A (5 µg/ml), Z (40 µM) or a combination thereof (A&Z). (Mean ± s.e.m., n= 3 biological repetitions). The double treatment (A&Z) leads to a very strong heat-stress dependent activation of *ONSEN* resulting
10 in up to 700 extrachromosomal *ONSEN* DNA copies.

Fig. 2 shows the stress-dependence of *ONSEN* mobilisation. The graph shows *ONSEN* copy numbers in *A. thaliana* seedlings after chemical treatment with A (5 µg/ml), Z (40 µM), the combinations of A and Z (A&Z) in WT, *nrpb2-3* and *nrpd1* plants following the CS. *ONSEN* copy number measured by qPCR (mean s.e.m., n=3 biological replicates, values relative to
15 *ACTIN2).* This result shows that the production of *ONSEN* extrachromosomal DNA is dependent on heat-stress.

Fig. 3 shows that simultaneous inhibition of methyltransferases and Pol II mimics the *nrpd1*-mutant. (a) Asymmetric methylation analysis of the *ONSEN* LTR and the *soloLTR* in untreated and A (5 µg/ml), Z (40 µM) or A&Z-treated seedlings of the WT and the *nrpd1*
20 mutant after CS. PCR products obtained from genomic DNA that was used undigested (input) or after digestion with the CHH-methylation sensitive restriction enzyme *Dde*I. ACTIN2 is included as a control for complete *Dde*I digestion. The A&Z double treatment with A (5 µg/ml) and Z (40 µM) resulted in a very strong reduction of DNA methylation at *ONSEN* and *soloLTR* comparable to the *nrpd1* mutant. (b) Northern blot indicating *ONSEN*-
25 transcription directly after CS, HS and HS plus treatment with A (5 µg/ml), Z (40 µM) or a combination of A&Z in WT and *nrpd1* plants. A Midori-stained agarose-gel is shown as a loading control. The level of the full length *ONSEN* transcript after heat stress and the double treatment with A (5 µg/ml) and Z (40 µM) is comparable to the *nrpd1*-mutant. (c) Accumulation of *ONSEN* DNA measured by qPCR in seedlings of WT, *rdr6, dcl2/3/4* and
30 *nrpd1* plants directly after CS, HS and HS plus treatment with A, Z or a combination of A&Z. This result shows that RNA pol II is active upstream of the DICER-like enzymes.

Fig. 4 shows the drug-induced mobilisation of *ONSEN* in wild-type Arabidopsis plants. (a) Transposon display confirming novel *ONSEN* insertions in the F2 generation of HS (HS control) and HS and A (5 µg/ml) and Z (40 µM) treated WT plants (HS+A&Z). Integrated
35 *ONSEN* copies were measured by qPCR (upper part) and detected by transposon display (lower part). *ONSEN* copy numbers of seven selected individual, non-related plants are

depicted. Copy numbers exceeding eight as measured by qPCR (upper part) and the observed additional bands on the transposon display (lower part) in the HS+A&Z treated Col WT plants indicate *novel* insertions of additional *ONSEN* copies. M is a 1kb size marker. (b), *ONSEN* copy number in the F1, F2 and F3 generation measured by qPCR (n=3 technical replicates, values relative to *ACTIN2*) Copy numbers >8 in lines 1-7 indicate insertions of additional *ONSEN* copies. c, and d, photographs of F2-plants containing novel *ONSEN* insertions showing both homogeneous and stress-dependent phenotypic variability induced by the HS+A&Z treatment when grown under long (c) and short day conditions (d). qPCR-Data for the F3-generation of line 6 in (b) as well as pictures of phenotypes in (c) and (d) are missing due to severe infertility and extinction of this line. Examples for phenotypes observed in of lines with novel *ONSEN* insertions (lines1-7) include higher biomass under short day conditions (line 3), early flowering under long day conditions (line 7) and reduced chlorophyll accumulation (line 1). In summary this dataset shows that A&Z-treatment leads to an efficient burst of ONSEN transposition. New *ONSEN* insertions are stably inherited over several generations and cause phenotypic changes.

Fig. 5 shows the increase in *ONSEN* copy numbers in F1 pools of heat-stressed and treated plants. Parental plants were treated and heat stressed in independent experiments (characters a-c) with a combination of A (5 µg/ml) and Z (40 µM). Pools with significantly increased *ONSEN*-copy numbers (>10) are highlighted in dark grey. *ONSEN* copy number measured by qPCR (mean ± s.e.m., n=3 technical repetitions, values relative to *ACTIN2*). Approximately 29.4% of tested F1 pools of heat stressed and A+Z treated wild type plants showed a significantly increased *ONSEN* copy number.

Fig. 6 shows a summary of confirmed novel *ONSEN* insertions. (a) Genome wide distribution and (b) Close-up of regions with new *ONSEN* insertions in the F2 generation of a selected HS+A&Z treated WT plant (line #3). Orientation of the *ONSEN* insertions is indicated central arrows.

Fig. 7 shows the drug-induced activation of the *Houba* retrotransposon in rice (*O. sativa*). Mobilome analysis of DNA extracted from seedlings after growth on control conditions (C), A (5 µg/ml), Z (40 µM), and the combination of A&Z. (a) Detail of the normalized depth of coverage compared to the untreated control plant obtained after aligning the sequenced reads on one *Houba* element. (b) Scheme of primers localization (*black bar*: *Houba* element, *arrows*: PCR primers, *box*: LTR). (c) extrachromosomal circular forms of *Houba* are specifically detected in plants treated with both A&Z using inverse PCR with primers shown in 4b. (d) Specific PCR on circular chloroplast DNA is shown as a loading control. Total DNA subjected to a rolling circle amplification was used as a template. These results demonstrate the efficient A&Z-dependent mobilization of the *Houba* transposon in rice.

Fig. 8 shows increased heat tolerance in the F2 generation of treated *Arabidopsis* seedlings. Tolerance to repetitive heat stress (42°C) in the F2 progeny of wildtype plants that were either only heat stressed (control) or heat stressed and treated with A (5µg/ml) and Z (40 µM) (#1-3). (a) Two biological replicates (I and II) are depicted, (b) Percentage of vital seedlings (Mean ± s.e.m., n=2 biological repetitions). F2 seedlings originating from Heat stressed and A&Z-treated plants show a significantly increased heat tolerance (>60% vital seedlings) compared to the F2 of a plant that was only heat stressed (10% vital seedlings). This demonstrates that the A&Z-dependent mobilization of a heat-stress responsive transposon can produce plants that are better adapted to heat stress.

Fig. 9 shows dose dependent accumulation of *ONSEN* extrachromosomal DNA upon pharmacological treatment and heat stress. *ONSEN* copy number was measured by qPCR in seedlings of Columbia (Col) wildtype directly after control stress (CS), heat stress (HS) and a treatment with α-amanitin in different concentrations given in µg/ml. (Mean ± s.e.m., n=3 technical repetitions). This shows that the number of mobilized transposons can be regulated by the amount of A used for the treatment.

Fig. 10 shows epigenetic changes at the DNA methylation level induced by the treatment of plants and human cells with A. (a) Midori stained Agarose gel showing reduction of DNA-methylation at the *ONSEN* LTR upon pharmacological treatment and heat stress in WT (Col) seedlings directly after control stress (CS; 24h 6°C), heat stress (HS; 24h 6°C and 24h 37°C) and a treatment with A (20 µg/ml), Z (10 µM) or a combination thereof (A&Z). Undigested DNA was used as a PCR-template for the loading control (Input). PCR on *Dde*I-digested DNA shows reduction in DNA-methylation after treatment with A, Z or a combination of A&Z. (b) CHH methylation state at the *ONSEN* LTR assessed by bisulfite sequencing performed on CS plants grown on medium with or without A. (c) LINE-1 DNA methylation levels assessed in human A549 cancer cells grown in control medium and medium supplemented with 0.5 µg/ml A. This shows that A can be used as a potent DNA demethylating agent in plants and human cells.

Fig. 11 shows accumulation of *ONSEN* extrachromosomal DNA upon combined pharmacological and flagellin-treatment. *ONSEN* copy number measured by quantitative PCR (qPCR) in seedlings of Col wild type directly after control stress, 5h after treatment with flagellin (flg22) alone or in combination with A (5 µg/ml), Z (40 µM) or a combination thereof (A&Z). (Mean ± s.e.m., n=3 technical repetitions).

Fig. 12 shows activation of *ATCOPIA17* upon combined pharmacological and flagellin-treatment. *ATCOPIA17* fold change was measured by quantitative PCR (qPCR) on total DNA in seedlings of the Col wild type directly after control stress, 5h after treatment with flagellin

(flg22) alone or in combination with A (5 µg/ml), Z (40 µM) or a combination thereof (A&Z). (Mean ± s.e.m., n=3 technical repetitions).

Fig. 13 shows accumulation of *ONSEN* extrachromosomal DNA upon pharmacological treatment and heat stress. *ONSEN* copy number was measured by quantitative PCR (qPCR) on total DNA in seedlings of WT directly after control stress (CS; 24h 6°C), heat stress (HS; 24h 6°C and 24h 37°C) and HS plus treatment with alpha-aminitin(A, 20 µg/ml), zebularine (Z, 10 µM) and HS plus the combination of A&Z. This result demonstrates the robustness of the treatments independent of the relative concentrations of A and Z.

Examples

The inventors have discovered a highly efficient method to activate and mobilize TEs in eukaryotes. The treatment involves drugs that target highly conserved eukaryotic mechanisms: DNA methylation and transcription.

Example 1

In order to investigate the role of Pol II on TE mobility the inventors chose the well-characterized heat-responsive copia-like *ONSEN* retrotransposon (Ito, H. *et al.,* Nature, 2011, 472: 115-119) of Arabidopsis. The inventors first tested if Pol II deficient plants showed enhanced TE activity. For that purpose, the inventors took advantage of the hypomorphic *nrpb2-3* mutant allele that accumulates reduced NRPB2 protein levels (Zeng, B. *et al.*, Genes Dev, 2009, 23: 2850-2860). Using real-time PCR, it was determined that challenging *nrpb2-3* seedlings by heat stress (called HS here) lead to a mild increase in *ONSEN* ecDNA compared to the wild type (Fig. 1a). This result was supported by the observed increase in *ONSEN* ecDNA after pharmacological inactivation of Pol II with 5 µg/ml α-amanitin (called A here), a potent Pol II inhibitor that does not affect Pol IV or Pol V (Haag, J.R. *et al.,* Mol Cell, 2012, 48: 811-818) (Fig 1a,b). Transcription by RNA Polymerase II (Pol II) is inhibited by α-amanitin, derivatives thereof or other Pol II inhibitors. Global inhibition of DNA methylation is achieved by treatments with zebularine or 5-aza-2'-deoxycytidine (and derivatives thereof). In order to test the interaction between Pol II-mediated repression of TE activation and DNA methylation the inventors grew wild-type and *nrpb2-3* plants on media supplemented with moderate amounts of zebularine (called Z here, 40 µM for wild-type plants, 10 µM for *nrpb2-3* plants to ensure the viability of *nrpb2-3* seedlings), an inhibitor of DNA methyltransferases active in plants (Baubec, T. *et al.,* Plant J, 2009, 57: 542-554) and submitted them to HS. The presence of Z in the medium during HS generally enhanced the production of ONSEN ecDNA. Notably, this induced increase in ecDNA accumulation was more distinct in the *nrpb2-3*-background (Fig. 1a). This indicated that both, DNA methylation and Pol II transcriptional activity contribute to the repression of *ONSEN* ecDNA production. Because

both DNA methylation and Pol II can be specifically inhibited by the addition of different drugs the inventors tested if treating wild-type plants with both A and Z at the same time could strongly activate and even mobilize *ONSEN* after a heat stress treatment. The inventors grew WT seedlings on MS medium supplemented with each drug individually and both
5 combined. In conformity with the strong activation of *ONSEN* in heat stressed and Z-treated *nrpb2-3*-seedlings, the combined treatment (A+Z) of the WT gave rise to a very high (Fig. 1b) and HS-dependent (Fig. 2) accumulation of ONSEN ecDNA comparable to the *nrpd1* mutant (Fig. 3c).

Example 2

10 In order to better understand the effect the drugs had at the DNA level underlying the increased activation of *ONSEN* after HS, the inventors assessed how they influenced DNA methylation at the long terminal repeat (LTR) of a selected *ONSEN* endogenous locus (*AT1TE12295*) and at an unrelated well characterized RdDM target (soloLTR). Treating plants with A or Z individually already resulted in reduced CHH methylation levels at the
15 *ONSEN* LTR after CS (Fig. 3a). Combining the two drugs lead to a loss of DNA methylation comparable to the *nrpd1* mutant. DNA methylation at the soloLTR showed a different response to the drug treatments as a reduction in DNA methylation levels was only observed in plants submitted to a combined A and Z treatment. The inventors then checked by Northern Blot whether the degree of reduction in CHH methylation would coincide with
20 increased ONSEN-transcript-levels directly after HS. The inventors found that treatment with Z alone already resulted in the highest *ONSEN*-transcript levels after HS (Fig. 3b). From this observation, the inventors concluded that these additional Z-induced transcripts were not suitable templates for the production of *ONSEN* ecDNA (compare Fig. 1 and Fig. 3b).

In *Drosophila*, it has been shown that Pol II-mediated antisense transcription results in the
25 production of TE-derived siRNAs in a Dicer-2 dependent manner (Russo, J. *et al.* Genetics, 2016, 202:107-21). Supporting this notion for Arabidopsis, a recent publication pointed out the importance of DCL3 in regulating ONSEN in the *ddm1* background (Panda, K. *et al.* Genome Biol, 2016, 17:1-19). To elucidate whether the effect of Pol II inhibition was also dicer-dependent, the inventors grew both the rdr6- and the dcl2/3/4-triple mutant (defective in
30 three of the four plant dicer-like enzymes, DCLs) on A, applied HS and measured *ONSEN* ecDNA. The inventors found that A was still enhancing ecDNA accumulation in *rdr6* whereas inhibition of Pol II had no effect in the *dcl2/3/4* triple-mutant (Fig. 3c). This finding supports the notion that Pol II acts upstream of the processing step catalyzed by the DCLs.

Example 3

35 Mobilization of endogenous TEs in plants has so far been very inefficient, thus limiting their use in basic research and plant breeding. We have previously not observed *ONSEN*

transposition in HS treated wild-type plants (Ito, H. *et al.* Nature, 2011, 472:115-119). Because the A&Z drug treatment resulted in an increased *ONSEN* ecDNA accumulation to a similar degree like in *nrpd1*, the inventors tested if the combined drug treatment could lead to an efficient *ONSEN* mobilization in wild-type plants. First, the inventors assessed by real-time PCR if, and at which frequencies, new *ONSEN* copies could be detected in the progeny of A&Z-treated and heat stressed plants. The inventors found new *ONSEN* insertions in 29.4% of the tested F1 pools (n=51) with mean copy numbers of the pools reaching up to 52 (Fig. 5). The inventors then confirmed stable novel *ONSEN* insertions in a subset of independent individual high copy plants by transposon display (Fig. 4a), real-time PCR (Fig. 4b) and sequencing of some insertions in a selected high-copy line (#3) (Fig. 6). The combination of HS, A and Z resulted in a similar extrachromosomal *ONSEN* copy number as has been previously observed in RdDM deficient plants. The inventors detected novel *ONSEN* insertions in the progeny of 27 % of the treated plants. According to qPCR measurements, up to 90 +/-6 inserted copies were detected in individual plants in the F2 and successive generations of A, Z and HS treated plants (Figure 4a). These insertions were further confirmed by transposon display. The inventors did not observe further increases in *ONSEN* copy numbers over three generations indicating that the new insertions were stable and that *ONSEN* was not transposing anymore (Fig. 4b).

TE insertions can interrupt genes or alter their expression by either recruiting epigenetic marks or by stress-dependent readout transcription from the 3'LTR into flanking regions (Lisch, D., Nat Rev Genet, 2013, 14: 49-61). To test this, the inventors grew the F2 generation of the aforementioned selected high copy lines under long and short day conditions. The inventors observed that many lines showed clear and homogenous phenotypes in response to the different growth conditions (plant size, chlorophyll content and flowering time, Fig. 4c and d).

Example 4

The inventors tested if Pol II plays a more general role in repressing TEs in plants. Due to its significantly different epigenetic landscape compared to Arabidopsis the inventors chose the genetically well characterized monocotyledonous rice crop *O. sativa* (Kawahara, Y. *et al.,* 2013, Rice, 6: 4-10). In order to capture drug-induced mobilized TEs, the inventors characterized the active mobilome in *O. sativa* seedlings that were grown on MS medium supplemented either with no drugs, A only, Z only or the combination of A and Z, using a method that allows to specifically sequence extrachromosomal circular DNA (eccDNA). eccDNA is a byproduct of the LTR retrotransposon life cycle. Using this approach, the inventors identified *Houba*, a copia like retrotransposon (Panaud, O. *et al.,* Mol. Genet. Genomics, 2002, 268:113-121), as highly activated only when plants were treated with both

A and Z (Fig. 7a). The sequencing data were confirmed by an eccDNA-specific PCR on the *Houba* LTRs (Fig. 7b and c).

Example 5

Because the treatment with A alone reduced DNA methylation (Fig. 3a) in Arabidopsis, the
5 inventors wanted to test the robustness and generality of this treatment. In order to confirm the robustness, plants were treated with A (20µg/ml), Z (10 µg/ml) and A&Z. A alone already strongly reduced DNA methylation at this higher concentration (Fig. 10a), this result was then further supported by the assessment of DNA methylation in the CHH context by bisulfite sequencing (average of 10 sequenced clones for each sample). Because A inhibits the
10 highly conserved RNA Pol II enzyme and that A is also active in human cells, the inventors tested the effect of A on DNA methylation in the A549 human cancer cell line. Global DNA methylation content in the cells was assessed and compared to untreated or Z-treated cells. Supplementation of the growth medium with A (0.5 µg/ml) resulted in a 40% reduction of DNA methylation. This reduction was comparable to a treatment with the DNA demethylating
15 agent Z (350 µM) (Fig. 10c). The authors then also assessed the DNA methylation levels at the *long interspersed element 1* (*LINE-1*) retrotransposon. At LINE-1 A had an even more pronounced effect on the reduction of DNA methylation than Z (40% versus 20% reduction, respectively). These results demonstrate that an inhibitor of transcription can be used as a potent DNA demethylating agent in eukaryotic cells.

20

Plants and growth conditions

After stratification for two days at 4°C, *Arabidopsis thaliana* plants (accession Col-0) were grown on sterile ½ MS medium with 1% sucrose and a pH of 5.8 (control medium) under long day conditions (16 h light) at 24°C (day) and 22°C (night), respectively. *Oryza sativa*
25 plants were grown on sterile ½ MS medium with 1% sucrose and a pH of 5.8 (control medium) 16 h at 28 °C (day) and 27 °C (night), respectively.

In order to analyze successive generations, seedlings were transferred to soil and grown under long day conditions (16 h light) at 24 °C (day) and 22 °C (night) (*A. thaliana*) in a Sanyo MLR-350 growth chamber until seed maturity.

30 For phenotyping, *A. thaliana* plants were grown under long day conditions (16 h light) at 24°C (day) and 22 °C (night) and short day conditions (10 h light) at 21 °C (day) and 18 °C (night).

The induction of epigenetic changes and the activation and stable integration of transposable elements in Arabidopsis seedlings was enhanced by germinating and growing them on ½ MS-medium that contained zebularine (final concentration: 10-40 µM), α-amanitin (final
35 concentration: 0.005-20 µg/ml) or a combination of both chemicals (inductive media).

In order to trigger the transposition of the heat-responsive retrotransposon *ONSEN*, seven days old seedlings were exposed to a cold shock for 24h at 6°C followed by a heat-stress for 24h at 37°C (heat stress, HS) under controlled conditions in a growth chamber (Sanyo). Control plants were transferred back to longday-conditions 24°C (day) and 22°C (night) after
5 the cold treatment at 6°C for 24h (CS, control stress, according to Ito et al., 2011).

In order to trigger a biotic-stress response, nine days old *Arabidopsis*-seedlings were grown for nine days on 5 ug/ul alpha-amanitin and 40 uM zebularine and sprayed with flg22 (10 µM). After 5h of incubation, total DNA from the aerial part of seedlings was extracted and TE copy number assessed by qPCR.

10 *qPCRs on total DNA to measure ONSEN and COPIA17 copy numbers*

Total DNA from seedlings and adult plants was isolated using a DNeasy Plant Mini Kit (QIAGEN).

