Functional characterization of

EFFECTOR OF TRANSCRIPTION (ET)

in *Arabidopsis thaliana*

Dissertation

zur Erlangung des akademischen Grades Doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät

Martin-Luther-Universität Halle-Wittenberg

Fachbereich Biologie

von

Le, Hong Diep

geboren am 20 April 1971 in Hanam, Vietnam

Gutachter:

1. Prof. Dr. Ulrich Wobus
2. Prof. Dr. Gunter Reuter
3. Prof. Dr. Holger Puchta

Halle (Saale), 14. Februar 2011

ERKLÄRUNG

Hiermit erkläre ich, daß diese Arbeit bisher von mir weder an der Mathematisch Naturwissenschaftlichen Fakultät der Martin-Luther-Universität Halle-Wittenberg noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, daß ich diese Arbeit selbständig verfaßt und keine anderen als die darin angegebenen Hilfsmittel benutzt habe.

DECLARATION

Hereby, I declare that all the work presented in this dissertation is my own, carried out solely with the help of literature and the aids cited.

Le, Hong Diep

Gatersleben, October 2010

List of abbreviations

ABA	ABSCISIC ACID
ABI	ABSCISIC ACID INSENSITIVE
ALP	Alkaline phosphatase
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
bHLH	basic Helix Loop Helix
BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
cDNA	Complementary DNA
CDR	Complementary determining region
CLSM	Confocal laser scanning microscope
Col	Columbia-0
DAPI	4’,6-diamidino-2-phenylindole
DEPC	Diethyl pyrocarbonate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	1,4-Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
EGTA	Ethyleneglycol tetraacetic acid
ELISA	Enzyme linked immunosorbent assay
EMSA	Electrophoretic mobility shift assay
FUS	FUSCA
GA	Gibberellin
GFP	Green fluorescent protein
GST	Gene-specific sequence tag
GUS	β-glucuronidase
HEPES	4-(2-hydroxyethyl)piperazine-1- ethanesulfonic acid
HIS	Hexahistidine
HRT	HORDEUM REPRESSOR OF TRANSCRIPTION
IPTG	Isopropyl-ß-D-thiogalactoside
KDa	Kilo Dalton
LEC	LEAFY COTYLEDON
LEA	LATE EMBRYOGENESIS ABUNDANT
mRNA	Messenger RNA
MS	Murashige and Skoog

NASC	Nottingham *Arabidopsis* Stock Centre
NBT	Nitro blue tetrazolium
Ni-NTA	Nickel-nitrilotriacetic acid
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PIPES	Piperazine-N,N'-bis (2-ethanesulphonic acid)
RNA	Ribonucleic acid
RT-PCR	Reverse transcription PCR
SDS	Sodium dodecyl sulfate
SSP	Seed storage protein
TAG	Triacyl glycerol
TBS	Tris buffer saline
T-DNA	Transferred DNA
X-gal	5-bromo-4-chloro-3-indolyl-ß-D-galactoside
X-gluc	5-bromo-4-chloro-3-indolyl ß-D-glucuronide
VP1	VIVIPAROUS1
Ws	Wassilewskaja-2
UV	Ultraviolet

Legends

ET2	indicates the gene (Capital, italic)
ET2	indicates the protein (Capital)
et2	indicates the mutant allele (Lowercase, italic)

Table of contents

1. Introduction

1.1. *Arabidopsis* embryogenesis

Embryogenesis is a crucial developmental period in the life cycle of flowering plants, allowing the connection between two distinct sporophytic generations to maintain the species. This process starts with the zygote and passes through a sequence of characteristic stages. During embryogenesis, various cellular processes such as rapid synthesis and accumulation of proteins and lipids occur simultaneously to transit zygote from morphogenetic phase to the maturation phase, resulting in seed formation. Embryogenesis can be conceptually divided into distinct phases: morphogenesis (early embryogenesis) and maturation (late embryogenesis) (Lotan *et al.*, 1998; Harada, 2001; Raz *et al.*, 2001). Morphogenetic phase involves the establishment of the basic body plan of embryo, whereas maturation phase implicates cell expansion and accumulation of storage macromolecules to prepare for desiccation, germination and early seedling growth (Braybrook *et al.*, 2006; Park and Harada, 2008).

1.1.1. Early embryogenesis

The early embryogenesis is initiated via a unique double fertilization process in flowering plants. In *Arabidopsis thaliana*, fertilization event comprises of karyogamy between a haploid sperm cell of the pollen and a haploid egg cell of the embryo sac leading to diploid zygote formation and subsequently the embryo. The second fertilization event by which another haploid sperm cell fuses with the homo-diploid central cell generates triploid primary syncytial endosperm nuclei and afterward the mature endosperm cells (West and Harada, 1993; Goldberg *et al.*, 1994; Berleth, 1998; Chaudhury *et al.*, 2001; Park and Harada, 2008). In cereals, the endosperm is a major site for reserve storage and is persistent in the mature dry seed, whereas this function is solely performed by the embryo in *Arabidopsis* and many other dicotyledonous plants (Hirner *et al.*, 1998; Berger, 1999; Berger and Gaudin, 2003).

Following fertilization, the zygote undergoes first asymmetrical division to generate an apical cell and a basal cell that are different in sizes and cytoplasmic densities. The apical cell forms an eight-cell embryo upon two rounds of longitudinal and one round of

transverse divisions. Each cell of the embryo properly undergoes a periclinal division to the surface generating a single outer layer (protoderm) and a dermatogen-stage globular embryo. At the same time, the offsprings of the basal cell divide transversely to form the suspensor and the uppermost cell, hypophysis. The suspensor anchors the embryo to the endosperm and serves as a nutrient conduit from endosperm for the developing embryo whereas the hypophysis gives rise to the root quiescent centre and the initial of the central root cap (West and Harada, 1993; Souter and Lindsey, 2000; Chaudhury *et al.*, 2001; Jurgens, 2001; Laux *et al.*, 2004; Park and Harada, 2008). At the late globular stage, the cell number of *Arabidopsis* embryo increases to more than a hundred cells (Berleth, 1998; Capron *et al.*, 2009).

By a series of cell divisions parallel to the surface, the *Arabidopsis* embryo expands in size and proliferates at two opposite positions in the apical region to transit the embryo from globular to heart stage. At the early heart stage, the embryo shows the first appearance of a bilateral symmetry and a remarkable arrangement of three basic tissue types, such as epidermis (from protoderm), cortex (ground tissue) and procambium. The shift in the embryo symmetry from radial at the globular stage to bilateral at the heart stage represents the initial delineation of the two major embryogenic organ systems, the cotyledon and axis. Subsequent to their formation, axis elongates rapidly as a result of cell division to generate root meristem. The shoot apical meristem is formed later from cell layers localized in the upper axis between the two cotyledons. The morphogenesis phase ends at the heart stage when all embryo structures have been formed (Mayer *et al.*, 1991; Park and Harada, 2008). The heart stage of *Arabidopsis* embryo is followed by the torpedo stage when further elongation of cotyledons, hypocotyls and extension of vascular tissues occur. Although the embryo continues to increase in size and exhibits several changes in shape, it retains the same pattern of basic body plan of shoot root axis and becomes clear at the end of the torpedo stage.

