EFFECTOR OF TRANSCRIPTION (ET):

Novel plant specific epigenetic regulators of reproductive processes act on DNA-methylation

Dissertation

zur Erlangung des akademischen Grades

***doctor rerum naturalium* (Dr. rer. nat.)**

vorgelegt der Naturwissenschaftlichen Fakultät I

-Biowissenschaften-

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Bui Thi Mai Huong

geb. 25. Dezember 1981 in Thai Binh, Vietnam

Gutachter:

1. Prof. Dr. Thomas Altmann
2. Prof. Dr. Dröge-Laser
3. Prof. Dr. Gunter Reuter

Halle (Saale), 17. Dezember 2015

Contents

Contents

Contents

List of figures

List of figures

List of Tables

List of abbreviations

FM	Functional megaspore
PMC	Pollen mother cells
TAG	Triacylglycerols
ABA	Abscisic acid
GA	Gibberellic acid
AP	APETALA
SEP	SEPALLATA
siRNA	small interfering RNA
Col 0	Columbia-0
Ws	Wassilewskija-2
bHLH	Basic Helix Loop Helix
DAPI	4, 6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
HRT	HORDEUM REPRESSOR OF TRANSCRIPTION
mRNA	Messenger RNA
MS	Murashige and Skoog
RNA	Ribonucleic acid
UV	Ultraviolet
EDTA	Ethylenediamine tetraacetic acid
SDS	Sodium dodecyl sulfate

List of abbreviations

DMSO	Dimethyl sulfoxide
cDNA	Complementary DNA
PCR	Polymerase chain reaction
RT-PCR	Reverse transcription PCR
SDS	Sodium dodecyl sulfate
T-DNA	Transferred DNA
MYB	Myeloblastosis
MET1	Methyltransferease1
CMT3	Chromomethylase3
HDA6	Histone deacetylase
SUVH	SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE
RDR2	RNA-dependent-RNA polymerase 2
DCL3	Dicer-like protein
HEN1	HUA ENHANCER 1
AGO4	ARGONAUTE4 protein
POL IV	Polymerase IV
DRM2	DNA methyltransferase
ET	EFFECTOR OF TRANSCRIPTION

DML	DEMETER-LIKE
DME	DEMETER
ROS1	REPRESSOR OF SILENCING1
BER	Base excision repair
MEA	MEDEA
X-gal	5-bromo-4-chloro-3-indolyl-ß-D-galactoside
IPTG	Isopropyl-ß-D-thiogalactoside
EGTA	Ethyleneglycol tetraacetic acid
MALDI-TOF	Matrix assisted laser desorption ionization-time-of-flight

Legends

AtET1,2,3	indicates the gene (Capital, italic)
AtET1,2,3	indicates the protein (Capital)
et1,2	indicates the mutant allele (Lowercase, italic)
et1et2	indicates the double mutant allele (Lowercase, italic)

Introduction

Plant seeds have evolved to nourish, protect, and distribute the next generation embryo. They are of utmost importance for human nutrition and provide basics for a bio-based economy and energy production. Plant developmental biology provides important insights for a basic understanding of the underlying reproductive processes.

Plants have a complex life cycle in which diploid and haploid generations alternate between the diploid, spores-producing sporophyte and the haploid, gametes-producing gametophyte. In angiosperms the female gametophyte, the embryo sac, is strongly reduced and deeply embedded in sporophytic tissue. It originates from a diploid megaspore mother cell which undergoes meiosis. Of the resulting tetrad of haploid megaspores a single cell survives and develops into a seven-celled embryo sac. Within the embryo sac, the haploid egg cell and the diploid central cell are fertilized independently and give rise to a diploid embryo and triploid endosperm, respectively. This double fertilization marks the transition between gametophyte and sporophyte and is vital for seed formation (Grossniklaus and Schneitz, 1998; Yadegari and Drews, 2004).

Meiotic products in animals differentiate without further cell division into egg- and sperm cells. In strong contrast, plant meiotic products start to proliferate and form a multicellular organism, the gametophyte. Therefore, the plant life cycle is characterized by an alternating sequence of a spore-producing (sporophyte) and a gamete-producing (gametophyte) generation. Meiosis and fertilization represent the transition points between both generations, respectively. Plant phylogeny is characterized by a gradual reduction of the gametophytic generation with the female gametophyte of angiosperms, designated as embryo sac, consisting of only a few cells surrounded by sporophytic tissue (Hofmeister, 1851). Thus, the small gametophyte is often called the “hidden generation”.

Aspects of seed development, like embryo and endosperm development as well as seed maturation processes (synthesis of storage compound, acquisition of desiccation tolerance and dormancy) depend on complex

metabolic and hormonal pathways, transcription factor controlled networks, cell cycle factors connecting growth and differentiation as well as epigenetic pathways (e. g. Grossniklaus et al., 2001; Weber et al, 2005; Berger et al., 2006; Holdsworth et al., 2008; Sabelli and Larkins, 2009; Berger and Chaudhury, 2009; Junker et al., 2010). Epigenetic regulation mechanisms such as DNA-methylation, histone modifications, small interfering and long non-coding RNAs have been recognised to be increasingly important for gene regulation (e. g. Schmitz et al, 2013; Weigel and Colot, 2012; Heard and Martienssen, 2014) and the former distinction between various groups of "classical" transcription factors such as MADS, MYB, MYC etc. (e.g.Riechmann et al., 2000; Latchman, 2008; AGRIS database, http://arabidopsis.med.ohio.state.edu/AtTFDB) and epigenetic gene regulators are fading.

Although crop plants are the final target of applied plant biology, choosing an appropriate model plant system is very important to efficiently perform basic research. Plant biologists prefer *Arabidopsis thaliana*, because of its small size, short life cycle, prodigious seed production, simple and established transformation methods, availability of the whole, small genomic sequence and large collections of precisely described mutants. The astonishing molecular similarity of basic developmental processes between Arabidopsis and crop plants such as wheat, barley, rice and corn establish the small cruciferous plant as a highly suitable model plant.

Plant reproductive processes can be divided into major developmental pathways such as female and male sporogenesis, female and male gametogenesis, double fertilisation, embryogenesis, endosperm formation, seed maturation including synthesis of storage compounds, acquisition of desiccation tolerance and dormancy as well as germination (Fig. 1).

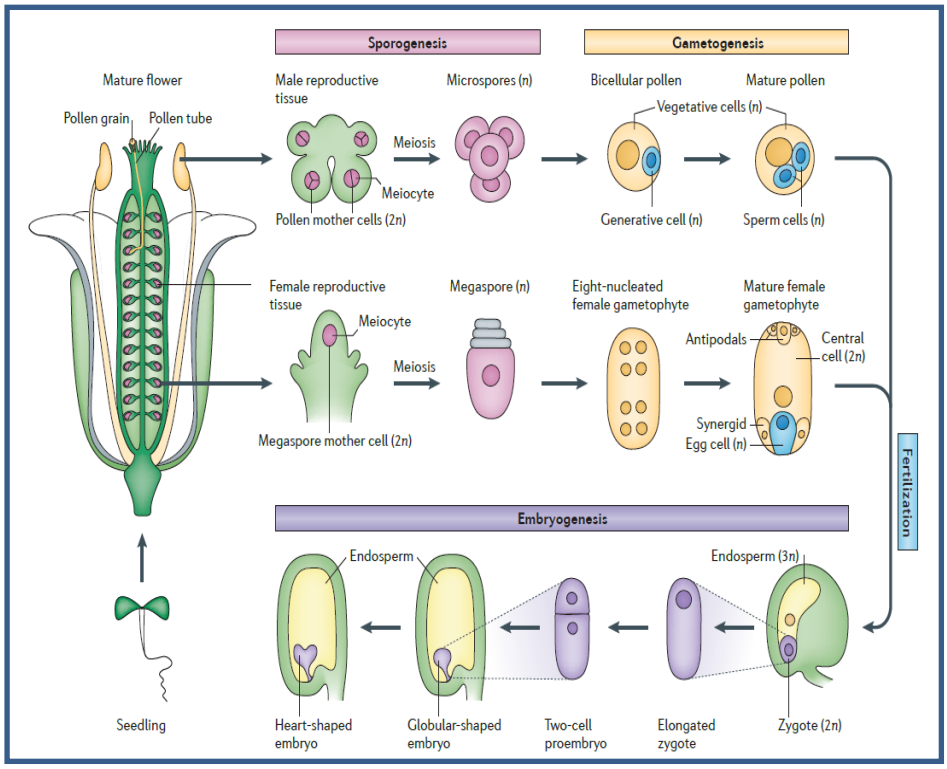

Fig.1. Scheme of sexual plant reproduction (Kawashima and Berger, 2014). Sporogenesis is initiated both in pollen mother cells (PMC) and megaspore mother cell (MMC) both undergoing meiosis. Whereas all four meiotic products survive in the male pathway, only the functional megaspore (FM) survives in the female pathway. The plant specific process of gametophyte development results in the formation of male gametes, the sperm cells and female gametes, the egg cell and the homodiploid central cell. In the plant specific double fertilization process two sperm cells fuse with the egg cell and central cell to produce the embryo proper and the endosperm, respectively.

1. Female gametophyte of Arabidopsis

The female gametophyte, also called embryo sac, develops in the ovules on the carpel´s ovary. The female gametophyte of Arabidopsis (Fig. 2) belongs to the most common monosporic polygonum type shared by more than 70% of flowering plants (Maheshwari, 1950; Wilemse and van Went, 1984; Haig, 1990; Huang and Russell, 1992; Yadegari and Drews, 2004). A special subepidermal cell, the megaspore mother cell, undergoes

meiosis. Three of the meiotic products perish by apoptosis with the most chalazal spore surviving as functional megaspore (FM). The FM undergoes three haploid mitosis generating an eight-nucleate embryo sac. Two polar nuclei fuse to the homodiploid nucleus of the central cell. Two synergids form and the egg cell form the egg apparatus at the micropylar end and three antipodal become positioned at the chalazal end and degenerate in the mature embryo sac which then ready for fertilisation.

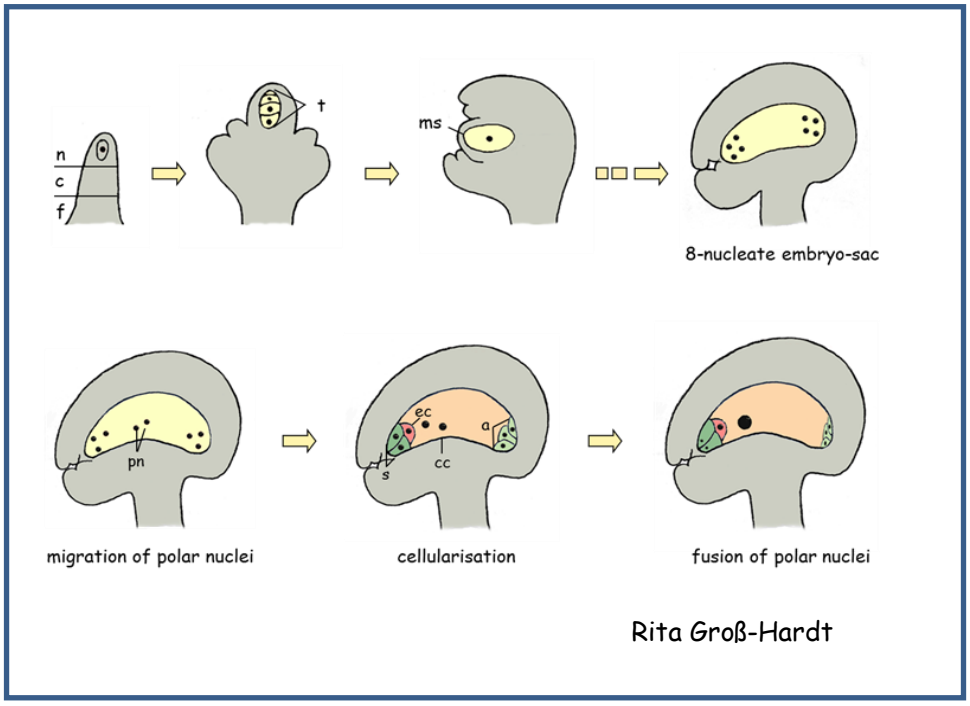

Fig. 2 Sporogenesis and gametophyte development of *Arabidopsis thaliana*

2. Male gametophyte of Arabidopsis

In plants, pollen is produced within the anthers of the flower. The reproductive cells give rise to the microspores whereas the non-reproductive cells form discrete anther tissues layers and include the epidermal, cortical and tapetal cell layers surrounding the sporogenous cells. During microsporogenesis the pollen mother cell (PMC) undergoes meiosis resulting in four surviving haploid microspores. These unicellular microspores undergo a first pollen mitosis to form two unequal cells, a large

vegetative cell and a small generative cell. The generative cell divides once more by a second mitosis to form the two sperm cells (Fig. 3).

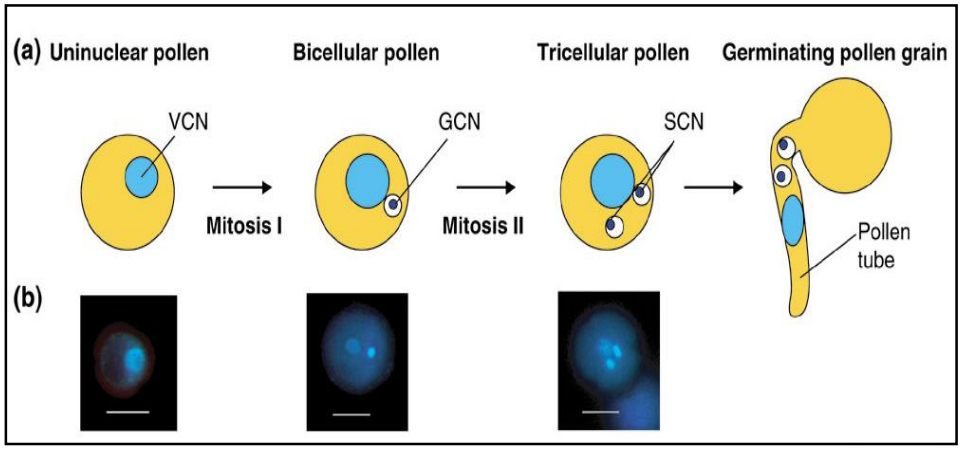

Fig. 3 Male gametophyte development of Arabidopsis (Park at al., 1998, Honys and Twell, 2004). The pollen mother cell undergoes meiosis to form four microspores which subsequently undergo two haploid mitotic steps leading to the mature microgametophyte which consists of a vegetative cell and two sperm cells. a) schematic drawing, b) fluorescence microscopy.

3. Embryogenesis

Embryogenesis of Arabidopsis is initiated by double fertilization. The vegetative cell of the male gametophyte triggers the growth of the pollen tube, interacts with the synergids of the female gametophyte and delivers the two sperm cells. One sperm cell fuses with the egg cell to form the zygote as initial for the development of the embryo proper. The second sperm cell fuses with the second homodiploid central cell to form the triploid endosperm, a storage tissue for embryo nutrition (Brown et al., 1999).

After zygote formation higher plant embryogenesis can be conceptually divided into two distinct phases (Fig. 4). The early phase is characterized by cell proliferation and morphogenesis with the basic body plan of shoot-root polarity being established (West and Harada, 1993; Goldberg et al., 1994; Laux and Jurgens, 1997). A later phase of maturation is characterized by storage compound synthesis mainly in the cotyledons,

desiccation tolerance and dormancy (West and Harada, 1993; Goldberg et al., 1994; Lotan et al., 1998; Harada, 2001; Raz et al., 2001).

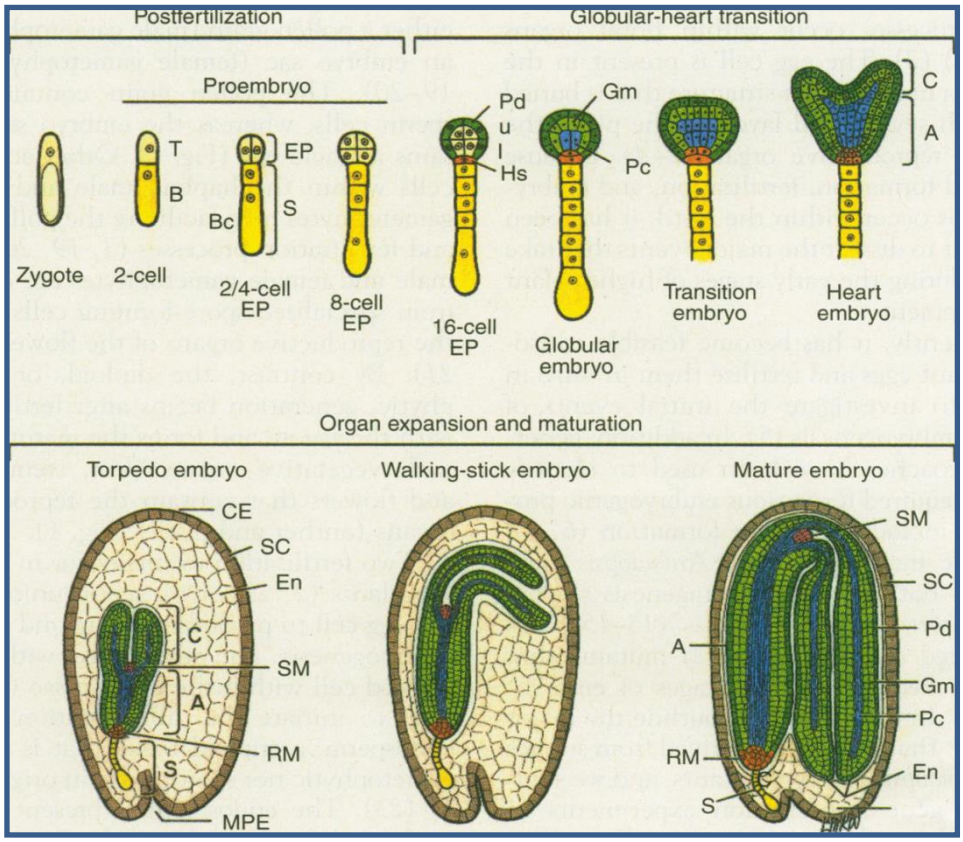

Fig. 4 Scheme of Arabidopsis embryogenesis (Goldberg et al., 1994). A generalized overview of plant embryogenesis. Characteristic stages of embryo development comprise the globular-, heart-, torpedo- and walking stick-stages with shoot- (SAM) and root- (RAM) apical meristems indicated. Abbreviations: T, terminal cell; B, basal cell; EP, embryo proper; S, suspensor; Bc, suspensor basal cell; Pd, protoderm; u, upper tier; I, lower tier; Hs, hypophysis; Pc, procambium; Gm, ground meristem; C, cotyledon; A, axis; MPE, micropylar end; CE, chalazal end; SC, seed coat; En, endosperm; SM, shoot meristem; and RM, root meristem.

The storage products of Arabidopsis comprise lipids, proteins and carbohydrates. Seed lipids are stored as triacylglycerols (TAG) in oil bodies (Murphy, 1993; Herman, 1995). TAGs synthesis occurs from the late heart stage and continues through the torpedo stage until the embryo desiccates.

The TAG core of an oil body is surrounded by a phospholipid monolayer and oleosins, which are special proteins involved in the preservation of the oleosome structure during seed desiccation (Huang et al., 1994; Mansfield and Briarty, 1992). They are associated with lipases which degrade the lipids and provide the main energy source of the growing seedling. Seed storage proteins are the primary source of carbon and nitrogen for the growing seedling. In developing Arabidopsis seeds, there are two types of seed storage proteins, the 12S globulins (cruciferins) and the 2S albumins (napins). They are synthesized at the rough endoplasmic reticulum and sorted into the protein storage vacuoles (Müntz, 1998). Starch is found in the plastids of embryo cells and seed coat cells (Focks and Benning, 1998). Later starch is mainly detected in the outer and inner cell layers of the outer integument, but not in the mature embryo (Western et al., 2000; Kim et al., 2005).

4. Germination

Seed maturation and germination is mainly regulated by the ratio of the two phytohormones abscisic acid (ABA) and gibberellic acid (GA) (Koornneef et al., 1998; White and Rivin, 2000). ABA concentration increases during late embryogenesis, reaches a peak in the maturation phase and decreases in mature seeds. In this phase ABA prevents germination. Therefore, mutants with affected ABA synthesis or response germinate precociously, fail to express maturation specific messengers and are intolerant to desiccation (Black, 1991). Gibberellins (GA) play an important role in the regulation of cell division and expansion as well as in seed germination (Olszewski et al., 2002). External GA application causes premature seed germination (Debaujon and Koornneef, 2000; White and Rivin, 2000).

5. Flowering control

Floral meristems initiate the formation of four different floral organs: sepals, petals, stamens and carpels (Coen and Carpenter, 1993) starting out from concentric rings, the whorls, around the flanks of the meristem. In

Arabidopsis the four whorls are arranged as follows: the first outermost whorl includes 4 green sepals; the second whorl consists of 4 petals with white color at maturity; the third whorl is composed of six stamens, two of which are shorter than the other four and the fourth whorl generates the gynoecium or pistil, which consists of an ovary with two fused carpels, each containing numerous ovules and a short style capped with a stigma (Fig 5).

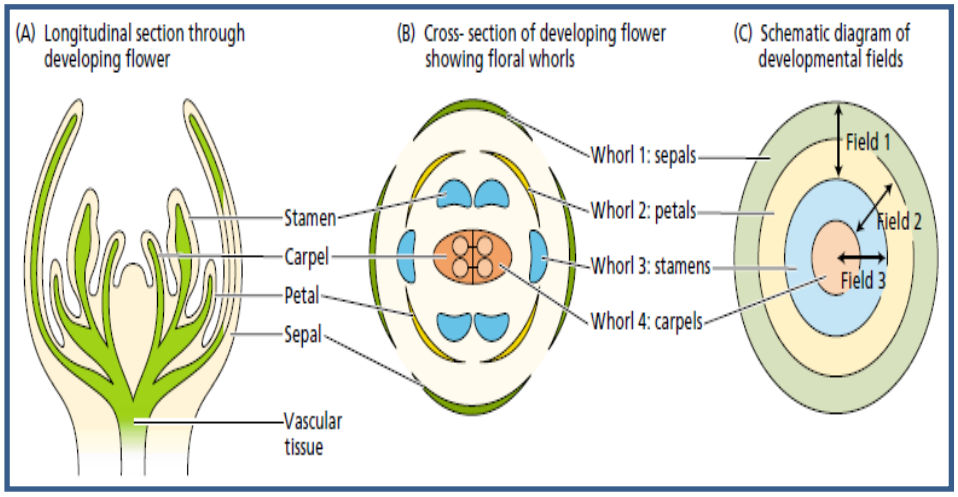

Fig. 5 Scheme of the Arabidopsis flower (Bewley et al. 2000). The floral organs are produced as successive whorls (concentric circles), starting with the sepals and progressing inward. According to the combinatorial ABC model, the functions of each whorl are determined by overlapping developmental fields. These fields correspond to the expression patterns of specific floral organ identity genes

Homoeotic mutants with changed organ identity define several transcription factors of the MADS box class. At least five MADS box genes are known to specify floral organ identity in Arabidopsis: APETALA1 (AP1), APETALA2 (AP2), APETALA3 (AP3), PISTILATA (PI) and AGAMOUS (AG) (Bowman et al., 1989; Weigel and Meyerowitz, 1994) which are summarized as A (AP1, AP2), B (AP3, PI) and C (AG) function with A expressed in the first and second whorl, B becomes active in the second and third whorl and C is activated in the third and the fourth whorl. The type E activity is encoded by

SEPALLATA (SEP),1, 2, 3 and 4 (Pelaz et al., 2000) and might be required for the combinatorial function of A, B and C (Soltis et al., 2007).