In preparation to the measurement of extrachromosomal DNA copies of *ONSEN* in CS/HS and untreated/treated seedlings, roots were dissected directly after the heat stress and
15 plants were immediately frozen in liquid nitrogen until DNA extraction.

To track *ONSEN* copy numbers in the F1-F3 generations of control and high copy lines, DNA from true leaves was extracted.

For the estimation of the *ONSEN* transposition frequency, total DNA of pools consisting of at least eight seedlings of the progeny of HS+A&Z-treated plants was isolated. The DNA
20 concentration was measured with a Qubit Fluorometer (Thermo Fisher Scientific).

The copy numbers of *ONSEN* were determined with qPCRs on total DNA using a TaqMan master mix (Life Technologies) in a final volume of 10 µl in the Light-Cycler 480 (Roche). *ACTOPIA17* copy number was measured by quantitative PCR (qPCR) in a Light-Cycler 480 (Roche), using XYBR 421 Green I Master Mix. Actin2 (At3g18780) served as a standard
25 gene for normalization. The sequences of the primers and probes for the qPCRs are listed in table 2.

For the mobilome-seq analysis DNA from the aerial parts of three *O. sativa* seedlings was extracted as previously reported (Mette, M. *et al.,* EMBOJ, 1999, 18: 241-248).

5µg of genomic DNA for each sample were purified using a Geneclean kit (MPBio, USA)
30 according to the manufacturer’s instructions. ecDNA was isolated from the GeneClean product using the PlasmidSafe DNase (Epicentre, USA) according to the manufacturer’s instructions, except that the 37°C incubation was performed for 17h. DNA samples were precipitated by adding 0.1 volume of 3M sodium acetate (pH 5.2), 2.5 volumes of ethanol and 1 µl of glycogen (Fisher, USA) and incubating overnight at -20°C. The precipitated

circular DNA was amplified by random rolling circle amplification using the Illustra TempliPhi kit (GE Healthcare, USA) according to the manufacturer's instructions except that the incubation was performed for 65h at 28°C. The DNA concentration was determined using the DNA PicoGreen kit (Invitrogen, USA) using a LightCycler480 (Roche, USA). One nanogram 5 of amplified ecDNA from each sample was used to prepare the libraries using the Nextera XT library kit (Illumina, USA) according to the manufacturer's instructions. DNA quality and concentration were determined using a high sensitivity DNA Bioanalyzer chip (Agilent Technologies, USA). Samples were pooled and loaded onto a MiSeq platform (Illumina, USA) and 2x250 nucleotides paired-end sequencing was performed. Quality control of 10 FASTQ files was evaluated using the FastQC tool (version 0.10.1). To remove any read originating from organelle circular genomes, reads were mapped 198 against the mitochondria and chloroplast genomes using the program Bowtie2 version 2.2.2 71 with --sensitive local mapping. Unmapped reads were mapped against the reference genome IRGSP1.0 (http://rgp.dna.affrc.go.jp/E/IRGSP/ Build5.html) using the following parameters: -- 15 sensitive local, -k 1. DNA from both mitochondria and chloroplast genomes integrated in nuclear genomes was masked (1,697,400 bp), The TE containing regions cover 194,224,800 bp in *O. sativa*. Finally, for each library, a .bam alignment file corresponding to enriched genomic regions was considered for statistical analysis and visualized with the Integrative Genomics Viewer (IGV) software (https://www.broadinstitute.org/igv/home)

20 _Table 2:_ Sequences of primers and probes that were used for the qPCRs (TaqMan, Life Technologies) to measure total number of extrachromosomal ONSEN DNA-copies. Actin 2 served as a control gene for normalization.

Primer	**Sequence 5`-> 3`**
SEQ ID No 001 (*ONSEN*_RT_fw)	CCACAAGAGGAACCAACGAA
SEQ ID No 002 (*ONSEN*_RT_rev)	TTCGATCATGGAAGACCGG
SEQ ID No 003 (*ONSEN* probe)	(FAM) AAGTCGGCAATAGCTTTGGCGAAGA (BHQ1)
SEQ ID No 004 (Actin2_RT_fw)	TGCCAATCTACGAGGGTTTC
SEQ ID No 005 (Actin2_RT_rev)	TTACAATTTCCCGCTCTGCT
SEQ ID No 006 (Actin2_probe)	(JOE)TCCGTCTTGACCTTGCTGGACG(BHQ-1)
SEQ ID No 007 (GenWalkAdaptator1)	GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTG GT
GenWalkAdaptor2	(PHOS)ACCAGCCC(AMINO)
SEQ ID No 008 (AP1)	GTAATACGACTCACTATAGGGC
SEQ ID No 009 (Copia78 3′ LTR)	AACACTTAAACACTTTCTCCA
SEQ ID NO 010	TTAGTATAAGGCTGAGCTGGAAACTG

706_ATCOPIA17 QT R	
SEQ ID NO 011	CAAGCCTAACCCTCAGCTACATG
705_ATCOPIA17 QT F	

Transposon display to confirm insertion of new ONSEN copies

The stable integration of additional copies of the *ONSEN* TE into the genome of heat stressed and treated plants was ascertained by a simplified transposon display based on the GenomeWalker Universal kit (Clontech Laboratories), according to Ito et al. 2011.

5 300 ng of total DNA from adult plants from the F2 generation of HS +/- A & Z was extracted with a DNeasy Plant Mini Kit (QIAGEN) and digested with a blunt cutter restriction enzyme (Dra I). After purification with a High Pure PCR Product Purification Kit (Roche) digested DNA was ligated to the annealed GenWalkAdapters 1 & 2. For the PCR, the adaptor specific Primer AP1 and the *ONSEN*-specific Primer Copia 78 3` LTR was used. The PCR products 10 were separated on a 2 % agarose gel that was stained with a Midori Green Nucleic Acid Staining Solution. For sequence information, see tables 2 and 3.

Cloning, sequencing and genotyping of new insertions

In order to identify the genomic region of new *ONSEN* insertions, the PCR product of the transposon display was purified using a High Pure PCR Product Purification Kit (Roche), 15 ligated into a pGEM-T Vector (Promega) and transformed into E. coli. After a blue white selection, positive clones were used for the insert amplification and sequencing (StarSEQ). The obtained sequences were analyzed with Geneious 8.2.1 and blasted against the Arabidopsis thaliana reference genome. The standard genotyping-PCRs to prove novel *ONSEN* insertions were performed with combinations of the *ONSEN*-specific primer “Copia 20 78 TE display 3`LTR" and primers listed in tables 2 and 3.

Table 3: Names, purpose and sequences of primers

Name / SEQ ID NO	Sequences 5' --> 3 '	Experiment
OnsenFull_F SEQ ID 12	AAGTGGTATCAGAGCTTGAAGATCC	Northern Blot
OnsenFull_R SEQ ID 13	CAACACCCCCTCTTAAACTTGATTTTGC	
M13F SEQ ID 14	CGCCAGGGTTTTCCCAGTCACGAC	Cloning and sequencing
M13R SEQ ID 15	TCACACAGGAAACAGCTATGAC	
286 OnsenBis F1 SEQ ID 16	GGTTGAAGGGTYAAAGAGTAAAT	Methylation analysis
287 OnsenBis R1 SEQ ID 17	CCTCCAAACTACAAAATATCTAAAA	
835 Chop PCR ACT2 F SEQ ID 18	TGTAGTGTCGTACGTTGAACAGAAAGC	

836 Chop PCR ACT2 R SEQ ID 19	TTGGCACAGTGTGAGACACACCA	
houba_F2 SEQ ID 20	ATCCTGGGAAGAACAAACCATTAA	PCR on circular rice TE and chloroplast control
houba_R2 SEQ ID 21	GAGTTCGAGTACCTTAGCCATGGT	
Chloroplast cyc F SEQ ID 22	ACAACCACTGATGAAGGATT	
Chloroplast cyc R SEQ ID 23	AGAAAGAAAAGCAACGACTG	
Prove TED 2_20 R SEQ ID 24	ACCTAGCTCTGAGTGATGAA	Genotyping of novel ONSEN insertions
Prove TED4_27 F SEQ ID 25	TGGATATACACATTGGTTGCA	
Prove TED 2_19 F SEQ ID 26	GGAGAAAGCTGAAAACTTGG	
Prove TED4_30_rev SEQ ID 27	CTAGGTTGGTGACTGATGAG	
Prove TED 2_17 F SEQ ID 28	AAGAATGGGAGCAGCATTAA	
Prove3_2R SEQ ID 29	GCAGTACTATAACCGGGACT	
prove TED3_1 Fw SEQ ID 30	GAACTTTCCGTTGTTACCGG	
Prove TED3 F SEQ ID 31	ATGAGACAGGGAGCTTATCT	
Prove TED1 R SEQ ID 32	GGTGTGAACCGAACCTAAAT	
Prove TED 4_25 F SEQ ID 33	AAACACCAGAAATCTTTCGC	

PCRs on extrachromosomal Houba DNA

The presence of circular *Houba*-copies was proven by an inverse PCR on 7 ng of the rolling-circle amplified template that was also used for sequencing. A PCR specific to a chloroplast
5 DNA served as a loading control. PCR products were separated on a 1% agarose gel that was stained with a Midori Green Nucleic Acid Staining Solution (Nippon Genetics Europe). Primer sequences are given in supplementary Table 4.

Table 4: Sequences of primers and probes that were used for the PCRs to measure total number of extrachromosomal *Houba* DNA.

Primer /SEQ ID NO	Sequence 5`-> 3`
286 OnsenLTRchopF SEQ ID 34	GGTTGAAGGGTYAAAGAGTAAAT
287 OnsenLTRchopR SEQ ID 35	CCTCCAAACTACAAAATATCTAAAA
Houba_F2 SEQ ID 36	ATCCTGGGAAGAACAAACCATTAA
Houba_R2 SEQ ID 37	GAGTTCGAGTACCTTAGCCATGGT

RNA analysis and Northern Blot

Total RNA from the aerial part of Arabidopsis seedlings was isolated using the TRI Reagent (Sigma) according to manufacturer's recommendations. RNA concentration was measured (Qubit RNA HS Assay Kit, Thermo Fisher), 15 µg of RNA was separated on a denaturing 1.5% Agarose gel, blotted on a Hybond-N+ (GE Healthcare) membrane and hybridized with 25 ng of a gel-purified and P32 labelled probe (Megaprime DNA Labelling System, GE Healthcare) specific to the full length *ONSEN* transcript (See table 3 for primer sequences).

DNA methylation analysis

20 ng of total genomic DNA isolated from Arabidopsis seedlings was digested with the methylation sensitive restriction enzyme, Dde1 (NEB) at 37 °C over night. Following heat inactivation at 60°C for 20 min, the digested DNA was used as a template for the chopPCR. Actin2 served as a control for the digest. Undigested DNA was used as a loading control. PCR products were separated on a 1% agarose gel and stained with Midori Green.

For the A549 human cancer cell line cells were grown in medium without treatment or supplemented with either Z (350 µM) or A (0.5 µg/ml), DNA was extracted by using the QiaAmp DNA mini Kit (Qiagen, France). Next, global DNA methylation was estimated by quantifying the presence of 5-methylcytosine 5-mC DNA ELISA Kit (Zymo Research) according to the manufacturers's instructions. DNA methylation at the LINE-1 transposons were assessed with the Global DNA Methylation Assay–LINE-1 kit (Active Motif).

Claims

1. A method for the mobilization of a transposable element in a eukaryotic cell, said method comprising the steps of:
   a) providing a eukaryotic cell comprising a transposable element
   b) contacting said cell with an an inhibitor of transcription,

   thereby yielding a eukaryotic cell with a mobilized transposable element.

2. The method according to claim 1, wherein the cell is additionally contacted with an inibitor of DNA methylation in step b).

3. The method according to claim 1 or 2, wherein the inhibitor of DNA methylation and/or the inhibitor of transcription is an exogenous compound, particularly an exogenous small molecule compound.

4. The method according to any one of the preceding claims, additionally comprising a step c):

   c) exposing said cell to an abiotic stress, biotic stress or chemical stress.

5. The method according to any one of the preceding claims, wherein the method comprises a subsequent step of isolating said cell and determining whether a phenotype of said cell has been changed.

6. The method according to any one of claims 4 to 5, wherein said eukaryotic cell is part of a multicellular organism, particularly a non-human organism, particularly a plant, and wherein subsequent to exposure of said cell to step c), the cell is cultivated to render a multicellular organism, and the phenotype of said multicellular organism, in particular a resistance to a stressor, wherein the stressor causes the stress applied in step c), is determined.

7. The method according to claim 6, wherein said plant cell is part of crop plants, particularly rice, sugar cane, maize, wheat, rye, barley, oat or millet.

8. A method for increasing genetic and/or epigenetic variation in a population of eukaryotic organisms comprising:
   i. providing an eukaryotic organism,
   ii. contacting said eukaryotic organism with
       - an inhibitor of DNA methylation, and/or
       - an inhibitor of transcription,
   iii. propagating said eukaryotic organism, yielding a population of eukaryotic organisms with increased genetic and/or epigenetic variation.

9. The method according to claim 8, additionally comprising a step ii.a following step ii.:
   ii.a exposing said eukaryotic organism to an abiotic stress, biotic stress or chemical stress.

10. The method according to any one of claims 7 or 8, wherein the method comprises the
5 subsequent step iv. comprising:
    a. determining any genetic changes in said organism, and/or
    b. determining any changes in the phenotype of said organism, particularly an increased resistance to a stressor applied in step ii.a

    wherein said genetic or phenotypical changes are determined in the individual
10 constituent eukaryotic organisms or for a representative sample of the population of eukaryotic organisms, or for all of the constituent eukaryotic organisms of the population.

11. The method according to any one of the preceding claims 4 to 7, wherein
    i. said abiotic stress is selected from heat, cold, drought, submergence/ water
15 excess, wind, UV-radiation, nuclear radiation, salinity, heavy metals, soil pH, tissue culture cultivation and starvation of phosphorous, nitrogen, light or CO_2, and/or
    ii. said biotic stress is selected from the negative impact of fungi, bacteria, viruses, insects, wounding by herbivores and biological competition, and/or
20 iii. said chemical stress is selected from herbicides, herbicide safener, insecticides, fungicides, plant secondary metabolites and synthetic or natural compounds that induce plant defense, in particular flagellin, more particularly flg22.

12. The method according to any one of the preceding claims, wherein said inhibitor of
25 DNA-methylation is selected from 5-azacytidine, 5-aza-2'-deoxycytidine, 5-fluoro-2'-deoxycytidine, 5,6-dihydro-5-azacytidine and zebularine.

13. The method according to any one of the preceding claims, wherein said inhibitor of transcription is an RNA polymerase inhibitor, in particular an RNA polymerase II inhibitor, an RNA polymerase IV inhibitor or an RNA polymerase V inhibitor, more
30 particular an RNA polymerase II inhibitor.

14. The method according to claim 13, wherein said RNA polymerase II inhibitor is selected from
    i. amatoxins, in particular alpha-amanitin,
    ii. derivatives of amatoxins, in particular alpha-amanitin oleate,

iii. nucleoside analogues, in particular 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB),

iv. actinomycin D,

v. flavopiridol,

5 vi. triptolide.

15. The method according to claim 14, wherein said amatoxin, in particular alpha-amanitin is used at a concentration of 0.0005 µg/ml to 50 µg/ml, in particular 0.001 µg/ml to 25 µg/ml, more particular 0.005 µg/ml to 20 µg/ml, even more particular 0.005 µg/ml to 5 µg/ml.

10 16. The method according to any one of the preceding claims, wherein said inhibitor of DNA-methylation, in particular zebularine, is used at a concentration of 5 µM to 100 µM, in particular 10 µM to 80 µM, more particular 10 µM to 40 µM, even more particular 20 µM to 40 µM.

17. The method according to any one of claims 8 to 16, wherein said increased genetic 15 and/or epigenetic variation in a plurality of eukaryotic organisms results in increased resistance of said organisms to said abiotic, biotic or chemical stress said organisms have been exposed to.

18. Use of a composition in a method according to claims 1 to 17, wherein the composition comprises an inhibitor of transcription and an inhibitor of DNA-20 methylation, and wherein the ratio of the molar concentrations of the inhibitor of transcription, in particular alpha-amanitin, to the inhibitor of DNA-methylation, in particular zebularine, is 0.000005 to 11, more particular 0.000125 to 2, even more particular 0.000125 to 0.125.

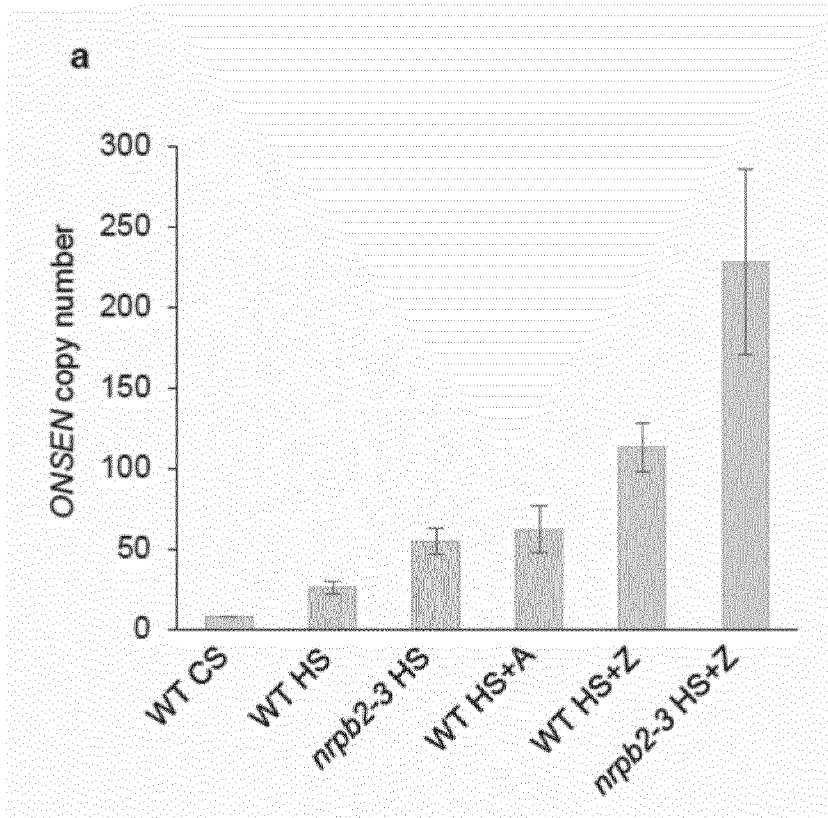
19. The use of a composition according to claim 18, wherein

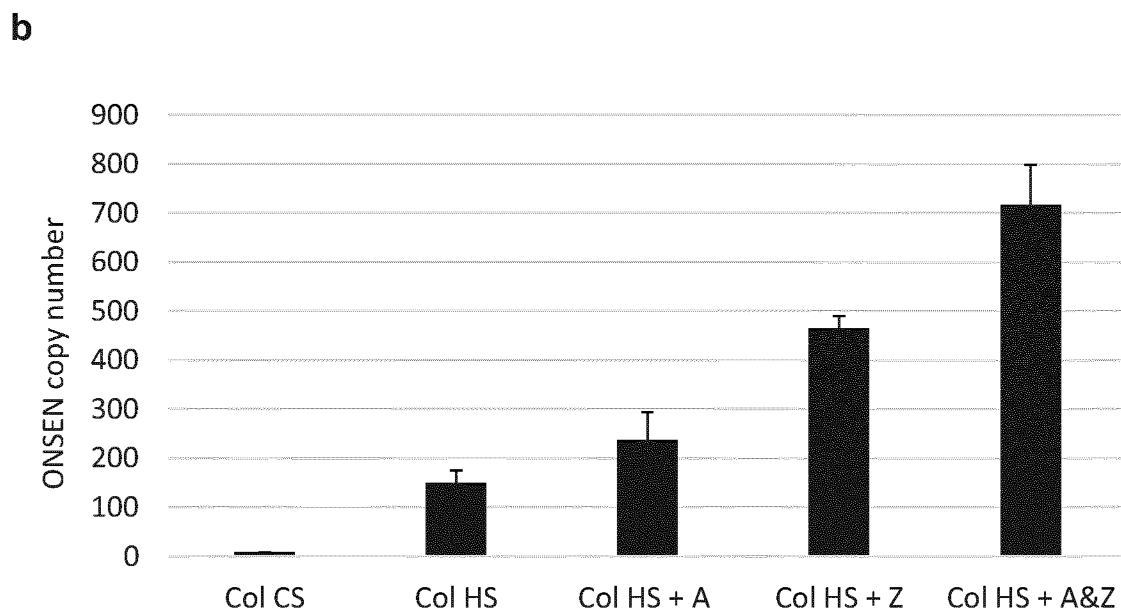
25 - said inhibitor of DNA-methylation, in particular zebularine, is used at a concentration of 5 µM to 100 µM, in particular 10 µM to 80 µM, more particular 10 µM to 40 µM, even more particular 20 µM to 40 µM

and / or

- said amatoxin, in particular alpha-amanitin is used at a concentration of 30 0.0005 µg/ml to 50 µg/ml, in particular 0.001 µg/ml to 25 µg/ml, more particular 0.005 µg/ml to 20 µg/ml, even more particular 0.005 µg/ml to 5 µg/ml.