1.1.2. Late embryogenesis

The termination of cell division in the early embryogenesis reveals completion of morphogenesis and establishment of the embryo body plan. In the following development, the *Arabidopsis* embryo enters the late embryogenesis (seed maturation) which is needed for a prolonged survival in a quiescent state. This period is characterized by the arrest of

tissue growth and development, the induction of dormancy, and the acquisition of desiccation tolerance (Parcy *et al.*, 1997; Raz *et al.*, 2001). Throughout late embryogenesis, the embryo is prevented from entering germination pathway and can only germinate after maturation phase is finished. The embryo growth interruption is a transient phase and is reversed upon germination when appropriate environmental conditions are provided and the dry seeds imbibe water (Parcy *et al.*, 1997; Raz *et al.*, 2001; Kagaya *et al.*, 2005; Donohue, 2009).

1.1.2.1. Embryo maturation

During this phase, the embryo grows mostly by cell elongation while cotyledons are programmed to accumulate storage products, such as lipids, carbohydrates and proteins that will be utilized as a food source by the seedling after germination. The deposition of storage reserves in the cotyledons of *Arabidopsis* embryos takes place during relatively short period of 72h beginning at sixth day after flowering, in parallel with cell elongation (Mansfield and Briarty, 1991).

As in most cruciferous plants, the *Arabidopsis* developing embryo mainly stores lipids in the form of triacyl glycerol (TAG) in spherical compartments referred to as spherosomes (Herman, 1995), oleosomes (Murphy, 1993) or most frequently oil bodies (Baud *et al.*, 2002; Siloto *et al.*, 2006). These organelles storing lipids arise from endoplasmic reticulum (ER) which contains the full complement of TAG biosynthesis enzymes (Murphy and Vance, 1999; Hsieh and Huang, 2004). Synthesis of TAGs starts in the late heart stage and continues through the torpedo stage, bent cotyledons until the embryo desiccates. Ultrastructural analysis reveals that oil bodies have a matrix of TAGs surrounded by a layer of phospholipids embedded with abundant structural proteins termed oleosins (Hsieh and Huang, 2004; Siloto *et al.*, 2006). The embedded oleosins modulate the size of oil bodies and are thought to stabilize them during desiccation of the embryo (Ting *et al.*, 1996; Voelker and Kinney, 2001; Hsieh and Huang, 2004).

Proteins are synthesized and accumulated during embryo maturation and mainly serve as sources of carbon, nitrogen and sulphur for the next generation. The most abundant storage proteins in *Arabidopsis* embryo are 2S albumins (referred to as napins) and 12S globumins (referred to as cruciferins). They are classified on the basis of their size and solubility in various solvents (Krebbers *et al.*, 1988). These proteins are initially

synthesized as precursors in the rough endoplasmic reticulum (ER) and then transported to the specialized vacuoles where they are quickly processed by processing enzymes to generate protein bodies (Hou *et al.*, 2005; Otegui *et al.*, 2006; Wan *et al.*, 2007). They are synthesized by small gene families, in which four genes encoding 12S globulins and five genes encoding 2S albumins are present in the embryo of *Arabidopsis* (Pang *et al.*, 1988; van der Klei *et al.*, 1993; Wan *et al.*, 2007). The activity of these genes is only in embryo at early and midstages of maturation under tight temporal and tissue-specific regulation (Lara *et al.*, 2003). The expression of 2S albumin and 12S globulin genes starts between day four and six and reach maximal level of transcripts approximately from nine to ten days after anthesis (Hirner *et al.*, 1998).

Soon after fertilization, starch accumulates temporarily at a very early stage of *Arabidopsis* seed development. This starch is detected in the plastids of embryo cells and also in seed coat cells on day three and reaches maximal amount on day seven after flowering (Focks and Benning, 1998). Later in maturation, starch is only detected in the outer and inner cell layers of outer integument but not in the embryo (Western *et al.*, 2000; Kim *et al.*, 2005). This is because the starch is remobilized during early stages of embryogenesis in order to facilitate rapid lipid and storage protein biosynthesis (King *et al.*, 1997; Vigeolas *et al.*, 2003; Fallahi *et al.*, 2008).

1.1.2.2. Embryo desiccation and dormancy

Preliminary to quiescence or dormancy, embryo desiccation occurs in which the water content decreases dramatically. Desiccation is therefore a normal programmed event in the final phase of seed development. As apart of the normal developmental program of orthodox seeds, most cellular water is lost during desiccation of embryo. This decrease in water content is acquired in the late embryogenesis and related to Late Embryogenesis Abundant (LEA) proteins that accumulate at high levels in plant embryos. Using reference sequences from *LEA* genes in cotton, Bies-Etheve and colleagues have found 50 *LEA* genes in *Arabidopsis* genome (Bies-Etheve *et al.*, 2008). Based on expression patterns, LEA mRNAs of *Arabidopsis* are subdivided into LEA-A and LEA-B classes, which respectively begin to accumulate about 13 and 18 days after pollination (Parcy *et al.*, 1994). The observation in *Arabidopsis* embryo clearly suggests that LEA transcripts and their products are most abundant in embryo just prior to desiccation and disappear

following germination. It has been proposed that, LEA proteins play an important role in maintenance of the enzyme activities, structure of vesicles and endomembranes, in replacement of water and functioning as molecular chaperones during cellular dehydration (Koag *et al.*, 2003; Grelet *et al.*, 2005; Reyes *et al.*, 2005). Some LEA proteins are induced in vegetative tissues in response to various conditions including cold, salt, water deficiency and hormone abscisic acid (ABA) (Skriver and Mundy, 1990; Grelet *et al.*, 2005; Hundertmark and Hincha, 2008). Overexpression of genes encoding LEA proteins in transgenic plants resulted in enhanced resistance to water deficit (Xiao *et al.*, 2007; Hundertmark and Hincha, 2008). All these evidences show that LEA proteins have important biological functions, not only in seed development but also in vegetative tissues where they play a role in cell stress tolerance. Besides LEA proteins, carbohydrates especially soluble sugars may play an essential role in the acquisition of desiccation. Carbohydrates are involved in the stabilization of proteins and retention of enzymic activity and protection of membranes during dehydration (Ooms *et al.*, 1993).

Following desiccation, the *Arabidopsis* embryo enters a period of quiescence or dormancy. Quiescent seeds germinate when provided with suitable conditions necessary for resumption of growth, whereas dormant seeds germinate only when some additional hormonal, metabolic, environmental, physical conditions are appropriate (Raghavan, 2002). Dormancy has been defined as the incapacity of a viable seed to germinate in the presence of favourable environmental conditions (Bewley, 1997; Foley, 2001). This phenomenon introduces a temporal delay in the germination process that provides additional time for seed dispersal over greater geographical distances and also enhances seedling survival by preventing germination under unfavourable conditions (Finkelstein *et al.*, 2008; Bentsink *et al.*, 2010). Two categories of seed dormancy have been recognized, embryo dormancy and seed coat-imposed dormancy. Embryo dormancy arises from a condition within the embryo itself and most likely due to germination inhibitors, especially ABA, as well as the absence of growth promoters including GA. Seed coat-imposed dormancy arises from seed coat and other enclosing tissues, such as endosperm, pericarp that in most common they are impermeable to the entry of moisture or gases (Bewley, 1997; Foley, 2001).