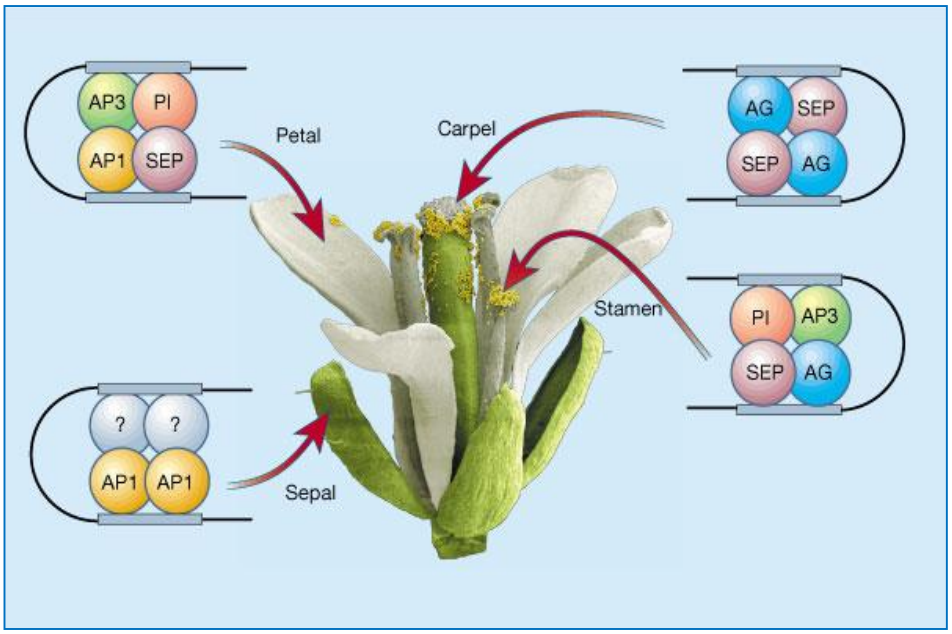

Fig. 6 Model of genetic control of floral organ identity based on the quartet model (Theissen and Saedler, 2001).

6. DNA methylation and epigenetics

Epigenetics refers to processes causing dynamic alterations in the transcriptional potential of a cell which are not caused by changes in the DNA sequence including for instance DNA methylation and histone modifications for instance leading to gene silencing. In principle, there are two different mechanisms of gene silencing known: the RNA-directed DNA methylation pathway and an RNA interference pathway for transcriptional gene silencing (He et al., 2011). DNA-methylation occurs by the addition of a methyl group to the cytosine bases of DNA to form 5-methylcytosine (He et al., 2011). Typically, DNA methylation is removed during zygote formation and re-established through successive cell divisions during development (Jaenisch and Bird, 2003). In animals, methylation occurs exclusively in the symmetric context CG and CHG, whereas in plants, methylation occures also in the assymmetic CHH context, where H stands for A, T or C (Lister et al., 2009). In Arabidopsis about 7% of the whole

genome is methylated- among this 24% CG, 6.7% CHG and 1.7% CHH (Cokus et al., 2008). DNA-methylation is considered to act as a protective mechanism to prevent the activation of retrotransposons but is also required for differential transcription regulation during differentiation and development (Köhler et al., 2012).

Three pathways for DNA methylation regulation have been described: a) *de novo* methylation, b) maintenance methylation and c) de-methylation (Fig.7).

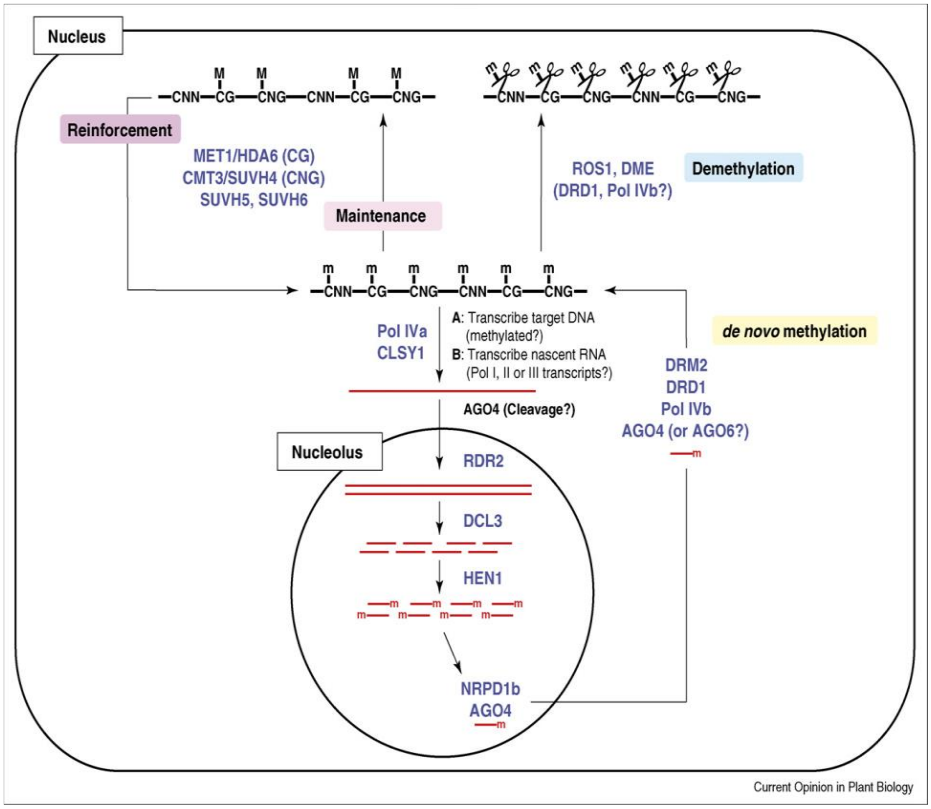

Fig. 7 A schematic model for the dynamic regulation of methylation (Matzke and Mosher, 2007):

Middle: The *de novo* methylation pathway: Pol IVa together with protein CLSY1 transcribes the target locus, which might already be lightly methylated (´m´) or associated with specific histone modifications (A). Alternatively, Pol IVa might transcribe a nascent RNA produced at the target locus by Pol I, II or III (B).

Top, left: The maintenance pathway: CG and CNG methylation can be maintained during DNA replication by MET1 and CMT3, respectively. Locus-specific histone modifications that are catalyzed by HDA6, SUVH4, SUVH5 and/or SUVH6 help to maintain cytosine methylation (´M´) and reinforce the silent state.
Top, right: The demethylation pathway: DNA methylation can be lost in nondividing cells by a base excision repair-type mechanism that involves DNA glycosylase/lyase proteins such as ROS1 and DME (Choi et al., 2002; Gehring et al., 2009).

6.1. De novo methylation

RNA-directed DNA methylation (RdDM) was first discovered in 1994 in viroid-infected tobacco plants (Viswanatha and Tian-Kang, 2009). A single-stranded transcript of polymerase IV (POL IV) is transferred from nucleus to the nucleolus by an unspecified mechanism, where it is copied into double stranded RNA by RNA-dependent-RNA polymerase 2 (RDR2). The double-stranded RNA is cleaved into 24-nt primary siRNAs by a Dicer-like protein DCL3 (Matzke and Mosher, 2007). The siRNAs are methylated at their ends by HUA ENHANCER 1 (HEN1) and then the siRNA is loaded on a RISC complex (RNAi-induced silencing complex). This complex contains the ARGONAUTE4 protein (AGO4), which interacts with the C-terminal domain of NRPD1b (largest subunit of POLIV). This complex moves out of the nucleolus into the nucleoplasm, where NRPD2a subunit is added to form functional POL IVb complex. In addition, nascent non-coding RNA transcripts produced by POL V have been suggested to serve as scaffolds for recruiting the AGO4-containing RdDM effector complex by base-pairing with guide siRNAs (Wierzbicki et al., 2008). The functional RdDM effector complex directs the *de novo* DNA methyltransferase DRM2 (Domains Rearranged Methyltransferase 2) to specific chromatin regions to catalyze new DNA methylation (Matzke and Mosher, 2007; He et al., 2011).

6.2. Maintenance methylation

In Arabidopsis about one-third of genes have CG methylation in their coding region, which can be maintained by methyltransferease 1 (MET1) Matzke et al., 2007). CHG methylation can be maintained by chromomethylase3 (CTM3) or SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 4 (SUVH4 also known as KYP) and SUVH5 and SUVH6 (Matzke and Mosher, 2007). Finally, the CHH methylation is maintained by CMT3 and DRM2 (Law and Jacobsen, 2010).

6.3. DNA demethylation

Active demethylation occurs in plants by DNA glycosylase activity, probably in combination with the base excision repair (BER) pathway. DNA glycosylases include DEMETER (DME) and REPRESSOR OF SILENCING1 (ROS1) as well as DEMETER-LIKE2 (DML2) and DEMETER-LIKE3 (DML3) (Choi et al., 2002; Gehring et al., 2009).

7. The EFFECTOR OF TRANSCRIPTION (ET) gene family

EFFECTOR OF TRANSCRIPTION (ET) genes were originally discovered by using South Western screens aiming to the isolating of transcription factors important for embryonic gene regulation (Ellerström et al., 2005; Ivanov et al., 2008). They represent a strictly plant specific class of gene regulators of barley, broad bean, rape seed and Arabidopsis (Raventós et al., 1998; Ellerström et al., 2005; Ivanov et al., 2008) designated as HORDEUM REPRESSOR OF TRANSCRIPTION (HRT) in monocots and EFFECTOR OF TRANSCRIPTION (ET) in dicots. ET proteins share variable numbers of highly conserved cysteine-histidine containing, zinc- and DNA binding repeats also found in lower plants such as the moss *Physcomitrella patens* demonstrating their evolutionary conservation (Ellerström et al., 2005; Ivanov et al., 2008).

Previous work has demonstrated that the nuclear barley factor HORDEUM REPRESSOR OF TRANSCRIPTION (HRT) binds to Gibberellin (GA)-

response elements of an endosperm expressed α-amylase gene promoter. Transient expression indicates that HRT acts as a repressor of the promoter activity (Raventós et al., 1998). Furthermore, ET factors the homologs or HRT in dicots, have been shown to be involved in GA response modulation. Thus, ectopic expression causes dwarf growth, late flowering, reduced germination, increased anthocyanin accumulation and reduced lignification as marker of terminal cell differentiation. Transient expression demonstrates its putative function as repressor of transcription of GA-controlled genes (Ellerström et al., 2005). Moreover, the AtET2 is involved in the control of various KNAT genes, required to maintain the undifferentiated state of cambium meristematic cells. Depending on the differentiation status ET factors have been located in the nucleus or the cytoplasm suggesting a shuttle process between both cellular compartments (Ivanov et al., 2008).

In Arabidopsis the ET family consists of three genes: *AtET1* (AT4G26170); *AtET2* (AT5G56780); *AtET3* (AT5G56770). *AtET1* is located on the fourth chromosome, while the other genes, *AtET2* and *AtET3* are located on the fifth chromosome (Ellerström et al., 2005; Ivanov, 2005; Ivanov et al., 2008). The *AtET1* and *AtET2* genes are intact coding sequence, whereas *AtET3* is a truncated version of *AtET2* due to the lack of the ET repeat domain.

Besides these DNA binding ET-repeats, ET factors share a characteristic DNA single strand cutting domain (GIY-YIG) with structural similarity to that of bacterial UVRC proteins and so called homing nucleases (Derbyshire et al., 1997; Aravind et al., 1999; Verhoeven et al., 2000; Stoddard, 2005). The bacterial UVRC protein is essential for DNA excision repair (Friedberg et al., 1995; Moolenaar et al., 1998). It is targeted to UV-induced DNA lesions like thymidine-dimers and introduces two single strand cuts 8 bp 5` and 4 bp 3` of the lesion. The two single strand cuts are processed by two structurally and functionally distinct domains. A C-terminal domain consisting of an Endonuclease V (EndoV) and Helix-hairpin-Helix (HhH) domain is required for the 5`cut, whereas the N-terminal GIY-YIG domain

inserts the 3`nick (Lin and Sancar, 1992; Derbyshire et al., 1997; Aravind et al., 1999; Kowalski et al., 1999; Verhoeven et al., 2000; Van Roey et al., 2002). The sequence similarity between plant ET factors and UVRC is only restricted to the single strand cutting GIY-YIG domain. This suggests that an ancestral bacterial GIY-YIG domain has been recruited by ET proteins and attached to the DNA-binding ET repeats to create a novel plant specific regulatory protein (Fig. 8). In all known ET genes the GIY-YIG domain is encoded by the separate second exon, consistent with a corresponding domain shuffling event during protein evolution. The functionality of the ET-derived single strand cutting domain was demonstrated by substituting the AtET2 GIY-YIG domain for the corresponding domain of the *E. coli* UVRC protein (Ivanov et al., 2008). This showed that the plant domain productively interacts with the C-terminal UVRC EndoV/HhH domain. Since the single amino acid exchange (R>A) of a highly conserved arginine residue in the known active centre of the nuclease domain results in the near complete loss of this activity, the single strand cutting activity of the GIY-YIG domain is most likely required for its molecular function. A conceivable hypothesis is that the nicking activity of the plant ET factor GIY-YIG domain may be involved in the catalysis of changes in higher order DNA structures, such as, for example, nucleosome sliding or the relaxation of supercoiled chromatin domains as a prerequisite for regulated gene expression (Choi et al., 2002; Xiao et al., 2003; Haince et al., 2006; Ju et al., 2006). Alternatively, the domain could be involved in active de-methylation processes as described for the plant regulators DEMETER (DME) and REPRESSOR OF SILENCING1 (ROS1) (Choi et al., 2002; Gong et al., 2002; Xiao et al., 2003; Choi et al., 2004; Morales-Ruiz et al., 2006).

Remarkably, the HhH domain as the separate second nicking domain of the UVRC protein, can be considered as the ancestor protein domain for both of these regulatory proteins. DME can introduce single strand nicks in the MEDEA (MEA) promoter as part of a DNA de-methylation pathway involved in the epigenetic imprinting of the MEA gene. ROS1 is described as protein that represses homology-dependent transcriptional silencing by

de-methylating the target promoter DNA (Gong et al., 2002). Thus, a possible evolutionary scenario is that plant-specific ET factors have recruited a single GIY-YIG domain from prokaryotic repair-related proteins by a domain shuffling process, joining this domain to the DNA-binding ET repeat. The resulting plant specific protein is no longer involved in repair processes but acts as a gene regulator (Fig. 8). The regulatory mechanism -in part analogous to the function of DME and ROS1- might include the insertion of nicks, with an impact on higher order structures of chromatin packed DNA or on the genomic DNA methylation pattern required for differentiation processes for instance during seed development.

A principally similar evolutionary process combining an ancient endonuclease domain with a DNA-binding domain has been described for the transcription factors family AP2/ERF (Magnani et al., 2004).

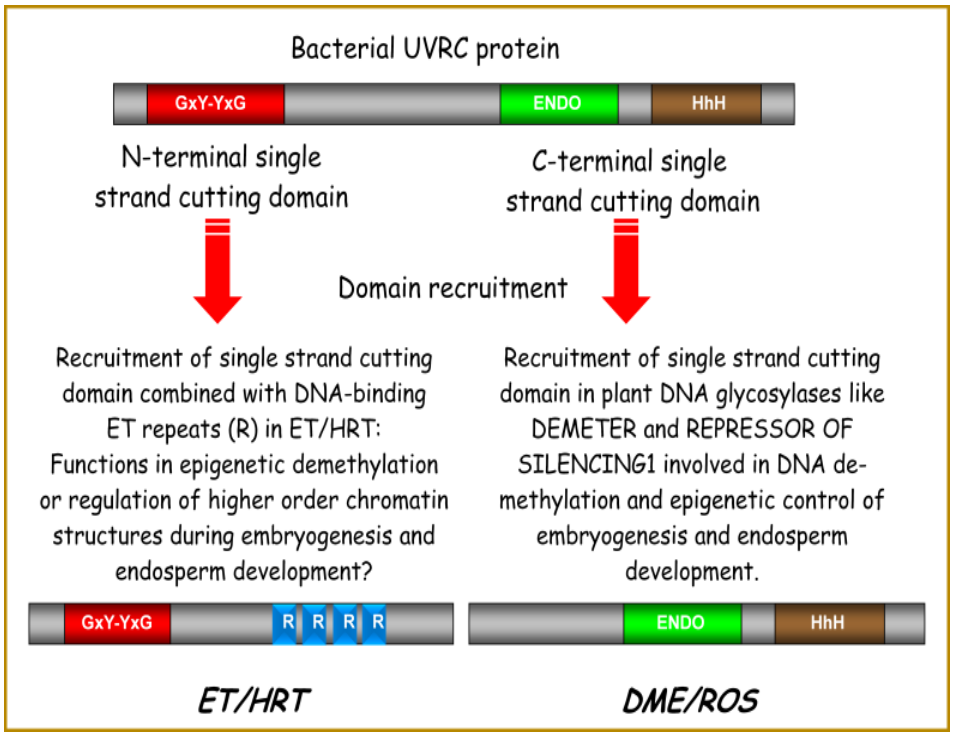

Fig. 8 Putative protein evolution scenario including the recruitment of two different DNA-single strand cutting domains of bacterial UVRC proteins by the plant regulatory proteins HRT/ET and DME/ROS. The HRT/ET factors

have adopted the N-terminal GxY-YxG domain whereas DME and ROS exploit the C-terminal single strand cutting domain of UVRC.

8. Aim of the dissertation

Since ET/HRT factors are most likely involved in the control of epigenetic DNA methylation processes, the current dissertation aims at the further functional characterisation of the ET regulator family and includes:

- The Isolation and functional characterization of mutant alleles of the ET gene family.
- The phenotypic description of the mutants focusing on gametophyte and seed development.
- The identification of putative target genes by transcriptome analysis using deep sequencing technology.
- The identification of differentially methylated genomic target regions using a whole genome methylation approach.

Materials and methods

1. Materials

1.1 Plant materials

Arabidopsis thaliana ecotypes Columbia-0 (Col) and *Wassilewskija-2* (Ws) were obtained from Gene Regulation Group (IPK, Gatersleben, Germany) and used throughout this study as wild type control experiments. T-DNA insertion lines have been received from Nottingham Arabidopsis stock center. From the genetic and molecular analysis of several SALK lines the following stable mutant lines have been obtained: *et1-1; et1-5; et2-3; et3-2; et3-3*. The line *et2-1* has been isolated from the Arabidopsis Knock-out Facility (AKF) at the University of Wisconsin Biotechnology center.

1.2 Bacterial strains

Several bacterial strains were used for different purposes such as DNA cloning, plasmid DNA amplification, sequencing etc..

Bacterial strains	Genotype/phenotype and reference
Escherichia coli XL1-Blue:	*rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd*R17, *supE44* lac [F *pro*AB, *lac*I^{q}ZΔM15, Tn10(tet^{R})], *rel*A1; (Stratagene, La Jolla, CA).
Escherichia coli DH5α:	F^{-}, ϕ80d/*lac*ZΔM15, *rec*A1, *end*A1, *gyr*A96, t*hi-1*, *hsd*R17(rK-, mK+), *sup*E44, *rel*A1, *deoR*, Δ(*lac*ZYA*arg*F) U169; (Grant et al, 1990)
Agrobacterium tumefaciens:	GV2260 (Deblare et al., 1985)

1.3 Enzymes, markers, antibiotics

Enzymes

EcoRI, T4 DNA ligase, *pfu* DNA polymerase, Dream Taq DNA polymerase, Shrimp Alkaline Phosphatase (SAP), Platinum Taq polymerase, RNase

inhibitor, Reverse transcriptase, 50X advantage ®2 DNA polymerase mix (Fermentas, Vilnius, Lithuania); DNase I, RNase I (Roche, Germany),

Markers

DNA Smart Ladder (Eurogentec, Seraing, Belgium); GeneRuler™ 1kb DNA Ladder Plus, PageRuler™ Prestained Protein Ladder (Fermentas, Vilnius, Lithuania).

Antibiotics

Ampicillin, Kanamycin, Rifampicin, Spectinomycin (Duchefa, Netherlands).

Other chemicals

X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside) and IPTG (Roche, Germany); Murashige-Skoog (MS) medium basal salt mixture including vitamins and microelements (Duchefa, The Netherlands); sucrose, glucose, malachite green, fuchsin acid, orange G, chloral hydrate, glacial acetic acid, nonidet P-40, DMSO, PIPES, EGTA, DAPI, formalin, sodium chloride, magnesium sulphate heptahydrate, potassium chloride, tris-base, yeast extract, glycerol, glycine (Carl Roth, Germany).

1.4 Commercial kits

GeneJET plasmid miniprep kit, GeneJET gel extraction kit, RevertAid first strand cDNA synthesis kit, DNA labelling kit (Fermentas, Vilnius, Lithuania); RNeasy kit, DNeasy plant mini kit, Epitect bisulfite kit, QIAquick PCR purification kit, QIAquick gel extraction kit, Qiagen plasmid purification kit mini (Qiagen, Hilden, Germany); TA cloning® kit, Zero Blunt® TOPO Cloning kit, BD SMART RACE cDNA Amplification kit (Takara, Japan).

1.5 Vectors

Various vectors were used for DNA amplification, cloning genes into plants and other purposes.

Vector	*Features*	*Reference or source*
pCR®II	Ampicillin^r, Kanamycin^r	Invitrogen, Carlsbad, CA
pCR®4Blunt-TOPO	Ampicillin^r, Kanamycin^r	Invitrogen, Carlsbad, CA
pDONR™/Zeo	Kanamycin^r, Zeocin^r	Invitrogen, Carlsbad, CA
pGV2260	Rifampicin, Spectrinomycin	Invitrogen, Carlsbad, CA

1.6 Primers and oligonucleotides

Primers for PCR and sequencing

Primer name	*Sequence 5'-3'*	*Tm (°C)*
ET1-1HuF	AAG AGA GAC GAC TAC ATT CGA ACT AAT C	68
ET1-1HuR	AGT ACC ATC TTC TAG TAA GAC TCC ACA AG	66
ET1-5F	CAT CGC CTA TCA AGT ATC AGC TTC CC	68
ET1-RACE1	AGG AGT AGT CCG CAA AAG TCT TGC GA	68
ET1-RACE2	GGG TTT ACG CAG AAA CAT AGA TCG GGC	72
ET2-3HuF	AAT ACC CGA TGA ACA GAT TTA CAT ATT	63
ET2-3HuR	GAG GTA AGT TCT GGA CTC TGT ATC TAC C	69
XR2	TGG GAA AAC CTG GCG TTA CCC AAC TTA AT	69

ET3F	GGA ATG AGA ATC ACC TAA CCT CTG C	66
ET3R	CTA CAC ATT GTC CGA CAT ATA CAC C	64
LBa1	TGG TTC ACG TAG TGG GCC ATC G	66
XR2-LB	CAT TTT ATA ATA ACG CTG CGG ACA TCT AC	66
LBb1	GCG TGG ACC GCT TGC TGC AAC T	68
Rba3	CGG CTT GTC CCG CGT CAT C	64
8409 – LB	ATA TTG ACC ATC ATA CTC ATT GC	57

Primers for RT - PCR

Primer name	***Sequence 5'-3'***	***Tm (°C)***
ET1–RACE 1	AGG AGT AGT CCG CAA AAG TCT TGC GA	68
ET1–RACE 2	GGG TTT ACG CAG AAA CAT AGA TCG GGC	72
ET2–RT–ACF	ATG GAA TTC GGC GAC GGC GTT TCC TTC G	73
ET2–RT–ACR	CTC GGA CTT TGG CGG TGT CTG TTT TTC G	72
AP3 F	CTA ACA CCA CAA CGA AGG AGA TC	63
AP3 R	GAA GGT AAT GAT GTC AGA GGC AG	63
ACT2-F	TCG GTG GTT CCA TTC TTG CT	57
ACT2-R	GCT TTT TAA GCC TTT GAT CTT GAG AG	55

Primers for Gateway®

Primer name	***Sequence 5'-3'***	***Tm (°C)***
PHP41893F	GGC GAC GAC ATG AAG ATG ACC T	64

PHP41893R	GGC CGC TTC AGA AGG GCA C	63
ET1-CompF	CTG AGA GAG GCG ATA GAG AGA CAC	67
ET1-CompR	GGA AAC TGA AGC TAA CAC TCC TCC	65

All oligonucleotides and primers were obtained from Metabion (Martinsried, Germany), MWG Biotech Company (Ebersberg, Germany) or Invitrogen (Carlsbad, CA).