Fig. 1a, b





5

Fig. 2

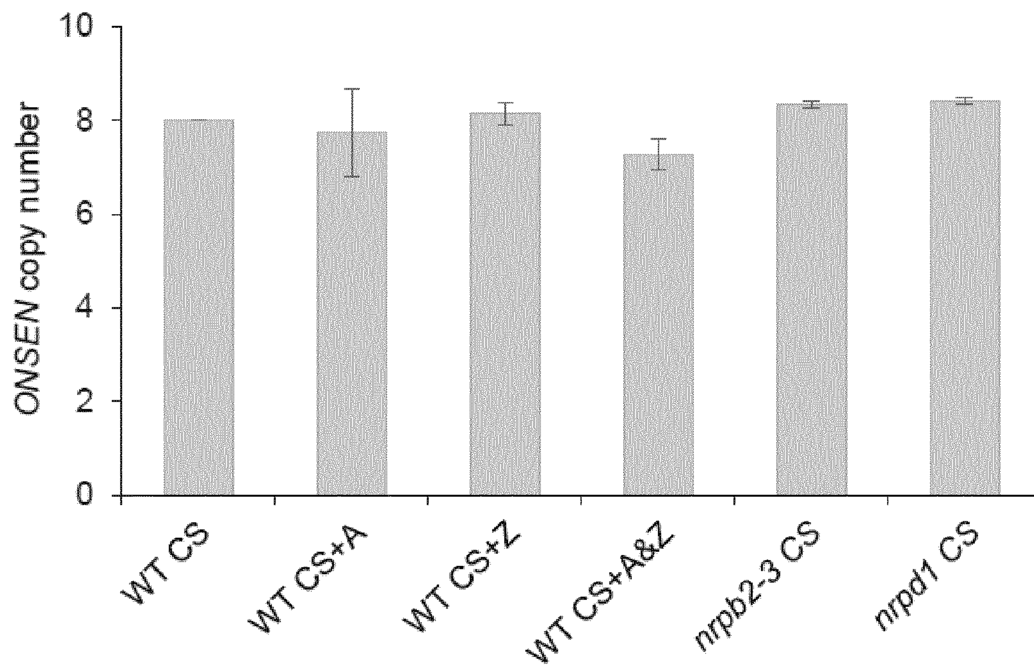
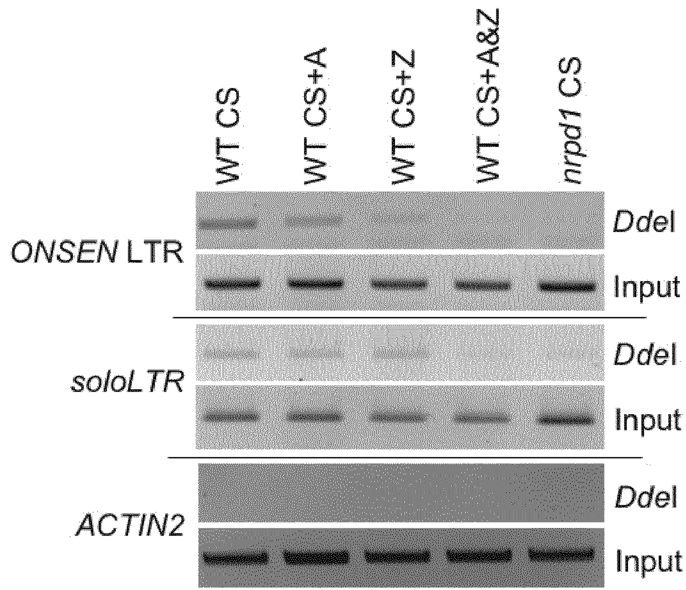


Fig. 3a, b



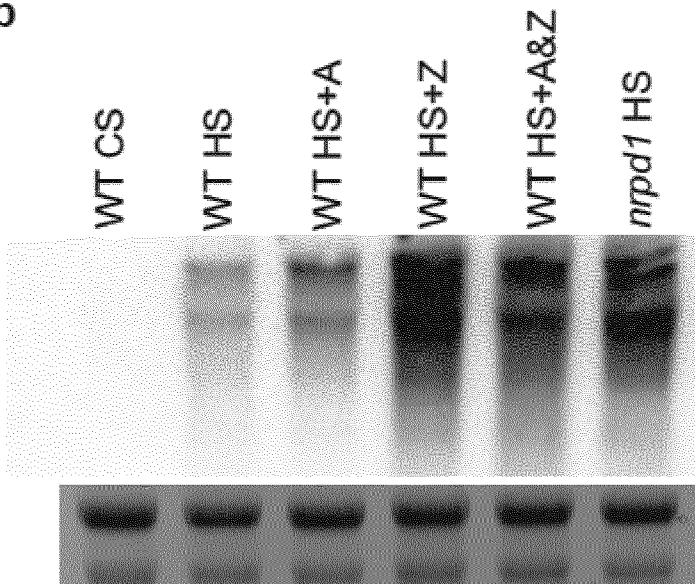


Fig. 3c

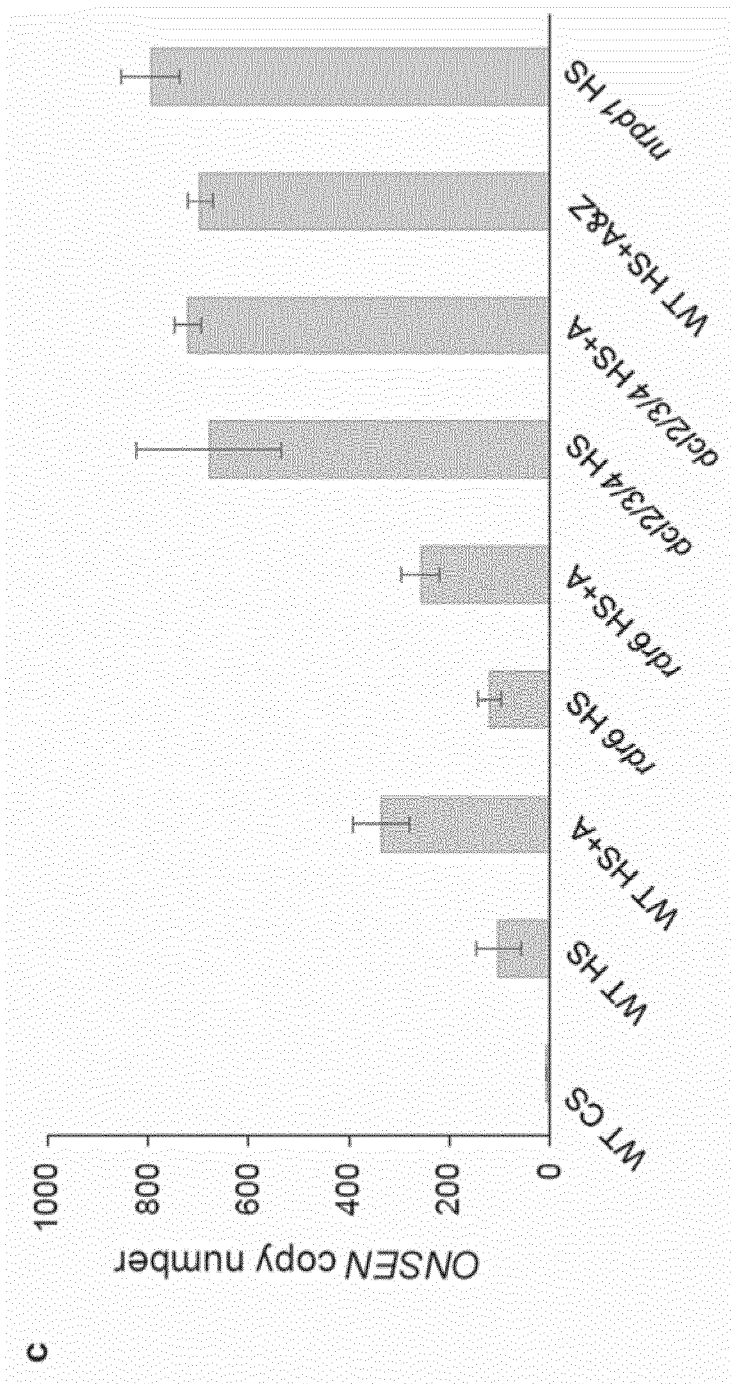


Fig. 4a

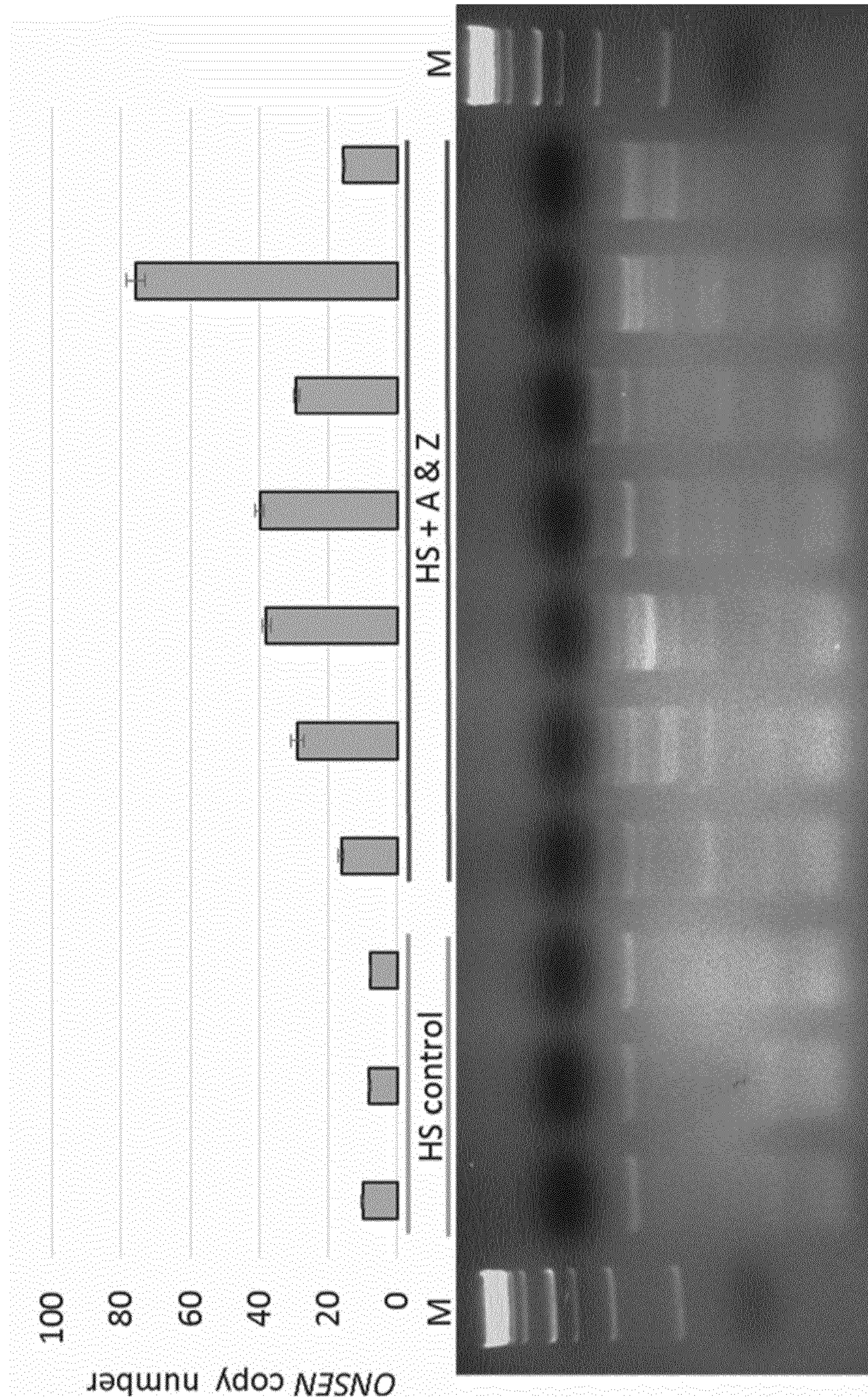


Fig. 4b

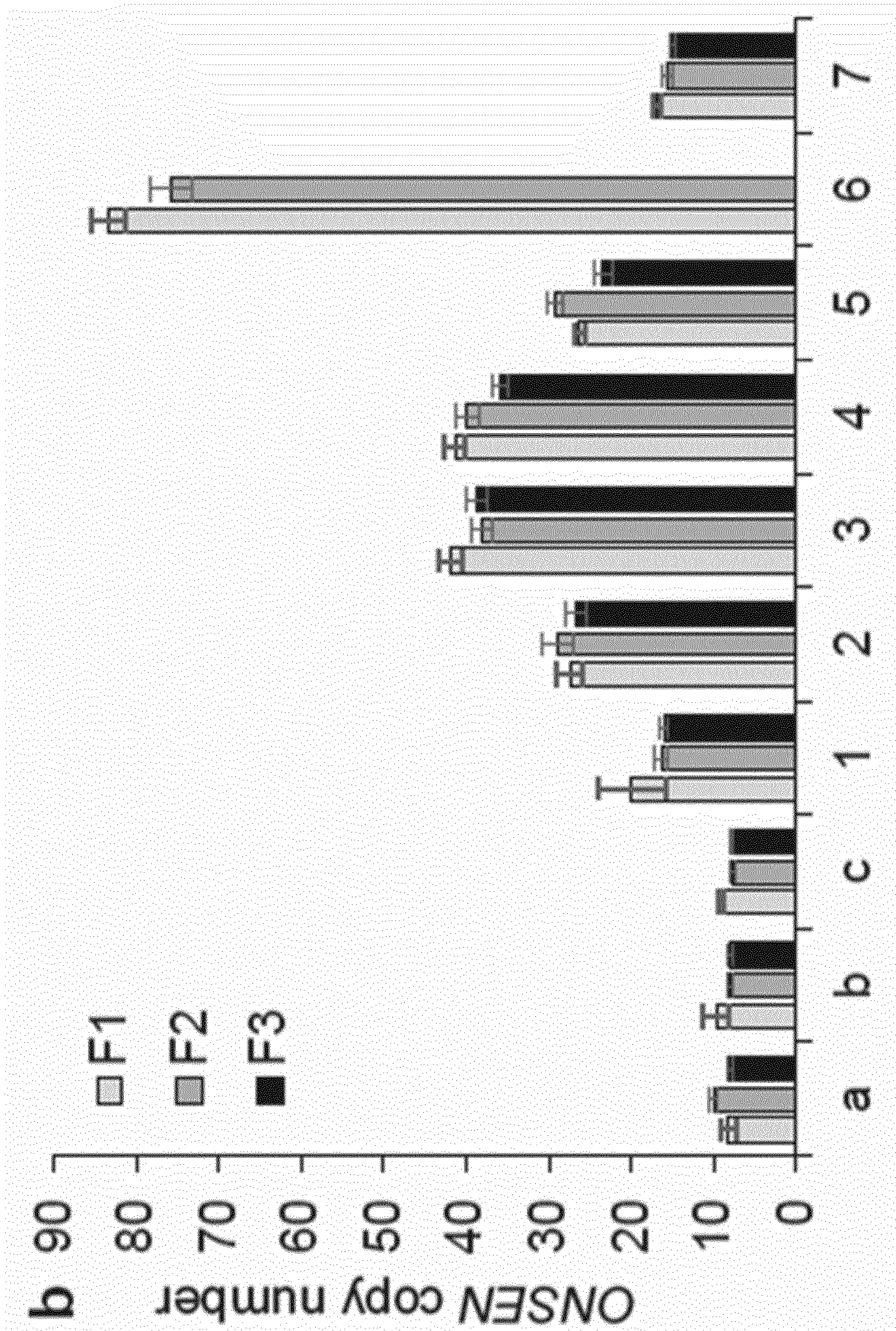
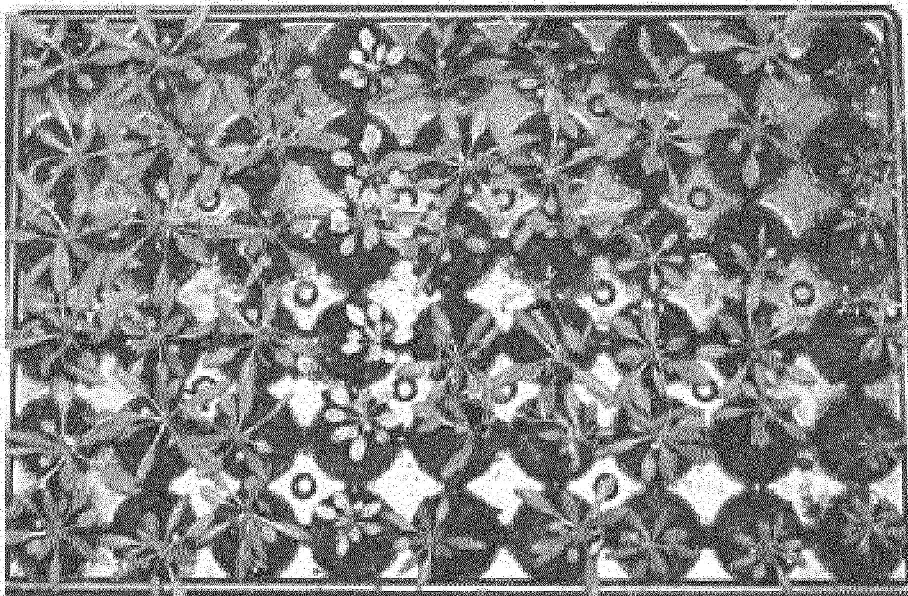