1.1.3. Genetic control of embryogenesis

Previous genetic and molecular studies have demonstrated that in *Arabidopsis*, ABA-INTENSITIVE 3 (*ABI3*) (Koornneef *et al.*, 1984; Giraudat *et al.*, 1992; Parcy *et al.*, 1997), FUSCA3 (*FUS3*) (Bäumlein *et al.*, 1994; Keith *et al.*, 1994) and LEAFY COTYLEDON1 (*LEC1*) (Meinke *et al.*, 1994; West *et al.*, 1994; Kagaya *et al.*, 2005) genes play central roles in controlling mid- and late embryogenesis. These genes (probably *LEC2* as well) have partially overlapping functions in the overall regulation of seed maturation (Parcy *et al.*, 1997). *ABI3*, *FUS3* and *LEC2* encode transcription factors comprising a B3 DNA binding domain, which was originally identified in *VP1* (VIVIPAROUS1), a transcriptional activator from maize (McCarty *et al.*, 1991; Giraudat *et al.*, 1992; Luerssen *et al.*, 1998; Stone *et al.*, 2001; Finkelstein *et al.*, 2002). The B3 domain presents in highly diverse forms of several protein families, including ABI3/VP1, HIS (High-level expression of Sugar-Inducible gene), RAV (Related to ABI/3VP1), ARF (Auxin Response Factor) and REM (Reproductive Meristerm) (Suzuki *et al.*, 1997; Ulmasov *et al.*, 1997; Kagaya *et al.*, 1999; Franco-Zorrilla *et al.*, 2002; Suzuki *et al.*, 2007). *LEC1* encodes other transcription factor sharing significant sequence similarity with the HAP3 subunit of CCAAT binding factor (also known as CBF) (Lotan *et al.*, 1998; Kwong *et al.*, 2003; Lee *et al.*, 2003).

In *Arabidopsis*, several genetically distinct *ABI* loci (*ABI1*, *ABI2*, *ABI3*, *ABI4* and *ABI5*) have been identified (Koornneef *et al.*, 1984; Finkelstein and Somerville, 1990; Finkelstein, 1994; Meyer *et al.*, 1994; Brady *et al.*, 2003). *ABI1* and *ABI2* genes encode protein serine/threonine phosphatase 2C (PP2C) (Bertauche *et al.*, 1996; Leung *et al.*, 1997; Rodriguez *et al.*, 1998). Some studies showed that they have minor influence on seed development, and mainly on aspects of ABA-regulated vegetative growth, such as cell elongation and/or stomatal regulation (Koornneef *et al.*, 1984; Finkelstein and Somerville, 1990; Finkelstein *et al.*, 2002). Mutations in these two loci (*abi1-1*, *abi2-1*) result in nondormant seed, pleotropic defects in vegetative ABA response, reduce phosphatase activity *in vitro* and have no effect on desiccation tolerance (Finkelstein and Somerville, 1990; Leung *et al.*, 1997; Finkelstein *et al.*, 2002). *ABI4* and *ABI5* genes encode proteins belonging to two distinct classes of transcription factors: APETALA2 (AP2) domain and basic leucine zipper (bZIP) domain, respectively (Finkelstein *et al.*, 1998; Lopez-Molina and Chua, 2000; Finkelstein *et al.*, 2002).

Physiological and genetic analyses have suggested that *ABI3, ABI4* and *ABI5* function primarily during seed development and are likely to participate in the same seed-specific signalling pathway. These genes can be expressed in specific tissues outside the seeds under certain conditions. When grown in the dark, *ABI3* is found to be expressed in the apex of *Arabidopsis* seedling after cell division is arrested. In addition, the 2S seed storage protein gene, a target of *ABI3* in seeds, is also induced in the arrested apex under the same conditions (Rohde *et al.*, 1999). *ABI4* and *ABI5* have been shown to have functions in both sugar and salt responses and in early seedling growth after germination (Lopez-Molina *et al.*, 2001). The ABI5 transcripts accumulate during seed development and limit to a narrow developmental window after germination. Mutations in the *ABI3, ABI4* and *ABI5* loci result in similar qualitative effects on seed development and a decreased responsiveness to ABA, but do not alter vegetative growth (Giraudat *et al.*, 1992; Finkelstein *et al.*, 1998; Finkelstein and Lynch, 2000). Severe mutations in *ABI3* gene (*abi3-4*, *abi3-5* and *abi3-6*) or double mutations combining the weak *abi3-1* with *aba1-1* have more severe defects in seed maturation than those in *ABI4* or *ABI5* gene (Finkelstein and Lynch, 2000). Null mutations in ABI3 produce over 1000 fold decrease in ABA sensitivity for germination inhibition (Ooms *et al.*, 1993) and entire loss of expression for several embryo-specific genes (Parcy *et al.*, 1994). In addition, overexpression of *ABI3, ABI4* or *ABI5* confers hypersensitivity to ABA and glucose, as well as producing similar and distinct effects on ABA-regulated gene expression (Brocard *et al.*, 2002; Finkelstein *et al.*, 2002; Kang *et al.*, 2002).

The *FUS3* gene encodes a predicted protein of 312 amino acid residues with a sequence similarity to the ABI3 and VP1 gene products. The homology of FUS3 is restricted to a stretch of more than 100 residues corresponding to the B3 domain which is conserved among VP1/ABI3-like protein family (Luerssen *et al.*, 1998; Wobus and Weber, 1999). The *FUS3* transcript in *Arabidospsis* is expressed 2 days after pollination and the expression level increases during the first-half of embryogenesis but the transcript peaks shortly after mid-embryogenesis (Luerssen *et al.*, 1998). It has been shown that *FUS3* and *LEC2* are involved in the control of gibberellin (GA) biosynthesis in *Arabidopsis*. During seed maturation, they repress the expression of AtGA3ox2, the product of which converts inactive GAs to biologically active forms (Curaba *et al.*, 2004; Gazzarrini *et al.*, 2004). In addition, ectopic expression of *FUS3* inhibits expression of AtGA3ox2 (Gazzarrini *et al.*, 2004). Promoter analysis indicates that the expression of AtGA3ox2 is directly regulated

by binding of LEC2 and FUS3 proteins with the RY motif (CATGCAT) on the promoter regions (Curaba *et al.*, 2004). Loss of *FUS3* function in *Arabidopsis* results in a complex phenotype specially affecting seed development. Particularly, *fus3* embryos are defective in producing the main seed proteins (12S and 2S) as well as storage lipids but accumulate large amounts of anthocyanin (Bäumlein *et al.*, 1994; Keith *et al.*, 1994; Luerssen *et al.*, 1998).

LEC genes are required for normal growth during both the morphogenesis and maturation phases of seed development. Evidences have been shown that *LEC1* participates from beginning of embryogenesis to the late maturation phase (Lotan *et al.*, 1998; Harada, 2001). For instance, *LEC1* is required to maintain the fate of embryonic cells that form the suspensor and to specify the identity of cotyledons and embryonic leaves in the early embryogenesis, to initiate and maintain the maturation phase as well as to inhibit premature germination in late embryogenesis (Meinke *et al.*, 1994; West *et al.*, 1994; Parcy *et al.*, 1997; Lotan *et al.*, 1998). The accumulation of *LEC1* mRNA is restricted to seed development from preglobular to bent cotyledon stage and degenerated during seed maturation, accumulation of storage reserves, and prevention of germination in immature seeds (West *et al.*, 1994; Lotan *et al.*, 1998; Vicient *et al.*, 2000). *LEC2* mRNA accumulates primarily during the maturation phase of seed development, although its transcript may be present at very low levels at other stages of life cycle (Stone *et al.*, 2001; Kroj *et al.*, 2003; Braybrook *et al.*, 2006).