1.7 Solutions and buffers

10 x TAE buffer

Tris-base	242 g
H_3BO_3	57,1 ml
EDTA	100 ml
Distilled water	up to 1000 ml

Extraction buffer for plant genomic DNA

Tris-HCl pH 7.5	0.20 M
NaCl	0.25 M
EDTA pH 8.0	25 mM
SDS	1%

Alexander staining solution

Ethanol 95%	10 ml
Malachite green solution (1% in 95% ethanol)	1 ml
Fuchsin acid (1% in water)	5 ml
Orange G (1% in water)	0.5 ml
Phenol	5 g
Chloral hydrate	5 g
Glacial acetic acid	2 ml
Glycerol	25 ml

Distilled water	50 ml

DAPI staining solution

Nonidet P-40	0.01%
DMSO	10%
PIPES	50 mM
EGTA	5 mM
DAPI	1 mg/ml

Clearing solution

Chloral hydrate	40 g
Water	10 ml
Glycerol	10 ml
Formalin	5 ml

Luria-Bertani-Medium (LB)

Trypton	10 g
Yeast extract	5 g
Sodium chloride	5,8 g
Mg sulphate heptahydrate	2,46 g
Agar	15 g
Distilled water	up to 1 l

SOC- Medium

Trypton	1 g
Yeast extract	0,5 g
5M NaCl	200 µl
1M KCl	250 µl
Distilled water	up to 1 l

Rich medium for Arabidopsis

MS salt mixture	4,3 g
Sucrose	10 g
Vitamin solution	10 ml
Agar (0,8%; for plates)	8 g
Distilled water	up to 1 l, pH 5,8

2. Methods

2.1 Extraction of genomic DNA

Genomic DNA extraction from plants was performed according to Edwards et al., 1991. Leaf tissue (~100 mg) was ground in liquid nitrogen into fine powder and suspended in 400 µl of extraction buffer. The suspension was centrifuged for 10 min at 13.000 rpm in a microcentrifuge. The supernatant was collected into a new tube and the DNA was precipitated with an equal volume of isopropanol. DNA was collected by centrifugation for 10 minutes, washed in 70% ethanol, dried and resuspended in 50 µl TE buffer. DNA concentration was determined by Nanodrop® ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA).

2.2 Screening and verifying T-DNA insertion mutants

T-DNA insertion lines *et1-1; et1-5; et2-3; et3-3; et3-2* in Columbia (Col) background were isolated from the Salk Institute collection of T-DNA lines transformed with pROK2 (http://signal.salk.edu/cgi-bin/tdnaexpress). The T-DNA specific primer LBa1 (O'Malley *et al.*, 2007) was used in combination with either forward or reversed gene specific primers. The line *et2-1* was isolated from the collection of the *Arabidopsis* Knock-out Facility (AKF) at the University of Wisconsin Biotechnology center, following a pool screening for insertion in *AtET2* gene in the *Wassilewskija* (Ws) background. The population lines were transformed with the T-DNA vector pD991-AP3 (Krysan *et al.*, 1999). The presence of T-DNA was verified by

PCR using T-DNA right border XR2 primer (Zhao *et al.*, 2002; Ivanov *et al.*, 2008) in combination with a gene specific primer.

The primer combinations were as follows:

Wild type *ET1*: ET1-1HUF/ET1-1HUR	T-DNA *et1-1*: ET1-1HuF/LBa1
Wild type *ET1*: ET1-5F/ET1-RACE1	T-DNA *et1-5*: ET1-RACE1/LBa1
Wild type *ET2*: ET2_RT_ACF/ET2_RT_ACR	T-DNA *et2-1*: ET2_RT_ACR/XR2
Wild type *ET2*: ET2-3HUF/ET2-3HUR	T-DNA *et2-3*: ET2-3HUR/LBa1
Wild type *ET3*: ET3F/ET3R	T-DNA *et3-2*: ET3R/8409-LB
Wild type *ET3*: ET3F/ET3R	T-DNA *et3-3*: ET3R/8409-LB

PCR conditions:

Initial denaturation:	95 °C, 5 min
	----------
	95 °C, 30 s
35 cycles	60 °C, 30 s
	72 °C, 40 s

Final extension:	72 °C, 5 min

2.3 Cloning methods and sequencing

Basic molecular methods such as enzymatic digestion, DNA ligation, DNA gel electrophoreses were performed according to standard protocols (Sambrook and Russell, 2001). DNA fragments were isolated and purified from agarose gel by QIAquick gel extraction kit (Qiagen, Hilden, Germany) and GeneJET gel extraction kit (Fermentas, Vilnius, Lithuania). DNA sequences were determined at the Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK, Gatersleben, Germany) or commercially by MWG Biotech Company (Ebersberg, Germany). Plasmid extractions and

purifications were done using Qiagen Plasmid kit and Fermentas GeneJET plasmid miniprep kit according to the protocol recommended by the manufactures.

2.4 Bacterial transformation

Transformations of *E. coli* and *A. tumefaciens* were carried out by using the heat shock procedure (Sambrook and Russell, 2001) and freeze-thaw method (Weigel and Glazebrook, 2002), respectively.

2.5 RNA extraction

Total RNA was isolated from 100 mg ground plant material using RNeasy Plant mini kit as described in the manufacture's protocol (Qiagen, Hilden, Germany). RNA was diluted in 30 µl DEPC-treated water and digested with RNase-free DNaseI (Roche, Mannhein, Germany) to exclude genomic DNA contamination. Total RNA concentration was quantified at a Nanodrop® ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA) at a wavelength of 260 nm.

2.6 cDNA synthesis and RT-PCR

First strand cDNA was synthesized by reverse transcription from total RNA using Revert Aid H Minus First strand cDNA synthesis kit (Fermentas, Vilnius, Lithuania). 1 µg of total RNA and 1 µl oligo (dT) primer were added to each tube to obtain a total volume of 11 µl. Priming was carried out at 70 ºC for 5 minutes, then 1 µl of ribonuclease inhibitor (20 units/µl), 2 µl of 10 mM dNTP mix and 4 µl of 5X RT buffer were added to each reaction tube. The reaction mixture was incubated at 37 ºC for 5 minutes and 1 µl of Reverse transcriptase (200units/µl) was added. The reaction mix was incubated for 1 h at 42 ºC for an hour, heated for 10 minutes at 70 °C and stored at -20 °C for further uses.

2.7 Reverse-transcription PCR

RT-PCR reaction to measure transcript amounts was performed using the primers ET1-1HUF/ET1-1HUR for ET1 transcript, ET2_RT_ACF/ET_RT_ACF for ET2 transcript and AP3F/AP3R for APETALA3 (AP3) transcript.

RT - PCR conditions: AP3

Initial denaturation:	95 °C, 5 min
	---
	95 °C, 30 s
25 cycles	60 °C, 30 s
	72 °C, 1 min

Final extension:	72 °C, 7 min

RT-PCR conditions: ET1, ET2

Initial denaturation:	95 °C, 5 min
	---
	95 °C, 30 s
35 cycles	60 °C, 30 s
	72 °C, 1 min

Final extension:	72 °C, 5 min

2.8 Seed germination and premature seed germination.

Seeds were collected from desiccated siliques and kept for one month in a dark and dry place. Seed were surface sterilized and spread on petri dishes with MS-agar. Germination rate was determined after 1 day. Premature seeds were collected from green siliques and grown on plates containing Murashige and Scoog medium (MS, Duchefa). Germination rates were determined for up to 12 days.

2.9 Generation of transgenic lines with central cell-specific marker

attB PCR conditions

Initial denaturation:	94 °C, 2 min
	----------
	94 °C, 30 s
35 cycles	61 °C, 30 s
	72 °C, 1 min

Final extension:	72 °C, 5 min

The fragments were cut and purified by Qiaquick kit and used for BP reaction.

The BP recombination reaction was performed as follows: 3µl attB-PCR product, 1 µl donor vector (pDONR/Zeo), 2 µl BP clonase II enzyme, 4 µl TE buffer, pH 8. The reaction was kept at room temperature overnight and then transformed into DH5α. Plasmid DNA was purified by Qiaquick kit and and resequenced. The LR recombination reaction was performed to transfer the gene of interest into an attR-containing destination vector to create an attB-containing expression clone. LR reaction conditions: 1.5 µl entry clone, 1.5 µl destination vector (pBGW), 4 µl 5X LR clonase reaction buffer, 13 µl TE buffer pH 8; incubation at 25 ºC for 1 h, addition of 2 µl of 2 µg/µl proteinase K, incubation at 37 ºC for 10 minutes and transformation of *E. coli*. Selected plasmid clones were purified and resequenced. Finally, the destination vector was transformed into the *Agrobacterium* strain GV2260. The culture was grown overnight at 28ºC for 2 days in YEB medium containing rifampicin, spectinomycin and carbenicillin. A stock culture was kept with glycerin 60% in -80ºC before transformation into Arabidopsis.

2.10 Methylation studies

Ten days old Arabidopsis seedlings were harvested and immediately frozen in liquid nitrogen. Total DNA was isolated from 100 mg ground plant material using DNeasy Plant Mini kit (Quiagen). DNA concentration was quantified at a Nanodrop ND-1000 spectrophotometer (NanoDrop

technologies Inc., USA). Total RNA was extracted from rosette leaves by using the Qiagen Plant RNeasy kit (Qiagen GmbH, Hilden, Germany). About 1 µg genomic DNA was split to 300 bp average size with a Covaris S2 instrument using the following settings for 120 s in frequency sweeping mode: intensity 5, duty cycle 10%, 200 cycles per burst. Then the DNA was purified by Qiaquick PCR purification columns. Libraries were generated by using the NEBNext DNA Sample Prep Reagent Set 1 (New England Biolabs) according to the Illumina Genomic Sample Prep Guide. After size selection, the non-methylated cytosine residues were converted to uracil by using the EpiTect Plus DNA Bisulfite kit (Qiagen) according to the manufacturer's protocol. Sequencing was done by an Illumina GAII instrument. Processing of genomic reads was performed by using the SHORE pipeline v.0.9.0 to trim and quality filter the reads (Ossowski, S. et al., 2008). The high quality sequences were aligned to Col-0 reference genome with Genome Mapper that supports the alignment of bisulphite converted reads (Schneeberger et al., 2009). The data processing was performed as described before (Becker et al., 2011). The determination of the differentially methylated genomic regions has been performed by Dr. C. Becker MPI Tübingen.

2.11 Histological methods

Pollen staining by Alexander

Inflorescences were collected from adult plants and fixed for 1-3 hours at 4 °C in acetic acid:ethanol (1:3). Anthers of mature flowers were isolated, transferred to a slide with a drop of Alexander solution (Alexander, 1969). Stained pollens were visualized under a Zeiss Axioplan 2 microscope to check pollen viability.

Pollen staining by DAPI

Analysis of mature pollen with DAPI was performed as previously described (Park at al., 1998). 5-10 flowers were incubated in 200 µl DAPI solution

overnight at 4 °C and examined by UV epi-illumination using a Zeiss Axioplan 2 microscope (Zeiss, Jena, Germany).

Clearing

Various plant tissues were collected and fixed in acetic acid:ethanol (1:3) at 4 °C overnight, dehydrated in an ethanol series (90%; 80%; 70%; 30 min per step) and stored overnight in 70% ethanol at 4 ºC. Ethanol was replaced with clearing solution. After 3 day at 4 °C tissue was observed using a Zeiss Axioplan 2 microscope (Zeiss, Jena, Germany).

2.12 Raster electron microscopy (REM)

Preparation and analysis of samples using REM were performed in cooperation with Dr. T. Rutten (Structural Cell Biology, IPK). Isolated flowers fixed overnight with 4% formaldehyde in 50 mM phosphate buffer pH 7,0. After washing with buffer and dehydration in an ethanol series, samples were critical point dried in a Bal-Tec critical point dryer (Bal-Tec AG, Balzers, Switzerland). Dried specimens were attached onto aluminium sample blocks and gold coated in an Edwards S150B sputter coater (Edwards High Vacuum Inc., Crowley, West Sussex, UK). Spikes were examined in a Hitachi S4100 SEM (Hisco Europe, Ratingen, Germany) at 5 kV acceleration voltage. Digital recordings were made and saved as Tif-files.

2.13 Confocal laser-scanning microscopy

Flowers of plants were emasculated, and whole-mount preparations of ovules were analyzed by microscopy 48h after emasculation. CFP fluorescence signal was studied with a Zeiss LSM 510 META or LSM780 confocal laser-scanning microscope (Zeiss, Jena, Germany). Fluorophore was detected with a 458 nm laser line in combination with a 480-520 nm band-pass (CFP). Identity of fluorophores was confirmed by photo spectrometric analysis with the help of the META-detector. This work has been performed in collaboration with Dr. T. Rutten, IPK Gatersleben.

2.14 Complementation assay

An ET1 genomic fragment including 1 kb upstream and 500 bp downstream sequence was PCR amplified using platinum Taq polymerase and resequenced. The gene fragment cloned into pDONR/Zeo using the BP reaction and further transferred into the pBGW destination vector using the LR reaction. Finally, the gene was transformed into *Agrobacterium tumefaciens* strain pGV2260 and used for Arabidopsis transformation (Col wild type and *et1-1* mutant). The presence of the transgene in *et1-1* plants was verified by PCR.

2. 15 RNA deep sequencing

cDNA libraries for Next Generation Sequencing were created following a slightly modified TruSeq RNA v2 protocol (Illumina). Starting from 0.4 – 4 µg total RNA with a RIN factor ≥8 (Agilent) in 50 µl DEPC treated water, polyA+ RNA was isolated via affinity chromatography on oligo-(dT) magnetic beads and fragmented at elevated temperature (94°C, 8 min) using divalent cations. First strand cDNA was synthesized (25°C, 10 min; 42°C, 50 min; 70°C, 15 min) using random hexamer primers and Superscript II reverse transcriptase (Invitrogen), followed by second strand synthesis (16°C, 60 min) and purification on magnetic AMPure XP beads (Beckman; PEG precipitation on bead surface, 2x EtOH wash, elution in 10 mM Tris-HCl, pH 8.5). After blunting of cDNA fragments (30°C, 30 min), purification on AMPure XP beads, A-tailing (37°C, 30 min), and ligation of Y-shaped adapters containing the respective index sequence as well as the P5 and P7 sequences for hybridization to the inner surfaces of an Illumina flowcell, the libraries were purified on AMPure XP beads and amplified using the P5 and P7 sequences as primers (98°C, 30 sec; 15x [98°C, 10 sec; 60°C, 30 sec; 72°C, 30 sec]; 72°C, 5 min). QiaQuick (Qiagen) purified libraries were applied to a 2% agarose gel stained with SYBR-Gold (Life Technologies; illumination with a Dark Reader [Clare Chemical Research]). After electrophoresis regions between 300 and 400 bp were cut from the gel and cDNA libraries were purified via MinElute spin columns (Qiagen). The average fragment length of cDNA libraries were determined on an

Agilent Bioanalyzer and their concentrations were calculated from qPCR reactions with cDNA libraries of known concentrations (known cluster densities on Illumina flowcells) as references. Libraries were denatured and diluted as recommended by Illumina, applied to a flowcell and sequenced. This work has been performed in collaboration with Dr. L. Altschmied, IPK Gatersleben.

2. 16 Plant transformation

Arabidopsis plants were transformed using the *Agrobacterium*-mediated floral dip method according to the protocol of Clough and Bent (1998). *A. tumefaciens* pGV2260 strain carrying the gene of interest was cultured in LB medium supplemented by 50mg/l kanamycin at 30^{o}C overnight. Bacterial cells were harvested by centrifugation and resuspended in the infiltration medium to obtain an OD_{600} of 0.7 to 0.9. To increase transformation efficiency, Silwet L-77 (Lehle Seeds, USA) was added to the bacterial suspension to a final concentration of 0.005%. Inflorescences were submerged into the *A. tumefaciens* suspension in a beaker for 5 seconds. Plants were placed on their side and covered with plastic wrap for 24 hour to maintain high humidity and could be set upright after a day. Seeds were harvested from dry siliques, sterilized and germinated onto selection medium. The insertion of transgene was control by PCR.

Results

1. Characterisation of *et* mutants

1.1 Mutants in *AtET1*

The ET gene family of the *Arabidopsis* genome comprises three genes, *AtET1* (At4g26170), *AtET2* (At5g56780) and *AtET3* (At5g56770). *AtET1* is located on the fourth chromosome, the two others on the fifth chromosome. *AtET2* is an intact gene whereas *AtET3* lacks the typical C-terminal ET repeat and is preliminarily considered as a non-functional pseudogene.

T-DNA insertion lines in *AtET1* were obtained from the SIGNAL T-DNA collection (http://signal.salk.edu/cgi-bin/tdnaexpress) and designated as *et1-1* (SALK_000422) and *et1-5* (SALK_006710).

About 50 plants for each original mutant line were genotyped in the first generation to verify the insertions and determine whether the line was heterozygous or homozygous. Genomic DNA was isolated and used for PCR with gene specific primers in combination with the T-DNA-specific primer LBa1. The sizes of PCR products were determined to be 655 bp and 1047 bp for *et1-1* and *et1-5,* respectively.

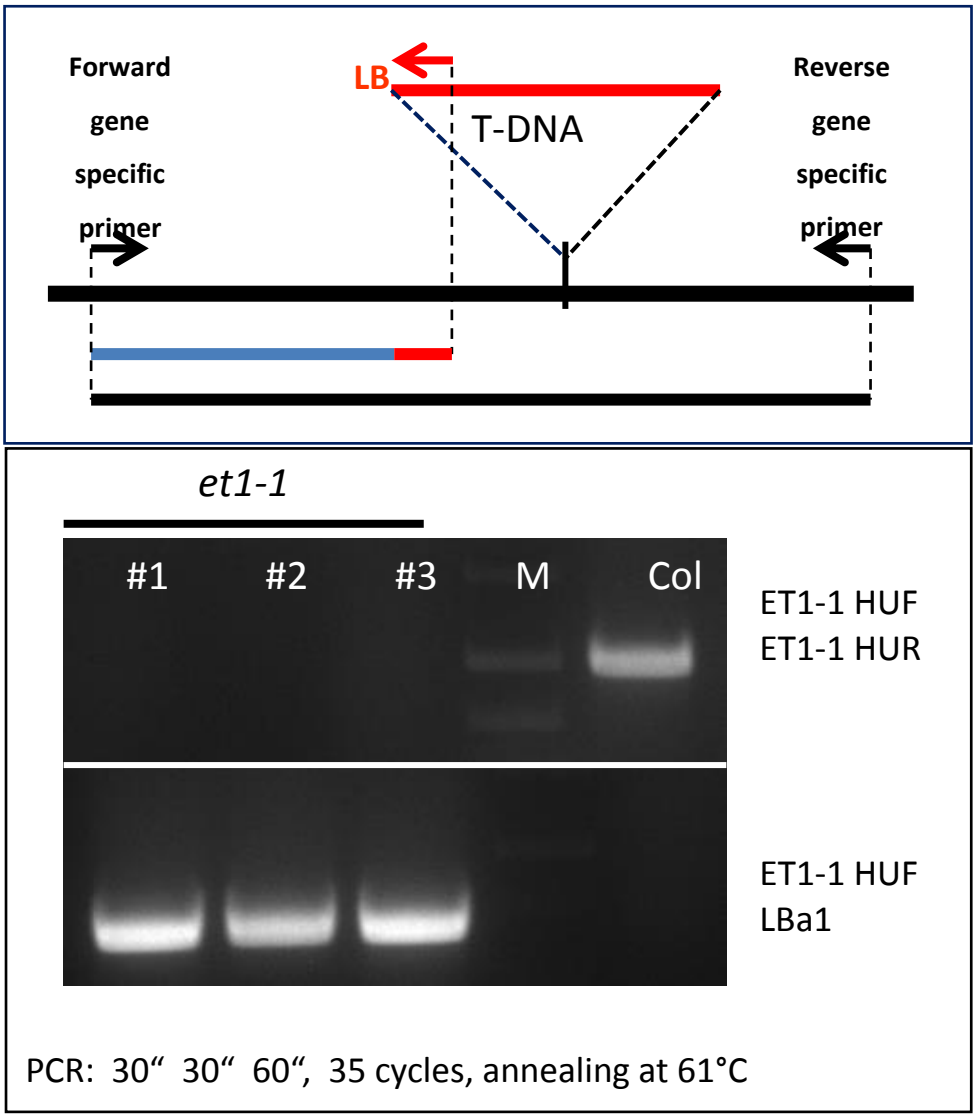

Fig. 9 Detection of the T-DNA insertion in *et1-1*. *Upper panel:* Principal strategy for T-DNA detection using a gene specific primer for the left T-DNA border (LB) and two gene specific primers to detect the wild type allele. *Lower panel:* PCR analysis of three homozygous *et1-1* lines (#) and wild type (Col) using ET1-1HUF and ET1-1HUR as gene specific primers and ET1-1HUF and LBa1 to detect the T-DNA insertion.

Connected to the verification of the T-DNA insertions the TAIR proposed gene model for AtET1 has been investigated by SMART-RACE-technique to amplify and sequence the 5´-terminal part of the transcript. The results suggest that the database predicted gene model of *AtET1* needs to be

corrected. The sequence of the RACE amplified fragment does not support the existence of the predicted small upstream exons and introns and the gene start needs to be shifted as shown in Fig. 10. Based on this new gene model, the T-DNA insertion in *et1-5* allele is now located far up in the 5`flanking region, 699 bp in front of the translation start and in the *et1-1* allele the T-DNA insertion is positioned in the second exon (Fig. 10).

Up to now it was not yet possible to isolate a homozygous *et1-5* mutant. Although more than 400 plants have been analysed, only heterozygous genotypes could be detected by now. This is rather unexpected since the position of the T-DNA insertion is rather far up in the 5´-flanking region, but it still might destroy the gene promoter activity. Currently, this is investigated further.

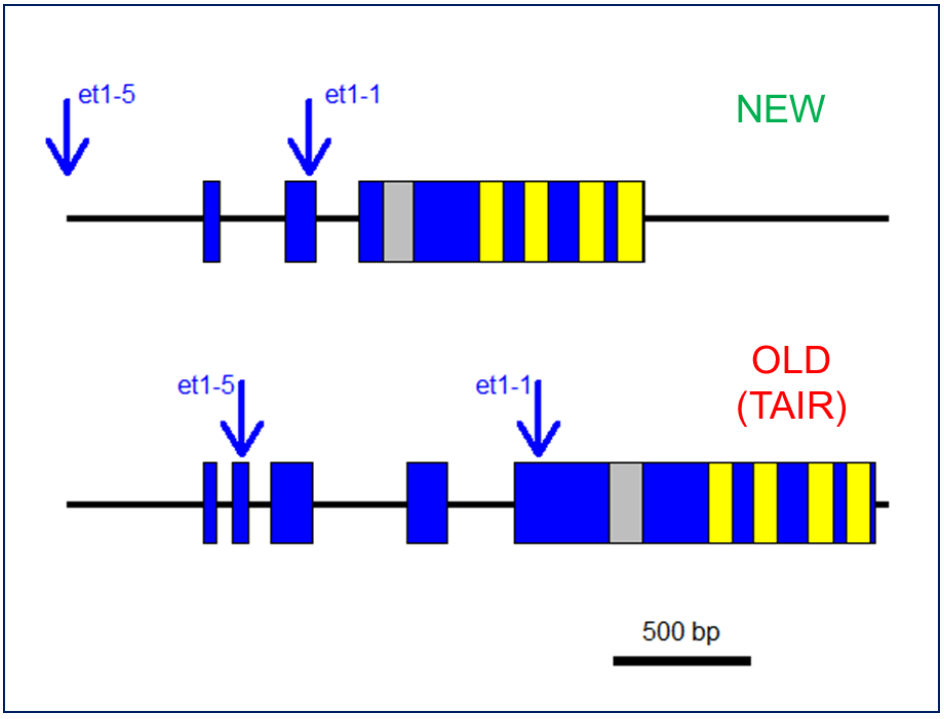

Fig. 10 Comparision of the TAIR data base predicted gene model and the new experimentally confirmed gene model.

The new model derived translation product is also supported by tryptic peptides identified by peptide mass fingerprinting using matrix assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry.

The AtET1 cDNA has been expressed in *E. coli* and the analysis of the isolated gene product resulted in the peptides given highlighted in red in Fig. 11. (These results have been provided by A. Matros and H. P. Mock, IPK). Therefore, we strongly favour the new gene model.