Fig. 4c,d



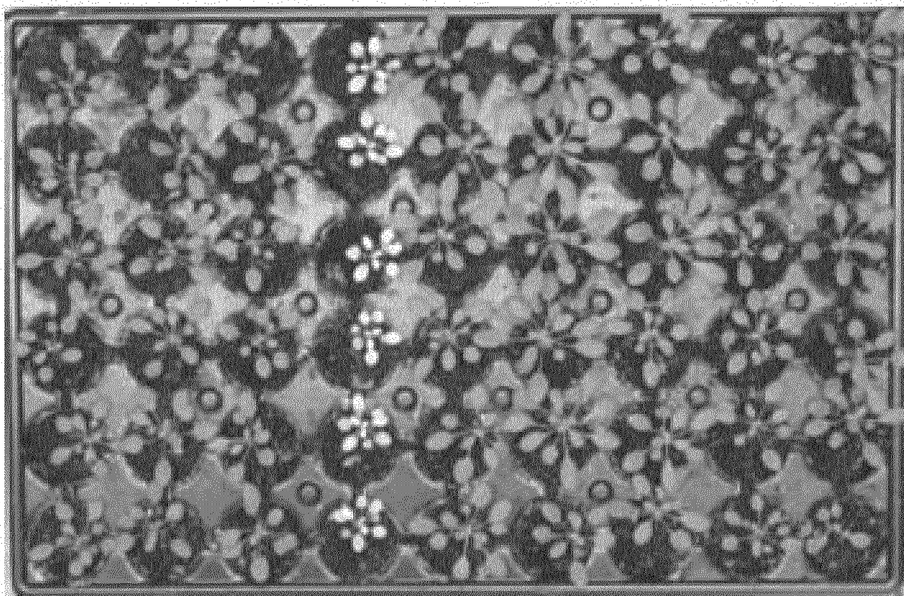


Fig. 5

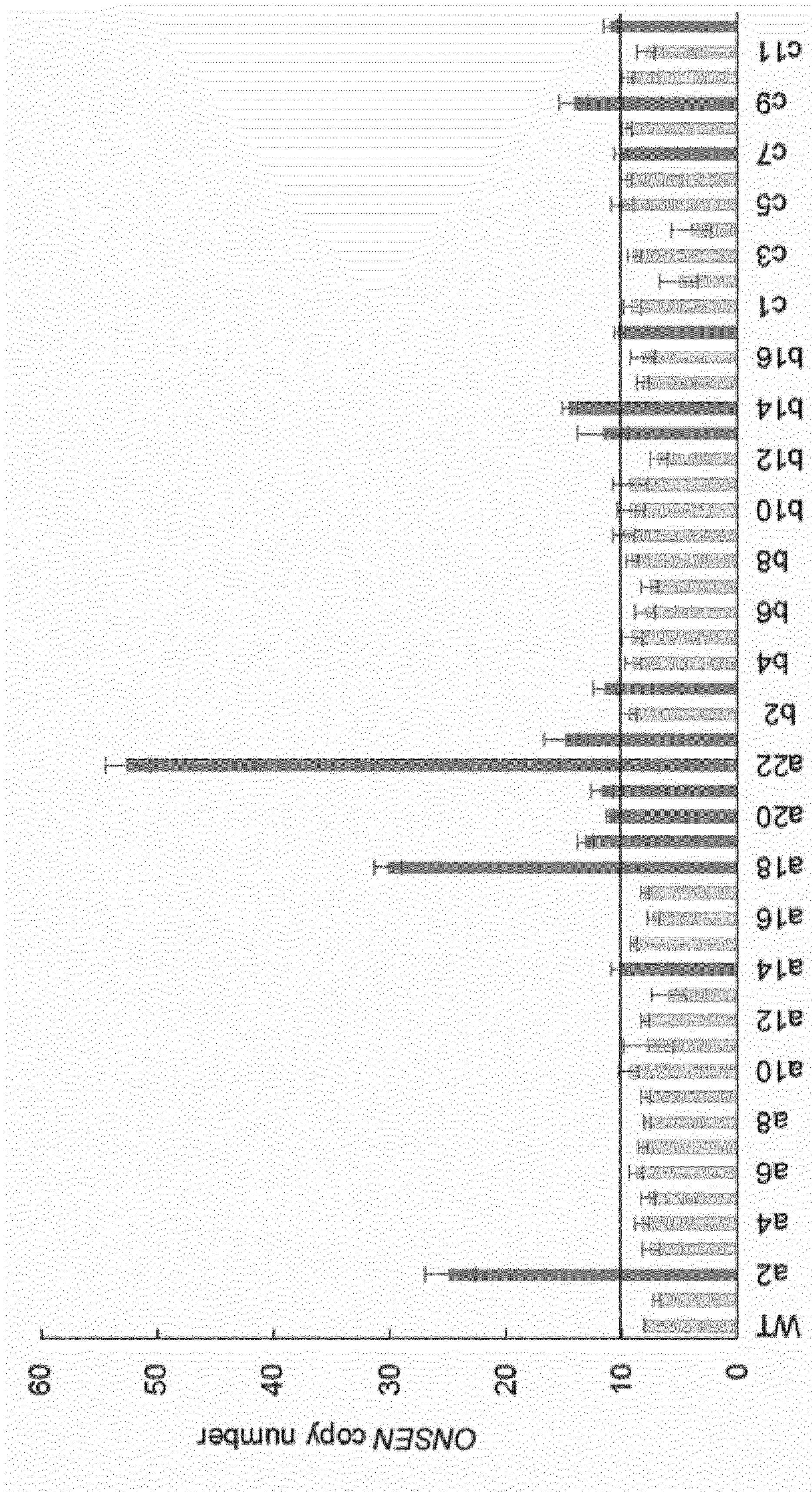


Fig. 6a

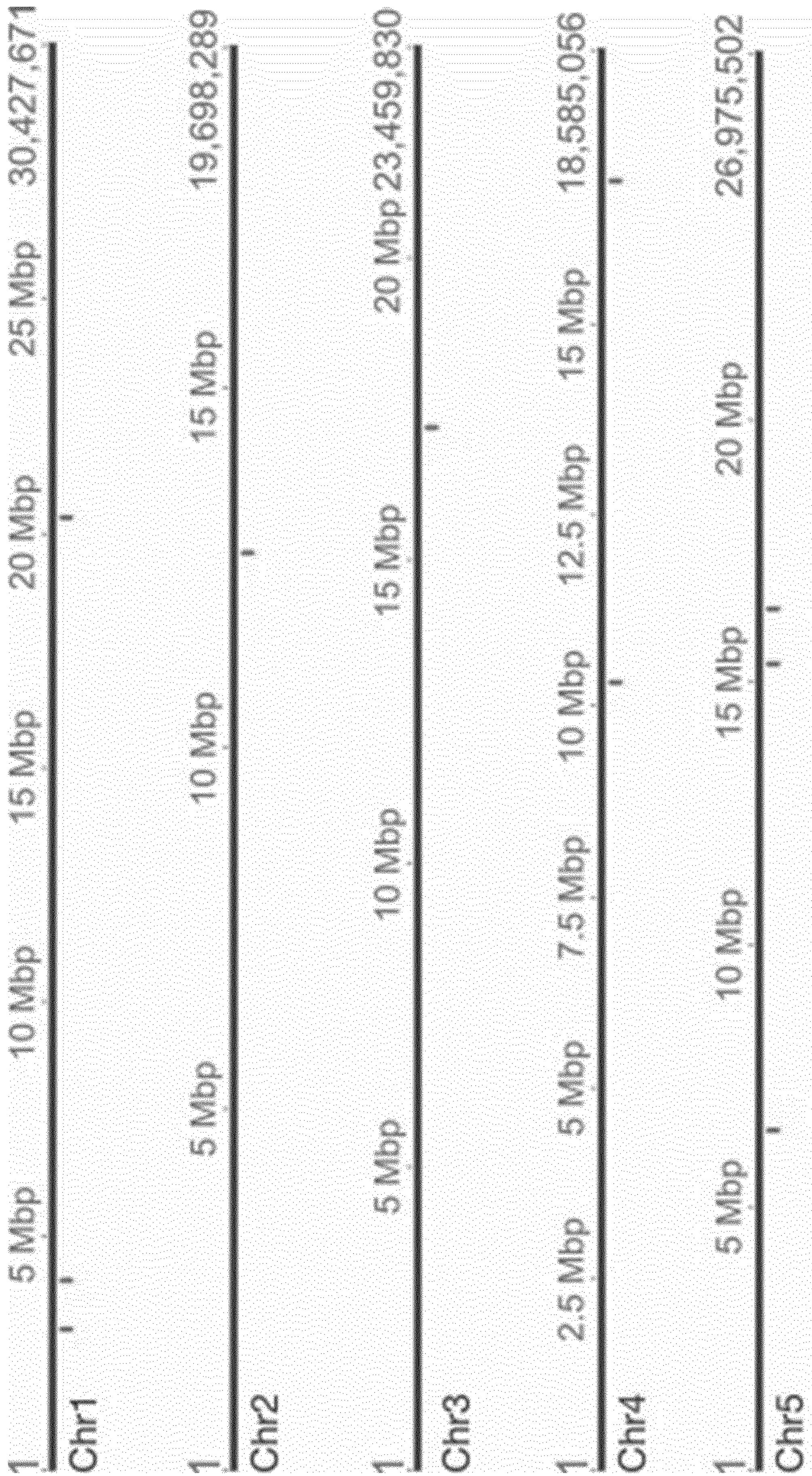


Fig. 6b

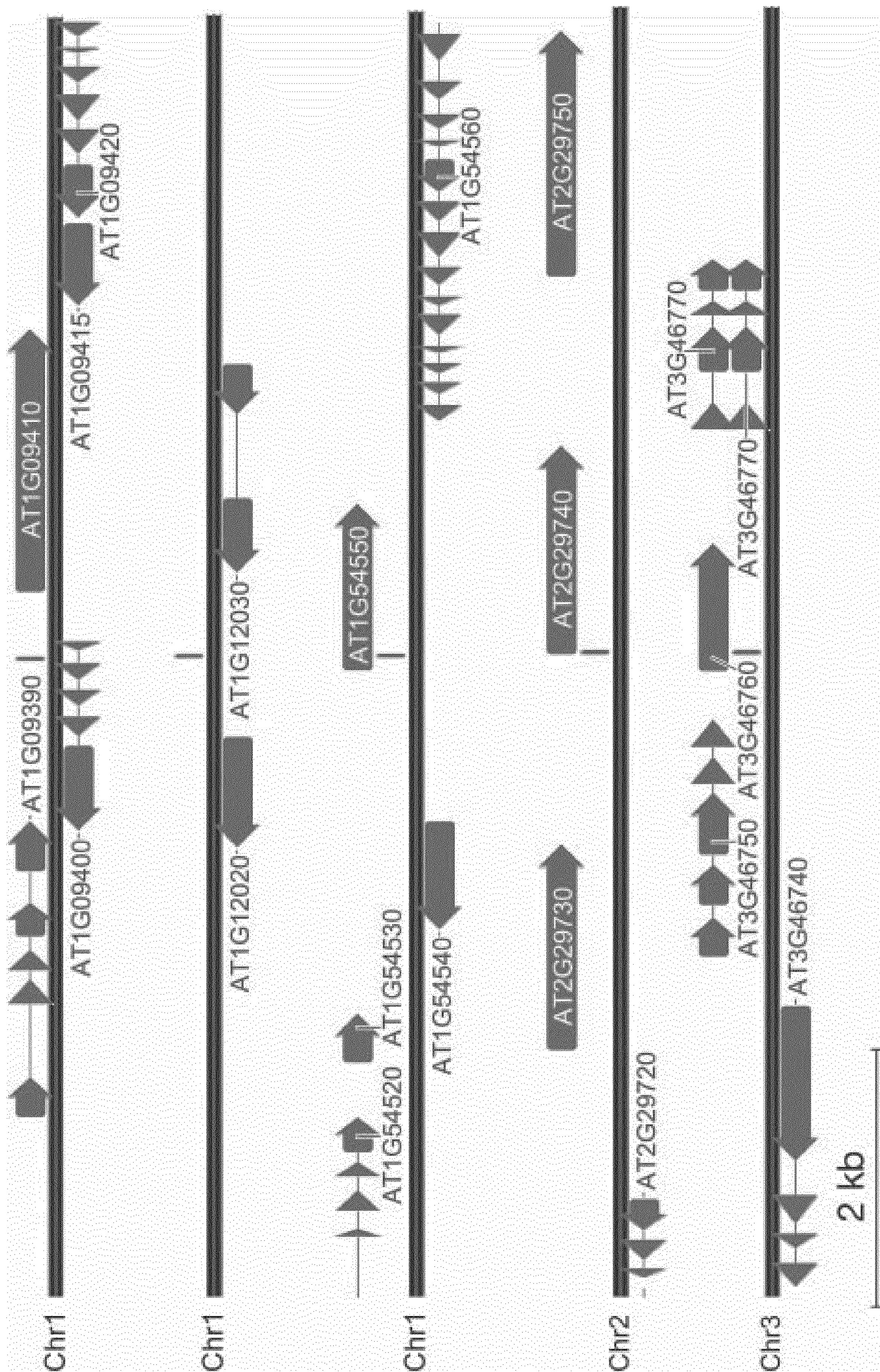


Fig. 6b continued

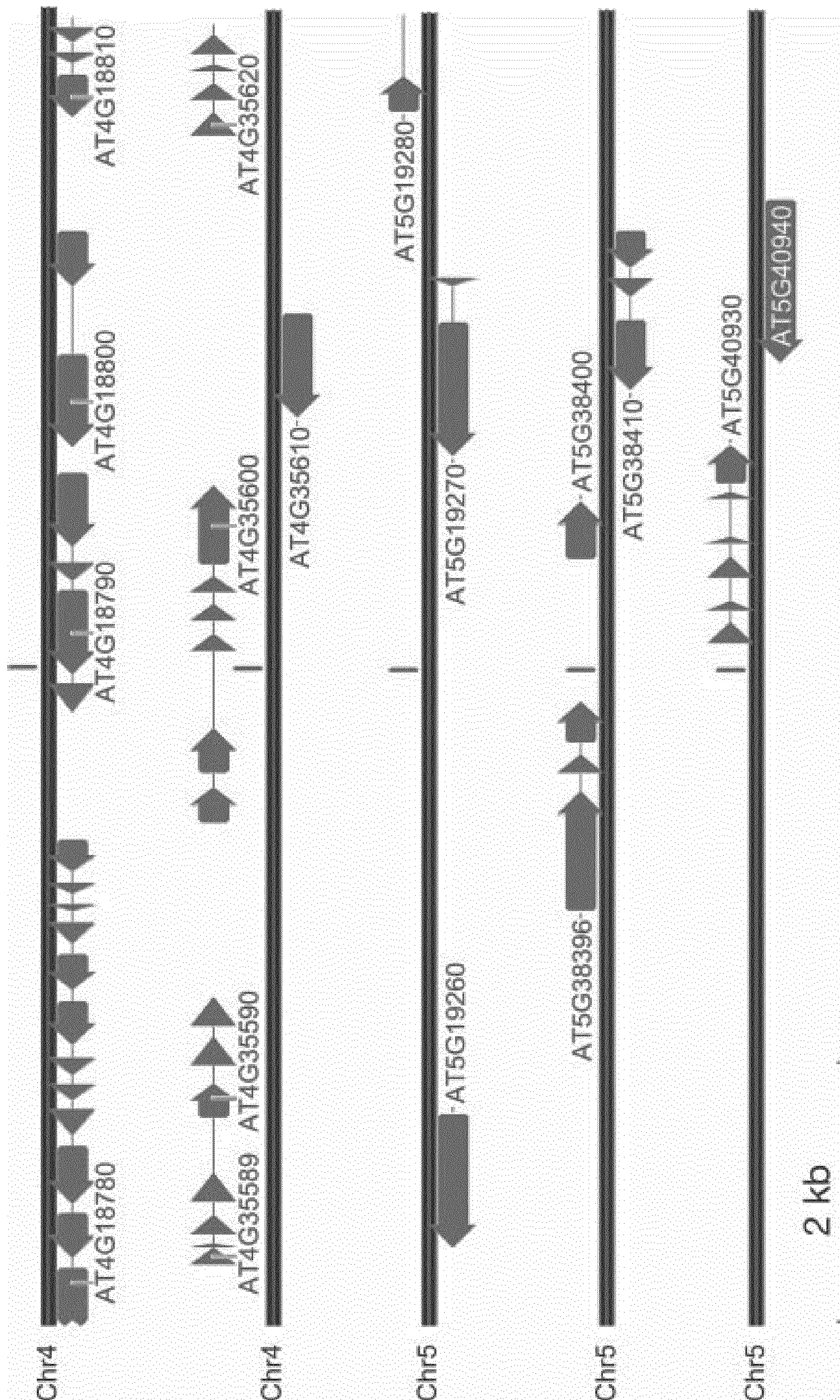


Fig. 7a

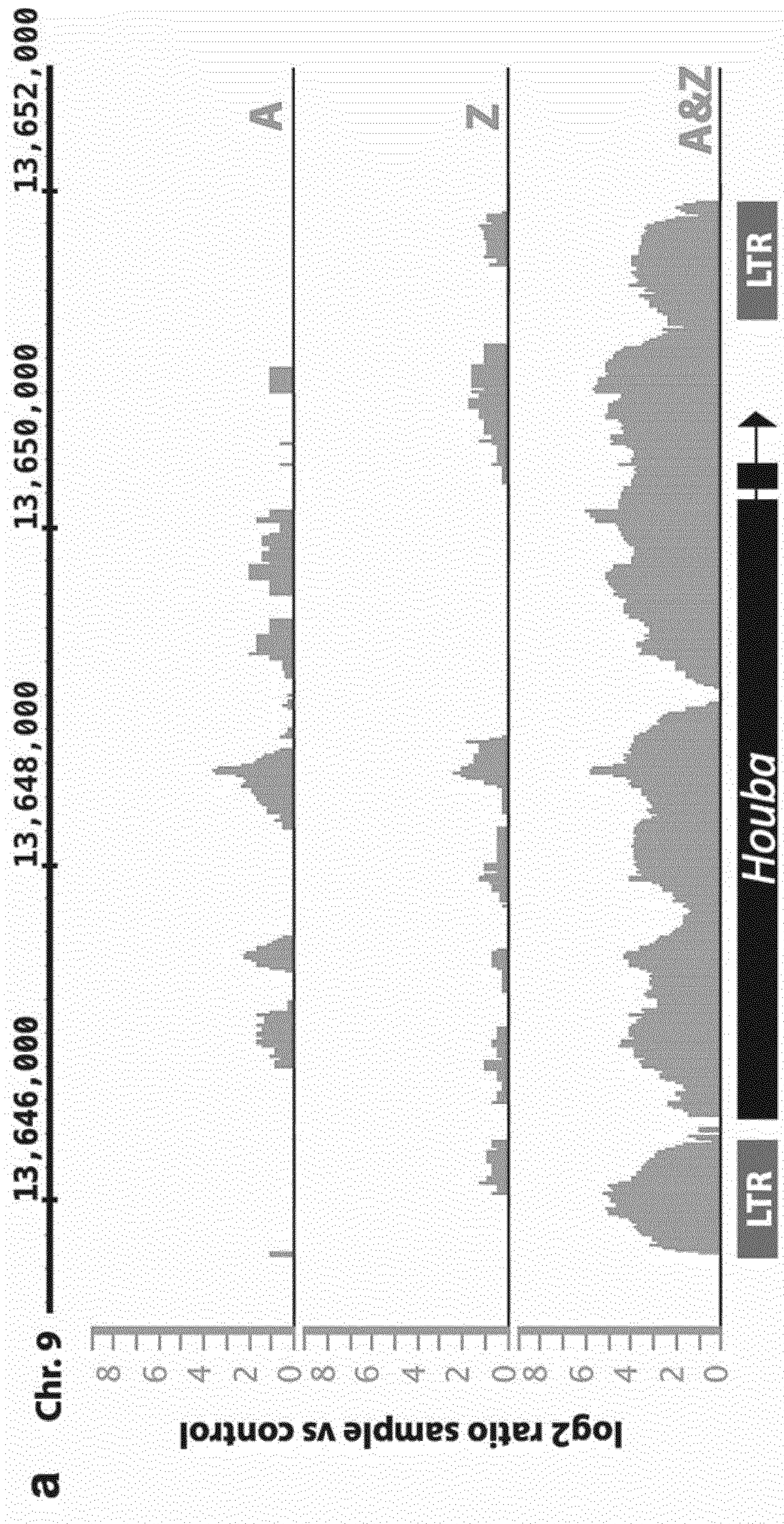
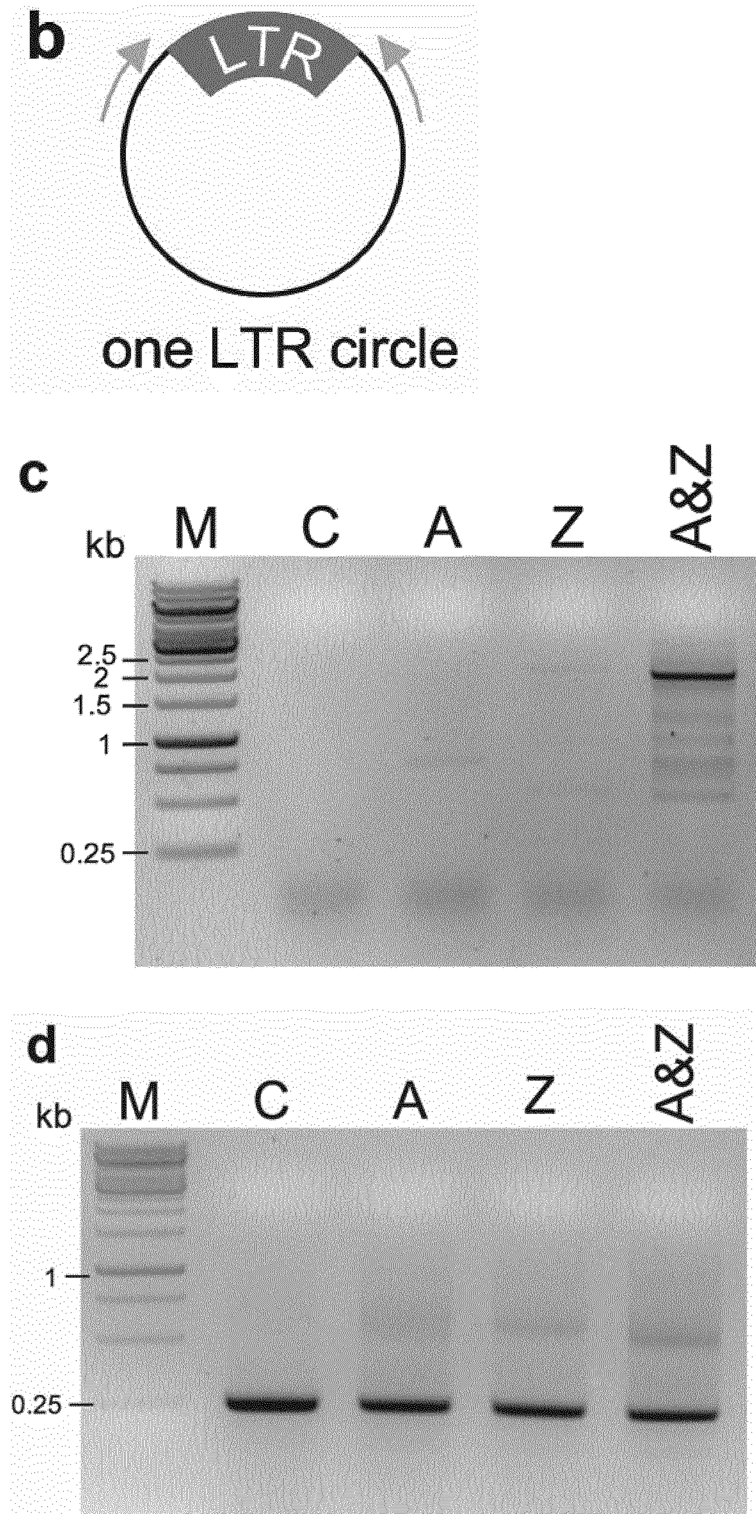