The functions of *LEC* genes in *Arabidopsis* are partly overlap and not completely redundant, acting as specific central regulators of embryogenesis. Their expressions are repressed outside the embryo by *PICKLE* (*PKL*), a chromatin remodelling factor (CHD3) acting in concert with GA to repress embryonic traits during and after germination. In contrast, loss of *PKL* function leads to elevated expression of seed-specific genes, including *LEC1*, *LEC2* and *FUS3* (Ogas *et al.*, 1999; Dean Rider *et al.*, 2003; Dean Rider *et al.*, 2004; Henderson *et al.*, 2004; Li *et al.*, 2005; Zhang *et al.*, 2008). In *pkl* mutants, primary roots are capable of expressing many embryonic traits after germination. The transcripts for *LEC1*, *LEC2* and *FUS3*, exhibit *PKL*-dependent repression and they are all expressed at levels more than 100-fold in *pkl* primary roots (Dean Rider *et al.*, 2003; Dean Rider *et al.*, 2004).

Ectopic expression of either *LEC1* or *LEC2* in vegetative tissues can trigger the formation of embryo-like structures (Lotan *et al.*, 1998; Stone *et al.*, 2001; Santos-Mendoza *et al.*, 2005). Moreover, the ectopic expression of *LEC1* also results in the expression of *FUS3* and *ABI3*, which preceded in the induction of SSP (Kagaya *et al.*, 2005). Additionally, ectopic expression of *LEC2* induces accumulation of seed storage proteins and oil bodies in vegetative and reproductive organ (Stone *et al.*, 2008) such as *S3* oleosin and *At2S3* albumin in *Arabidopsis* leaves (Santos-Mendoza *et al.*, 2005). Similarly, induction of *LEC2* under the control of CaMV35S promoter is sufficient to cause the transformation of unfertilized ovule integuments and roots into storage tissues for lipids and proteins (Stone *et al.*, 2008). More detailed analyses revealed that the expression of *LEC1, FUS3* and *ABI3* is also induced by *LEC2* activation (To *et al.*, 2006; Stone *et al.*, 2008).

The phenotypes of *lec* mutants (*lec1* and *lec2*) are similar in several ways to *fus3* and *abi3* such as cotyledons partially converted into leaves, viviparous embryos and reduced desiccation tolerance (Bäumlein *et al.*, 1994; Keith *et al.*, 1994; Meinke *et al.*, 1994; West *et al.*, 1994). In addition, embryonic leaves or cotyledons of *lec1* mutants possess trichromes on the cotyledons, epidermal hairs, which are normally produce only on leaves and stems of *Arabidopsis* (West *et al.*, 1994; Meinke, 1995). Furthermore, the expression of *FUS3* and *ABI3* is found to be down-regulated in developing siliques of the *lec1* mutant (Kagaya *et al.*, 2005). Loss of function mutations in *LEC2* during late embryogenesis in *Arabidopsis* produces cotyledon tips that do not accumulate storage reserves nor acquire desiccation tolerance, indicating defects in the initiation and maintenance of the maturation phase (Stone *et al.*, 2008).

Due to pleiotropic, and partially overlapping functions, *ABI3, FUS3, LEC1* and *LEC2* have been considered as master regulators of seed development (Kroj *et al.*, 2003; Kagaya *et al.*, 2005; Santos-Mendoza *et al.*, 2005; Verdier and Thompson, 2008).

1.2. Plant ET gene family

1.2.1. Discovery and isolation of ETs

EFFECTOR OF TRANSCRIPTION (ET) genes were discovered by using South Western screens with the aim of isolating transcription factors important for embryonic gene regulation (Ellerström *et al.*, 2005; Ivanov *et al.*, 2008). Several independent screenings were performed with seed-specific cDNA libraries prepared from barley (*Hordeum vulgare*), rapeseed (*Brassica napus*) and broad been (*Vicia faba*) (M. Ellerström, T. Wohlfarth, P. Wycliffe, L. Rask, H. Bäumlein, unpublished). The oligonucleotides used as probes in these approaches comprised GARE (Gibberellic Acid Response Element) region from *H. vulgare*, or sequences of *napA* promoter from *B. napus*, *USP* (unknown seed protein) and *leB4* (legumin B4) promoters from *V. faba*. The screenings resulted in isolation of three proteins from *H. vulgare*, *B. napus*, and *V. faba* named as HRT (*Hordeum* Repressor of Transcription), BnET (*Brassica napus* ET) and VfET (*Vicia faba* ET), respectively. The sequencing and alignment results revealed that these factors show low but significant similarity especially in their carboxy terminal regions. They share highly conserved cysteine-containing structural sequences, designated ET domains which are present twice in the *V. faba* protein (EMBL/GenBank accession number X97909), four times in the *B. napus* protein (EMBL/GenBank accession number AY533506) and three times in the *H. vulgare* protein. The conserved structure of ET domains contains a common pattern C-$X_{8/9}$-C-X_9-R-C-X_2-H-K (Figure 1). BLAST searches also confirmed that, these families are unique to the plant species and have so far not been found outside plant kingdom (Raventos *et al.*, 1998; Ellerström *et al.*, 2005; Ivanov, 2005; Ivanov *et al.*, 2008). In addition to ET domains, ET factor families do not show any other domain or sequence homology.

Database searches showed that members of ET family appear in species belonging to monocots and dicots as well as lower plants such as moss. A protein found in *Physcomitrella patens* (moss) (acession number: PPP_3786_C1 OSMOSS database) seems to contain only a single domain structure (Figure 1). Even though no functional data are available for the *Physcomitrella* protein, the presence of a similar domain also in such a distantly related phylum as mosses suggests an evolutionary conservation and therefore most likely important function in photosynthetic organisms (Ellerström *et al.*, 2005).