MFK**RDDYIR**TNHDPFFSKWQGFARSMFLR**KPISETAELRKTFADYSL**
ISRDLGPKPK**ILIGANEKENFR**EGKDLVGRNR**VQGAFQGLYELSHDH**
GRKDDVLVANLGQPESIRSRLRSYSR**SFAHHDLLKQGLSQTILPTTQ**
NKSDNQTEEKKSDSEEEREVSSDAAEK**ESNSLPSILR**LSR**SRPQPVS**
EKHDDIVDESDSASACGVLLEDGTTCTTTPVKGRKRCTEHKGKRLSR
VSPGIHIPCEVPTVRECEETENICGVILPDMIRCRSKPVSRRKRCED
HKGMR**VNAFFFLLNPTER**DKAVNEDKSKPETSTGMNQEGSGLLCEAT
TKNGLPCTRSAPEGSKRCWQHKDKTLNHGSSENVQSATASQVICGFK
LYNGSVCEKSPVKGRKRCEEHKGMRITS

Fig.11 Tryptic peptides given in red obtained by MALDI-TOF analysis confirms the newly proposed gene model.

1.2 Mutants in *AtET2*

The *et2-1* mutant was selected by pool screening from the collection of *Arabidopsis* Knock-Out-Facility (AKF), University of Wisconsin (Ivanov, 2005). This collection has been transformed with a derivative of the T-DNA vector pD991 in the *Wassilewskija 2* ecotype (Ws) (Krysan et al., 1999; Sussman et al., 2000) and was designated as *et2-1* (Ivanov et al., 2008). This mutant was backcrossed repeatedly into ecotype Col at least four times before used for further analyses and renamed *et2-1*. The homozygous mutant status was confirmed by PCR analysis using gene specific primers (ET2_RT_ACF and ET2_RT_ACR) as well as the T-DNA right border primer XR2 in combination with ET2_RT_ACR. The expected fragment length is 851 bp (Fig. 12).

The SALK_151861 line was screened from SIGNAL T-DNA collection and after confirmation and precise localization by sequencing was renamed *et2-3*. The homozygous *et2-3* also was identified by PCR analysis using the same primer of *et2-1* as shown Fig.12. The expected fragment length was 1284 bp and 1369 bp for the mutant and wild type allele, respectively.

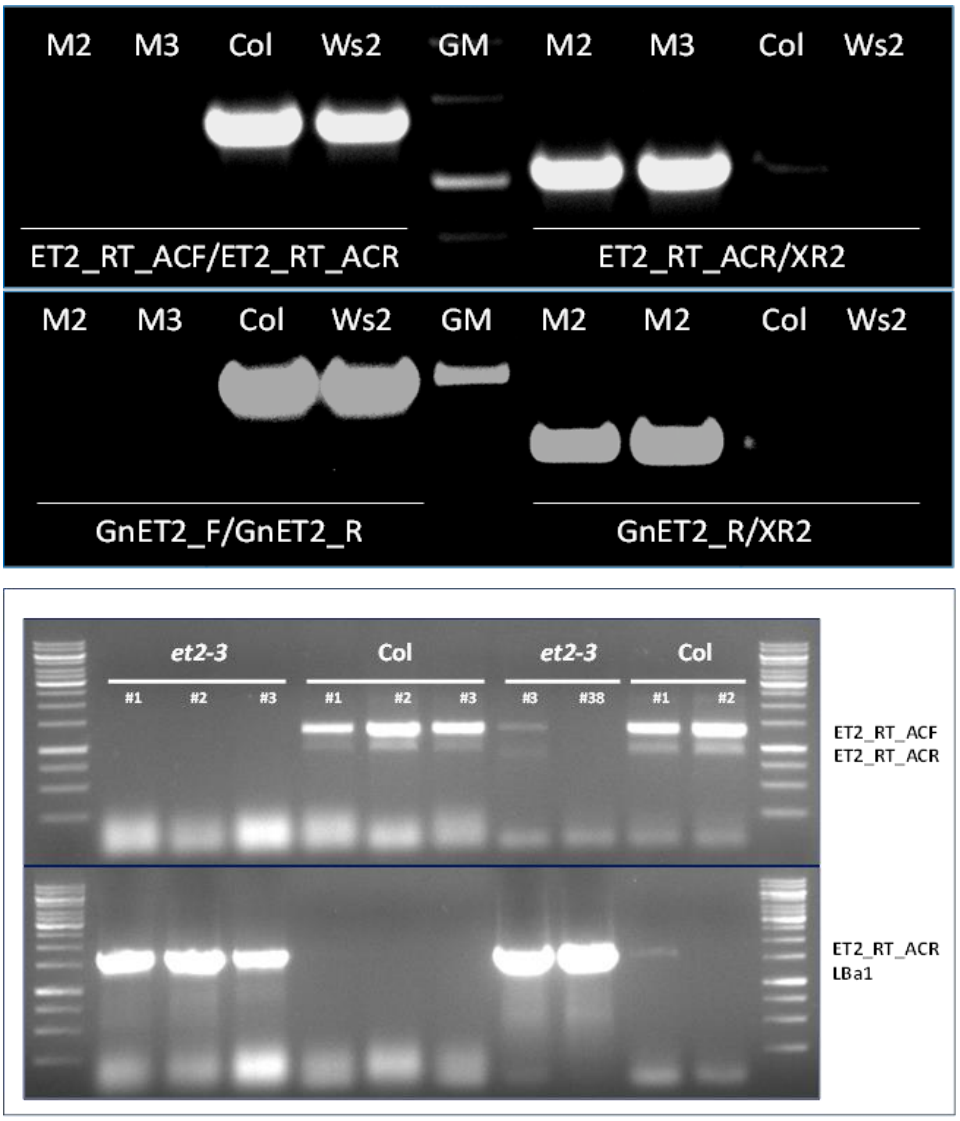

Fig. 12 Detection of T-DNA insertions in mutants AtET2. Upper panel: Genotyping of *et2-1* using to different primer sets (RT_ACF/RT_ACR and GnET2_F/GnET2_R). XR2, T-DNA right border primer of pD991; M2, M3, mutant lines; Col, Ws2 ecotypes Columbia and *Wassilewskija2*; GM, size marker. Lower panel: Genotyping of *et2-3* using gene specific primers ET2_RT_ACF/ET2_RT_ACR and ET2_RT_ACR/LBa1.

1.3 Mutants in *AtET3*

Two T-DNA insertion lines have been identified and characterized for *AtET3*. Lines CS423803 and CS431900 were obtained from the SIGNAL-

collection (see above) and after confirmation and precise localization by sequencing were renamed into *et3-3* and *et3-2*, respectively. Both mutants were shown to be homozygous. The insertions in both lines are located close to each other within the 5`-flanking gene region (Fig. 13).

In total there are 6 well characterized insertion lines available as summarise in Fig. 13. The single mutant *et1-1* was combined with the single mutants *et2-3* and *et2-1* to generate homozygous double mutants (see below).

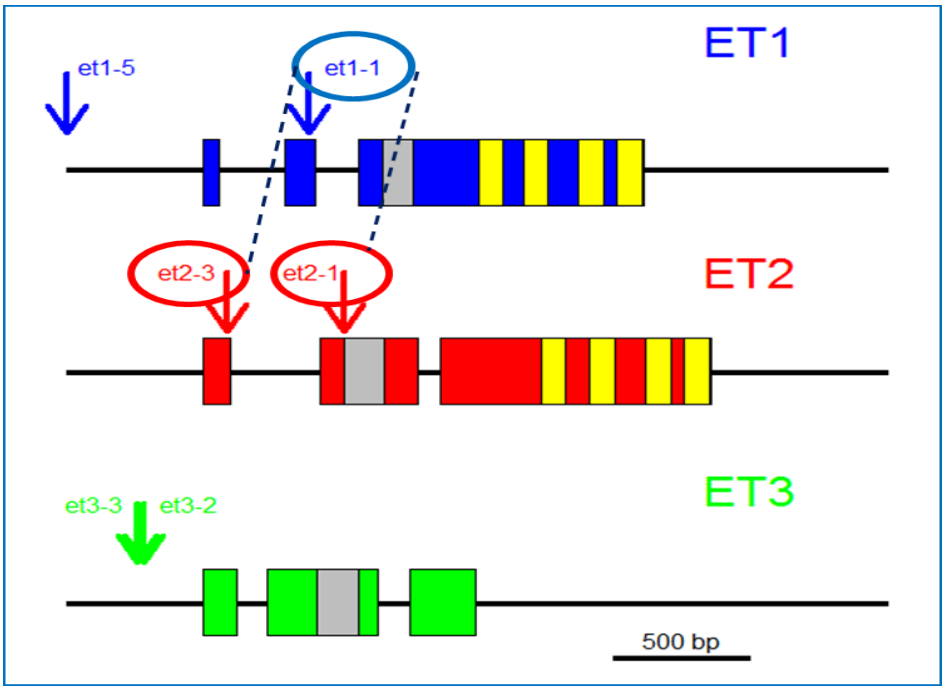

Fig.13 Gene model and positions of T-DNA-insertions of the *ET* gene family. The positions of T-DNA insertions are indicated by arrows. ET repeats and the GIY-YIG single strand cutting domain are given in yellow and grey, respectively. The dashed lines indicate the alleles which have been combined as homozygous double mutants.

Gene	Mutant	Position	Insert relative to ATG
ET1	*et1-5*	promoter	-699
	et1-1	Exon2	383
ET2	*et2-3*	Exon1	85
	et2-1	Exon2	518
ET3	*et3-3*	promoter	-239
	et3-2	promoter	-216

Tab. 1 Nucleotide positions of 6 T-DNA insertion mutants relative to the ATG start.

1.4 Generation of double *et1 et2* mutants

AtET1 and *AtET2* are two closely related proteins. They share an overall amino acid identity of 40%, especially in the ET repeats (58%). To analyze possible functional redundancy, as indicated by similar phenotypes (see below), two double mutants have been generated by crossing the homozygous mutant *et1-1* both with homozygous *et2-1* and *et2-3* mutants. Homozygous double mutants *et1-1 et2-1* and *et1-1 et2-3* have been selected and characterized in the F2 generation (Fig. 14).

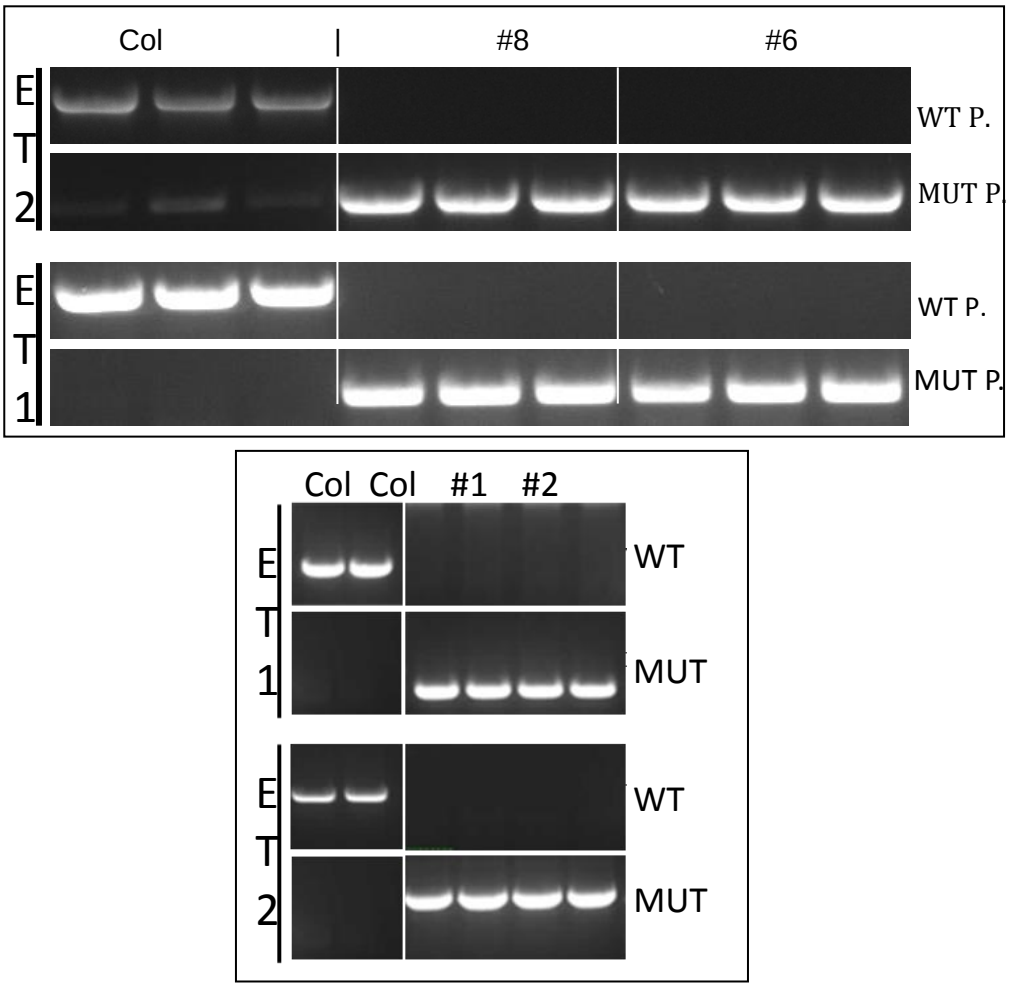

Fig. 14 Genotyping of *et1 et2* double mutants. *Upper panel*: Genotyping of *et1-1 et2-1* double mutant demonstrating the homozyogous double mutant status of lines #8 and #6. *Lower panel:* Genotyping of *et1-1 et2-3* double mutant demonstrating the homozygous mutant status of lines #1, #2, #3, #4.

1.5. Loss of transcripts in single mutants *et1-1, et2-1* and *et2-3*

A final step to verify the homozygous status of single mutants (*et1-1, et2-3, et2-1*) requires to demonstrate the absence of the wild type *AtET1* and *AtET2* mRNA, respectively. Therefore, RT-PCR on total RNA from wild type and homozygous mutant plants was performed and actin primers were used to control RNA quality and quantity. The *AtET1* and *AtET2* products, could be amplified only from the wild type sample, demonstrating the destruction of the intact *AtET1, AtET2* mRNA (Fig.15) in the homozygous

mutants. The data demonstrate that the *AtET1* insertion line *et1-1* and the *AtET2* insertion lines *et2-3* and *et2-1* can be used as suitable tools for functional studies of the ET gene family in Arabidopsis.

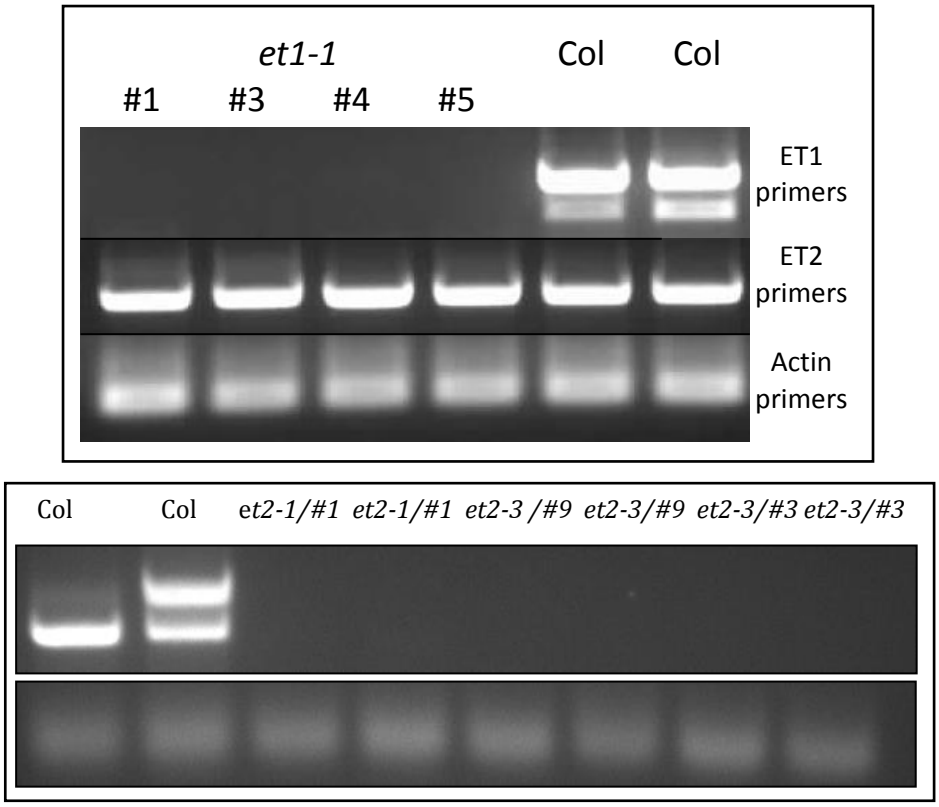

Fig. 15 RT-PCR to demonstrate the knock out character of single mutants *et1-1, et2-1* and *et2-3. Upper panel: et1-1* mutant and wild type probed with *ET1* specific primers which span the insertion site show the loss of the *ET1* transcript. Priming with *ET2* specific primers demonstrate that the *ET2* transcript is not affected. *Lower panel: ET2* specific primers which span the insertion site were used to show the loss of the *ET2* transcript. Actin gene was used in both cases as loading control.

2. Phenotypic characterisation of *et* mutants

A phenotypical characterization of these *et* mutants revealed several highly interesting observations, including effects on, flower organ identity, gametophyte development, endosperm development, immature seed germination, pollen development and seed development.

2.1. Homeotic transformation of flower organs in single and double mutants.

Phenotypic inspection of mutants and wild type reveals unusual numbers of flower organs such as sepals, petals and stamens (Fig. 16).

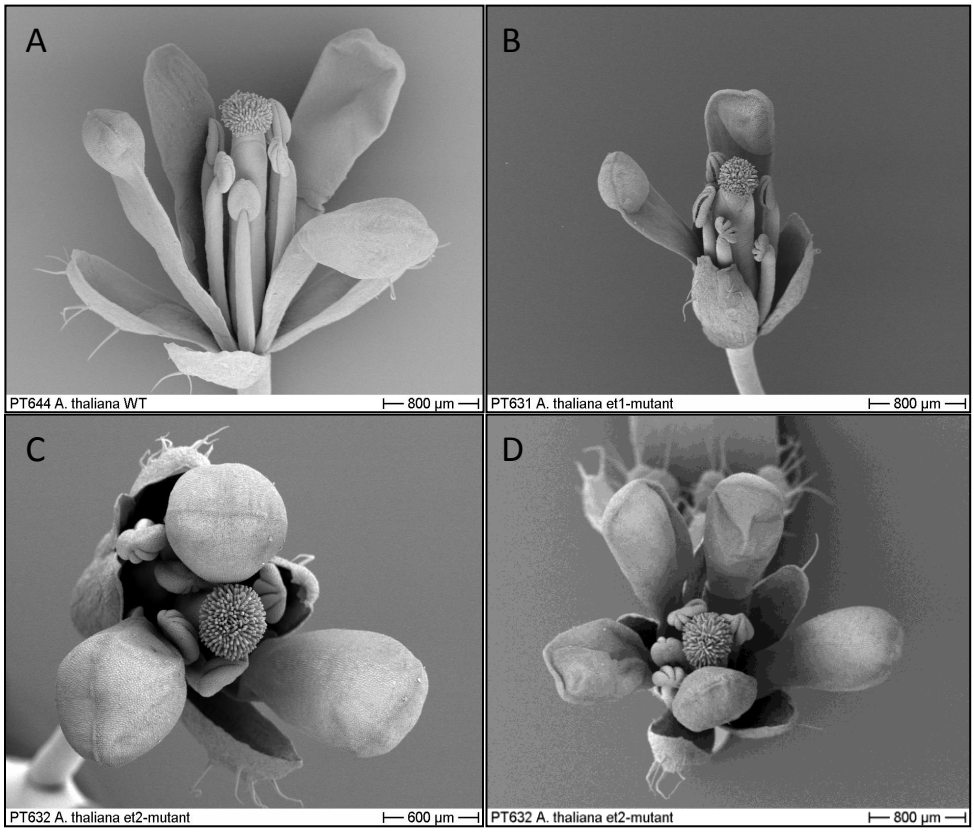

Fig.16 Changed numbers of flower organs in *et* mutants. REM pictures of wild type (A) and selected mutant flowers with two petals and two sepals in *et1-1* (B), with three petals in *et2-3* (C) and five petals in *et2-3* (D).

Remarkably, *et* mutants exhibit homoeotic transformations of anthers into carpel-like structures including the occurrence of stigma-like structures as well as ectopic ovules as shown in Fig. 17.

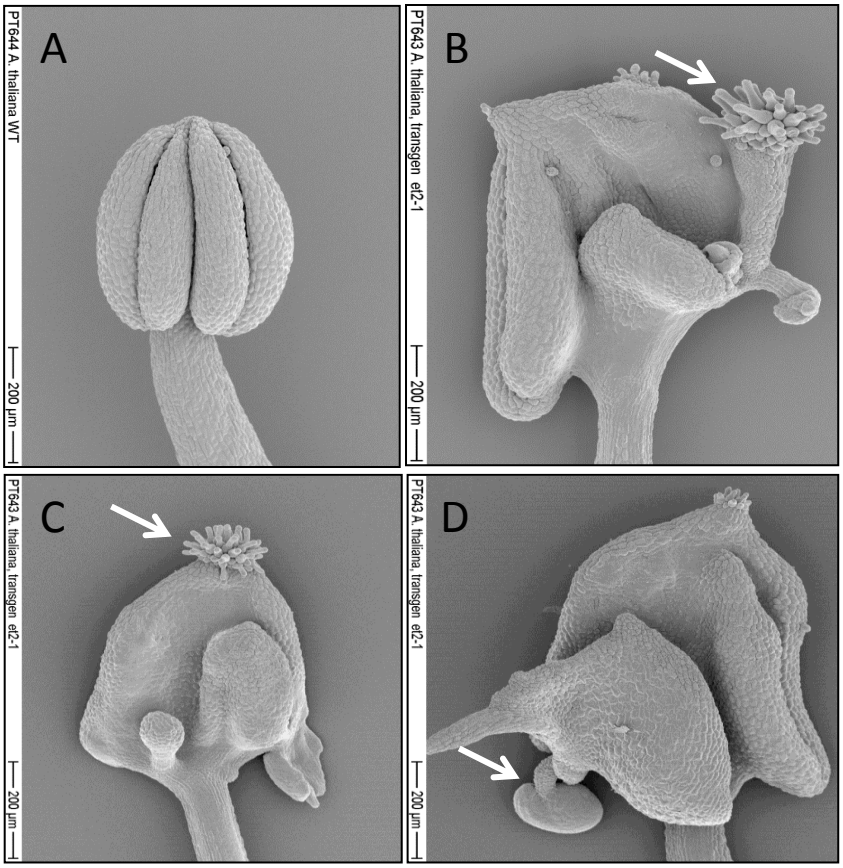

Fig. 17 Homoeotic transformation of anthers into carpel-like structures in *et2-1* mutants. REM pictures of wild type anthers (A) and various homoeotic transformations of anthers into carpel-like structures (B, C, D) including stigma and ovule formation (arrows).

To further characterize the stamen-derived ovules in more detail the tissue was cleared and analyzed by DIC microscopy. As shown in Fig. 18 the ectopic ovules contain a nearly normal gametophyte with fully developed egg cell, synergids and fused polar nucleus of the central cell. However, the normal polarity with synergids next to the micropyle followed by egg cell and central is distorted in these ovules.