Fig. 7b-d



5

Fig.8

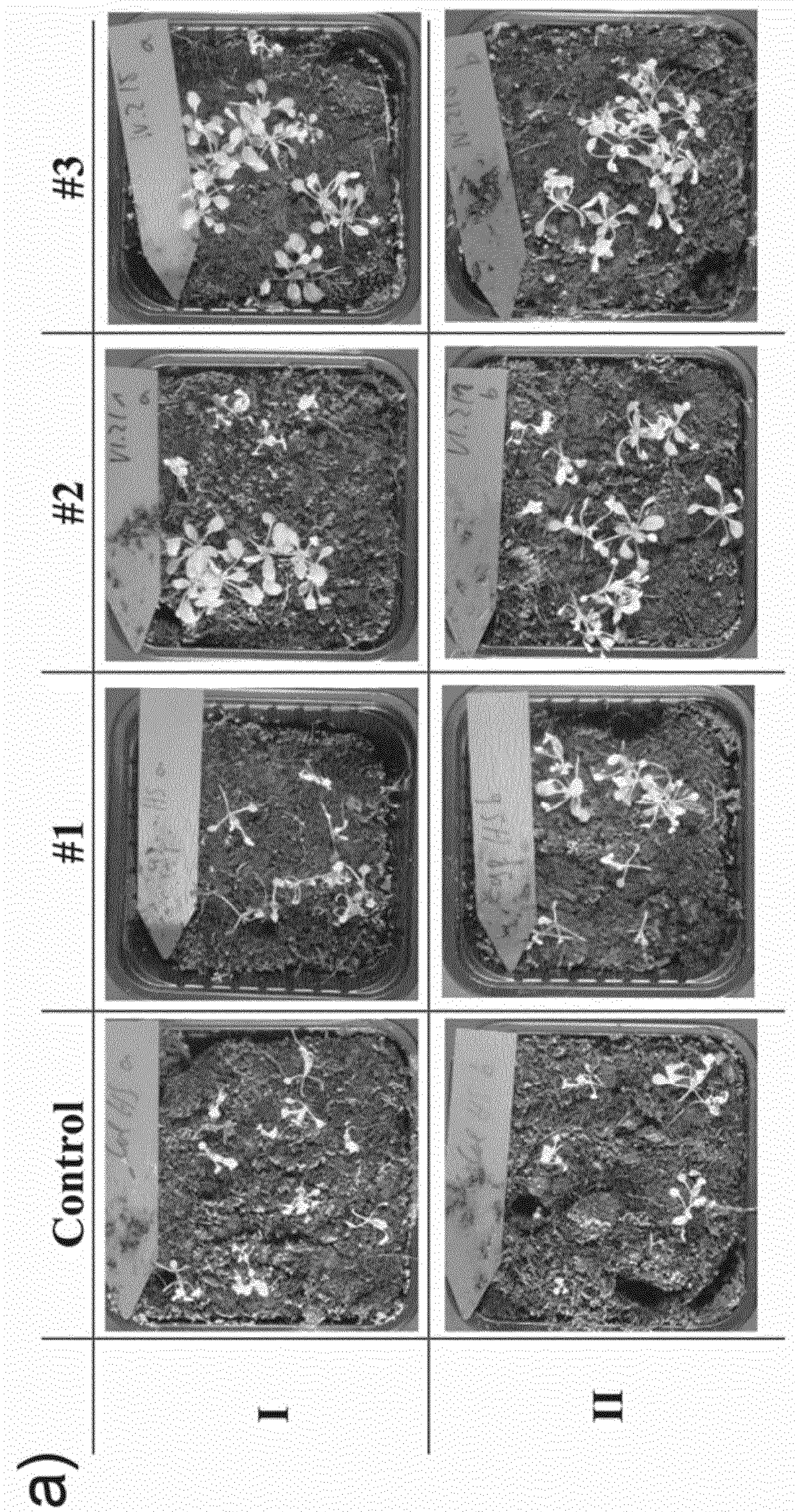


Fig. 8 continued

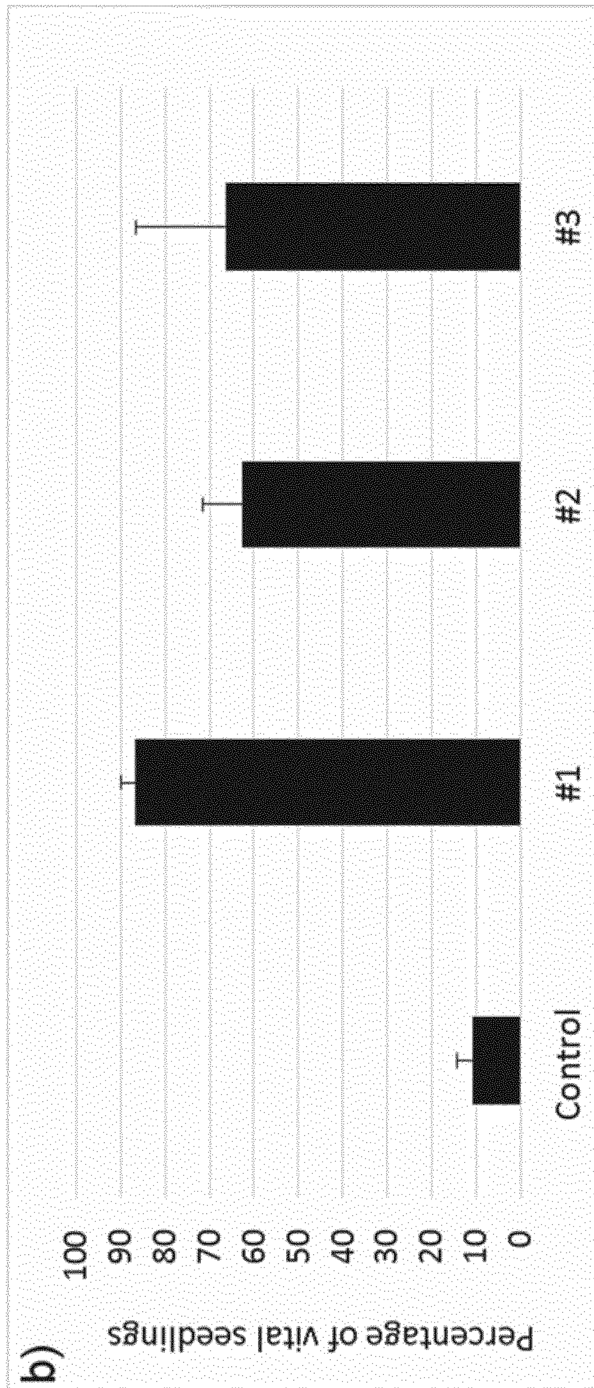


Fig. 9

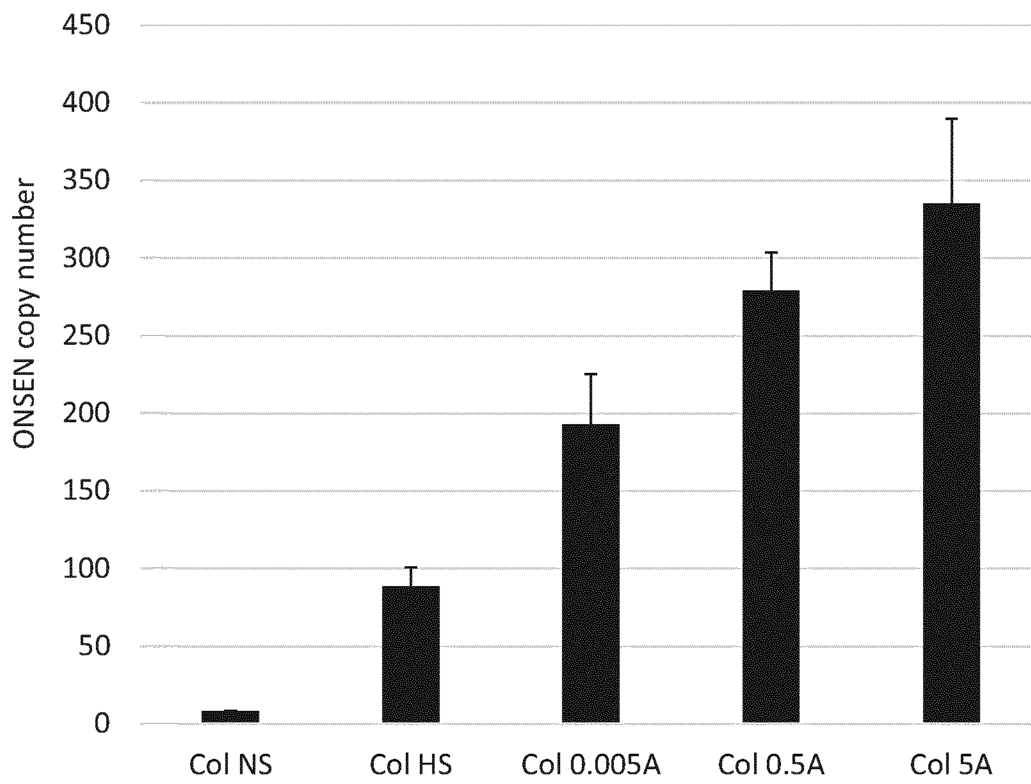
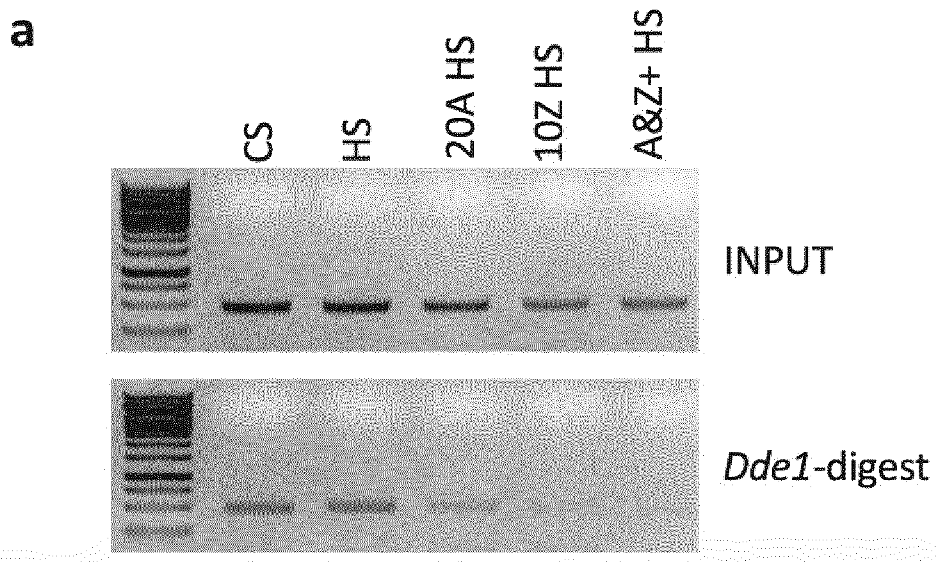
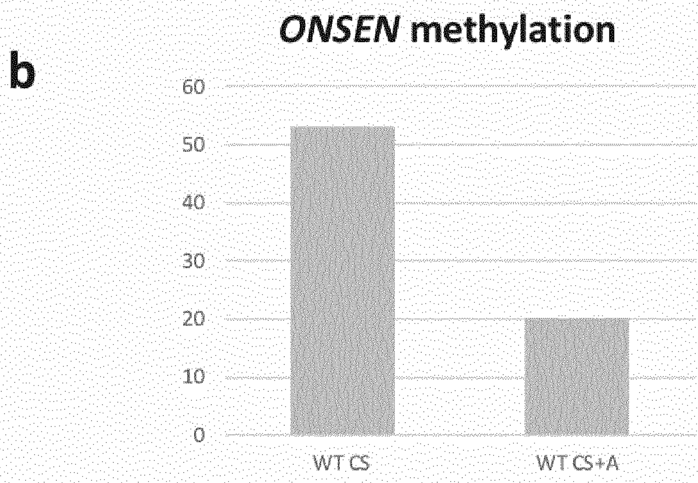