1.2.2. *Arabidopsis ET* genes

Similar investigations on the *Arabidopsis* genome eventuated in identification of three supplemental genes with homology to the previously detected *HRT*, *BnET* and *VfET*. They were denominated as *AtET1* (At4g26170), *AtET2* (At5g56780) and *AtET3* (At5g56770). As their given code numbers, *AtET1* is located on the fourth chromosome, while the other, *AtET2* is detected on the fifth chromosome close to *AtET3* (Ellerström *et al.*, 2005; Ivanov, 2005; Ivanov *et al.*, 2008). The sequencing data showed that the Columbia-0 (Col) and Wassilewskaja-2 (Ws) ecotypes are polymorphic with respect to *AtET1*. The *AtET2* gene is an intact coding sequence and identity in both Col and Ws, whereas *AtET3* is truncated version of AtET2 due to lack of the coding region for the Zn and DNA binding C-terminal cysteine repeats (ET domains). Despite this deficiency, some *AtET3* transcript can be detected by RT-PCR indicating that it might be a functional gene (Ivanov, 2005; Ivanov *et al.*, 2008).

```
PpET          ICGLKLLDGT VCPDPPRPDR KRCEAHKGLR

OsETa         VCGVMLEDGS SCLDHPVQGR KRCELHKGRR
OsETb         ICEAKKSDNS ACTNKVISGS KKCQLHNGCK
OsETc         ICEALT-DNR CRETIPMAGR ERCDAHEGIK
OsETd         ICGARASDGS PCKNQPIAGR KRCAMHKGQR

HRTa          VCGVMLEDGS SCLEDPMEGR KRCELHKGRR
HRTb          LCGVVTDNG- YCKLEPVIGR ERCEEHRGIE
HRTc          VCGARASDGS PCKNQPIARR KRCALHKGQR

VfETa         ICGVILDDGS ICSKMPVGKR VRCNEHKGMR
VfETb         ICGIVLEDGS TCRKEPVKGR KRCHEHKGKR

BnETa         VCGVLQEDGT TCLTAPVTGR KRCTEHKGQR
BnETb         ICGVILPEMV RCRSKPVSGR KRCEDHKGMR
BnETc         ICEATTKNGL PCTRSAPNGS KRCWQHKDET
BnETd         VCGVKLHNGS VCEKTPVKGR KRCQEHKGMR

AtET1a        ACGVLLEDGT TCTTTPVKGR KRCTEHKGKR
AtET1b        ICGVILPDMI RCRSKPVSRR KRCEDHKGMR
AtET1c        LCEATTKNGL PCTRSAPEGS KRCWQHKDKT
AtET1d        ICGFKLYNGS VCEKSPVKGR KRCEEHKGMR

AtET2a        VCGVLLEDGG CCIRSPVKGR KRCIEHKGKR
AtET2b        VCGVILPDME PCNKRPVPGR KRCEDHKGMR
AtET2c        FCEATTKNGL PCTRSSPKGS KRCWQHKEKT
AtET2d        ACGVKLGNGL ICERSPVKGR KRCEEHKGMR

Consensus     -C-------- -C-------- -RC--HK---
```

Figure 1. Sequence alignments of all known ET domains.

Pp: *Physcomitrella patens*; *Os*: *Oryza sativa*; *Vf*: *Vicia faba*; *HRT*: Hordeum Repressor of Transcription from *Hordeum vulgare*; *Bn*: *Brassica napus*; *At*: *Arabidopsis thaliana*. The order of sequences from amino to carboxyl terminals in every domain was indicated by a to d. The consensus pattern C-$X_{8/9}$-C-X_9-R-C-X_2-H-K was highlighted in yellow and underlined (Ellerström *et al.*, 2005; Ivanov *et al.*, 2008).

Further analysis of *AtET3* elucidated a mis-annotation of the exon-intron between Col and Ws ecotypes. The third exon of the Ws allele starts 20 bp earlier than that in Col ecotype, as well as 155 bp shoter coding region in this allele. In addition, the presence of 4 bp duplication at position 602 of the Ws allele creates a frameshift followed immediately by a stop codon (Ivanov *et al.*, 2008). Additional comparison of all three *AtET* genes revealed the similarity in their genomic organization. The structure of *AtET* genes comprises three exons separated by two introns on the similar places. Full lengths of *AtET1* and *AtET2* genes span approximately 1.6 and 1.9 kb, respectively, while *AtET3* is considerably smaller with 1.0 kb in length (Figure 2).

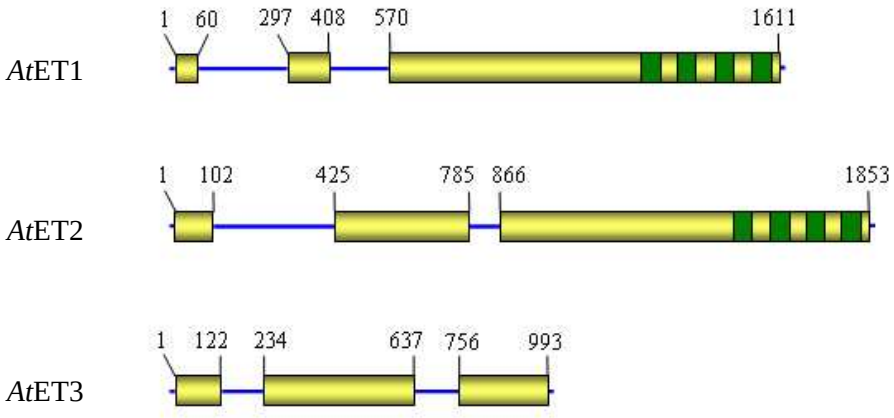


Figure 2. Structure of the *AtET* gene family.

Schematic representation of the three *AtET* genes: exons were indicated by yellow boxes and the ET repeat regions were shown in green. Both *AtET1* and *AtET2* contain four ET repeats at the C-terminal coding regions, while AtET3 lacks of ET repeats (Ivanov *et al.*, 2008). The starting and ending of exons were indicated by above numbers.

Furthermore, the last exons of both *AtET1* and *AtET2* genes are analogous in their relative lengths and in arrangement of the ET repeats. Among these three *AtET* members, *AtET1* shares the highest level homology to the *BnET* (Ellerström *et al.*, 2005; Ivanov, 2005; Ivanov *et al.*, 2008). As corresponding ET factors, their proteins except *AtET3* also specify the conserved pattern C-$X_{8/9}$-C-X_9-R-C-X_2-H-K at C-terminal regions (Figure 1).

1.2.3. Function of *ET* genes in plants

Functional analyses in plant cells indicated that HRT targets to the nuclei and can act as a transcriptional repressor. By interaction with *cis*-acting elements, it represses the expression from several GA-responsive promoters including *α-amylase Amy1/6-4 and Amy2/32* promoters (Raventos *et al.*, 1998). The latter evidences on *BnET* also confirmed the role of ET factors in gibberellin signalling modulation and cell differentiation. Transient expression of *BnET* in *Arabidopsis* protoplasts showed the repression of GA induced promoter activity. Further supporting information has been obtained with a similar transient expression system based on *Arabidopsis* protoplasts. The co-expression of *AtET2* driven by a constitutive CaMV35S promoter down-regulates GA-responsive GASA4 promoter. These results suggested that ET factors involve in modulation of GA responses (Ellerström *et al.*, 2005; Ivanov, 2005; Ivanov *et al.*, 2008).

Overexpression of *BnET* in either tobacco or *Arabidopsis* plants leads some deficiencies in early stages of development. The germination of green seeds is reduced and delayed in comparison with wild type seeds. In addition, the transgenic plants display severe lack of lignin, accumulation of excessive anthocianin, postpone flowering and dwarf phenotype due to short internodes (Ellerström *et al.*, 2005). In contrast, immature seeds isolated from *Arabidopsis* green siliques of *et2-1* knock out mutant line show strong precocious germination comparable to the control seeds of *fus3-5* and wild type Ws (Ivanov, 2005). Furthermore, loss of function of *AtET2* gene leads to the reduction of lignin, a reliable maker for differentiation of xylem tissue. The *et2-1* mutant line contained about 30% less lignin than did wild type, both in leaves and in the stems, suggesting that *AtET2* is required for differentiation of xylem cells (Ivanov *et al.*, 2008).