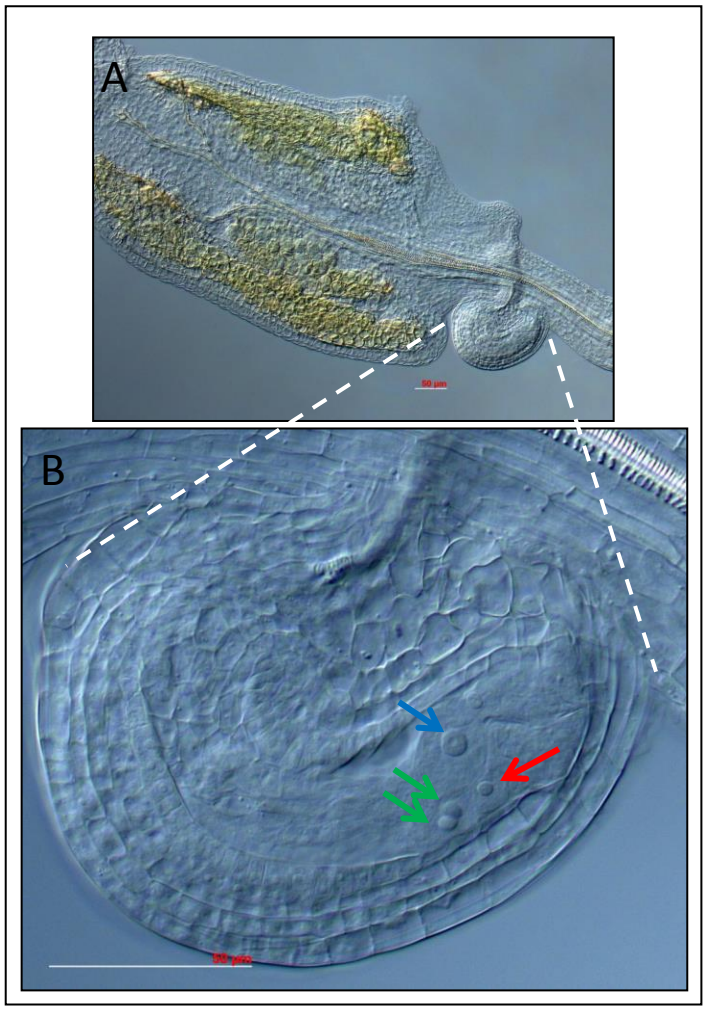

Fig. 18 Stamen-derived ovules of the *et2-1* mutant contain a fully developed gametophyte with egg cell (red arrow), two synergids (green arrows) and a homodiploid central cell nucleus (blue arrow). However, the normal polarity of the gametophytic cell types is partially distorted.

The severity of the homoeotic transformation is quantitatively and qualitatively further increased in the *et1-1 et2-1* double mutant as shown in Fig. 19. The double mutant exhibits multiple ovule- and stigma-like structures.

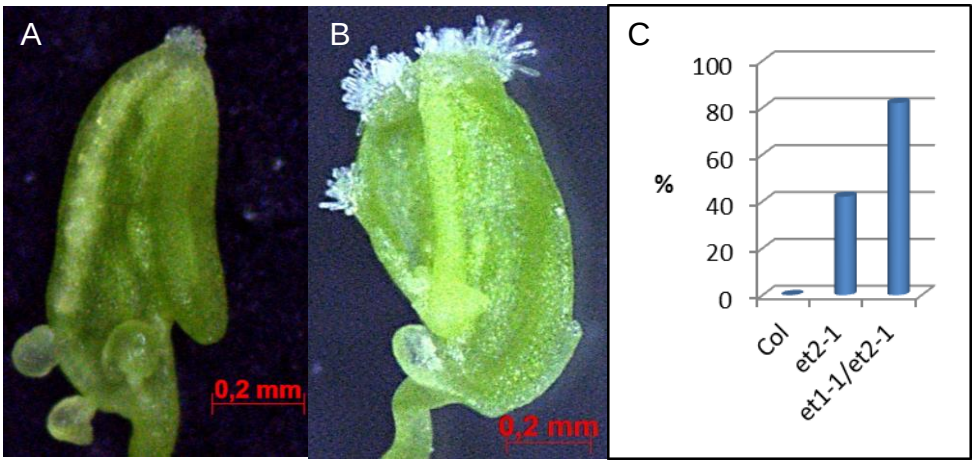

Fig. 19 The homoeotic transformation of stamen into carpel-like structures in the *et2-1* mutant (see above) is further enhanced in the *et1-1 et2-1* double mutant with multiple ovule and stigma formation (A, B). The effect has been quantified in 180 flowers each (C).

2.2 Distortions of gametophyte development.

The Arabidopsis female gametophyte, the embryo sac, develops within the ovule consists of two synergids, one egg cell, one central cell and three antipodal with the latter degenerating at the mature stage. In wild type the homodiploid nucleus of the central cell results from the fusion of the two polar nuclei. As shown in Fig. 20 all *et* mutants exhibit non fused central cell nuclei. In general et mutant embryo sacs show up to 20% distortions such as multiple non-fused polar nuclei or missing egg cells or synergid cells.

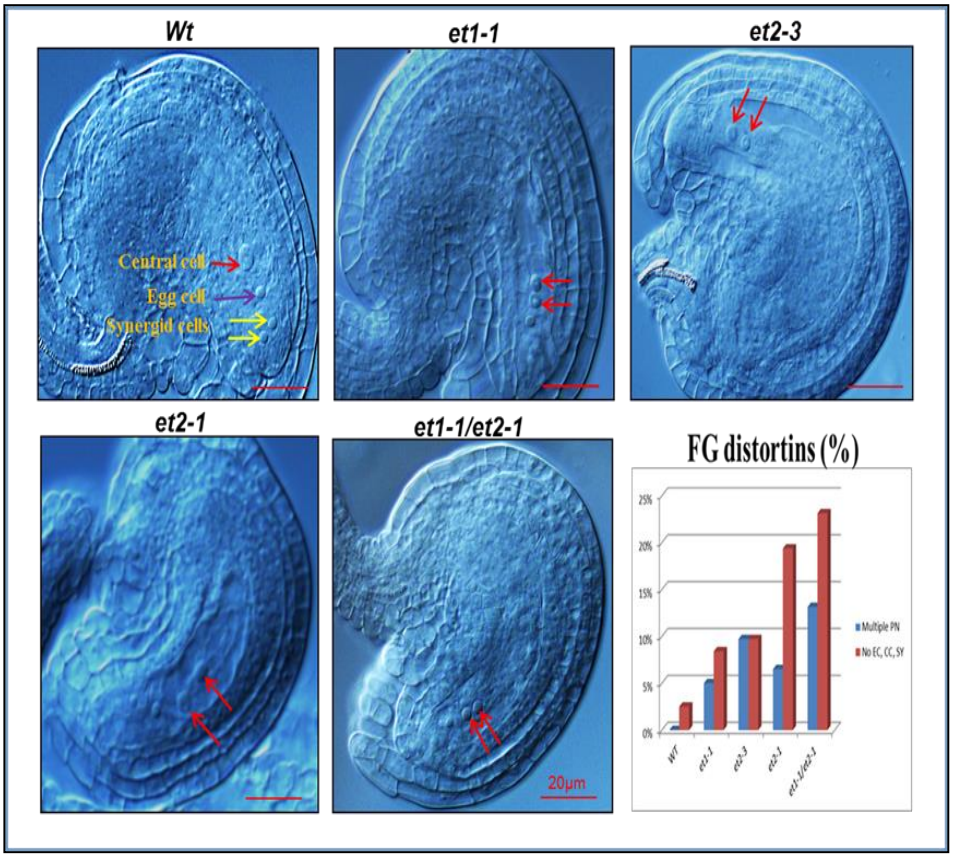

Fig. 20 Distorted embryo sac development in *et* mutants. In the mature wild type embryo sac the two polar nuclei fuse. In all *et* mutants this does not occur and the two polar nuclei remain non-fused (blue arrows). The gametophytic distortions have been quantified. PN, polar cell; EC, egg cell; CC, central cell, SY; synergids, Scale bar: 20 µm.

To further characterize the polar nuclei fusion, a central cell specific reporter construct has been used. The construct consists of the central cell-specific DD65 promoter controlling the AmCyan fluorescent protein gene. As shown in Fig. 21 the construct specifically labels the central cell in the wild type, whereas the marker signal is missing in the *et2-1* mutant.

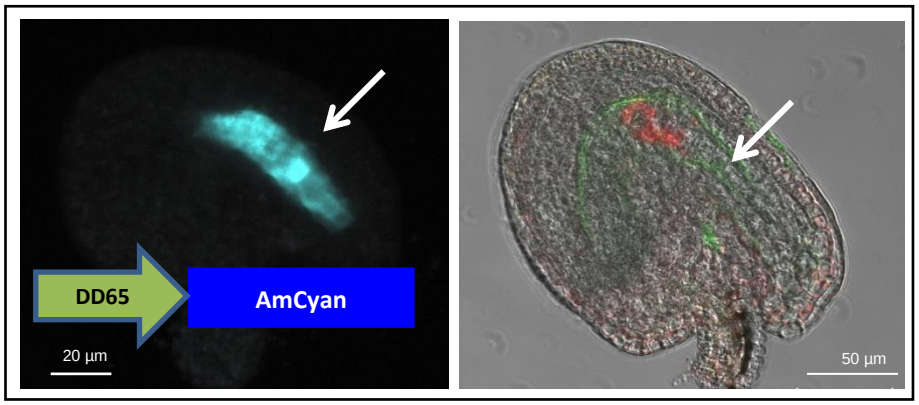

Fig. 21 Distorted central cell differentiation in the *et2-1* mutant. *Left panel:* The marker line is controlled by the central cell-specific promoter DD65 and specifically labels the central cell in wild type. *Right panel:* The signal is missing in the collapsed embryo sac of the *et2-1* mutant.

2.3. Distortions of endosperm development.

Another aspect of the *et* mutant phenotype concerns the endosperm differentiation. The endosperm nuclei exhibit a characteristically changed morphology with greatly enlarged nuclei in all *et* mutants, possibly indicating an enhanced synthetic activity (Fig. 22).

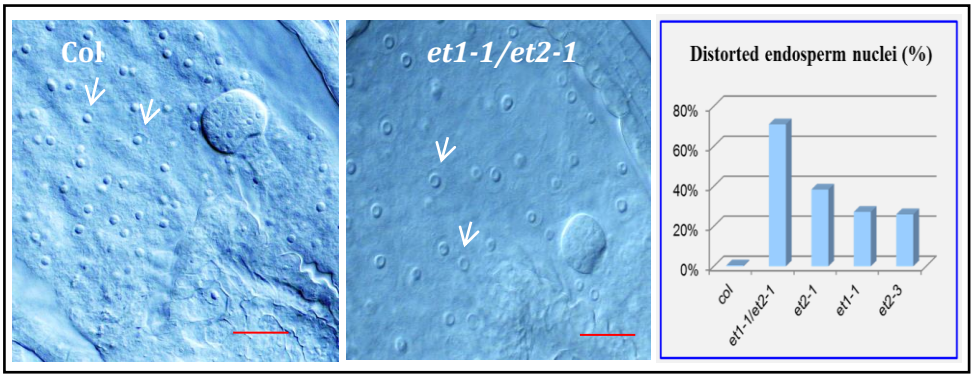

Fig. 22 Affected endosperm differentiation in *et* mutants. The nucleoli of mutant endosperm nuclei are greatly enlarged in the *et1-1 et2-1* double mutant in comparison to wild type. Enlarged nucleoli are considered to reflect an increased synthetic activity. The morphological effect has been quantified in the mutants (right panel), Scale bar: 20 µm.

2.4. Precocious germination of *et* mutants

Germination is initiated by the penetration of the radical through the surrounding seed coat. However, *et* mutants exhibit a strong phenotype of precocious germination. Immature seeds start to germinate already within the siliques. Remarkably, germination does not occur as in mature wild type seeds with the root tip in front, but in the mutant the seedling permeates the seed coat at the side with the cotyledons appearing first (Fig. 23). A similar behavior was observed when immature seeds were germinated *in vitro* (Fig. 24).

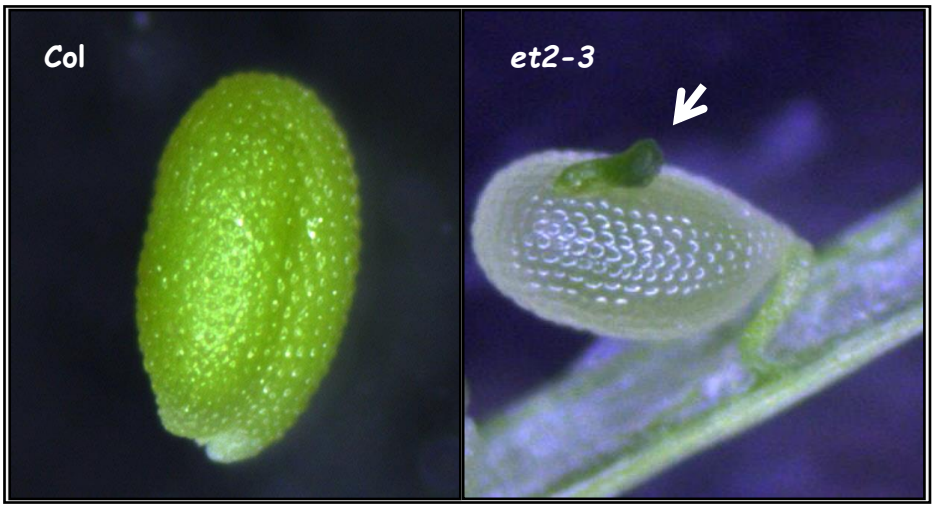

Fig. 23 Precocious germination of *et* mutants. The immature seed of the *et2-3* mutant start to germinate already within the silique. The seedling permeates the seed coat at the side of the seed with the cotyledon appearing first.

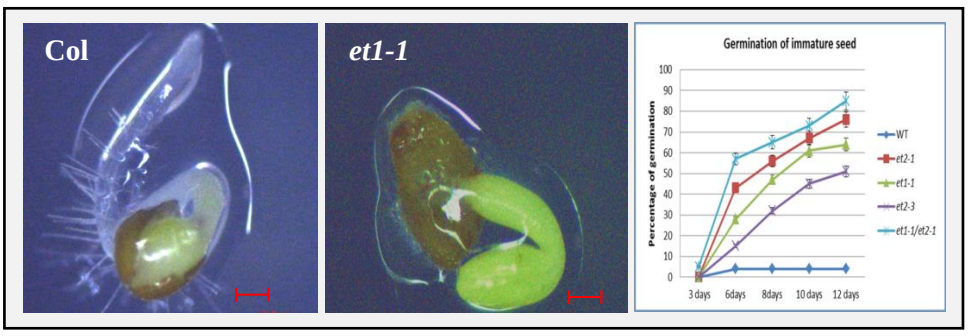

Fig. 24 Precocious germination of *et* mutants *in vitro*. Germinating wild type and *et1-1* mutant seedlings: In wild type the radicle appears first, whereas

the cotyledons show up first in the *et1-1* mutant. The phenotype has been quantified (*right panel*) with 200 seeds each. Scale bar: 0.2 mm.

2.5. Distortions of pollen development.

The male gametophyte, the pollen grain or microgametophyte, develops within the anther and consists of two sperm cells encased within a vegetative cell. Pollen of wild type and *et* mutants was analyzed using DAPI staining. In wild type the vegetative nucleus and the two generative nuclei were clearly distinguishable, whereas significant numbers of abnormal and collapsed pollen grains were detected in *et* mutants (Fig. 25, Fig. 26).

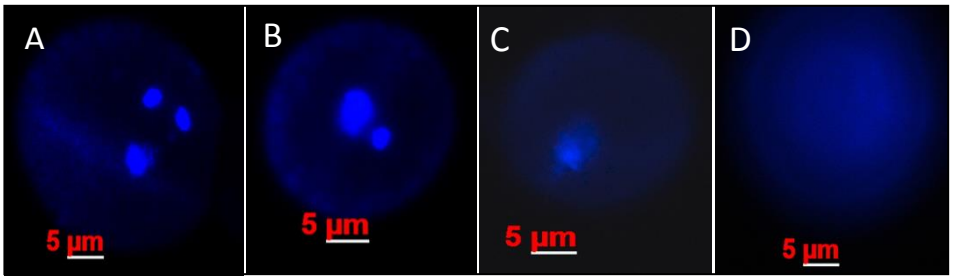

Fig. 25 DAPI stained pollen nuclei. Wild type pollen show the larger vegetative nucleus and the two generative nuclei (A). Various distortions of pollen differentiation including only one generative nucleus (B), one most likely vegetative nucleus (C) and completely collapsed pollen (D) are shown for the *et1-1* mutant.

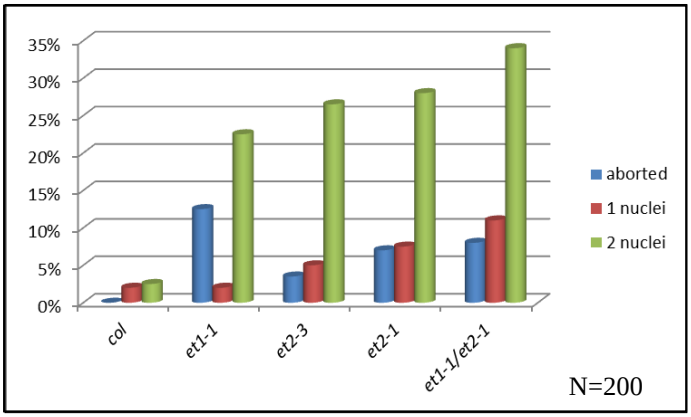

Fig. 26 Quantification of pollen nuclei distortions in *et* mutants. 200 pollen each have been analyzed.

2.6. ET influence on seed set

As a likely consequence of the distorted embryo sac and pollen development a high degree of seed sterility is detected in *et* mutants. Approximately 80% of the *et1-1 et2-1* double mutant seeds are sterile with less than 3% sterility in wild type. The effect on seed set was less pronounced for the two single mutants *et1-1* and *et2-3* (Fig. 27).

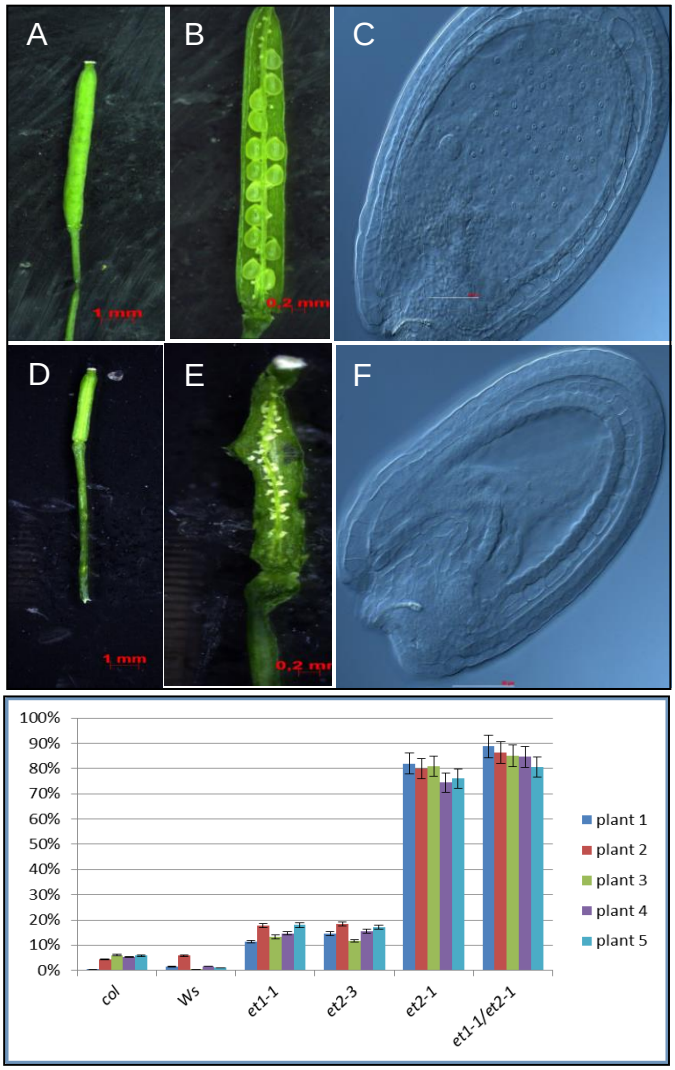

Fig. 27 Seed set in wild type and *et* mutants. *Upper panel:* Fertile silique of *et2-1* mutant (A, B) and early embryogenesis in cleared seeds (C); sterile silique of *et2-1* (D, E) and collapsed embryogenesis in cleared seeds (F). *Lower panel:* Quantification of sterility in wild type and *et* mutants.

3. Comparative transcriptome analysis of *et* mutants

To further investigate the molecular basis of *AtET1* and *AtET2* action and to identify putative target genes we have performed a deep RNA sequencing analysis of wild type compared to *et1-1* and *et2-3*. As tissue we have initially used 10 days old seedlings. In total 288 differential transcripts have been identified. Their differential expression relative to wild type has been mapped on the genome. As shown in Fig. 28 about 2/3 of these genes are downregulated genes in *et* mutants.

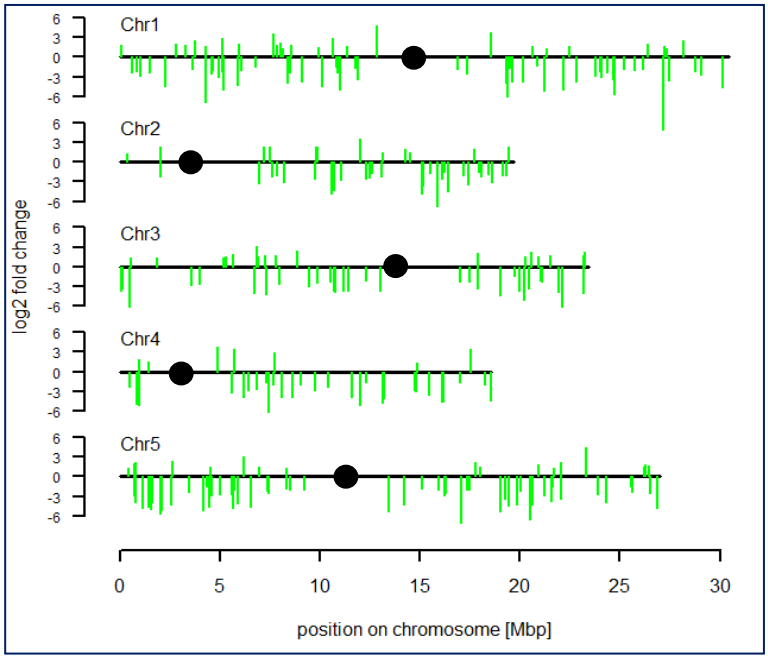

Fig. 28 Positions and levels of expression of differentially expessed genes in *et* mutants compared to wild type. About 2/3 of the genes are downregulated in *et* mutants.

A list of top **up**- and **down**-regulated genes of *et1-1* line and *et2-3* line are given in table 2.

et1_1

AGI	***log2R***	***Annotation***
AT1G40101	8,90	TE
AT1G38360	8,70	TE
AT5G33434	8,35	TE

AT4G06718	10,48	TE
AT1G41680	10,43	TE
AT1G35140	5,22	Phosphate-responsive protein
AT5G57560	4,12	Xyloglucan endotransgl.
AT3G45970	3,94	Expansin-like A1
AT4G08950	3,56	Phosphate-responsive protein
AT2G28190	3,41	Superoxide dismutase 2
AT4G13340	3,21	LRR family protein
AT2G18050	-3,02	Histone H1-3
AT4G21990	-3,07	APS reductase 3
AT2G46680	-3,02	Homeobox 7
AT1G68600	-3,10	Malate transporter
AT4G04610	-3,07	APS reductase 1
AT1G02930	-3,15	Glutathione S-transferase 6
AT3G44990	-3,27	Xyloglucan endo-transglyc.
AT2G05510	-3,31	Glycine-rich protein family
AT5G03190	-3,37	Conserved open reading frame
AT4G12480	-3,52	2S albumin superfamily protein
AT3G53980	-3,65	2S albumin superfamily protein
AT3G01420	-3,69	Peroxidase superfamily protein
AT4G16260	-3,70	Glycosyl hydrolase protein
AT4G31970	-3,83	Cytochrome P450, subfamily C
AT1G66270	-3,79	Glycosyl hydrolase protein
AT3G28270	-3,78	Unknown function (DUF677)
AT1G21310	-4,08	Extensin 3
AT1G26390	*-4,34*	FAD-binding Berberine protein
AT1G32900	-4,28	UDP-Glycosyltransf. protein
AT4G12500	-4,47	2S albumin superfamily
AT4G30170	-4,46	Peroxidase family protein
AT3G21670	-4,60	Major facilitator protein
AT2G39330	-4,65	Jacalin-related lectin 23
AT5G15960	-4,65	Stress-responsive protein (KIN1)
AT4G14090	-4,69	UDP-Glycosyltransferase protein
AT3G17790	-4,80	Purple acid phosphatase 17
AT5G66400	-4,76	Dehydrin family protein
AT4G12490	-5,23	2S albumin superfamily protein
AT5G13930	-5,29	Chalcone synthase family protein
AT5G17220	-5,83	glutathione S-transferase phi 12
AT4G22880	-5,90	leucoanthocyanidin dioxygenase
AT1G70260	-6,92	nodulin MtN21 transporter family
AT5G42800	-6,64	dihydroflavonol 4-reductase
AT5G54060	-6,57	UDP-glucose:flavonoid gluc.transf.
AT5G59320	-7,35	lipid transfer protein 3
AT5G59310	-7,40	lipid transfer protein 4

et 2-3

AT5G57560	3,38	Xyloglucan endotransglucosylase
AT2G28190	3,23	Copper/zinc superoxide dismutase 2
AT3G19680	3,08	Protein of unknown function (DUF1005)
AT3G01420	-3,36	Peroxidase superfamily protein
AT1G66270	-3,37	Glycosyl hydrolase superfamily
AT4G14090	-3,76	UDP-Glycosyltransferase superfamily
AT3G53980	-4,03	2S albumin superfamily protein
AT5G42800	-4,35	Dihydroflavonol 4-reductase
AT5G17220	-4,36	Glutathione S-transferase phi 12
AT5G13930	-4,52	Chalcone synthase family protein
AT4G22880	-4,82	Leucoanthocyanidin dioxygenase
AT5G59320	-4,94	Lipid transfer protein 3
AT5G59310	-5,08	Lipid transfer protein 4

Tab. 2 List of up- (given in red) and down- (given in blue) regulated genes in *et1-1 (upper part)* and *et2-3 (lower part)* mutant relative to wild type. Genes with overlapping regulation in both mutants are underlined. Note the strong upregulation of various transposable elements exclusively in the *et1-1* mutant.