Fig. 10 a,b,c





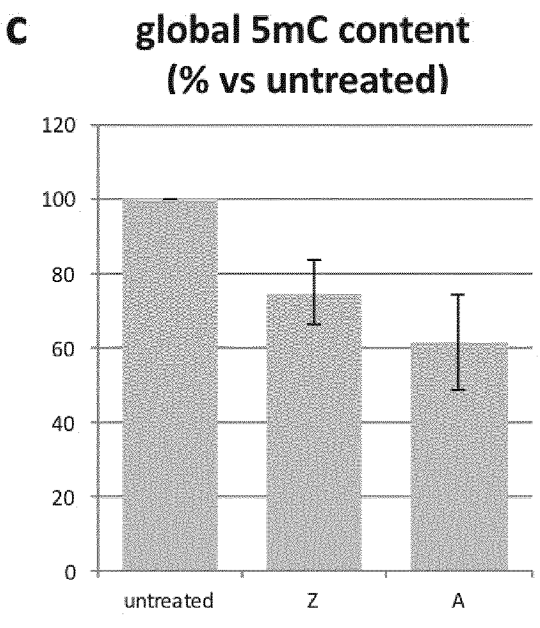


Fig. 10d

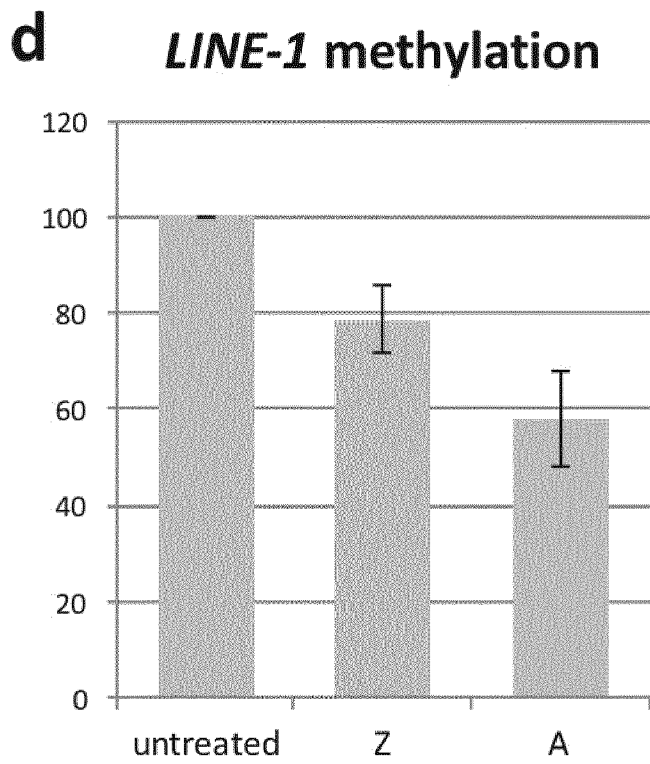


Fig. 11

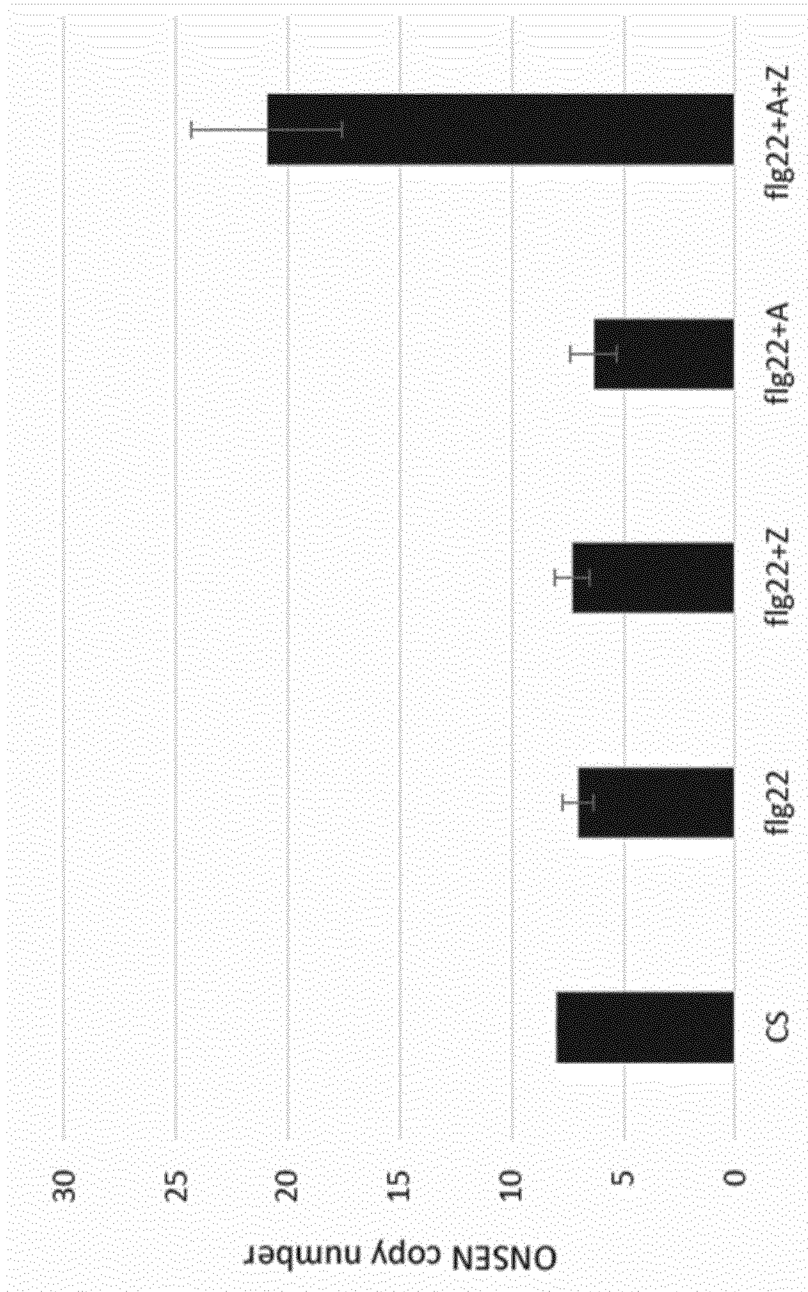


Fig. 12

5

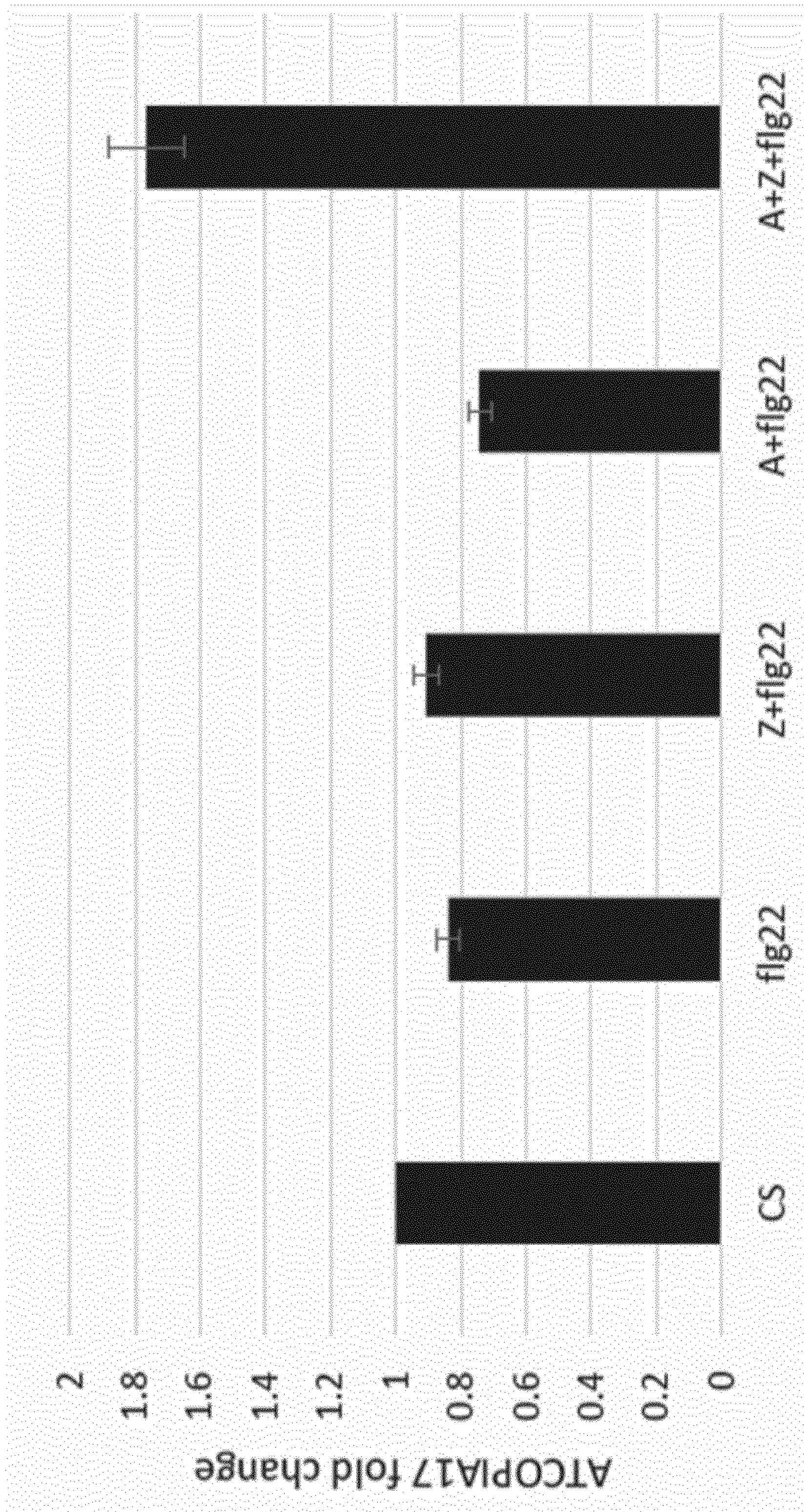
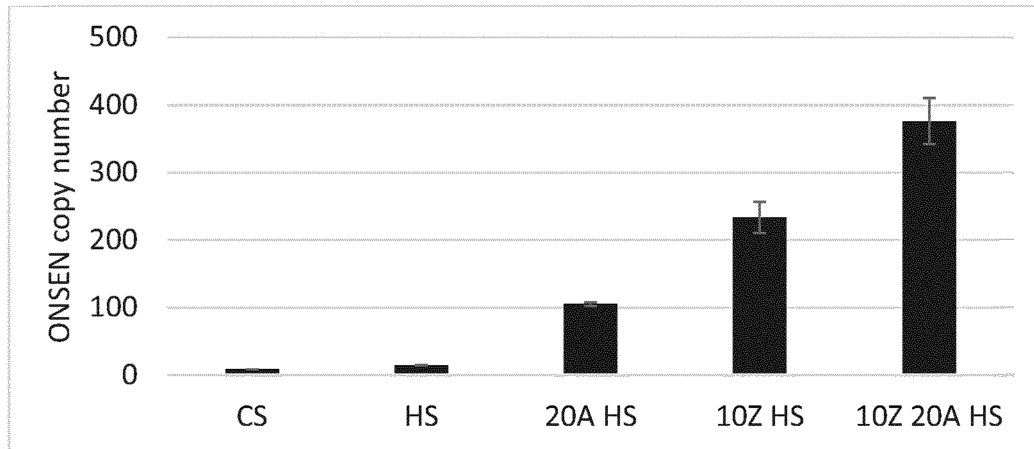


Fig. 13



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/079276

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/01 A61K31/7068 C07H19/06 A61L31/16 A01H1/06
A01H3/04
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A01H A61L A61K C07H C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, CHEM ABS Data, WPI Data, BIOSIS, Sequence Search, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DHAVAL VARSHNEY ET AL: "SINE transcription by RNA polymerase III is suppressed by histone methylation but not by DNA methylation", NATURE COMMUNICATIONS, vol. 6, 23 March 2015 (2015-03-23), page 6569, XP055272268, DOI: 10.1038/ncomms7569 abstract; pages 5-8, 10; Table 4 -----	1-3, 12-16, 18,19

-/--

[X] Further documents are listed in the continuation of Box C.

[ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 February 2017	07/03/2017

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Kurz, Birgit

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TUNCAY BAUBEC ET AL: "Effective, homogeneous and transient interference with cytosine methylation in plant genomic DNA by zebularine", THE PLANT JOURNAL, vol. 57, no. 3, 1 February 2009 (2009-02-01), pages 542-554, XP055271538, GB ISSN: 0960-7412, DOI: 10.1111/j.1365-313X.2008.03699.x Abstract; pages 543,546-548, 550; figure 6 -----	8,12,16
A	VLADIMIR V. CAVRAK ET AL: "How a Retrotransposon Exploits the Plant's Heat Stress Response for Its Activation", PLOS GENETICS, vol. 10, no. 1, 30 January 2014 (2014-01-30), page e1004115, XP055271558, DOI: 10.1371/journal.pgen.1004115 the whole document -----	1-19
A	W. MATSUNAGA ET AL: "The effects of heat induction and the siRNA biogenesis pathway on the transgenerational transposition of ONSEN, a copia-like retrotransposon in Arabidopsis thaliana", PLANT AND CELL PHYSIOLOGY, vol. 53, no. 5, 14 December 2011 (2011-12-14), pages 824-833, XP055271560, UK ISSN: 0032-0781, DOI: 10.1093/pcp/pcr179 the whole document -----	1-19
A	MARJORI A. MATZKE ET AL: "RNA-directed DNA methylation: an epigenetic pathway of increasing complexity", NATURE REVIEWS GENETICS, vol. 15, no. 6, 8 May 2014 (2014-05-08), pages 394-408, XP055271562, GB ISSN: 1471-0056, DOI: 10.1038/nrg3683 the whole document -----	1-19
A	WATARU MATSUNAGA ET AL: "Role of RNA polymerase IV in plant small RNA metabolism.", FRONTIERS IN PLANT SCIENCE, vol. 104, no. e129, 9 February 2015 (2015-02-09), page 4536, XP055271567, DOI: 10.1073/pnas.0611456104 the whole document -----	1-19
	-/--	

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BUCHER ETIENNE ET AL: "Epigenetic control of transposon transcription and mobility inArabidopsis", CURRENT OPINION IN PLANT BIOLOGY, vol. 15, no. 5, 2012, pages 503-510, XP028958933, ISSN: 1369-5266, DOI: 10.1016/J.PBI.2012.08.006 the whole document -----	1-19
A	S. VISPE ET AL: "Triptolide is an inhibitor of RNA polymerase I and II-dependent transcription leading predominantly to down-regulation of short-lived mRNA", MOLECULAR CANCER THERAPEUTICS, vol. 8, no. 10, 1 October 2009 (2009-10-01), pages 2780-2790, XP055272275, US ISSN: 1535-7163, DOI: 10.1158/1535-7163.MCT-09-0549 the whole document -----	1-19
A	H. SAZE ET AL: "DNA Methylation in Plants: Relationship to Small RNAs and Histone Modifications, and Functions in Transposon Inactivation", PLANT AND CELL PHYSIOLOGY, vol. 53, no. 5, 1 February 2012 (2012-02-01), pages 766-784, XP055272283, UK ISSN: 0032-0781, DOI: 10.1093/pcp/pcs008 abstract; pages 774, 776, 777 -----	1-19

1

11.3 Appendix III: Newsletter article: Thieme, M. (2017) Putting plants in shool: On the potential of epigentic memory in crop breeding. *Plant Science News* 32: 4-5. (2 pages)

Putting plants in school: On the potential of epigenetic memory in crop breeding

Next generation of cold-stressed soybeans growing in the Botanical Garden, University of Basel. © Michael Thieme

Michael Thieme

Conventional and organic plant breeding is based on the presence of naturally occurring, or induced random changes in the DNA sequence of an organism, so called mutations. These mutations result in both genetic and phenotypic diversity, and can be used to select advantageous genotypes or traits during the breeding process. Figuratively, breeding for desired traits can be compared with a laborious search for new and meaningful sentences in a book where single letters or words were randomly erased or exchanged.

Innovative approaches: stimulating the short term memory

As plants are not equipped with legs or wings to escape life-threatening situations, they had to evolve a great diversity of mechanisms to keep up with evolution. One of the most fascinating skills of plants is their distinct ability to remember situations they experienced during their life cycle. Surprisingly, this knowledge of the plant's past not only influences the individual plant itself, but can under certain circumstances also be passed on to the plant's progeny to prepare it for similar situations.

This concept was already been proposed in the 19th century by Jean-Baptiste de Lamarck. The phenomenon that information other than the genetic sequence of an organism is passed on to the next generation is nowadays described as «epigenetic memory».

Metaphorically speaking, this additional, epigenetic information tells the plant where exactly in its large «genetic encyclopedia», on which page or in which sentence, it can read to overcome a threatening situation. Several studies provide proof of evidence for such a transgenerational memory in plants. For example, plants that were in contact with pathogens revealed an increased pathogen-resistance in the next generation, compared to plants that were sheltered from harm.

The formation of this epigenetic memory is based on two biochemical mechanisms: DNA methylation and histone modification. In order to shape the specific epigenetic landscape of an organism, methyl groups are added to or removed from specific nucleotides of the genetic code. This modifies their meaning and, for example, suppresses the activity of nearby genes. One can compare this mechanism with highlighting certain words in a book with a text marker. During histone modification, important structural modifications are made, which makes certain regions of the genetic code more or less easy to read. Again, compared with a book, this mechanism is analogous to changing the font size of certain paragraphs.

However, similar to humans, plants apparently tend to forget without permanent training. Hence, even if the plant's memory can be passed on to two or three generations, this is hardly enough to face future challenges in agriculture.

The total recall: plants use «cheat sheets»

Besides using text markers to highlight certain words or sentences, plants have an even more powerful natural genetic resource that can be harnessed for breeding purposes. Referring to the analogy as the genome being a huge encyclopedia, this genetic resource, so called retrotransposons (retroTEs), can be pictured as «cheat sheets». Just like during a school exam, if placed at the correct position in the notebook, these highly informative pages can positively influence the performance of a plant. Fascinatingly, it seems that these mobile genetic elements are particularly effective in regulating flanking genetic regions under stress conditions. They function as natural genetic switches and direct linkers

of the environment to the genome. They are considered as hot-spots of epigenetic memory.

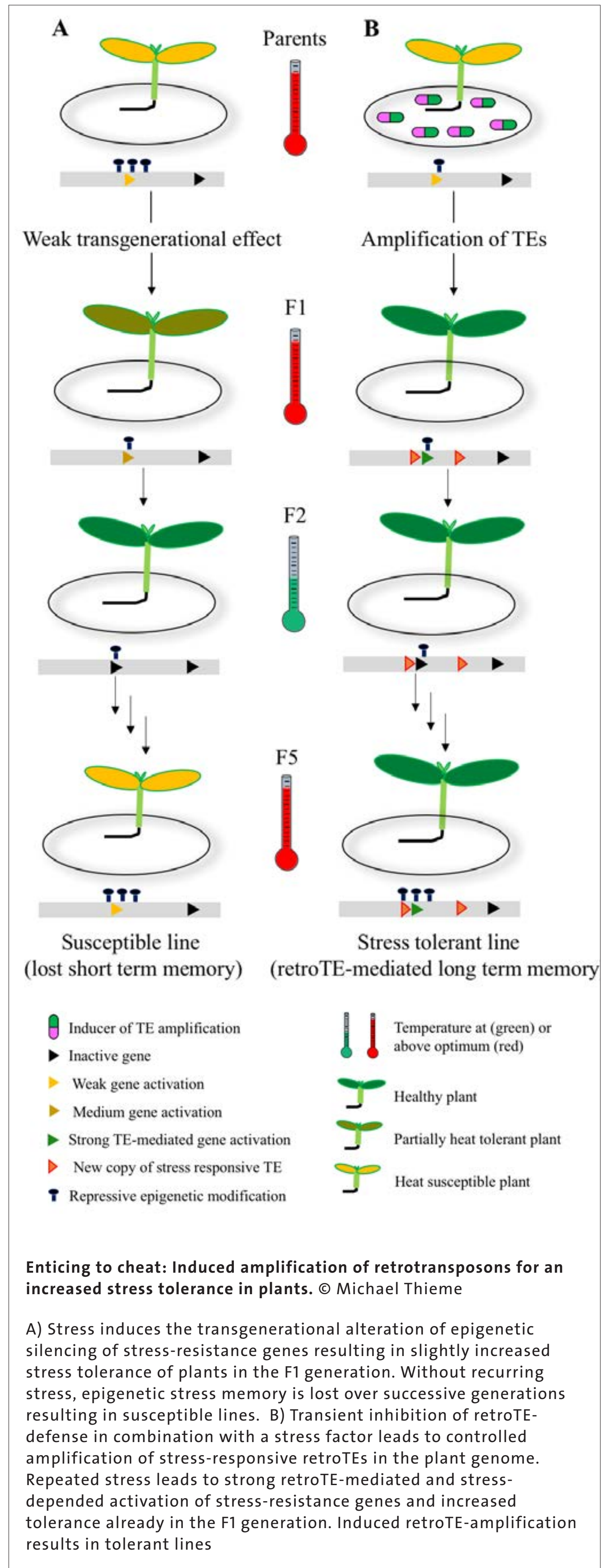
The best part about these «cheat sheets» however is that they can be copied and firmly integrated somewhere else in the encyclopedia. The huge amount of such retroTEs in plants strongly suggests that they serve as a genetic backup resource that can be used in various different situations.

Reinforcing the memory: encouraging plants to copy their «cheat sheets»

So far, the major challenge in making use of retroTEs for crop breeding was that plants would normally avoid the risk of producing too many copies of their «cheat sheets». Hence, to avoid their uncontrolled proliferation under optimal growth conditions these elements are normally strictly repressed by the plant. During my PhD project I discovered a key mechanism at the origin of retroTEs repression that can easily be targeted with a simple transient application of drugs. I was able to show that plants grown on two specific compounds produce huge amounts of a specific «cheat sheet», called *ONSEN* (jap. «hot spring») in response to heat stress. Most importantly, I demonstrated that such treatments efficiently result in new stably inserted copies of these heat responsive elements in the progeny of treated and heat stressed plants (Thieme *et al.* 2017).

Its potential in crop breeding

As this transient controllable process of «enticing to cheat» involves only the stimulation of natural processes and does not include controversial genetic engineering, it has the potential to revolutionize breeding for the organic sector. Based on what is already known about the contribution of retroTEs to plant diversity and evolution, I now believe that the term «hidden treasure» (Mirouze and Vitte, 2014) would actually be more appropriate than only describing them as a collection of ordinary «cheat sheets». By multiplying these retroTEs in the genome, they can, amongst other advantageous effects, cause valuable «gain of function» mutations and contribute to the evolution of new gene regulatory networks. In case of future challenges related to global warming, imagine having a crop with a «cheat sheet» saying «how to adapt to heat» next to a gene important for heat-resistance.



Enticing to cheat: Induced amplification of retrotransposons for an increased stress tolerance in plants. © Michael Thieme

A) Stress induces the transgenerational alteration of epigenetic silencing of stress-resistance genes resulting in slightly increased stress tolerance of plants in the F1 generation. Without recurring stress, epigenetic stress memory is lost over successive generations resulting in susceptible lines. B) Transient inhibition of retroTE-defense in combination with a stress factor leads to controlled amplification of stress-responsive retroTEs in the plant genome. Repeated stress leads to strong retroTE-mediated and stress-depended activation of stress-resistance genes and increased tolerance already in the F1 generation. Induced retroTE-amplification results in tolerant lines

References

Mirouze, M. and Vitte, C. (2014) Transposable elements, a treasure trove to decipher epigenetic variation: insights from *Arabidopsis* and crop epigenomes. J Exp Bot 65: 2801-2812.

Thieme, M., Lanciano, S., Balzergue, S., Daccord, N., Mirouze, M. and Bucher, E. (2017) Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. Genome Biol 18: 134.

This research project was supported by a PSC PhD fellowship as part of the Innovative Doctoral Program IDP BRIDGES, a Marie Curie Action of the European Union.

11.4 Appendix IV: Transcript of presentation for "Fachtagung Dialog Grün 2016": Thieme, M. (2017) Der mobile Teil des Pflanzengenoms als Ressource für den zukünftigen Pflanzenschutz. *Neue Technologien in der Pflanzenforschung-eine Alternative zu Pflanzenschutzmitteln?*: 36-39 (4 pages)

Der mobile Teil des Pflanzengenoms als Ressource für den zukünftigen Pflanzenschutz

Michael Thieme

Sogenannte springende Gene oder Transposons (TEs) wurden bereits Mitte des 20. Jahrhunderts erstmals von Barbara McClintock beschrieben, wofür für sie im Jahre 1983 mit dem Nobelpreis ausgezeichnet wurde (McClintock, 1950). Bei ihrer Arbeit mit Mais fiel McClintock auf, dass es oft zu Doppelstrangbrüchen im Genom kam. Außerdem war mit bloßem Auge festellbar, dass manche Maiskörner eine dunklere Färbung aufwiesen als andere. Bei genauerer Untersuchung stellte sie fest, dass sich mobile genetische Elemente, in diesem Fall durch einen «cut and paste» Mechanismus im Genom bewegen können. Je nachdem in welcher Entwicklungsstufe ein Farbengen, durch ein TE unterbrochen wird, entstehen eher hellere oder dunklere Maiskörner. Heute weiß man, dass es neben diesen sogenannten DNA-TEs, welche sich durch einen «cut and paste» Mechanismus bewegen auch Elemente gibt (sog. Retrotransposons, retroTEs), die sich durch einen «copy and paste» Mechanismus im Genom vermehren können. Alle mobilen Elemente zusammen werden als das sog. «Mobilome» bezeichnet.

Betrachtet man die Genome unserer Nutzpflanzen genauer, so stellt man fest, dass sie im Prinzip nichts anderes sind als riesige Sammlungen von TEs. Dabei machen retroTEs in den meisten Fällen den grössten Anteil aus. Zwar besitzen längst nicht mehr alle TEs in Pflanzengenomen die Fähigkeit sich zu bewegen oder sich zu vermehren, jedoch machen sie immerhin bis zu rund 85 % (Mais) des Genoms aus. Aufgrund der Tatsache, dass TEs durch eine unkontrollierte Bewegung im Genom natürlich auch erheblichen Schaden anrichten können, werden sie von der Pflanze streng überwacht. Bis vor wenigen Jahren wurden TEs v.a. als parasitäre, eigennützige Elemente abgestempelt, die ihrem Wirt nur Schaden zufügen können. Dank neuer Labortechniken (z. B. Oxford Nanopore Sequencing) die es erlauben TEs besser zu untersuchen, ist heute zunehmend ein Meinungswechsel in der Fachwelt zu beobachten. So wird der grosse Anteil von TEs im Pflanzengenomen als eine genetische Schatztruhe bezeichnet (Mirouze und Vitte, 2014).

Es gibt inzwischen eine Reihe schöner Beispiele welche die wichtigen Funktionen von TEs für die Pflanze verdeutlichen. So können TEs eine Art Bindeglied zwischen dem Genom und der Umwelt darstellen. Am Beispiel der Blutorange soll dies ge-

nauer erläutert werden: Klassische gelbe Orangen besitzen eine natürliche Mutation neben einem Gen welches für die rote Färbung des Fruchtfleischs verantwortlich ist. Aufgrund dieser Mutation ist das Gen für die Färbung des Fruchtfleisches inaktiv und das Fruchtfleisch der Orange bleibt gelb. Die sizilianische Blutorange «Tarocco» besitzt neben dem entsprechenden Gen eine Insertion eines retroTEs, welches selbst Kälte wahrnehmen kann. Interessanterweise führt die Wahrnehmung von Kälte durch dieses retroTE zu einer Aktivierung des benachbarten Gens für den roten Farbstoff der Blutorange (Butelli et al., 2012). Das bedeutet, dass TEs Umwelteinflüsse, in diesem Fall Kälte, wahrnehmen und an benachbarte Gene vermitteln können. Kurz gesagt: **TEs können benachbarte Gene stressresponsiv machen**.