1.3. Objectives of the thesis

The initial identification of ET factors from *Hordeum vulgare*, *Vicia faba* and *Brassica napus* supposed a function of ET during embryogenesis (Raventos *et al.*, 1998; Ellerström *et al.*, 2005), and subsequently it was suggested that an ET factor controls both maintenance of meristem identity and normal vascular bundles in plants (Ivanov *et al.*, 2008). These findings indicated that expression of ET was not restricted to seed tissues only. ET transcripts were detectable from various plant organs showing the ubiquitous expression pattern of these factors throughout plant development (Ivanov *et al.*, 2008). Therefore, I have attempted to elucidate the complex mechanisms by which AtET proteins act as putative transcription factors during growth and development of *Arabidopsis thaliana*.

Specific objectives in our research included were:

1. To delineate the structure of AtET factors and their domains responsible for the molecular functions. Using purified AtET proteins, I attempted to produce specific antibodies and to characterize binding properties of AtET proteins *in vitro.*
2. To determine expression pattern of AtET during plant growth and development by analyses of promoter activities and transient expression in protoplasts.
3. Finally to characterize the function of AtET factors by reverse genetic approaches and by gain of function. The knock-out of AtETs and the RNAi lines, as well as lines that constitutively express AtET were studies in detail for this purpose.

2. Materials and methods

2.1. Materials

2.1.1. Plant materials

Arabidopsis thaliana Columbia-0 (Col) and Wassilewskija-2 (Ws) ecotypes were obtained from Gene Regulation Group (IPK, Gatersleben, Germany) and used throughout this study as wild type controls and for all gene transfer experiments.

2.1.2. Bacterial strains and phages

Several bacterial strains and phages were used for different purposes as DNA cloning, plasmid DNA amplification, protein expression, plant transformation, and antibody production.

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^qZΔM15, Tn10(tet^R)], *rel*A1; (Stratagene, La Jolla, CA).
Escherichia coli SOLR	*e14$^-$(McrA$^-$) Δ(mcrCB-hsdSMR-mrr)171 sbcC recB recJ uvrC umuC::Tn5(Kanr) lac gyrA96 relA1 thi$^-$1 endA1 λ R [F′ proAB lacIqZΔM15] C Su$^-$* ; (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)
Escherichia coli M15 [pREP4]	Nal^s, Str^s, Rif^s, Thi^-, Lac^-, Ara^+, Gal^+, Mtl^-, F^-, Uvr^+, Lon^+; (Qiagen, Hilden, Germany).
Escherichia coli TG1	*K12, Δ(lac-pro), supE, thi, hsdD5/FtraD36, pro A^+ B^+, lac1^q, lacZΔM15.*
Escherichia coli HB2151	*K12, ara, Δ(lac-pro), thi/F‘ pro A^+ B^+, lac1^q ZΔM15*.
Escherichia coli DB3.1	F^- *gyr*A462 *end*A1 Δ(sr1-*rec*A) *mcr*B *mrr* *hsd*S20(r_B-, m_B-) *sup*E44 *ara*-14 *gal*K2 *lac*Y1

	*pro*A2 *rps*L20(Sm^R) *xyl*-5 λ- l*eu mtl*1 (Invitrogen, Carlsbad, CA)
Agrobacterium tumefaciens	(Deblaere *et al.*, 1985)
pGV2260	
Agrobacterium tumefaciens pGV3101 (pMP90-pSOUP)	(Hilson *et al.*, 2004)

Phages

Helper phage M13KO7	GE Healthcare

2.1.3. Enzymes, markers, antibiotics and others

Enzymes:

- EcoRI, HindIII, SphI, SalI, T4 DNA ligase, *pfu* DNA polymerase, DreamTag DNA polymerase, Klenow fragment, Shrimp Alkaline Phosphatase (SAP) (Fermentas, Vilnius, Lithuania). Phusion polymerase (Finnzymes, Keilaranta, Finland).
- Dnase I, RNase I (Roche, Germany).

Markers

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

Anibiotics:

- Ampicillin, carbenicillin, chloramphenicol, gentamicin, kanamycin, rifampicin, spectinomycine, tetracycline (Duchefa, The Netherlands).
- Hygromycin B (Roche, Germany).

Other chemicals

- Dexamethasone, p-nitrophenyl phosphate, Imidazole, Anti-poly histidine monoclonal antibody, Anti-Mouse IgG alkaline phosphatase, Anti-mouse IgG (whole molecule) peroxidase conjugate (Sigma-Aldrich, Germany).
- X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside) (Roche, Germany).
- Ni-NTA agarose (Qiagen, Hilden, Germany).
- Salmon Sperm DNA (Stratagene, La Jolla, CA).

- NBT (nitro blue tetrazolium chloride), BCIP (5-bromo-4-chloro-3-indolyl-phosphate-p-toluidine-salt); Coomassie brillant blue G250, R250, Dithiothreitol (DTT) (Serva, Heidelberg, Germany).
- IPTG, Murashige - Skoog (MS) medium basal salt mixture including vitamins and microelements (Duchefa, The Netherlands).
- [α-^{32}P]-ATP, [α-^{32}P]-dCTP (3000 Ci/mmol) (Hartmann Analytic, Braunschweig, Germany).
- Biorad protein assay reagent (Biorad Laboratories, München, Germany).
- GelCode blue stain reagent (PIERCE, USA).
- Silwet L-77 (Lehle Seeds, USA).
- Sucrose, glucose, rotiphorese gel 30, triethylamine, sodium hypochlorite, tris-base, yeast extract, glycerol, glycine (Carl Roth, Germany).

Membranes

- Nitrocellulose membranes (Schleicher & Schuell, Germany).
- Hybond N+ membrane (Amersham Biosciences, United Kingdom).
- Spectra molecularporous membrane tubing (Spectrum Laboratories, Inc., USA).

2.1.4. Commercial kits:

- GeneJET plasmid miniprep kit, GeneJET gel extraction kit, RevertAid first strand cDNA synthesis kit, DNA labelling kit (Fermentas, Vilnius, Lithuania).
- ECL Western blotting detection regeagents kit (Amersham, United Kingdom).
- RNeasy kit, QIAquick PCR purification kit, QIAquick gel extraction kit, Qiagen plasmid purification mini and midi kit (Qiagen, Hilden, Germany).
- TA cloning® kit dual promoter, Zero Blunt® TOPO Cloning kit, pENRT™ Directional TOPO Cloning kit, Gateway® BP clonase II Enzyme mix, Gateway® LR clonase II Enzyme mix, SuperScript II kit, SuperScript III kit (Invitrogen, Carlsbad, CA).
- SYBR-GREEN PCR master mix (Applied Biosystems, United Kingdom)
- Dynabeads mRNA direct kit (Dynal Biotech, Norway)

2.1.5. Vectors

Various vectors were used for DNA amplification, protein expression in *E. coli*, 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
pENTR™/D-TOPO	Kanamycin^r	Invitrogen, Carlsbad, CA
pQE30	Ampicillin^r	Qiagen, Hilden, Germany
pBlueScriptSK	Ampicillin^r	Stratagene, La Jolla, CA
pDONR201	Kanamycin^r, gateway donor vector	Invitrogen, Carlsbad, CA
pDONR207	Gentamicin^r, gateway donor vector	Invitrogen, Carlsbad, CA
pDONR223	Spectinomycin^r, gateway donor vector	Invitrogen, Carlsbad, CA
pKGWFS7.0	Kanamycin^r, gateway binary vector	VIB, Ghent, Belgium
pGKGWG (N9831)	Kanamycin^r, gateway binary vector	NASC, Nottingham, UK
pGBGWG (N9837)	Basta^r, gateway binary vector	NASC, Nottingham, UK
pAGRIKOLA	Basta^r, gateway binary vector	(Hilson *et al.*, 2004)
p35S::R1R2::ΔGR	Kanamycin^r, gateway binary vector	(Baudry *et al.*, 2004)