4. Identification of differentially methylated genomic regions in *et* mutants

The above mentioned working hypothesis concerning the molecular function of ET factors suggests that they might act as novel regulators of epigenetic methylation patterns. To test this, the methylation status of the whole genome of *et* mutants has been analysed with wildtype as control. In total 527 differentially methylated regions (DMR) could be detected by using the iPlant visualization tool. This experiment has been performed by Dr. C. Becker, MPI Tübingen. The positions of these DMRs have been mapped on the Arabidopsis genome (Fig. 29).

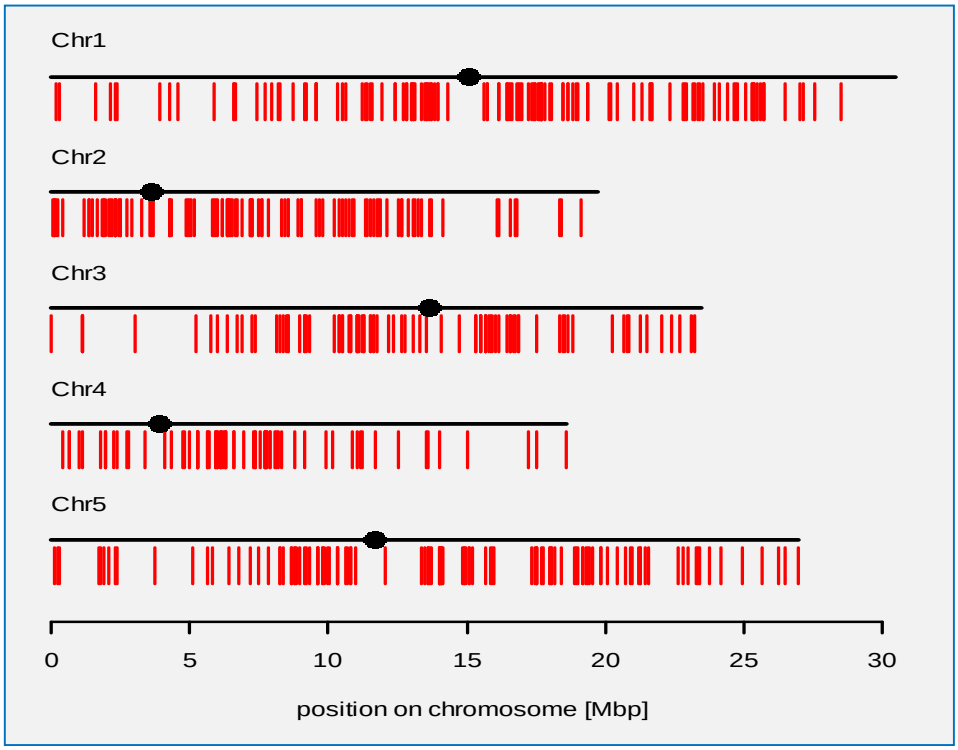

Fig. 29 Positions of differentially methylated regions in *et* mutants compared to wild type.

A few differentially methylated regions have been analyzed in more detail and some examples are shown in the following figures.

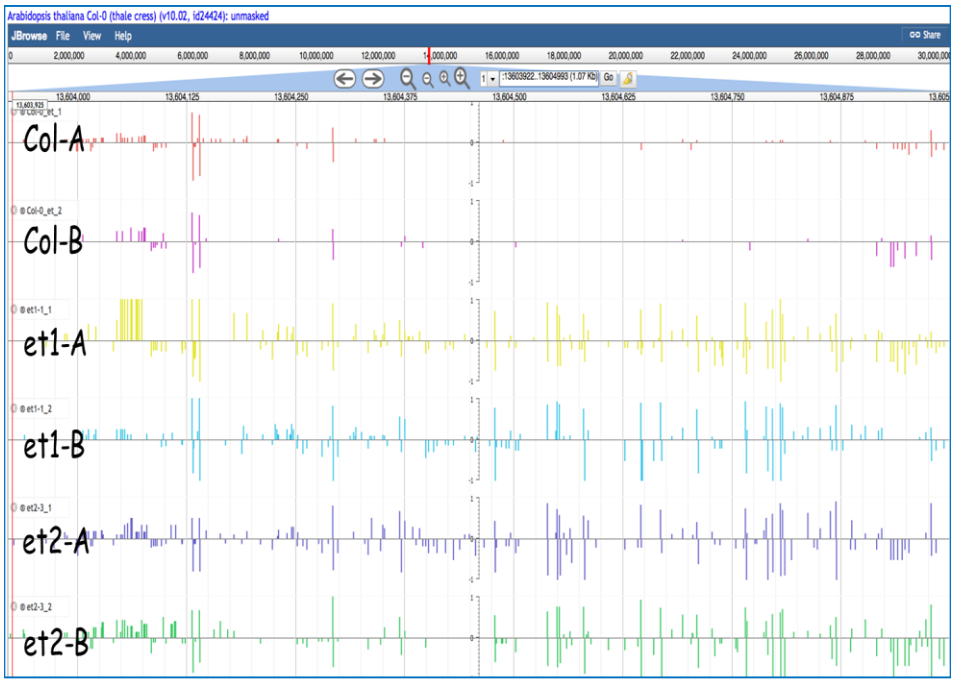

Fig. 30 Methylation pattern at a selected position on chromosome 1. The chosen position is 1.07 kb long between position 13603922 and 13604993. Two independent biological replicates have been analyzed (A, B) and demonstrate the high reliability of the technique.

Another example concerns the region designed as diffM133 (Fig 31). This region is methylated in wild type but unmethylated in both mutants *et1-1* and *et2-3*. The position of the differential methylation pattern in wild type and mutants precisely coincides with the position of the transposable element Helitron 1 in the promoter of the gene AT1G75950. This gene encodes for an E3-SCF protein also known as ASK1 (Arabidopsis SKP1 homologue 1) and has been described to be involved in flower development. In cooperation with UFO and LEAFY the gene product regulates the B function required for flower development in Arabidopsis (Zhao et al., 2001).

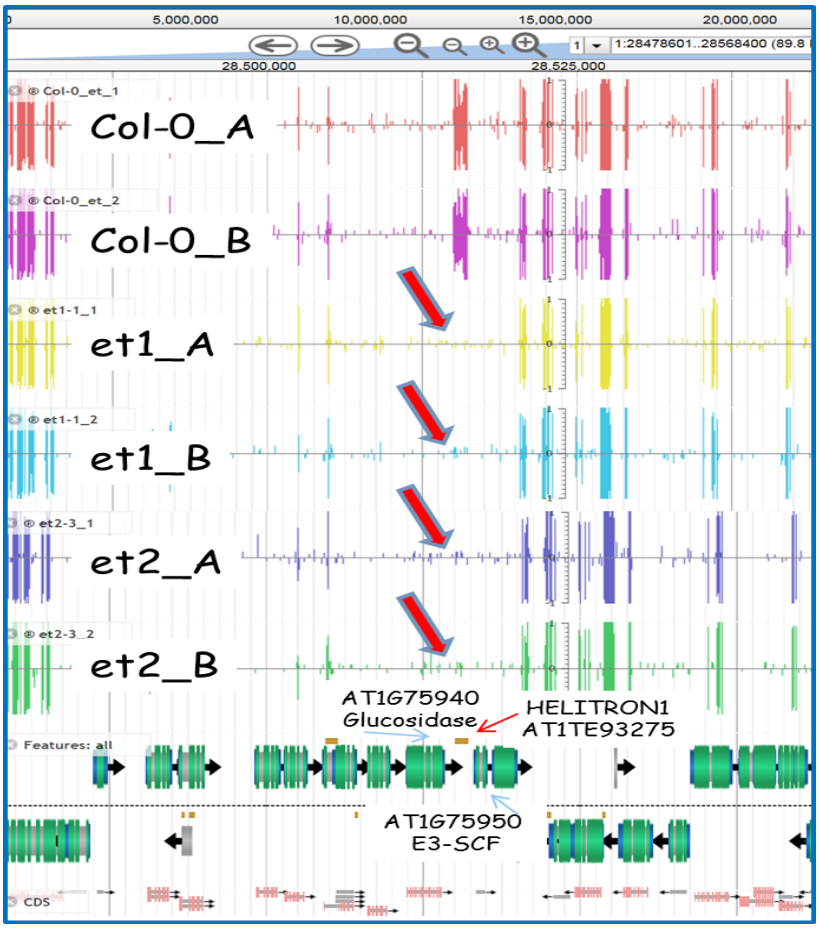

Fig. 31 Differential methylation pattern at the DMR diffM133. The transposable element Helitron1 (AT1TE93275) positioned in the putative promoter region of the gene AT1G75950 is unmethylated in both mutants. The gene product is known to be involved in flower development. Note the high reproducibility of the two independent biological replicates A and B. The basic figure is derived from the iPlant Visualisation Tool.

Another selected example shows the methylation pattern in the DMR diffM105. This region exhibits different methylation with the gene AT1G61810 encoding a β glucosidase. The gene is methylated in wild type and the *et2-3* mutant but not in *et1-1* mutant. This preliminarily indicates the existence of ET1 and ET2 specific targets. An adjacent gene (AT1G61800) encodes a glucose 6-phosphate translocator and exhibits a differentially expression with a log2R of -3.8.

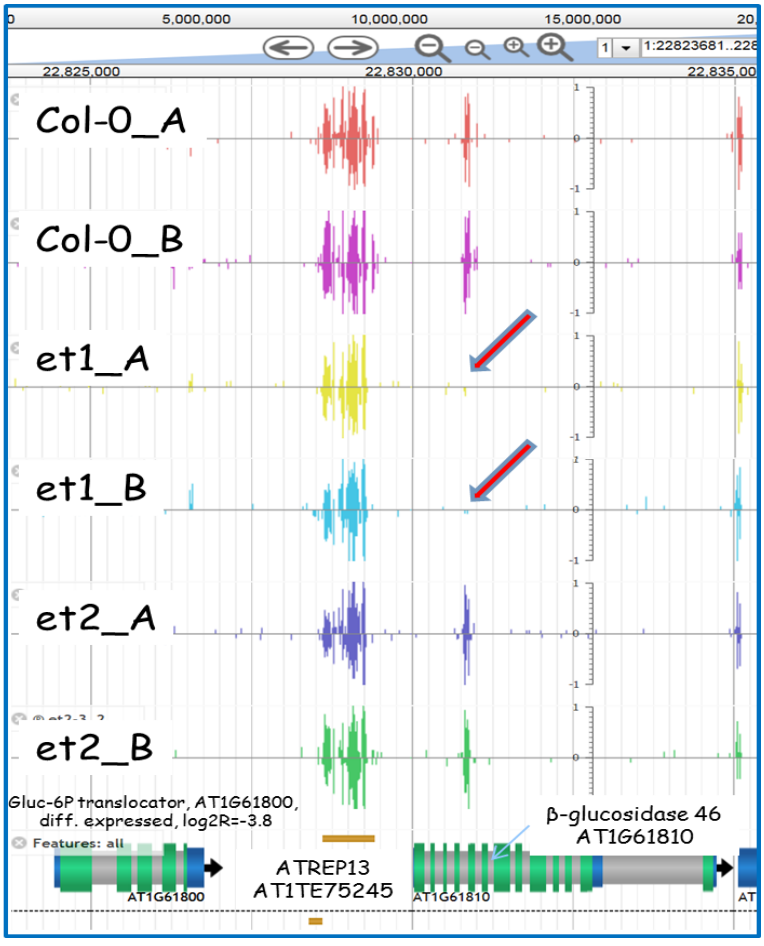

Fig. 32 Differential methylation pattern at the DMR diffM105. The methylation is missing in *et1-1* but occurs in *et2-3* and wild type. The adjacent gene for glucose 6-phosphate translocator has been found to be differentially expressed.

The analysis of various other DMR demonstrates that all combinatorial possibilities of methylation between wild type and the two mutants do occur. Thus the diffM024 region is methylated in wild type and *et2-3*, but not in *et1-1*. At diffM048 methylation is detected in wild type and *et1-1* with methylation missing in *et2-3.* Finally, in diffM001 no methylation is found in wild type and *et2-3* whereas the region is methylated in *et1-1* (Fig. 33).

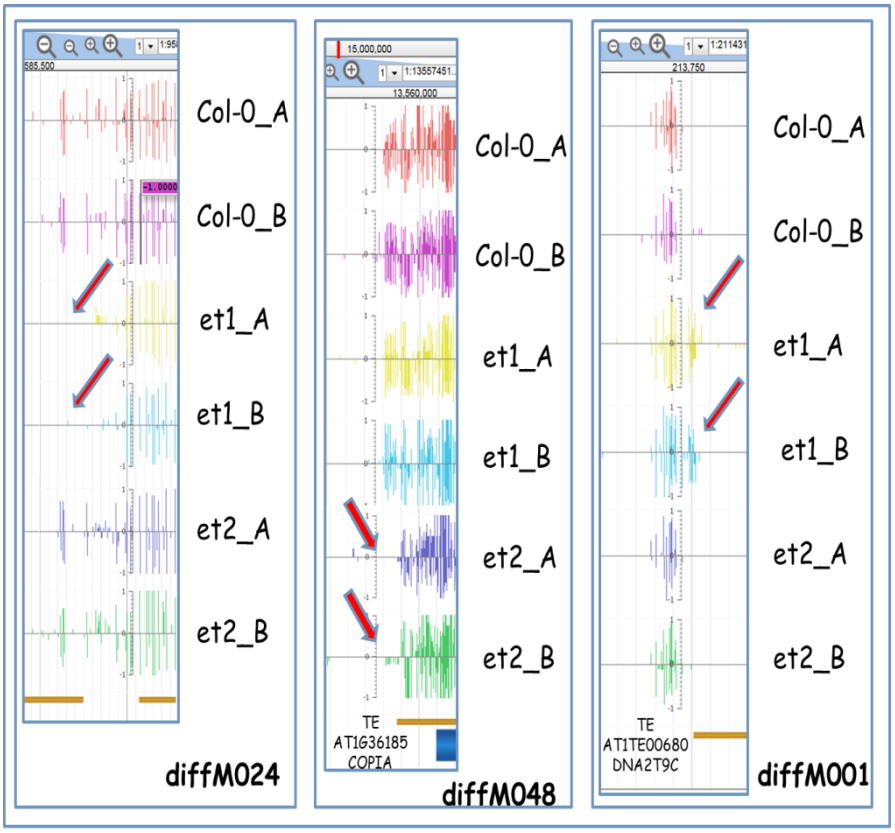

Fig. 33 Differential methylation in selected DMRs reveals the occurrence of various combinations between wild type and the two mutants.

Taken together, all possible combinations concerning the methylation patterns in wild type and the two mutants are possible and have been

detected as summarized in table 3. The methylation context CG is the most frequent one, followed by CHG with very rare cases of the CHH context. Up- and down methylation in mutants occurs with similar frequency. The observation suggests that ET1 and ET2 have common and gene specific target regions.

Col	*et1-1*	*et2-3*	
+	-	-	√
+	+	-	√
+	-	+	√
-	+	+	√
-	+	-	√
-	-	+	√

Tab. 3 All combinations of methylation patterns between wild type and both *et* mutants can be found. ET1 and ET2 have **common AND gene specific** target regions.

Summarizing the described results it is concluded that ET factors represent novel plant specific epigenetic regulators of reproductive tissue development acting on DNA-methylation.

Discussion

EFFECTOR OF TRANSCRIPTION (ET) factors have been originally isolated as DNA-binding proteins using seed specific gene promoter sequence motifs. However, extensive interaction studies failed to identify a specific sequence motif. Ectopic expression of ET factor genes resulted in severe growth distortions including dwarf growth, late flowering, reduced germination rates, strong anthocyanin accumulation, reduced lignification etc.. Together, these observations indicated a putative function as repressors of transcription of GA-regulated genes including KNAT genes involved in cell differentiation. This is also supported by studies on a barley homologue, HORDEUM REPRESSOR OF TRANSCRIPTION (HRT), which was shown to repress the expression of amylase genes in barley aleuron cells. Remarkably, ET/HRT factors are exclusively found in plant genomes including phylogenetically old species such as the moss *Physcomitrella*. These findings strongly suggest that ETs are involved in the regulation of plant specific processes. All ET factors share a variable number of characteristic, highly conserved cystein-histidine containing ET repeats involved in zinc and DNA binding. In addition to these repeats ETs have a so called GIY-YIG domain in common. These domains are found in bacterial repair proteins such as UVRC and are known to function in the insertion of single strand cuts in DNA. Remarkably, functionally analogous domains have been recruited by plant specific regulators of DNA methylation such as DME and ROS1 leading to the speculation that ET might be also involved in the regulation of the methylation status of genomic DNA. But, despite the plethora of these observations (Wohlfarth, 1996; Raventós et al., 1998; Ellerström et al., 2005, Ivanov et al. 2008), the molecular mode of action and the principal functional importance of ET factors in plant differentiation and development remain obscure. For a more detailed functional analysis of the ET gene family, a genetic approach has been applied. Thus, T-DNA insertion mutants of Arabidopsis have been genotyped followed by detailed phenotypic and molecular analysis including deep RNA sequencing and whole genome methylation studies. Together, these data provide strong indications that ETs are novel epigenetic regulators of the DNA methylation status in plant genomes

leading to pleiotropic developmental effects which include gametophytic distortions, homoeotic transformations of flower organs, affected endosperm differentiation as well as precocious ectopic germination.

1. Verification of genotype and phenotype correlation

Genetic studies often rely on the analysis of phenotypic consequences of certain genotypic mutant alleles. Prior to a more detailed discussion of the phenotypic effects of the mutants, it is required to demonstrate the relation between genotypes and phenotypes. Principal possibilities include the analysis of multiple alleles and/or the phenotypic complementation of the genotype. The first option has been applied for *ET2* by the analysis of the mutant alleles, *et2-1* and *et2-3*, which obviously exhibit similar phenotypes. The second approach has been used for *ET1*. In this case, a 4.5 kb genomic wild type fragment has been PCR amplified, re-sequenced and transformed into the *et1-1* mutant. The phenotypic characterisation of the transformed lines indicates the partial correction of the mutant phenotypes (Fig. 34).

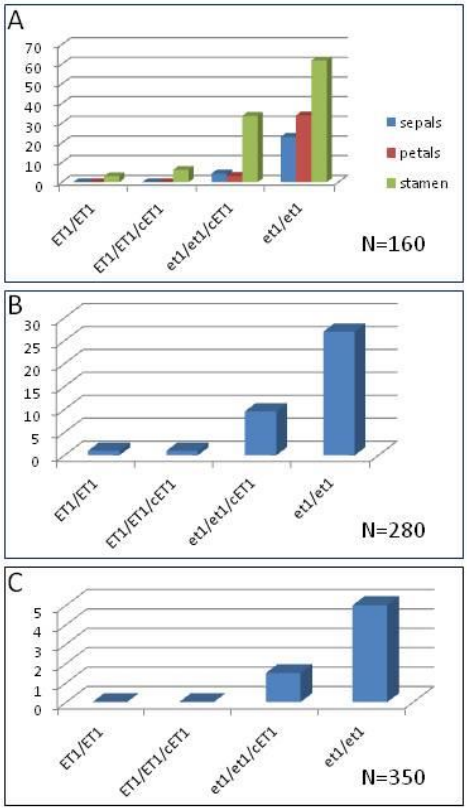

Fig. 34 Partial phenotypic complementation of the *et1-1* mutant after transformation of a 4.5 kb genomic wild type fragment. Partial correction of effects on distorted flower organ numbers (A), partial complementation of the described effects on enlarged nucleoli in endosperm nuclei (B) and partial correction of the effects on non-fused polar nuclei in the gametophyte (C).

2. Gametophytic cell differentiation

The typical mature gametophyte of Arabidopsis consists of two synergids, the egg cell and the homodiploid central cell with the two nuclei fused together. The most obvious gametophytic phenotype observed in *et* mutants concerns the distorted fusion of the two polar nuclei. A relatively large numbers of mutations described previously share this phenotype. Thus, defects in genes encoding BiP1 and BiP2 -molecular chaperons in

the ER- exhibit non-fused polar nuclei (Maruyama et al. 2010). Polar nuclei fusion also fails in mutants with miss-specification of gametophytic cells fate, such as *clotho* and *lachesis* (Gross-Hardt et al., 2007). Remarkably, non-fused polar nuclei are also observed in mutants affecting mitochondrial genes such as *nuclear fusion defective (nfd1)* (Portereiko et al., 2006), *gametophytic factor 2 (gfa2)* (Christensen et al., 2002) and *syco1* (Kägi et al., 2010). Finally, there are two MADS-domain proteins, AGAMOUS-LIKE80 (AGL80) and AGAMOUS-LIKE61 (AGL61) which are expressed in the central cell and exhibit similar defects (Bemer et al., 2008, Steffen et al., 2008). Failure in polar nuclei fusion might be caused by mis-differentiation of gametophytic cell types, such as synergids, egg cell and central cell. The gametophytic cell differentiation is thought to be triggered by an auxin gradient, although this view was currently challenged (see below). Nevertheless, genes involved in auxin synthesis and signalling (YUCCA10 is a paternally expressed gene (PEG) and AUXIN RESPONSE FACTOR 17 is a maternally expressed gene (MEG) have been identified as targets of imprinting processes (Köhler et al., 2012). This invites the speculation that ETs mediate changes in the methylation status of such genes and could contribute to the phenotype of non-fused polar nuclei. The failure of polar nuclei fusion in *et* mutants also resembles the phenotype of mutants with defects in the gene encoding the glucose 6-phosphate translocator (GPT1). Thus, *gpt1* mutation also affects the fusion of the polar nuclei during embryo sac development (Niewiadomski et al., 2005). Interestingly, a *GPT* gene was preliminarily found to be down-regulated in *et* mutants (not shown). Currently it is not known whether the above mentioned genes anyhow interact with ET-factors – all questions which obviously need further investigations.

3. Homoeotic transformation of flower organs

A rather remarkable phenotype of *et* mutants includes the homoeotic transformation of anthers into carpel-like structures including the occurrence of ectopic ovules and stigmata. Following the predictions of the ABC model (Theißen et al., 2001) described in the introduction, this

phenotypic transition of stamens into carpels is best explained by the assumption that ET-factors would either inhibit the activity of the B-function AP3/PI or enhance the C-function AG (see Fig. 35).