Neben dieser erstaunlichen Eigenschaft, als stress-abhängiger genetischer Schalter im Genom zu fungieren, können insbesondere retroTEs durch ihren «cut and paste» Mechansimus einen auftretenden Stress auch für ihre eigene Vermehrung verwenden. In der Modellpflanze Ackerschmalwand (*Arabidopidis thaliana*) konnte gezeigt werden, dass das retroTE-*ONSEN* (japanisch für heiße Quelle) nach einem Hitzestress bei 37 °C neue DNA-Kopien von sich selbst herstellen kann (Ito et al., 2011). Normalerweise werden diese neu gebildeten freien *ONSEN*-Kopien wieder abgebaut und das retroTE-*ONSEN* stillgelegt. Wie bereits angedeutet, liegt der Grund hierfür in der strengen und ausgefeilten Regulation von TEs durch die Pflanze. Als Konsequenz dieser strickten Überwachung ließen sich TEs bisher nicht effizient für die Züchtung nutzbar machen.

Induktion von Hitzetoleranz

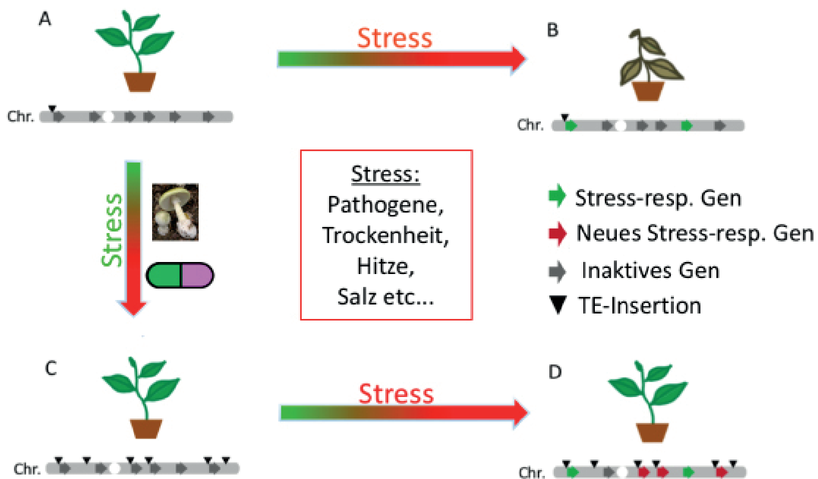
Um eine gezielte Vermehrung von retroTEs ohne Gentechnik zu erreichen, verfolgen wir den Ansatz einer transienten Inhibierung des Regulationsmechanismus. Unter Verwendung zweier Inhibitoren, ist es uns gelungen die Abwehrmechanismen gezielt und für eine kurze, definierte Zeitspanne zu umgehen (Thieme et al., 2017). Bei Zugabe der Inhibitoren während eines auftretenden Hitzestresses konnten wir die Anzahl der neu gebildeten freien *ONSEN*-Kopien etwa verfünffachen. In den Nachkommen konnten wir durch genetische Analyse erstmals neue stabil integrierte *ONSEN*-Kopien detektieren. Somit ist es uns gelungen, mit hoher Effizient und ohne Verwendung gentechnischer Methoden, das hitze-repsonsive retroTE *ONSEN* im Wildtyp zu vermehren. Durch Untersuchung der Kinder-, Enkel- und Urenkel-Generationen der erzeugten Linien konnten wir nachweisen, dass die Vermehrung von *ONSEN* einmalig während der Behandlung stattfindet und die neuen *ONSEN*-Kopien über Generationen sehr stabil im Genom integriert bleiben. Erwartungsgemäß konnten wir zudem einen starken Effekt dieser neuen *ONSEN*-Kopien

auf den Phänotyp nachweisen. Interessanterweise konnten wir beobachten, dass die Wachstumsbedignungen, ähnlich wie oben für die Blutorange beschrieben, einen starken Einfluss auf die Phänotypen der erzeugten Linien hatten. So zeigten manche Pflanzen abhängig von Temperatur und Tageslänge eine erhöhte Biomasse oder verfrühte Blühzeit im Vergleich zu den Kontrollpflanzen.

Durch die Amplifikation von *ONSEN* lässt sich eine enorme genetische und phänotypische Diversität erzeugen. Vorläufige Daten weisen zudem darauf hin, dass einige der generierten Linien mit mehr *ONSEN*-Kopien eine erhöhte Hitzetoleranz im Vergleich zum Wildtyp aufweisen.

Da sich Ergebnisse mit der Modellpflanze *Arabidopsis* nur beschränkt auf unsere Nutzpflanzen übertragen lassen, haben wir in einer Kollaboration mit Forschenden in Montpellier den Effekt der Behandlung auf Reis-Keimlinge untersucht. Auch in dieser fern verwandten, sehr wichtigen Nutzpflanze konnten wir mit Hilfe gezielter DNA-Sequenzierungen aller mobilen Elemente nach Behandlung mit den beiden Substanzen die Aktivierung eines retroTEs nachweisen. Diese Daten sprechen dafür, dass unsere Methode universell in praktisch allen Pflanzen anwendbar ist.

Abb. 1: Schematische Darstellung der stressabhängigen Aktivierung von Retrotransposons (retroTEs) für die Pflanzenzüchtung. A) Normales Wachstum der Pflanze unter optimalen Bedingungen. B) Absterben der Pflanze durch Stress. C) Vorübergehende Inhibierung der Transposonabwehr in Kombination mit Stress führt zur gezielten stabilen Vermehrung stress-responsiver retroTEs im Pflanzengenom. D) Erneutes Auftreten desselben Stresses führt, vermittelt durch neue retroTE-Insertionen, zur Aktivierung zusätzlicher Gene und dadurch erhöhten Stresstoleranz der Pflanze.



Retrotransposons als genetische Ressource

Unter der Annahme, dass Pflanzen für eine breite Palette von Stressarten eine entsprechendes TE als genetisches Backup bereit halten, welches wir dank unserer Behandlung jetzt gezielt aktivieren und im Falle von retroTEs vermehren können, stellen TEs und insbesondere retroTEs eine sehr vielversprechende genetische Ressource für die Pflanzenzüchtung dar.

In gewisser Weise gleicht die stressabhängige Vermehrung von retroTEs mit dem Ziel eine erhöhte Stresstoleranz für genau denselben Stress zu erreichen einer gerichteten Evolution nach der Theorie von Jean-Baptiste de Lamarck. Die Erinnerung an einen erlebten Stress wird demnach indirekt durch neue retroTE-Insertionen stabil an die Folgegeneration weiter gegeben.

Auch wenn dieser Ansatz noch in der Entwicklung ist, ergibt sich daraus ein erhebliches Potential für die Einsparung von Pflanzenschutzmitteln. Während beim klassischen Ansatz versucht wird, durch Pflanzenschutzmittel einen bestimmten Stress, zum Beispiel Pathogene zu bekämpfen, verwendet unsere Methode genau diesen Stress, um natürliche Prozesse in der Pflanze zu stimulieren, die in den Nachkommen zu einer gesteigerten und stabilen Stresstoleranz führen. Derartig verbesserte Nutzpflanzen mit neuen stress-responsiven Genen sind in geringerem Ausmass auf die grossflächige Anwendung von Pflanzenschutzmitteln angewiesen.

LITERATUR

MCCLINTOCK B. (1950): The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences of the United States of America*, 36, 344–355.

BUTELLI E., LICCIARDELLO C., ZHANG Y., LIU J., MACKAY S., BAILEY P., REFORGIATO-RECUPERO G., MARTIN C. (2012): Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell*, 24, 1242–1255.

ITO H., GAUBERT H., BUCHER E., MIROUZE M., VAILLANT I., PASZKOWSKI J. (2011): An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature*, 472, 115–119.

MIROUZE M. und VITTE C. (2014): Transposable elements, a treasure trove to decipher epigenetic variation: insights from Arabidopsis and crop epigenomes. *J Exp Bot*, 65: 2801–2812.

THIEME M., LANCIANO S., BALZERGUE S., DACCORD N., MIROUZE M., BUCHER E. (2017): Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. *Genome Biol*, 18: 134.

11.5 Appendix V: Multiple choice test (World-Café) correct answers marked in red (2 pages)

FiBL

Universität Basel

Fragebogen zum Thema:

"Stressinduzierte Mobilisierung von Retrotransposons (retroTEs) als natürliche Ressource für die Pflanzenzüchtung"

Frage 1/7: **RetroTEs bewegen sich durch:**

- [ ] *Cut & paste*
- [x] *Copy & paste*

Frage 2/7: **Viele retroTEs können durch diverse Stresse aktiviert werden. Dies kann folgend(e)n Effekt(e) haben:**

- [x] *Stressabhängige Aktivierung benachbarter Regionen im Genom;*
- [ ] *Ausschneiden von Kopien des retroTEs aus dem Genom;*
- [x] *Synthese neuer Kopien des retroTEs;*

Frage 3/7: **DNA-Methylierung ist ein wichtiger epigenetischer Mechanismus zu Regulation von retroTEs. Eine <u>erhöhte</u> Methylierung führt in der Regel zu einer Verstärkung der Aktivität von retroTEs:**

- [ ] *Richtig*
- [x] *Falsch*
- [ ] *Weiß ich nicht*

Frage 4/7: **Mit Hilfe der beiden Stoffe α-Amanitin (A) und Zebularin (Z) können retroTEs aktiviert werden. Zutreffende Aussagen bitte ankreuzen:**

- [ ] *A und Z sind beides Naturstoffe;*
- [ ] *A wird in die DNA eingbaut;*
- [x] *Z wird in die DNA eingebaut;*
- [x] *Gleichzeitige Behandlung mit A und Z und einem Stress kann zur Transposition von retroTEs führen;*
- [ ] *Um einen Effekt zu erzielen muss die Pflanze bis zur Samenreife mit A&Z behandelt werden;*
- [x] *A und Z sind prinzipiell auch im Menschen wirksam;*

Frage 5/7: **Retrotransposition führt zu einer Erhöhung der genetischen Diversität in Nutzpflanzen. Zutreffende Aussagen bitte ankreuzen:**

- [ ] *Aktivierbare retroTEs müssen nur einmal von außen als Fremd-DNA in die Zelle eingbracht werden;*
- [x] *Durch retroTE erzeugte Diversität spielte schon immer eine Rolle in der Züchtung;*
- [ ] *Bedingt durch ihre Größe beruht die durch retroTE-Insertionen erzeugte phänotypische Diversität lediglich auf knock-out Mutationen;*
- [ ] *RetroTEs inserieren ausschließlich in unmittelbarer Nähe ihres Ursprungslocus;*
- [x] *RetroTEs erzeugen „markierte Mutationen";*

Frage 6/7: **Genetische Stablilität spielt in der Züchtung eine entscheidende Rolle. Welche der folgenden Aussagen ist Aufgrund unserer Beobachtungen in *Arabidopsis* als zutreffend zu bewerten?**

- [x] *A&Z-behandelte Linien weisen über drei Generationen nur schwache Fluktuationen der retroTE-Kopienzahl auf, was auf heterozygote Insertionen zurückzuführen ist;*
- [ ] *Werden high-copy Linien erneut einem Stress ausgesetzt kommt es zu einer sehr starken Aktivierung von retroTEs (sogenannter TE-burst);*
- [ ] *retroTE-induzierte Phäntotypen lassen sich oft nur in einer Generation beobachten, da neue Insertionen oft instabil sind;*

Frage 7/7: **In Hinblick auf DNA-Methylierung und retroTE-Verteilung im Genom stellt *Arabidopsis* im Vergleich zu anderen Nutzpflanzen eine Ausnahme dar. Es ist daher nicht zu erwarten, dass die beiden Stoffe A&Z in anderen Nutzpflanzen ebenfalls wirksam sind.**

- [ ] *Richtig*
- [x] *Falsch*
- [ ] *Weiß ich nicht*

Vielen Dank für Ihre Teilnahme!

11.6 Appendix VI: Questionnaire (World-Café) (1 page)

FiBL

Universität Basel

Fragebogen zum Thema:

Akzeptanz des Einsatzes von retroTEs in der biologischen Landwirtschaft

Ich habe teilgenommen an:

☐ *Vortrag von Michael Thieme* ☐ *World Cafe zum Thema*

Mein Bezug zur Biolandwirtschaft ist: ___________________________________

____

Frage 1/5: Sollten Ihrer Meinung nach Sorten, die mit Retrotransposons gezüchtet wurden, in der biologischen Landwirtschaft verwendet werden dürfen?
(Antwort 1- 5; 1: gar nicht, 5: unbedingt) ____

Frage 2/5: Was sind Ihrer Meinung nach die potentiellen Risiken welche von der Nutzung von Retrotransposons ausgehen könnten?

____

____

____

Frage 3/5: Sollten Ihrer Meinung nach in der Züchtung für die biologische Landwirtschaft <u>natürliche Substanzen</u>, welche Retrotransposons aktivieren, verwendet werden dürfen? (Antwort 1- 5; 1: gar nicht, 5: unbedingt) ____

Frage 4/5: Was ist Ihrer Meinung nach das Potential dieser Züchtungsmethode für die biologische Landwirtschaft? (Antwort 1- 3; 1: keines , 3: hoch) ____

Frage 5/5: Verfahrenspatente sind die Grundlage für Innovation in der Landwirtschaft. (Antwort 1- 5; 1: auf keinen Fall , 5: absolut richtig) ____

Kommentare und Ergänzungen:

Diese Daten dienen dazu die Bandbreite der verschiedenen Meinungen abzubilden und werden nicht für Statistiken verwendet.

11.7 Appendix VII Questions for guiding the discussion of World-Café (1 page)

Guiding-Questions World Cafe-Diskussion

Chancen und Risiken von stressinduzierter Mobilisierung von Retrotransposons

- ➔ Was sind die grossen Herausfoderungen in der biologischen Züchtung?
- ➔ Wie könnte die retroTE-Methode zur Lösung beitragen?-> Erwähnen dass Genomsequenz nicht unbedingt bekannt sein muss.
- ➔ Wo sind die Grenzen der Methode
- ➔ Welche Risiken bestehen: zufällige Insertion -> unerwünschte Nebenwirkungen? Toxizität bei der Anwendung? Random SNPs durch Z.
- ➔ Auswirkung auf Lebensmittel: Auch in Menschen Wirksam-> Angst vor Transfer in Nahrung?
- ➔ Auskreuzung? Ausbreitung „egoistischer Gene“? Gain of function führt zur Evolution von «Superweeds?»

Akzeptanz von chemisch versus gentechnisch induzierter Mutationsauslösung

- ➔ Wie ist die Akzeptanz von chemisch ausgelösten Punktmutationen mittels EMS oder Polyploidisierung mittels Colchicin im Biolandbau?
- ➔ Vergleich synthetisch hergestelltes Colchicin/Amanitin versus Colchicin/Amanitin aus der Herbstzeitlose/Fliegenpliz?
- ➔ Ethische Aspekte-> Zu grosser Eingriff ins Genom?
- ➔ Eingriff erfolgt an ganzer Pflanze, in vitro Kultur oder Einzelzelle oder direkter Eingriff in die Zelle (Gentechnik)
- ➔ Gezielte Mutationsauslösung mittels CRISPR Cas9 (Gene-editing), weniger Nebenwirkungen, deswegen sogar besser? Auch in Hinblick auf random SNPs durch genotoxic Effekt von Z. Fliessende Übergänge von chemischer und gentechnischer Veränderung bei Infiltration von CRISPR-Cas9 Konstrukten→ Identifizierung neuer TEs mit Behandlung und Einbringen in andere Pflanze mittel CRISPR Cas9;
- ➔ Kann man genotoxische Effekte, die durch Zebuarin ausgelöst werden, in Kauf nehmen wenn man gleichzeitig Gain of function phenotypes erzeugen kann? Cisgene Pflanzen mit TEs aus anderen Sorten vertretbar?
- ➔ Allg. Verwendung von (eventuell)limitierten Naturstoffen im Bio-Sektor vertretbar?

Vor- und Nachteile von Patenten in der Pflanzenzüchtung

- ➔ Methodenpatent: Inhibierung der freien Wissenschaft oder Förderung von Innovationen?
- ➔ Auch TEs sind potentiell patentierbar -> Konflikt mit kein Patent auf Leben?
- ➔ Produktpatent: Die entstandenen Pflanzen werden ebenfalls leicht zu schützen sein..
- ➔ Welchen Einfluss hat das auf die Verfügbarkeit von Sorten für den Landwirt?
- ➔ Welchen Einfluss hat es für die Züchter?
- ➔ Welchen Einfluss hat es konkret für den Biolandbau?

11.8 Appendix VIII: Presentation given prior to the discussion in the World-Café at FiBL (pictures were removed) (5 pages)

Stressinduzierte Mobilisierung von Retrotransposons als natürliche Resource für die Pflanzenzüchtung

Michael Thieme

Zürich-Basel Plant Science Center
Botanisches Institut, Universität Basel

Bucher Lab

FiBL

FiBL, 7. September 2017

Twitter: @gene_hop

Entdeckung springender Gene: „Transposons“

1. Cut and paste: DNA-TEs

Barbara McClintock
Nobelpreis 1983

Das „Mobilom“

2. Copy and paste: retroTEs

2

Pflanzengenome: Sammlungen von TEs

Bis zu 85% des Pflanzengenoms ist TE-DNA!

-> strenge epigenetische Regulation!

Vitte et al., 2014

3

TEs: nur parasitäre, eigennützige DNA?

Transposable elements, a treasure trove to decipher epigenetic variation: insights from Arabidopsis and crop epigenomes

Controlled activation of retrotransposition for plant breeding
Jerzy Paszkowski

The necessary junk: new functions for transposable elements

4

retroTEs: Bindeglied von Genom und Umwelt

Ruby — Helle Orange

Ruby — Blutorage "Tarocco"

Kälteresponsives retroTE

Butelli et al., 2012

-> retroTEs können benachbarte Gene stress-responsiv machen!

5

Weindiversität durch retroTEs

Cabernet

Insertion

Chardonnay

LTR retroelement

LTR recombination

Reversion

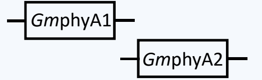
Ruby Okuyama

(Lisch, 2013)

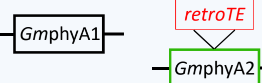
6

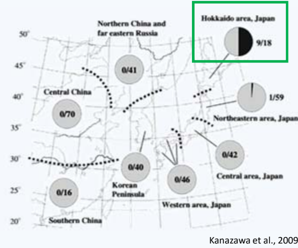
Adaptive Evolution durch retroTEs

Photoperiod sensitive *G.max*



Photoperiod insensitive *G.max*





Kanazawa et al., 2009

- Blüte im Langtag
- Schnellere Fruchtreife

Adaptive Evolution!

7

retroTEs für die Pflanzenzüchtung?

Markierte Mutationen

Gain of function phenotypes

Alternative splicing

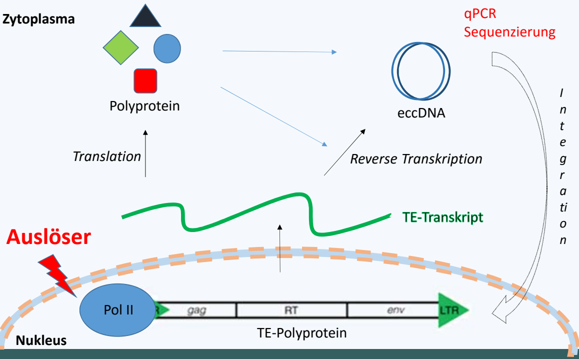
Transport von Genfragmenten

Erzeugung epigenetischer Diversität (Epiallele)

Neue regulatorische Netzwerke

8

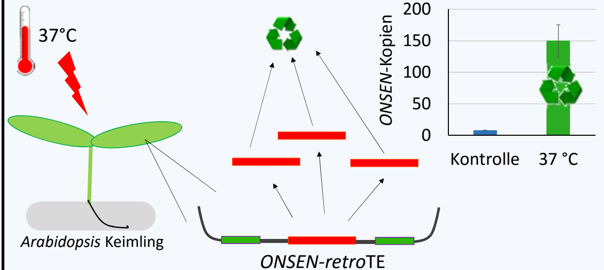
Der Lebenszyklus von retroTEs



9

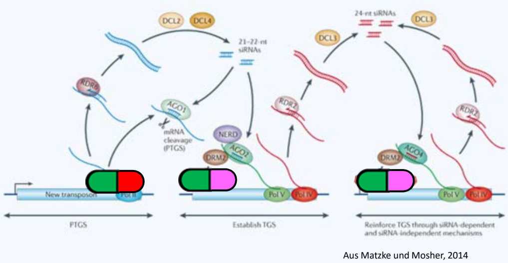
retroTEs können durch Stress aktiviert werden

Normalerweise werden neue Kopien wieder abgebaut und *ONSEN* stillgelegt..