2.1.6. Primers and oligonucleotides

Oligonucleotides for binding experiments (EMSA)

Oligonucleotide name	*Sequence 5'-3'*
O.1	GCATGAACGTCACGTGGACAAAGGTA
O.2	TTCTGTCACACGTGTTACTCTCTAAGCT
O.3	TTGCTGCTACACGTATATAAGAAAAGCT
O.4	GCATAGCTGGCAAATGCTCATAGGTA
O.5	GCATGAACGTAGCGCAGACAAAGGTA
O.6	GCATGGACAAGACGTGGACATAGGTA
O.7	TCTCCTGCTACGCCTATATAAGACCA
O.8	TCATCCTGCCGGTCGCGCTCAGGCT

O.9	GTATGTAGCGTTCGTTGGAGCTAGGA
O.10	GCATAGCTGGACATACCTCATAGGTA

Primers for PCR and sequencing

Primer name	***Sequence 5'-3'***	***Source* or *Tm* (^{o}*C*)**
ACT2-F	TCGGTGGTTCCATTCTTGCT	56.8
ACT2-R	GCTTTTTAAGCCTTTGATCTTGAGAG	54.7
Agri51	CAACCACGTCTTCAAAGCAA	53.8
Agri56	CTGGGGTACCGAATTCCTC	54.8
Agri64	CTTGCGCTGCAGTTATCATC	54.3
Agri69	AGGCGTCTCGCATATCTCAT	55.8
GET1-F1	GTTTCCTTCGCCGTCGTGTT	58.0
GET1-F2	TGTTCACCATGTTCAAGAGAGACG	57.0
GET1-R1	GACCCTGGAAGGCTCCTTGG	60.0
GET1-R2	GAAACATAGATCGGGCGAAACC	56.5
GET2-F	ATGGAATTCGGCGACGGCG	(Ivanov *et al.*, 2008)
GET2-R	TGCTCTTCACATCTCTTACGTCCTTTTA	57.6
LBa1	TGGTTCACGTAGTGGGCCATCG	61.5
XR2	TGGGAAAACCTGGCGTTACCCAACTTAAT	61.9
ET1-RT-F1	ATGTTCAAGAGAGACGACTACATTC	54.4
ET1-RT-F2	CTGATGCTGCTGAGAAGGAG	55.3
ER1-RT-R1	GAAACATAGATCGGGCGAAACC	56.0
ET1-RT-R2	ATCACGTTCTGTTGGGTTCA	54.3
pQET1-F	ACATGCATGCTTCAAGAGAGACGACTACATTC	61.3
pQET1-R	ACGCGTCGACAGATGTGATTCTCATCC	62.0
pQET2-F	ACATGCATGCGAATTCGGCGACGGCGTT	68.0
pQET2-R	ACGCGTCGACGGTGATTCTCATTCCCTT	64.9
Uvp1	GCTGATGTCAAAATCATCATG	(Ivanov *et al.*, 2008)

Uvp2	TCAATGTTTCAACGACCAGAAG	(Ivanov *et al.*, 2008)
Uvp3	CCCGGGCTTGATAATGTCTCCGCA	(Ivanov *et al.*, 2008)
Uvp4	ATGTCTTGTCCGGGTCTGTATGAG	(Ivanov *et al.*, 2008)
Uvp5	GATATCGTTAAGGTTGTTAACAT	(Ivanov *et al.*, 2008)
Uvp6	CAGATCCAGCAAATTGATGTA	(Ivanov *et al.*, 2008)
Uvp7	TCAATGTTTCAACGACCAGAAG	(Ivanov *et al.*, 2008)

Primers for gateway cloning

Primer name	***Sequence 5'-3'***	***Tm (°C)***
ET1-F	ATGTTCAAGAGAGACGACTACATTCGA	57.3
ET1-R	GAGATGTGATTCTCATCCCCTTGTG	57.0
ET2-F	ATGGAATTCGGCGACGGCGT	61.0
ET2-R	GGGTGATTCTCATTCCCTTATGCTC	58.3
pfET1-F	CAAGTGAATAATCAATGCCTGGTTCAGA	57.4
pfET1-R	AGATGTGATTCTCATCCCCTTGTG	53.2
pfET2-F	TGAAGGAAGAGACAATGGTGTG	57.5
pfET2-R	GGTGATTCTCATTCCCTTATGCTC	57.3
proRT2-R	TTCGATAAAACCGATGATATAGTG	52.5
iET1-F	AAGCCAAAGATTCTGATAGGAGCCA	58.1
iET1-R	TGGAAGCGAATTACTCTCCTTCTCA	57.5
i2ET1-F	GATGGTACTACTTGCACTACAACT	54.3
i2ET1-R	GGATTTGTCCTCATTGACGGCTTTA	57.4
iET3-F	GCATTCTCCAAATGGCGGAATC	56.0
iET3-R	TGAGTATCCTTCCGAGAATATGTC	53.6

Every forward primer and reverse primer contains the additional attB1-recombination site ggggacaagtttgtacaaaaaagcaggct and the attB2-recombination site ggggaccactttgtacaagaaagctgggt at the 5' ends, respectively.

Primers for realtime PCR

Primer name	*Sequence 5'-3'*	*Tm (°C)*
qET1-F	TGCGAGGATCACAAGGGAATGAGA	60.0
qET1-R	CGGCTTTATCACGTTCTGTTGGGT	59.6
qET2-F	GTTTCCTTCGCCGTCGTC	56.5
qET2-R	CACCATCTTTTCCGTTTTTGA	52.6
UBQ10-F	CACACTCCACTTGGTCTTGCG	58.5
UBQ10-R	TGGTCTTTCCGGTGAGAGTCTTCA	59.6

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

2.1.7. Solutions and buffers

EMSA

2X binding buffer	HEPES pH 7.9	20 mM
	Tris-HCl pH8.0	16 mM
	EDTA pH 8.0	0.5 mM
	Glycerol	20%
	DTT	2 mM
6X EMSA loading buffer	Bromophenol blue	0.25% (w/v)
	Xylene cyanol	0.25% (w/v)
	Sucrose	40% (w/v)
TES buffer	Tris-HCl	10 mM
	EDTA	1 mM
	NaCl	300 mM
	Adjust to pH 7.9	
10X TBE buffer	Tris-base	108 g
	H_3BO_3	55 g
	EDTA	8.8 g
	Distilled water	up to 1000 ml

GUS expression

GUS-staining solution	Sodium phosphate buffer (pH7.2)	0.1 M
	$Fe(CN)_2$	0.5 mM
	$Fe(CN)_3$	0.5 mM
	Tween 20	0.1%
	X-Gluc	2 mM

Extraction of plant genomic DNA

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

RNA gel electrophoresis

10x MOPS buffer	MOPS	0.2 M
	NaOAc	0.05 M
	EDTA	0.01 M
	pH 5.5-7.0	
Loading buffer		
	Formamide, deionized	50%
	Formaldehyde	6.5%
	Glycerol	20%
	Xylencyanol	0.2%
	Bromophenol blue	0.2%
	Ethidium bromide	0.005%
	In 1x MOPS buffer	