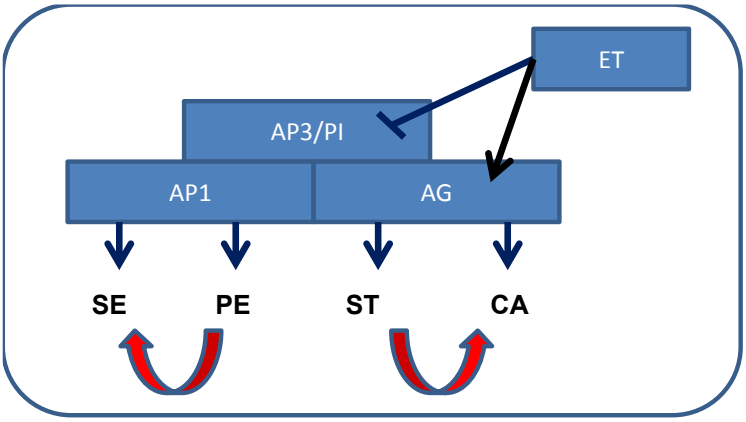

Fig. 35 Schematic ABC model with proposed ET function to either inhibit the B-function (AP3/PI) or to enhance the C-function (AG).

Preliminary measurements of AP3/PI-transcript levels in *et* mutants indeed show a reduced level of the transcript. Further support for this view comes from the observation that the phenotype of the *ap3-3* mutant closely resembles that of *et* mutants (Krizek and Meyerowitz, 1996). Finally, the observed phenotypic differences in flower organs in *et* mutants are also similar to flower phenotypes observed in homozygous mutants (*dme-1*) of the *DEMETER* gene (Choi et al., 2002).

It was surprising to observe that the ectopic ovules on the stamen derived carpel-like structures contain a rather well developed embryo sac including two synergids, an egg cell and a fused homodiploid polar nuclei. This shows that the ovule differentiation -once its initiation is triggered- follows a mainly autonomous developmental pathway including the sporophyte-gametophyte transition and gametophyte formation. However, a closer inspection reveals that the normal polarity along the embryo sac is partially distorted with the usual order of synergids, egg cell and polar nucleus being scrambled (Fig. 18). The differentiation of gametophytic cell fates was proposed to be determined by an auxin gradient within the embryo sac. High auxin levels would control the differentiation of synergids, a lower

auxin concentration would then trigger the egg cell differentiation followed by even lower auxin levels and the occurrence of the polar nuclei (Pagnussat et al., 2009). But this view has recently been challenged (Lituiev et al. 2013) and re-challenged (Panoli et al., 2015). Although the existence and relevance of an autonomous gametophytic auxin gradient is currently not clear, most authors agree on the importance of a phyothormone-mediated control of gametophytic cell fate. Since genes involved in auxin synthesis and signaling are targets of imprinting (Köhler et al., 2012), one might assume that ET-mediated methylation processes indirectly influence the gametophytic cell fate and function (see also above).

4. Endosperm differentiation

Another aspect of the *et* mutant phenotype concerns the endosperm differentiation (Fig. 22). The primary endosperm nucleus divides and forms a syncytium of free nuclei. At early heart stage the endosperm starts to become cellular before it degenerates during the cotyledon stages with the aleuron layer being a remnant of the endosperm in the mature seed. At a globular stage the nuclei of the early endosperm of *et* mutants contain unusually large nucleoli, known to be the site of ribosomal RNA synthesis. Enlarged nucleoli are usually connected to an increased synthetic cellular activity (Shaw and Doonan, 2005; Baker, 2013). The observation invites the speculation that in *et* mutants the early endosperm cells are precociously activated, perhaps for cell proliferation or a precocious initiation of early storage compounds synthesis or even for their early mobilisation.

5. Precocious seed germination

Finally, *et* mutants show a precocious germination phenotype both *in vivo* and *in vitro* (Fig. 23, Fig. 24). The germination starts already within the silique, a phenotype which resembles the phenomenon of pre-harvest sprouting in cereals (Gao and Ayele, 2014). Usually germination is initiated when the seedling penetrates the seed coat with the radicula ahead. In

contrast, the precocious germination of *et* mutants occurs with the still green cotyledons first. The cotyledon-first- phenotype is also retained when isolated seeds are allowed to germinate *in vitro*. Similar phenotypes of germination with the cotyledons ahead have been described for ABA-immunomodulated tobacco seeds (Phillips et al., 1997). Seeds of plants which express anti-ABA scFV antibodies underwent a form of precocious germination when removed from the capsules and germinate by the emergence of the cotyledons first. Further, species of the genus *Aethionema* exhibit a seed dimorphism possibly controlled by an epigenetic pathway. One type germinates normally with the radicula coming out first and another type germinates-similar to the here described Arabidopsis *et* mutants- with the cotyledons ahead (G. Leubner, University of London, personal communication). Again one might speculate that ETs are involved in epigenetic pathway involved in the regulation of the ratio between the phytohormones GA and ABA. This view fits with the well established knowledge that the ratio of both hormones is crucial for the maintenance of dormancy *versus* initiation of seed germination, with ABA favouring dormancy and GA triggering germination. Assuming that ET acts as an epigenetic repressor of GA activity (see above), indeed one would predict an early germination phenotype for *et* mutants as observed.

6. RNA deep sequencing

Based on the assumption that ETs act as regulators of genomic DNA methylation with further influence on gene transcription, an RNA deep sequencing approach has been performed. In a comparison of transcripts between wild type and *et1-1* and *et2-3*, in total 288 differentially expressed genes could be identified and their expression was mapped on the genome (Fig. 28). About 2/3 of these differentially expressed genes are down-regulated in the mutants. This corresponds well with the initial hypothesis that ET factors might act as de-methylators. It is broadly accepted that hypo-methylated promoters lead to stronger expressed genes, whereas hyper-methylated promoters lead to suppressed gene expression. Thus, a missing or reduced de-methylation in the *et* mutants would lead to hyper-

methylation and suppression of gene transcription. Following this interpretation, the 1/3 up-regulated genes might be seen as secondary effects in a dynamically changing methylation and transcription pattern.

As shown in Tab. 2 up- and down-regulated genes strongly overlap between both mutants. With only one exception -a functionally unknown gene encoding a DUF1005 domain- all genes found to be differentially expressed in *et2-3* are also found as differentially expressed genes in the *et1-1* mutant. Additional genes with differential expression have been found in the *et1-1* mutant. Among the genes most down-regulated in the *et1-1* mutant are genes encoding compounds typical for storage processes in the late embryo maturation phase such as lipid transfer proteins and 2S albumin storage proteins. The most conspicuous difference in the transcriptome of *et1-1* and *et2-3* mutants (see Tab. 2) concerns the expression of transposable elements (AT1G40101, AT1G38360, AT5G33434) and functionally unknown pseudogenes also annotated by TAIR as transposable elements (AT4G06718, AT1G41680) strongly up-regulated in the *et1-1* mutant with log2 ratios of about 10 and 8, respectively. The three transposable elements are retrotransposon which belong to the Athila subfamily of the Ty3/Gypsy family (for a review see Wicker et al., 2007). They represent over 2.7% of the total Arabidopsis genome and are a major building block of the centromere (Slotkin, 2010). Athila retrotransposons share a large internal region of up to 10.5 kb flanked by two about 1.8 kb long terminal repeats (LTR). The internal region encodes the capsid structural protein *gag* and the protein *pol* which carries the protease, reverse transcriptase and integrase domains essential for element duplication. The silencing of Athila retrotransposons has come to the forefront of Arabidopsis small RNA regulation, the control of centromere core as well as potentially playing a role in speciation (Slotkin et al., 2009). Taken together, the data demonstrate that ET1 and ET2 have both overlapping as well as gene-specific functions with ET1 being specific for the regulation, most likely the suppression, of retrotransposons especially of the Athila subfamily.

7. Genome wide methylation

In collaboration with Dr. C. Becker at MPI Tübingen it became possible to compare the genome-wide methylation status between wild type and both mutants *et1-1* and *et2-3*. Using the iPlant visualisation tool in total 527 differentially methylated regions could be detected. As shown in Fig. 29 these regions are rather equally distributed along the five chromosomes. At first, the data show a remarkable reproducibility for detected genomic regions in two completely independent experimental replicates. Despite of the obviously high reliability of the described results, it is currently difficult to interprete the results in detail. The methylation context CG is the one which is mostly affected followed by the other symmetric context CHG whereas the CHH context is very rarely affected. Up- and down methylation in mutants occurs with similar frequency. Concerning the methylated *versus* non-methylated status of different genomic regions compared in wild type, *et1-1* and *et2-3* mutants all combinations have been detected (Tab. 3) indicating a high complexity of the underlying regulatory mechanisms. As revealed by the transcriptome analysis, ET1 and ET2 have both common and gene specific target regions in terms of methylation.

A putatively interesting special case deserving a more detailed discussion concerns the differentially methylated region diffM133 (Fig. 31), which is non-methylated in both mutants. The differentially methylated region precisely overlaps with the position of a HELITRON transposable element. The transposon is located in the putative promoter region of a gene encoding an E3-SCF protein also known as Arabidopsis SKP1 Homologue (ASK1). Remarkably, this gene product -in cooperation with UFO and LEAFY- is involved in flower development where it is required to regulated the B function (Zhao et al., 2001). Thus, this invites the speculation that this differentially methylated region contributes to the above described flower phenotypes of *et* mutants.

Another selected example concerns the differentially methylated region diffM105 (Fig. 32) containing a β-glucosidase encoding gene. This gene is methylated in wild type and *et2-3* but not in the *et1-1* mutant further indicating that ET1 and ET2 have specific target genes. In preliminary

results, an adjacent gene encoding a glucose-6-phosphate translocator was found to be differentially expressed (shown by qRT-PCR, but not detected by deep sequencing).

8. Correlation between transcriptome and methylom

There is broad agreement that transcriptional activation correlates with a hypomethylated status of a certain genomic region. Clearly, this view is not easily compatible with the genome-wide analysis both of gene transcripts and the methylation status of the corresponding genomic region and probably requires the detailed analysis of in a gene-specific manner.

Previous work has shown that the gene encoding a FAD-binding Berberine protein (AT2G26400) is significantly hyper-methylated in the *et1-1* mutant (M. Kuhlmann, IPK, pers. comm.). This observation is consistent with a down-regulation of the corresponding gene transcript in the *et1-1* mutant by a log2 ratio of -4.34 (Tab. 2).

Moreover, preliminary results indicate that the DEMETER transcript level is strongly increased in the *et1-1* mutant, but not in the *et2-3* mutant. This invites the speculation that ET1 is involved in the regulation of *DME* expression, possibly related to transposon inactivation, but this requires further analysis.

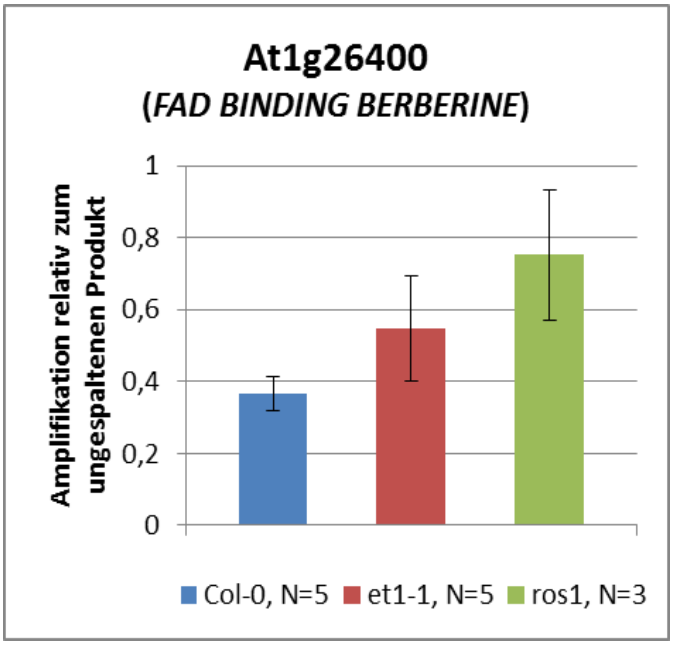

Fig. 36 Increased transcript level of DEMETER in the *et1-1* and *ros1* mutant.

9. Transposons and cell specification

Recently, it has been proposed that companion cells such as the vegetative cell of the male gametophyte and the central cell of the female gametophyte undergo active de-methylation followed by the activation of transposable elements. Subsequently, transposon-derived small interfering RNAs can move to the gametes, the sperm cell and the egg cell, to reinforce silencing of transposons in gametes, zygote and the derived next generation. This has been indicated as a fundamental biological process of reproductive biology in plant, and likely in animals as well (Slotkin et al., 2009; Calarco and Martienssen, 2011). In the female gametophyte DEMETER is required for this active demethylation process (Choi et al, 2002); a corresponding gene product in the male gametophyte has been predicted, but is still unknown. This might lead to the proposal that ETs acting as regulators of DNA methylation might be involved in this basic reproductive pathway.

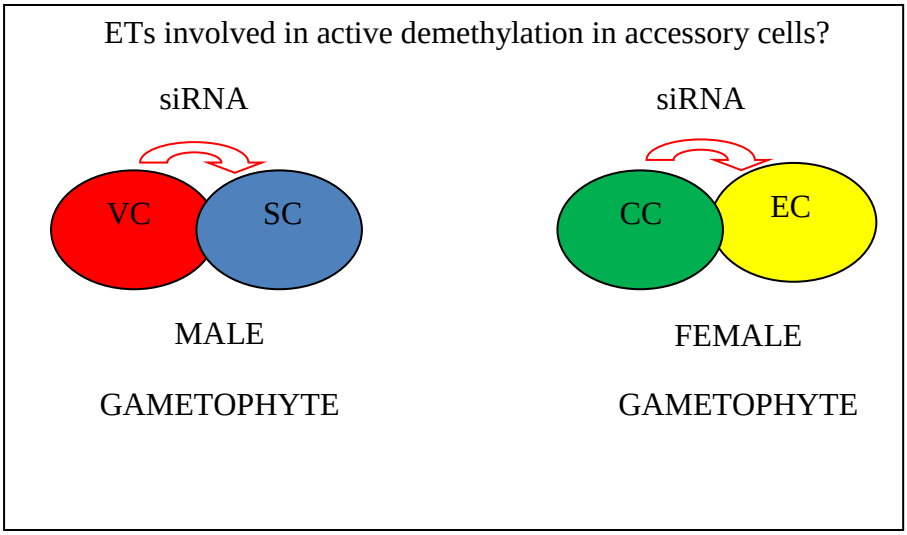

Fig. 37 Hypothetical function of ETs in accessory reproductive cells in plant gametophytes. VC-vegetative cell, SC-sperm cell, CC-central cell, EC-egg cell.

10. Conclusion and outlook

In summary, the described data assign the EFFECTORS OF TRANSCRIPTION (ET) an essential function as epigenetic regulators. The data are consistent with the view that ETs act as mainly DNA de-methylating factors to control the dynamics of the genomic methylation status to trigger reproductive cell differentiation and development.

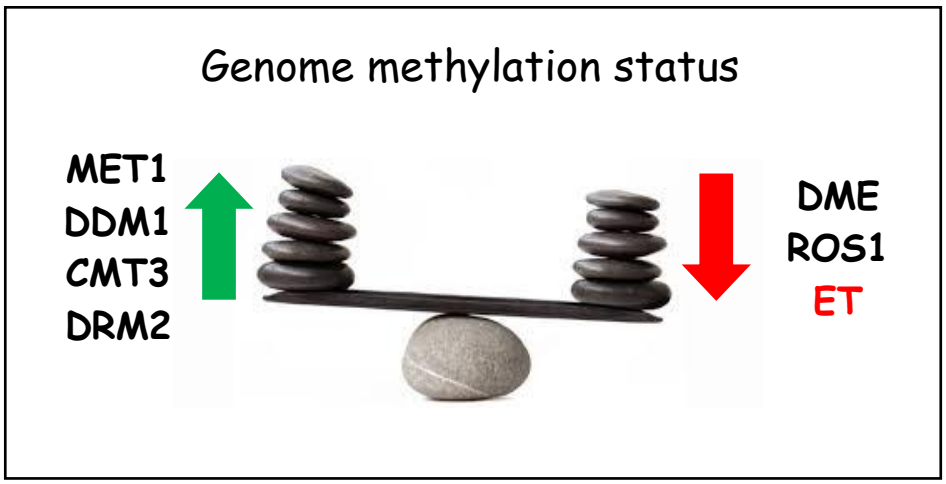

Fig. 38 Scheme to illustrate the current ideas concerning the function of ET factors. Available data are consistent with a function related to DNA de-methylation.

Further work is required to understand the function of ETs in more detail. Planned experiments include the further analysis of new double mutants, the confirmation of genomic data (transcripts, methylation) for selected genes, analysis of the molecular mechanism, bioinformatic and experimental approaches to identify putative ET binding sites, inclusion of the truncated, but transcribed *ET3* gene and extension of the analysis to the functional homologue HRT in the crop plants barley and wheat.

Summary

EFFECTOR OF TRANSCRIPTION (ET) factors share a variable number of characteristic, highly conserved cystein-histidine containing ET repeats involved in zinc- and DNA-binding. Ectopic expression of ET factor genes resulted in severe growth distortions including dwarf growth, late flowering, reduced germination rates, strong anthocyanin accumulation, reduced lignification etc.. Together, these observations indicated a putative function as repressor of transcription of gibberellic acid (GA)-regulated genes including KNAT genes involved in cell differentiation. A barley homologue, HORDEUM REPRESSOR OF TRANSCRIPTION (HRT), was shown to repress the expression of amylase genes in barley aleuron cells. ET/HRT factors are exclusively found in plant genomes including the phylogenetically old species such as the moss *Physcomitrella*. In addition, to the ET repeats ETs have a so called GIY-YIG domain in common. The latter is also found in bacterial repair proteins such as UVRC and is known to function by the insertion of single stand cuts in DNA. Functionally analogous domains have been recruited by plant specific regulators of DNA methylation such as DEMETER and ROS1 leading to the suggestion that ETs might be involved in the regulation of the methylation status of genomic DNA. Despite the plethora of previous observations, the molecular mode of action and the principal functional importance of ET factors in plant differentiation and development remained obscure.

Here I describe a more detailed functional analysis of the ET gene family using a genetic approach. T-DNA insertion mutants of Arabidopsis have been isolated and genotyped. A thorough phenotypic description of single and double mutants reveals pleiotropic developmental effects. This includes for instance the failed fusion of the two polar nuclei as prerequisite for double fertilisation and endosperm development. The *et* mutants exhibit a conspicuous homoeotic transformation of flower organs with anthers transformed into carpels containing rather well developed stigmata and ovules with nearly normal embryo sac formation. The endosperm nuclei of the mutants exhibit unusually large nucleoli probably indicating a high synthetic activity. Finally, the mutants germinate precociously when still

attached to the mother plant with the cotyledons and not the root tip penetrating the seed coat first. Searching for putative target genes a comparative molecular analysis including deep RNA sequencing and genome-wide methylation studies have been performed. Together, the results provide strong evidence for the conclusion that ETs are novel epigenetic regulators of reproductive processes and act *via* the regulation of the DNA methylation status of plant genomes.

Zusammenfassung

EFFECTOR OF TRANSCRIPTION (ET)-Faktoren haben eine unterschiedliche Zahl charakteristischer, Zink- und DNA-bindender ET Domänen gemeinsam. Überexpression der ET-Gene führt zu Wachstumsstörungen wie Zwergwachstum, später Blühzeitpunkt, verminderte Keimungsrate, Akkumulation von Anthocyanin, reduzierter Lignifizierung etc.. Zusammen weisen diese Beobachtungen auf eine mögliche Funktion als Repressoren Gibberellin-regulierter Geneexpression, insbesondere der an der Zelldifferenzierung beteiligten KNAT-Gene, hin. Ein homologer Faktor in Gerste, HORDEUM REPRESSOR OF TRANSCRIPTION (HRT), wirkt als ein Repression der Expression von Amlyase-Genen in Aleuronzellen. ET/HRT-Faktoren werden nur in pflanzlichen Genomen gefunden, einschließlich im Genom des phylogenetisch alten Moos *Physcomitrella*. Neben den ET-Domänen besitzen alle ET-Faktoren eine GIY-YIG Domäne. Diese wird auch in bakteriellen, an der DNA-Reparatur beteiligten Proteinen wie z. B. UVRC gefunden. Die Domäne fügt dabei Einzelstrangschnitte in die DNA ein. Funktionell analoge Domänen wurden von anderen pflanzenspezifischen Regulatoren der DNA-Methylierung, wie DEMETER und ROS1 rekrutiert. Dies führte zu der Vermutung, dass auch ET-Faktoren an der Regulation des Methylierungsstatus der genomischen DNA beteiligt sind. Trotz der umfangreichen früheren Beobachtungen sind die molekulare Funktion und die funktionelle Bedeutung der ET-Faktoren weitgehend unbekannt.

Hier beschreibe ich eine detaillierte funktionelle Analyse der ET-Genfamilie mit Hilfe eines genetischen Ansatzes. In Arabidopsis wurden T-DNA-Insertionsmutanten isoliert und genetisch charakterisiert. Phänotypische Analyse von Einzel- und Doppelmutanten offenbart eine ganze Reihe pleiotroper Veränderungen. So ist z. B. in den Mutanten die Fusion der beiden Polarkerne - eine Voraussetzung für die doppelte Befruchtung und die Differenzierung des Endosperms- gestört. Auch zeigen die *et*-Mutanten eine auffällige homöotische Transformation der Blütenorgane, wobei Antheren in Karpelle transformiert sind. Letztere bilden gut entwickelte Stigmata sowie Ovulen mit nahezu normaler Embryosack-Entwicklung aus.

Die Endospermkerne der Mutanten besitzen einen sehr großen Nucleolus, was auf eine hohe synthetische Aktivität hindeutet. Schließlich keimen die unreifen Samen der Mutanten schon auf der Mutterpflanze, wobei die Samenschale zuerst von den Kotyledonen und nicht von der Wurzelspitze durchstoßen wird. Auf der Suche nach potentiellen Zielgenen wurde eine vergleichende molekulare Analyse durch Hochdurchsatz-RNA-Sequenzierung sowie eine genomweite Methylierungsanalyse durchgeführt. Zusammen führen die Ergebnisse zur der Schlussfolgerung, dass ET-Faktoren den DNA-Methylierungsstatus des Genoms kontrollieren und dadurch als neue, epigenetische Regulatoren an der Kontrolle reproduktiver Prozesse beteiligt sind.

References

Alexander, M. P. (1969). Differential staining of aborted and nonaborted pollen. *Stain Technol.* 44: 117-122.

Aravind, L., D. R. Walker and E. V. Koonin (1999). Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res.* 27: 1223-1242.

Baker, N. E. (2013). Developmental regulation of the nucleolus size during *Drosophila* eye differentiation. *PLoS One* 8, e58266.

Becker, C., J. Hagmann, J. Müller, D. Koenig, O. Stegle, K. Borgwardt and D. Weigel (2011). Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* 480: 245-249.

Bemer, M., M. Wolters-Arts, U. Grossniklaus and G. C. Angenent (2008). The MADS domain protein DIANA acts together with AGAMOUS-LIKE80 to specify the central cell in Arabidopsis ovules. *Plant Cell* 20: 2088-2101.

Berger, F. and A. Chaudhury (2009). Parental memories shape seeds. *Trends in Plant Sci.* 14: 550-556.

Berger, F., P. E. Grini and A. Schnittger (2006). Endosperm: An integrator of seed growth and development. *Curr. Opin. Plant Biol.* 9: 664-670.

Bewley, J. D., F. D. Hempel, S. McCormick and P. Zambryski (2000). Reproductive Development. In: Biochemistry and Molecular Biology of Plants. *American Society of Plant Biologists*, pp. 988–1034.

Black, M. (1991). Involvement of ABA in the physiology of the developing and mature seeds. In: Abscisic acid physiology and biochemistry. *Oxford Bios. Scientific Publishers*. 99-124.

Bowman, J. L., D. R. Smyth and E. M. Meyerowitz (1989). Genes directing flower development in Arabidopsis. *Plant Cell* 1: 37–52.

Brown, R. C., B. E. Lemmon, H. Nguyen and O. A. Olsen (1999). Development of endosperm in A*rabidopsis thaliana*. *Sexual Plant Reproduction* 12: 32–42.

Calarco, J. P. and R. A. Martienssen (2011). Genome reprogramming and small interfering RNA in the Arabidopsis germline. *Curr. Opin. Genet. Dev.* 21: 134-139.

Choi, Y., M. Gehring, L. Johnson, M. Hannon, J. J. Harada, R. B. Goldberg, S. E. Jacobsen and R. L. Fischer (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. *Cell* 110: 33-42.