10

TEs werden streng überwacht



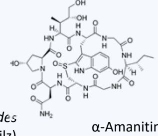
Aus Matzke und Mosher, 2014

-> Bisher konnten retroTEs nicht gezielt für die Züchtung nutzbar gemacht werden..

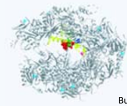
11

Inhibitoren von TE-silencing: α-Amanitin

Amanita phalloides (Knollenblätterpilz)



α-Amanitin



Bushnell *et al.*, 2001

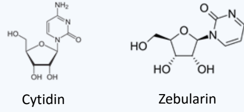
Inhibierung von RNA-Polymerase II

- Naturstoff (Peptid) gewonnen aus Knollenblätterpilz
- Inhibiert RNA-Polymerase II in Eukaryonten
- Hoch spezifisch und wirksam: 0,1 mg·kg−1 (LD50, Maus)

-> vgl. Colchicin 5,89 mg·kg^{-1}

12

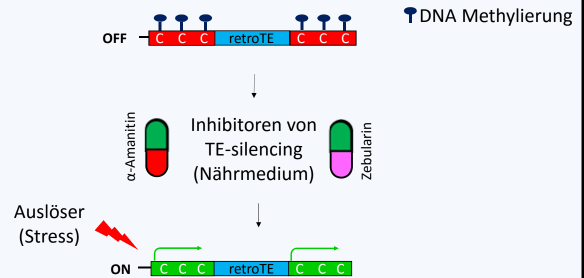
Inhibitoren von TE-silencing: Zebularin



- Nukleosidanalogon, in Krebstherapie eingesetzt
- Wird in DNA eingebaut
- Inhibiert DNA-Methyltransferasen in Eukaryonten
- Kann genotoxischen Effekt haben (Liu *et al.*, 2015)

13

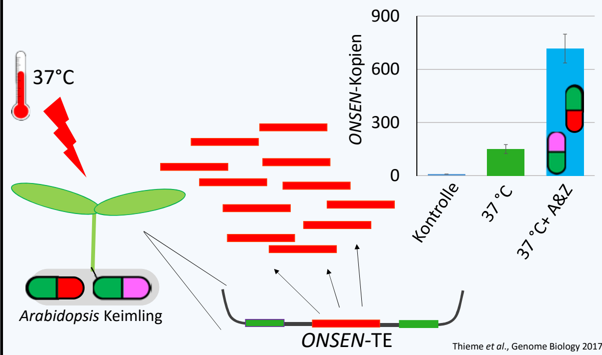
Induzierte retroTE-Aktivierung durch Inhibitoren



Thieme *et al.*, Genome Biology 2017

14

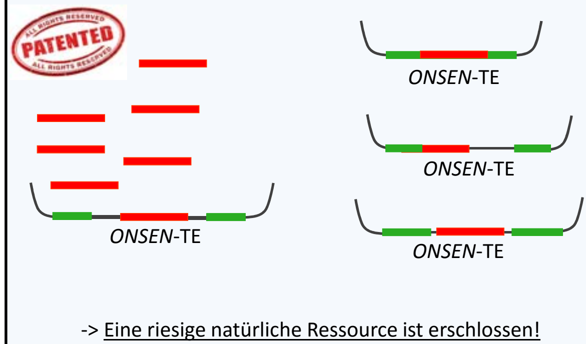
retroTE-Aktivierung kann verstärkt werden



Thieme *et al.*, Genome Biology 2017

15

Innovation: *ONSEN*-Kopien integrieren!

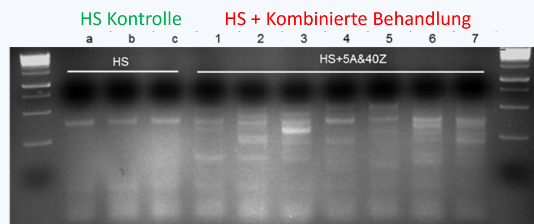


-> Eine riesige natürliche Ressource ist erschlossen!

16

Neue *ONSEN*-Insertionen im Wildtyp

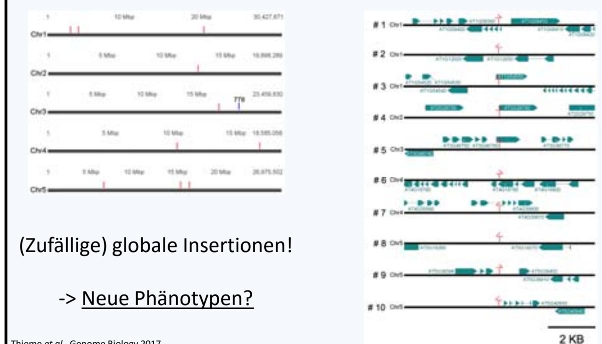
TE-display: S2 Generation nach HS



Thieme *et al.*, Genome Biology 2017

17

ONSEN integriert in der Nähe von Genen



(Zufällige) globale Insertionen!

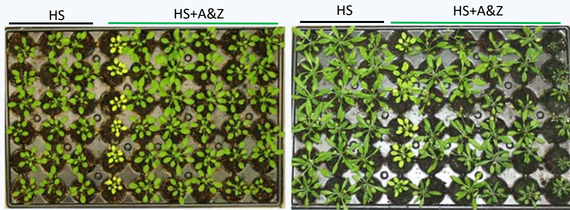
-> Neue Phänotypen?

Thieme *et al.*, Genome Biology 2017

18

ONSEN erzeugt Diversität I

S2 Generation von "high-copy" Linien



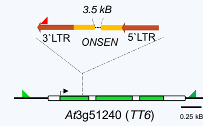
Kurztag   Langtag

-> Umweltabhängige Phänotypen!

Thieme *et al.*, Genome Biology 2017

19

ONSEN erzeugt Diversität II



Flavanone 3-hydroxylase
(key factor in flavonoid biosynthesis)

-> *ONSEN* verursacht "markierte" Mutationen!

Thieme *et al.*, Genome Biology 2017

20

Wie stabil sind neue Insertionen?

DNA-TE: **Cut** and paste

-> Instabile Phäntoypen

RetroTEs: **Copy** and paste

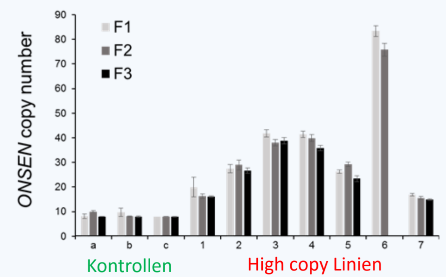
-> Unkontrollierbare Amplifikation von *ONSEN*?

21

Stabililtät der Linien: Kontrolle

Thieme *et al.*, Genome Biology 2017

qPCR on total DNA from individual plants in sucessive generations

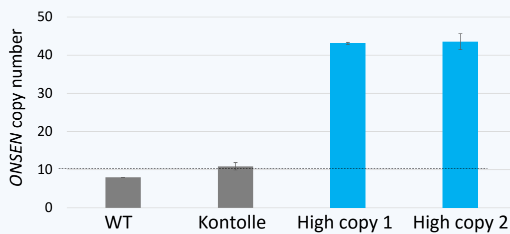


Kontrollen   High copy Linien

-> *ONSEN* stabil, Fluktuationen wegen heterozygoten Kopien

22

Stabililtät der Linien: Kontrolle



WT   Kontolle   High copy 1   High copy 2

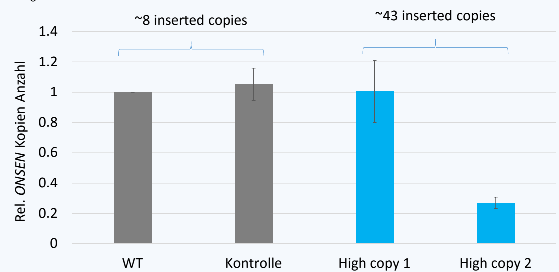
Hitzestress 24h 37°C (S3 Generation)

Thieme *et al.*, Genome Biology 2017

23

Stabililtät der Linien: Unter Stress

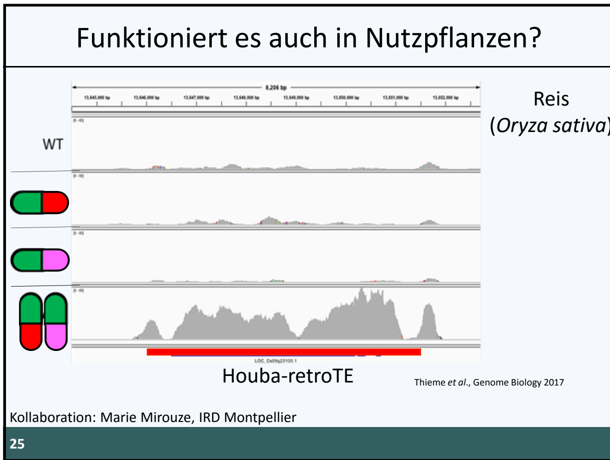
n=3 biolog. relicates



→ Keine unkontrollierte Amplifikation von *ONSEN* nach HS

Thieme *et al.*, Genome Biology 2017

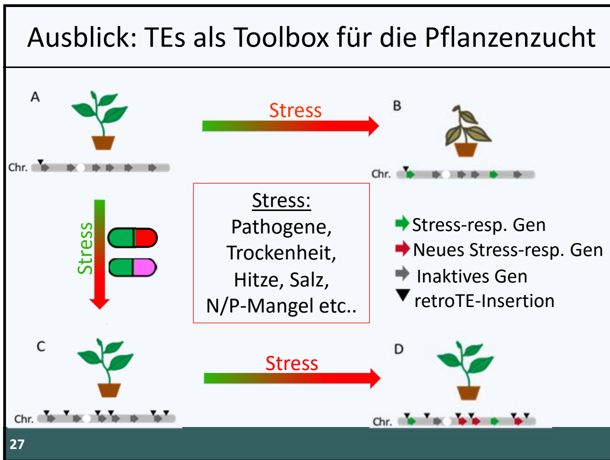
24



Gerichtete Evolution für den Pflanzenschutz

- Für (jeden) Stress gibt es ein entsprechendes retroTE
- retroTEs können durch diesen Stress aktiviert und vermehrt werden
- Neue retroTEs generieren regulatorische Netzwerke und (epi)genetische Diversität

26



Vielen Dank!

INRA SCIENCE & IMPACT

Etienne Bucher & Bucher Lab

FiBL

Monika Messmer

Universität Basel

Thomas Boller & Botanical Institute Basel

MARIE CURIE ACTIONS

IRD Institut de recherche pour le développement

Marie Mirouze
Sophie Lanciano

unitectra

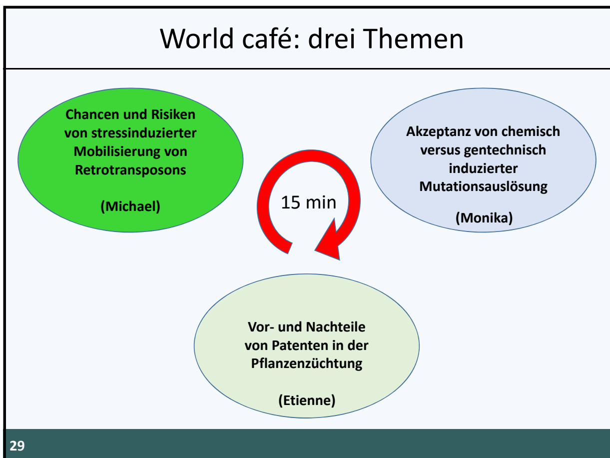
Peter Eckard

IDP BRIDGES

Zürich-Basel Plant Science Center

..für die Aufmerksamkeit!

28 Twitter: @gene_hop



11.9 Appendix IX Article published by "pflanzenforschung.de" "Hitze lässt *ONSEN* hüpfen-Retrotransposonen kontrolliert zum Springen bringen" (2 pages)

Pflanzenforschung.de

Hitze lässt ONSEN hüpfen

Retrotransposonen kontrolliert zum Springen bringen

15.08.2017 | von Redaktion Pflanzenforschung.de

Es ist noch nicht lange her, da waren Retrotransposonen „für die Tonne“. Wie alle transposablen Elemente wurden sie als DNA-Schrott oder Junk-DNA bezeichnet. Heute weiß man es besser. Die Möglichkeit, diese kontrolliert zur Transposition zu bewegen, könnte helfen, mehr über die virenähnlichen DNA-Abschnitte zu erfahren und sie in der Pflanzenzucht einzusetzen.

Der Grüne Knollenblätterpilz (*Amanita phalloides*) ist der Serienmörder unter den Pilzen. 90 % aller Pilzvergiftungen mit Todesfolge in Europa gehen auf sein Konto. Für viele von uns müssen daher schon gute Gründe vorliegen, um aus freien Stücken ausgerechnet mit dem Hauptgift des Wulstlings, ?-Amanitin, zu hantieren.

Was passiert bei einer Vergiftung?

Stellt man sich eine eukaryotische Zelle als Schweizer Uhrwerk vor, dann ist es so, als würde plötzlich ein wichtiges Zahnrad fehlen, was dazu führt, dass die Uhr stehen bleibt. Die verheerende Wirkung von ?-Amanitin ist vergleichbar. Es hindert durch Anlagerung an die RNA-Polymerase II diese an ihrer Aktivität, die Transkription. Der Prozess, bei dem die Information der DNA in RNA übersetzt wird. Ohne diese Übersetzungsleistung ist eine Zelle funktionsuntüchtig, da keine Informationen mehr vom Zellkern aus übertragen werden. Diese aber werden von den Ribosomen zur Produktion von Proteinen benötigt. Die Folge: der Stoffwechsel bricht zusammen, die Zelle kollabiert. In der Onkologie wird dieser Effekt bereits genutzt, um Tumorzellen zu bekämpfen.

Mykotoxin im Nährmedium

Nun wurde genau dieses Prinzip bei jungen Arabidopsiskeimlingen (*Arabidopsis thaliana*) angewandt. Man ließ sie auf Nährmedien wachsen, die neben den üblichen Nährstoffen ?-Amanitin enthielten. Der Grund ergibt sich im Zusammenhang mit den eingangs erwähnten Retrotransposonen. Diese zählen als eigene Klasse zu den Transposonen. Jenen genetischen Elementen, die ihre Position im Genom verändern können (Transposition).

Wie das Virus so das Retrotransposon

Das Besondere an Retrotransposonen ist, dass der Informationsfluss für die Transposition wie bei Retroviren durch eine reverse Transkription der RNA erfolgt. Die Folge ist, dass das Original an seiner ursprünglichen Stelle verbleibt und eine Kopie an anderer Stelle im Genom integriert wird – „*copy* and paste“ statt „*cut* and paste. Retrotransposonen sind somit auf fremde Hilfe angewiesen.

Seit einigen Jahren vermutet man, dass die RNA-Polymerase II die Rolle des Dienstleisters übernimmt. Sie ist zugleich aber auch darin involviert, die Aktivität bzw. Transposition zu unterdrücken. Letzteres ist deshalb wichtig, weil die Insertion eines Retrotransposons durchaus Folgen haben kann, wenn z. B. eine Kopie inmitten eines codierenden Gens landet. Um die genetische Integrität und Stabilität zu bewahren, haben Pflanzen daher Mechanismen entwickelt, um die Aktivität der sonderbaren DNA-Elemente einzuschränken. Die Hemmung der Transkription mithilfe unsere Polymerase ist das beste Beispiel.

ONSEN heißt auf Japanisch „heiße Quelle“

Dann ist noch zu wissen, dass viele Retrotransposonen vor allem unter Stress aktiv werden. Dies gilt auch für ONSEN, einem hitzeempfindlichen Retrotransposon, das bei Hitze bzw. Hitzestress aktiv wird. Dass dieses hier namentlich erwähnt wird, kann als Beweis gewertet werden, dass seine Aktivität erstens durch die

RNA-Polymerase II reguliert wird und zweitens von außen regulierbar ist. Schließlich stand das Retrotransposon mit dem fernöstlichen Namen im Zentrum der Machbarkeitsstudie (Proof of Concept).

Um es auf den Punkt zu bringen: Wird die Enzymaktivität der RNA-Polymerase II mithilfe des Pilzgifts ?-Amanitin unterbunden, wird ONSEN aktiv und breitet sich im Genom aus, sobald man den Thermoregler auf 37 °C hochdreht und die Pflanze unter Hitzestress gerät.

Von A wie ?-Amanitin bis Z wie Zebularin

Wen die Vorstellung, Nährmedien zu experimentellen Zwecken mit einem Mykotoxin zu versetzen, schlucken lässt, wird kaum erraten, womit man Nährmedien noch befüllen kann: Zebularin zum Beispiel. Ein außergewöhnlicher Wirkstoff, der in der Humanmedizin u. a. eingesetzt wird, um traumatische Ereignisse aus dem menschlichen Gedächtnis zu löschen. Erhalten Arabidopsispflanzen über das Nährmedium eine Extraportion Zebularin, dann hat dies einen ähnlichen Effekt wie das Pilzgift. Auch hier wird ONSEN bei Hitzestress aktiv und vervielfältigt sich. Kombiniert man ?-Amanitin mit Zebularin, erhöht sich sogar der Effekt. Jedoch steckt bei Zebularin ein anderer Mechanismus dahinter.

Auch die Epigenetik spielt mit

Die Wirkung basiert auf der Hemmung der Aktivität von Methyltransferasen. Enzyme, die Methylgruppen (-CH3) an DNA-Abschnitte heften. Dies hat zur Folge, dass die Transkription dieser Abschnitte gehemmt wird. Damit verfügen Pflanzen mit diesem epigenetischen Mechanismus über eine weitere Schutzmaßnahme, um Retrotransposonen wie ONSEN zu bändigen. Wird ihnen nun auch diese Fähigkeit genommen, läuft dies im Endeffekt auf das gleiche Ergebnis hinaus wie im Versuch mit ?-Amanitin. Wie es aussieht, sichern sich Pflanzen lieber doppelt ab, um die sprunghaften Retrotransposonen in die Schranken zu weisen.

Neue Möglichkeiten für die Pflanzenforschung

Mit der Entdeckung der externen Steuerungsmöglichkeiten der Retrotransposonenaktivität öffnen sich neue Möglichkeiten und Perspektiven, sich zukünftig mit dem Einfluss von Retrotransposonen genauer und systematisch zu beschäftigen. Auch steht der Weg prinzipiell offen für den Einsatz in der Pflanzenzucht. Nicht nur weil die zusätzlichen ONSEN-Kopien über mehrere Generationen stabil erhalten bleiben, sondern auch weil die Mechanismen zur Kontrolle der Retrotransposonenaktivität – mittels RNA-Polymerase II und Methylierung – bei Nachfolgegenerationen wieder greifen.

Mitnahmeeffekte nutzen

So könnte man sich z. B. einen weiteren Nebeneffekt zunutze machen, der bei Transpositionen auftritt. Es ist nicht ungewöhnlich, dass bei der Transkription eines Retrotransposons nicht nur das Retrotransposon selbst, sondern auch Gene in direkter Nachbarschaft abgelesen und aktiviert werden. Da Retrotransposonen vor allem bei Stress aktiv werden, könnte man sich diesen Mitnahmeeffekt zunutze machen, wenn sich in direkter Nachbarschaft Gene befinden, die die Stressresistenz erhöhen.

Quelle:
Thieme, M. et al. (2017): Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. In: Genome Biology, Vol. 18, (7. Juli 2017), doi.org/10.1186/s13059-017-1265-4.

Zum Weiterlesen auf Pflanzenforschung.de:

- Neues Apple-Produkt vorgestellt
- Ganz ohne Zellkultur
- Sprunghafte Entwicklungen

Titelbild: Bei 37 °C beginnt das Retrotransposon ONSEN im Genom von Arabidopsiskeimlingen, aktiv zu werden und sich zu vermehren. (Bildquelle: © CSIRO/ CC BY 3.0)