Choi, Y., J. J. Harada, R. B. Goldberg and R. L. Fischer (2004). An invariant aspartic acid in the DNA glycosylase domain of DEMETER is necessary for transcriptional activation of the imprinted MEDEA gene. *Proc. Natl. Acad. Sci.* 101: 7481-7486.

Christensen, C. A., S. W. Gorsich, R. H. Brown, L. G. Jones, J. Brown, J. M. Shaw and G. N. Drews (2002). Mitochondrial GFA2 is required for synergid cell death in Arabidopsis. *Plant Cell* 14: 2215-2232.

Clough, S. J. and A. F. Bent (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* 16: 735-743.

Coen, E. S. and R. Carpenter (1993). The metamorphosis of flowers. *Plant Cell* 5: 1175–1181.

Cokus, S. J., S. Feng, X. Zhang, Z. Chen, B. Merriman, C. D. Haudenschild, S. Pradhan, S. F. Nelson, M. Pellegrini and S. E. Jacobsen (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452: 215-219.

Debeaujon, I. and M. Koornneef (2000). Gibberellin requirement for Arabidopsis seed germination is determined both by testa characteristics and embryonic abscisic acid. *Plant Physiol.* 122: 415-424.

Derbyshire, V., J. C. Kowalski, J. T. Dansereau, C. R. Hauer and M. Belfort (1997). Two-domain structure of the *td* intron-encoded endonuclease I-*Te*vI correlates with the two-domain configuration of the homing site. *J. Mol. Biol.* 265: 494-506.

Edwards, K., C. Johnstone and C. Thompson (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* 19: 1349.

Ellerström, M., W. Reidt, R. Ivanov, J. Tiedemann, M. Melzer, A. Tewes, T. Moritz, H. P. Mock, F. Sitbon, L. Rask and H. Bäumlein (2005). Ectopic expression of EFFECTOR OF TRANSCRIPTION perturbs gibberellin-mediated plant developmental processes. *Plant Mol. Biol.* 59: 663-681.

Focks, N. and C. Benning (1998). Wrinkled1: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.* 118: 91-101.

Friedberg, E. C., G. C. Walker and W. Siede (1995). Nucleotide excision repair in prokaryotes. DNA Repair and Mutagenesis. ASM Press, Washington, DC.

Gao, F. and B. T. Ayele (2014). Functional genomics of seed dormancy in wheat: Advances and prospects. *Frontiers in Plant Science* 5: 1-11.

Gehring, M., W. Reik and S. Henikoff (2009). DNA demethylation by DNA repair. *Trends Genet.* 25: 82–90.

Goldberg, R. B., G. Paiva and R. Yadegari (1994). Plant embryogenesis: Zygote to seed. *Science* 266: 605–614.

Gong, Z., T. Morales-Ruiz, R. R. Ariza, T. Roldán-Arjona, L. David and J. K. Zhu (2002). ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* 111: 803-814.

Gross-Hardt, R., C. Kägi, N. Baumann, J. M. Moore, R. Baskar, W. B. Gagliano, G. Jürgens and U. Grossniklaus (2007). LACHESIS restricts gametic cell fate in the female gametophyte of Arabidopsis. *PLoS Biol.* 5: e47.

Grossniklaus, U. and K. Schneitz (1998). Genetic and molecular control of ovule development and megagametogenesis. *Seminars in Cell & Devl. Biol.* 9: 227-238.

Grossniklaus, U., C. Spillane, D. R. Page and C. Köhler (2001). Genomic imprinting and seed development: Endosperm formation with and without sex. *Curr. Opin. Plant Biol.* 4: 21-27.

Haig, D. (1990). New perspectives on the angiosperm female gametophyte. *Bot. Rev.* 56: 236–274.

Haince, J. F., M. Rouleau and G. G. Poirier (2006). Gene expression needs a break to unwind before carrying on. *Science* 312: 1752-1753.

Harada, J. J. (2001). Role of Arabidopsis LEAFY COTYLEDON genes in seed development. *Plant Physiol.* 158: 405-409.

He, X. J., T. Chen and J. K. Zhu (2011). Regulation and function of DNA methylation in plants and animals. *Cell Research* 21: 442-465.

Heard, E. and R. A. Martienssen (2014). Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* 157: 95-109

Herman, E. M. (1995). Cell and molecular biology of seed oil bodies. In: Seed development and germination. *New York*. Pp: 195-214.

Hofmeister, W. (1851) Entfaltung und Fruchtbildung höherer Kryptogamen (Moose, Farn, Equisetaceen, Rhizocarpeen und Lycopodiaceen) und die

Samenbildung der Coniferen. Vergleichende Untersuchungen der Keimung. Leipzig.

Holdsworth, M. J., L. Bentsink and W. J. Soppe (2008). Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytol.* 179: 33-54.

Honys, D. and D. Twell (2004). Transcriptome analysis of haploid male gametophyte development in Arabidopsis. *Genome Biology* 5: R85.1-R85.13.

Huang, K. H., C. C. Akoh and M. C. Erickson (1994). Enzymatic modification of melon seed oil: Incorporation of eicosapentaenic acid. *Journal of Agricultural and Food Chemistry* 42: 2646-2648.

Huang, B.Q. and S. D. Russell (1992). Female germ unit: Organization, isolation, and function. *Int. Rev. Cytol.* 140: 233–292.

Ivanov, R., J. Tiedemann, A. Czihal, A. Schallau, L. H. Diep, H. P. Mock, C. Bernhard, A. Tewes and H. Bäumlein (2008). EFFECTOR OF TRANSCRIPTION2 is involved in xylem differentiation and includes a functional DNA single strand cutting domain. *Developmental Biology* 313: 93-106.

Ivanov, R. (2005). Molecular characterization of the EFFECTOR OF TRANSCRIPITION (ET) gene family in Arabidopsis and its role in plant development. PhD thesis, Halle, Germany, Martin-Luther-Universität Halle-Wittenberg.

Jaenisch, R. and A. Bird (2003). Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nature Genetics* 33: 245-254.

Ju, B. G., V. V. Lunyak, V. Perissi, I. Garcia-Bassets, D. W. Rose, C. K. Glass and M. G. Rosenfeld (2006). A topoisomerase IIß-mediated dsDNA break required for regulated transcription. *Science* 312: 1798-1802.

Junker, A., A. Hartmann, F. Schreiber and H. Bäumlein (2010). An engineer´s view on seed development. *Trends in Plant Sci.* 15: 303-307.

Kawashima, T. and F. Berger (2014). Epigenetic reprogramming during male gametogenesis. *Nature Reviews Genetics* 15: 613–624.

Kägi, C., N. Baumann, N. Nielsen, Y. D. Stierhof and R. Groß-Hardt (2010). The gametic central cell of Arabidopsis determines the lifespan of adjacent accessory cells. *Proc. Natl. Acad. Sci.* 107: 22350-22355.

Kim, Y. C., M. Nakajima, A. Nakayama and I. Yamaguchi (2005). Contribution of gibberellins to the formation of *Arabidopsis* seed coat through starch degradation. *Plant Cell Physiol.* 46: 1317-1325.

Koornneef, M., K. Leon-Kloosterziel, S. Schwartz and J. Zeevaart (1998). The genetic and molecular dissection of abscisic acid biosynthesis and signal transduction in Arabidopsis. *Plant Physiol. Biochem.* 36: 83-89.

Kowalski, J. C., M. Belfort, M. A. Stapleton, M. Holpert, J. T. Dansereau, S. Pietrokovski, S. M. Baxter and V. Derbyshire (1999). Configuration of the catalytic GIY-YIG domain of intron endonuclease I-*Tev*I: Coincidence of computational and molecular findings. *Nucleic Acids Res.* 27: 2115-2125.

Köhler, C., P. Wolff and C. Spillane (2012). Epigenetic mechanisms underlying genomic imprinting in plants. *Annu. Rev. Plant Biol.* 63: 331-352.

Krizek, B. A. and E. M. Meyerowitz (1996). Mapping the protein regions responsible for the functional specificities of the Arabidopsis MADS domain organ-identity proteins. *Proc. Natl. Acad. Sci.* 93: 4063-4070.

Krysan, P. J., J. C. Young and M. R. Sussman (1999). T-DNA as an insertional mutagen in Arabidopsis. *Plant Cell* 11: 2283-2290.

Latchman, D. S. (2008). Eukaryotic transcription factors. Academic Press, Fifth Edition.

Laux, T. and G. Jurgens (1997). Embryogenesis: A New Start in Life. *Plant Cell* 9: 989–1000.

Law J. A. and S. E. Jacobsen (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11: 204–220.

Lin, J. J. and A. Sancar (1992). Active site of (A)BC excinuclease. I. Evidence for 5' incision by UvrC through a catalytic site involving Asp399, Asp438, Asp466, and His538 residues. *J. Biol. Chem.* 267: 17688-17692.

Lituiev, D. S., G. K. Krohn, B. Müller, D. Jackson, B. Hellriegel, T. Dresselhaus and U. Grossniklaus (2013). Theoretical and experimental evidence indicates that there is no detectable auxin gradient in the angiosperm female gametophyte. *Development* 140: 4544-4553.

Lister, R., M. Pelizzola , R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren and J. R. Ecker (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462: 315-322.

Lotan, T., M. Ohto, K. M. Yee, M. A. L. West, R. Lo, R. W. Kwong, K. Yamagishi, R. L. Fischer, R. B. Goldberg and J. J. Harada (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93: 1195-1205.

Magnani, E., K. Sjölander and S. Hake (2004). From Endonucleases to transcription factors: Evolution of the AP2 DNA binding domain in plants. *Plant Cell* 16: 2265–2277.

Maheshwari, P. (1950). An Introduction to the Embryology of Angiosperms. (New York: McGraw-Hill).

Mansfield, S. G. and L. G. Briarty (1992). Cotyledon cell development in *Arabidopsis thaliana* during reserve deposition. *Canadian Journal of Botany-Revue Canadienne De Botanique* 70: 151-164.

Matzke, M. A. and R. A. Mosher (2007). RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15: 394-408.

Maruyama, D., T. Endo and S. Nishikawa (2010). BiP-mediated polar nuclei fusion is essential for the regulation of endosperm nuclei proliferation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* 107: 1684-1689.

Moolenaar, G. F., R. S. Uiterkamp, D. A. Zwijnenburg and N. Goosen (1998). The C-terminal region of the *Escherichia coli* UvrC protein, which is homologous to the C-terminal region of the human ERCC1 protein, is involved in DNA binding and 5´-incision. *Nucleic Acids Res.* 26: 462-468.

Morales-Ruiz, T., A. P. Ortega-Galisteo, M. I. Ponferrada-Marín,, M. I. Martínez-Macías, R. R. Ariza and T. Roldán-Arjona (2006). DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc. Natl. Acad. Sci.* 103: 6853-6858.

Murphy, D. J. (1993). Structure, function and biogenesis of storage lipid bodies and oleosins in plants. *Prog Lipid Res.* 32: 247-280.

Müntz, K. (1998). Deposition of storage proteins. *Plant Molecular Biology* 38: 77-99.

Niewiadomski, P., S. Knappe, S. Geimer, K. Fischer, B. Schulz, U. S. Unte, M. G. Rosso, P. Ache, U. Flügge and A. Schneider (2005). The Arabidopsis plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. *Plant Cell* 17: 760–775.

Olszewski, N., T. P. Sun and F. Gubler (2002). Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell* 14: S61–S80.

O'Malley, R.C., J. M. Alonso, C. J. Kim, T. J. Leisse and J. R. Ecker (2007). An adapter ligation-mediated PCR method for high-throughput mapping of T-DNA inserts in the Arabidopsis genome. *Nat. Protoc.* 2: 2910-2917.

Ossowski, S., K. Schneeberger, R. M. Clark, C. Lanz, N. Warthmann, D. Ossowski, S., K. Schneeberger, R. M. Clark, C. Lanz, N. Warthmann and D. Weigel (2008). Sequencing of natural strains of *Arabidopsis thaliana* with short reads. *Genome Res.* 18: 2024-2033.

Park, S. K., R. Howden and D.Twell (1998). The *Arabidopsis thaliana* gametophytic mutation *gemini pollen1* disrupts microspore polarity, division asymmetry and pollen cell fate. *Development* 125: 3789-3799.

Pagnussat, G. C., M. Alandete-Saez, J. L. Bowman and V. Sundaresan (2009). Auxin-dependent patterning and gamete specification in the Arabidopsis female gametophyte. *Science* 324: 1684-1689.

Panoli, A., M. V. Martin, M. Alandete-Saez, M. Simon, C. Neff, R. Swarup, A. Bellido, L. Yuan, P. C. Pagnussat and V. Sundaresan (2015). Auxin import and local auxin biosynthesis are required for mitotic divisions, cell expansion and cell specification during female gametophyte development in *Arabidopsis thaliana*. *PLoS One 10*: 1-23.

Pelaz, S., G. S. Ditta, E. Baumann, E. Wisman and M. F. Yanofsky (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 405: 200–203.

Phillips, J., O. Artsaenko, U. Fiedler, C. Horstmann, H. P. Mock, K. Müntz and U. Conrad (1997). Seed-specific immunomodulation of abscisic acid activity induces a developmental switch. *EMBO J.* 16: 4489-4496.

Portereiko, M. F., L. Sandaklie-Nikolova, A. Lloyd, C. A. Dever, D. Otsuga and G. N. Drews (2006). NUCLEAR FUSION DEFFECTIVE1 encodes the Arabidopsis RPL21 M protein and is required for karyogamy during female gametophyte development and fertilization. *Plant Phys.* 141: 957-965.

Raventós, D., K. Skriver, M. Schlein, K. Karnahl, S. W. Rogers, J. C. Rogers and J. Mundy (1998): HRT, a novel zinc finger, transcriptional repressor from barley. *J. Biol. Chem.* 273: 23313-23320.

Raz, V., J. H. Bergervoet and M. Koornneef (2001). Sequential steps for developmental arrest in Arabidopsis seeds. *Development* 128: 243-252.

Riechmann, J. L. and O. J. Ratcliffe (2000). A genomic perspective on plant transcription factors. *Curr. Opinion Plant Biol.* 3: 423-434.

Sabelli, P. A. and B. A. Larkins (2009). The contribution of cell cycle regulation to endosperm development. *Sex. Plant Reprod.* 22: 207-219.

Sambrook, J. and D. W. Russell (2001). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Schneeberger, K., J. Hagmann, S. Ossowski, N. Warthmann, S. Gesing, O. Kohlbacher and D. Weigel (2009). Simultaneous alignment of short reads against multiple genomes. *Genome Biol.* 10: R98.1-R98.12.

Schmitz, R. J., M. D. Schultz, M. A. Urich, J. R. Nery, M. Pelizzola, O. Libiger, A. Alix, R. B. McCosh, H. Chen, N. J. Schork and J. R. Ecker (2013). Patterns of population epigenomic diversity. *Nature* 495: 193–198.

Shaw, P. and J. Doonan (2005). The nucleolus. Playing by different rules?. *Cell cycle* 4:102-105.

Slotkin, R. K. (2010). The epigenetic control of Athila family of retrotransposon in Arabidopsis. *Epigenetics* 5: 483-490.

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

Soltis, D. E., A. S. Chanderbali, S. Kim, M. Buzgo and P. S. Soltis (2007). The ABC Model and its applicability to basal angiosperms. *Annals of Botany* 100: 155–163.

Steffen, J. G., I. H. Kang, M. F. Portereiko , A. Lloyd and G. N. Drews (2008). AGL61 interacts with AGL80 and is required for central cell development in Arabidopsis. *Plant Physiol.* 148: 259-268.

Stoddard, B. L. (2005). Homing endonuclease structure and function. *Q. Rev. Biophys.* 38: 49-95.

Sussman, M. R., R. M. Amasino, J. C. Young, P. J. Krysan and S. Austin-Phillips (2000). The Arabidopsis Knockout Facility at the University of Wisconsin–Madison. *Plant Physiol.* 124: 1465-1467.

Theissen, G. and H. Saedler (2001). Floral quartets. *Nature* 409: 469–471.

Theißen, G., T. Münster, L. U. Wingen, W. Faigl, S. Werth and H. Saedler (2001). Characterization of three GLOBOSA-like MADS-box genes from maize: Evidence for ancient paralogy in one class of floral homeotic B-function genes of grasses**.** *Gene* 262: 1-13.

Van, R. P., J. Meehan, C. Kowalski, M. Belfort and V. Derbyshire (2002). Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-*Tev*I. *Nat. Struct. Biol.* 9: 806-811.

Verhoeven, E. E., M. V. Kesteren, G. F. Moolenaar, R. Visse and N. Goosen (2000). Catalytic sites for 3' and 5' incision of *Escherichia coli* nucleotide excision repair are both located in UvrC. *The Journal of Biological Chemistry* 275: 5120–5123.

Viswanathan, C. and Z. Jian-Kang (2009). RNA-directed DNA methylation and demethylation in plants. *Sci. China in Life Sci.* 52: 331–343.

Weber, H., L. Borisjuk and U. Wobus (2005). Molecular physiology of legume seed development. *Ann. Rev. Plant Biol.* 56: 253-279.

Weigel, D. and E. M. Meyerowitz (1994). The ABCs of floral homeotic genes. *Cell* 78: 203–209.

Weigel, D. and J. Glazebrook (2002). Arabidopsis: *A Laboratory Manual*. New York, Cold Spring Harbor Laboratory Press.

Weigel, D. and V. Colot (2012). Epialleles in plant evolution. *Genome Biology* 13: 1-6.

West, M. and J. J. Harada (1993). Embryogenesis in higher plants: An overview. *Plant Cell* 5: 1361–1369.

Western, T. L., D. J. Skinner and G. W. Haughn (2000). Differentiation of mucilage secretory cells of the Arabidopsis seed coat. *Plant Physiol.* 122: 345-356.

White, C. N. and C. J. Rivin (2000). Gibberellins and seed development in maize. II. Gibberellin synthesis inhibition enhances abscisic acid signaling in cultured embryos. *Plant Physiol.* 122: 1089-1097.

Wicker, T., F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel and A. H. Schulman (2007). An unified classification system for eukaryotic transposable elements. *Nature Rev. Genet.* 8: 973-982.

Wierzbicki, A. T., J. R. Haag and C. S. Pikaard (2008). Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 135: 635–648.

Willemse, M. T. M. and J. L. van Went (1984). The female gametophyte. *In: Embryology of Angiosperms*. (Berlin: Springer-Verlag), pp. 159–196.

Wohlfarth, T. (1996). Regulation samenspezifischer Gene der Ackerbohne *Vicia faba*: Analyse von cis-Elementen sowie Isolierung und Charakterisierung von Transkriptionsfaktoren. Dissertation, Martin-Luther-Universität Halle-Wittenberg.

Xiao, W., M. Gehring, Y. Choi, L. Margossian, H. Pu, J. J. Harada, R. B. Goldberg, R. I. Pennell and R. L. Fischer (2003). Imprinting of the MEA polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. *Developmental Cell* 5: 891-901.

Yadegari, R. and G. N. Drews (2004). Female gametophyte development. *Plant Cell* 16: S133-S141.

Zhao, D., Q. Yu , M. Chen and H. Ma (2001). The ASK1 gene regulates B function gene expression in cooperation with UFO and LEAFY in Arabidopsis. *Development* 128: 2735-2746.

Zhao, Y., A. K. Hull, N. R. Gupta, K. A. Goss, J. Alonso, J. R. Ecker, J. Normanly, J. Chory and J. L. Celenza (2002). Trp-dependent auxin biosynthesis in Arabidopsis: Involvement of cytochrome P450s CYP79B2 and CYP79B3. *Genes Dev.* 16: 3100-3112.

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor Dr. habil. Helmut Bäumlein, Head of Gene Regulation group, Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben for giving me the opportunity to pursue my PhD-work in his research group. He taught me how to question thoughts and encouraged me to express my ideas. I am greatly indebted to him for his encouragement and guidance throughout my graduate study.

I would like to gratefully and sincerely thank to my mentor Prof. Dr. Udo Conrad, Head of Phytoantibody group, IPK, Gatersleben for his scientific advice and many considerable suggestions and discussions along this research.

I am very thankful to Dr. Markus Kuhlmann and the members of his group for their help and support, especially for providing access to the equipment in their lab.

I am also grateful to all my colleagues in Gene Regulation Group for various support, Andreas Czihal, Sabine Skiebe, Annett Busching for excellent technical assistance, Elke Liemann for taking care of plants. Dr. Francesca Tedeschi and Paride Rizzo for discussions and reading the manuscript.

I would like to thank Dr. Twan Ruttten for preparation of microscopy pictures, Dr. Lothar Attschmied for transcriptome analysis and Dr. Claude Becker for methylom analysis. Many thanks to Dr. Britt Leps for all her support during my study.

I would like to thank Prof. Dr. Thomas Altmann with numerous insightful comments and consistent discussions.

I would like to thank the Vietnam Forestry University for allowing me to be away such a long time. And special thanks to Ministry of Education and Training of Vietnam (MOET) for the financial support during my stay at IPK, Germany.

Acknowledgements

Last but not least, my deepest gratitude goes to my family: my parents, my son, my husband, my brother and my sister for their love, support and encouragement which helped me to overcome many difficulties and finish this dissertation.

Curriculum Vitae

Personal information

Name:	Bui Thi Mai Huong
Gender:	Female
Birth:	December 25, 1981 in Thai Binh, Vietnam
Nationality:	Vietnamese
Marital status:	Married, 1 child.
Residential address in home country:	No.100, Tan Binh Quarter, Xuan Mai Town, Chuong My District, Hanoi City.
Residential address in Germany:	Selkeweg 4A, 06466 Stadt Seeland, OTGatersleben, Germany.
Email:	huong@ipk-gatersleben.de thohuongminh21@gmail.com

Academic qualification and employment

- 2000-2004: Bachelor degree in Biology, Vietnam National University, Hanoi, Vietnam.
- 2004-2006: Master degree in Biology, Vietnam National University, Hanoi, Vietnam.
- 2005-2011: Lecturer, Vietnam Forestry University, Hanoi, Vietnam.
- 2011-present: PhD student at Gene Regulation group, Department of molecular genetics, Leibniz-Institute of plant genetics and crop plant research (IPK), Gatersleben, Germany.

- 4 / 2015: Participate in the scientific (technical) course “Histological and ultrastructural analysis of plant tissue using Light, Scanning and Transmission Electron Microscopy” within the IPK Graduate Program.

Publication

Nguyen Van Mui, Bui Thi Mai Huong. (2004). Genetic diversity of some white tea species Camellia at Tam Dao national park by RAPD technique. *Fundamental Research in Life Sciences*. Pp. 529-534.

Bui Thi Mai Huong. (2009). Classizying the yellow tea species (Camellia) of Vietnam by molecular method. *Journal of Agriculture and Rural Development*. ISSN 0866: 34-37.

Posters and presentations

Bui Thi Mai Huong, Twan Rutten, Paride Rizzo, Francesca Tedeschi, Andreas Czihal, Markus Kuhlmann, Lothar Altschmied, Claude Becker and Helmut Bäumlein. (2014). EFFECTOR OF TRANSCRIPTION (ET): Novel epigenetic regulators of reproduction are involved in DNA-methylation. IPK day, Gatersleben (Poster).

Bui Thi Mai Huong, Twan Rutten, Paride Rizzo, Francesca Tedeschi, Andreas Czihal, Markus Kuhlmann, Lothar Altschmied, Claude Becker and Helmut Bäumlein. (2015). EFFECTOR OF TRANSCRIPTION (ET): Novel epigenetic regulators of reproduction are involved in DNA-methylation. PSSC student conference at Halle University (talk).

Declaration about personal contributions

The data in this dissertation were carried out by me with the great help and contribution from other people as follows:

- Preparation and analysis of samples using REM were performed in cooperation with Dr. T. Rutten, IPK Gatersleben.
- Transcriptome analysis by deep RNA sequencing were done by Dr. Lothar Altschmied, IPK Gatersleben.
- The determination of the differentially methylated genomic regions has been performed by Dr. C. Becker, MPI Tübingen.

Declaration

Declaration about originality

I hereby declare that all the work in this dissertation is my own. All literature and data provided from other sources were cited and declared and acknowledged (for contribution of collaboration partners see page 93).

Bui Thi Mai Huong

Gatersleben, July 29, 2015