Purification and dialysis of proteins

Suspension buffer	Tris-HCl pH8.0	20 mM
	EDTA pH 8.0	1 mM
Guanidine lysis buffer	Guanidine hypochloride	6 M
	NaH_2PO_4	20 mM
	NaCl	500 mM
	Adjust to pH 7.8	
Urea lysis buffer	Urea	8 M
	NaH_2PO_4	20 mM
	NaCl	500 mM
	Adjust to pH 7.8	

Native washing buffer	Imidazole	20 mM
	NaCl	500 mM
	Tris-HCl	20 mM
	Adjust to pH 8.0	
Native elution buffer	Imidazole	250 mM
	NaCl	500 mM
	Tris-HCl	20 mM
	Adjust to pH 8.0	
10X PBS buffer	KH_2PO_4	0.02 M
	Na_2HPO_4	0.08 M
	NaCl	1.5 M
	Adjust to pH 7.5	

SDS-PAGE and Western blot

SDS running buffer	Tris-base	25 mM
	SDS	3.5 mM
	Glycine	192 mM
Transfer buffer	SDS running buffer supplemented 20% methanol	
5X SDS loading buffer	Tris-HCl pH8.0	250 mM
	Glycerol	25% (w/v)
	SDS	7.5% (w/v)
	Bromophenolblue	0.25 mg/ml
	Mercaptoethanol	12.5% (v/v)
Roti-Block buffer	(Carl Roth, Germany)	
4X Marvel buffer	Tris-HCl, pH 7.8	80 mM
	NaCl	720 mM
Coomassie stain solution	Coomassie brilliant blue R250	0.2% (w/v)
	Coomassie brilliant blue G250	0.006% (w/v)
	Acetic acid	10% (v/v)
	Methanol	30% (v/v)
	Ethanol	17.5% (v/v)
Coomassie destaining solution	Methanol	40% (v/v)
	Acetic acid	10% (v/v)

TBS buffer	Tris-HCl pH8.0	20 mM
	NaCl	180 mM
	Adjust to pH 7.4	
TBST buffer	TBS supplemented with 0.1% Tween 20	
ALP substrate buffer	Tris-HCl	100 mM
	NaCl	100 mM
	$MgCl_2$	5 mM
	Adjust to pH 9.5	

Macroarray

Church-Gilbert buffer	NaH_2PO_4	0.5 M
	Na_2HPO_4	0.5 M
	EDTA	1 mM
	SDS	7%
	BSA	1%
	Adjust to pH 7.2	
5x RT buffer	Tris-base	250 mM
	KCl	250 mM
	$MgCl_2$. $6H_2O$	50 mM
	Adjust to pH 8.3	
Elution buffer	EDTA	2 mM
	Adjust to pH 8.0	
20xSSC buffer	NaCl	0.3 M
	Tri-sodium citrate	0.03
	Adjust to pH 7.0-8.0	

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

2.1.8. Media

For bacteria and phage: All media were sterilised by autoclaving.

- M9 medium	NaH_2PO_4	0.6%
	KH_2PO_4	0.3%
	NaCl	0.05%
	NH_4Cl	0.1%
	Adjust to pH 7.4	

After autoclaving the following sterile solutions were added:

	1M $MgSO_4$	0.1%
	20% glucose	1%
	1M $CaCl_2$	0.01%

- SOC medium	Yeast extract	5 g
	Tryptone	20 g
	NaCl	0,58 g
	KCl	0,186 g
	Distilled water	up to 1000 ml
	Adjust to pH 7.0	

After autoclave add 1 ml of 2M glucose solution

- LB medium:	Yeast extract	5 g
	Tryptone	10 g
	NaCl	10 g
	Distilled water	1000 ml
	Adjust to pH 7.5	

Add 15g Bacto agar per litre for solid medium

- YEB medium	Beef extract	5 g
	Yeast extract	1 g
	Peptone	5 g
	Sucrose	5 g
	Distilled water	up to 1000 ml
	Adjust to pH 7.0	

After autoclaving, add 2 ml of sterile 1M $MgSO_4$ solution.

- 2x TY medium	Tryptone	16 g
	Yeast extract	10 g
	NaCl	5 g
	Distilled water	up to 1000 ml
	Adjust to pH 7.0	

Add 15g Bacto agar per litre for solid medium

- TYE medium	Tryptone	10 g
	Yeast extract	5 g
	NaCl	8 g
	Distilled water	1000 ml
	Adjust to pH 7.0	

Plant culture and transformation:

- MS medium:

	MS including vitamins	4.6 g
	Sucrose	30 g
	Distilled water	up to 1000 ml
	Adjust to pH	5.8

Add 15 g Bacto agar per litre for solid medium

- Infiltration medium for *Arabidopsis* plant transformation:

	MS including vitamins	2.30 g
	Sucrose	50 g
	Silwet L-77	0.5 ml
	Distilled water	up to 1000 ml

Medium was prepared freshly and not necessary to be sterilised

2.2. Methods

2.2.1. Basic cloning methods and sequencing

The basic molecular cloning methods such as enzymatic digestion, DNA ligation, DNA and RNA gel electrophoreses were performed according to the 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 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.

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.

2.2.2. Domain swapping

The wild-type UvrC promoter and gene were amplified from *Escherichia coli* DH5α using Uvp1 and Uvp3 primers and were cloned into pCR2.1 vector. To express the chimeric UvrC protein containing the AtET2 GIY-YIG domain, the *E. coli* UvrC promoter was amplified using Uvp1 and Uvp2 primers. The amplicon was inserted into pCR2.1 and the resulting plasmid was linearized by SmaI restriction. The coding sequence for the AtET2 GIY-YIG domain, either with or without the R>A mutation, was amplified by Pfu polymerase with Uvp4 and Uvp5 primers to ensure a blunt ended product. The amplicon was inserted into the linearized vector downstream of the UvrC promoter. The resulting construct was linearized by EcoRV digestion and was ligated to the remainder of the UvrC coding sequence amplified with Pfu polymerase using Uvp6 and Uvp7 primers. The fidelity of both constructs was confirmed by resequencing.

2.2.3. Complementation assay

E. coli strain SOLR (Stratagene, La Jolla, CA) was used for the complementation test. Cells carrying either the empty vector, the *E. coli* UvrC gene, the chimeric UvrC protein containing either the wild-type or R>A mutated AtET2 GIY-YIG domain were grown until

the mid-exponential phase (OD=0.6) in LB containing 50 µg/ml ampicillin. Cells were harvested by centrifugation at 5000 rpm for 10 min at 4°C, washed twice and resuspended in M9 medium. A volume of 4 ml cell suspension was transferred to a 70-mm diameter Petri dish, producing a <2-mm-deep liquid layer. Irradiation was applied with a 4-W UV lamp (254 nm) from a distance of 90 cm in a dark room for 5, 10, 20, 40 or 60 seconds. At each time point, 100 µl cell suspensions were diluted in M9 medium, and survival rate after irradiation was compared to that of the non-irradiated sample.

2.2.4. Bacterial expression and purification of fusion proteins

2.2.4.1. Protein expression

The sequences encoding for AtET1 and AtET2 proteins were amplified by PCR-based method from full length cDNAs and cloned into the expression vector pQE30 (Qiagen, Hilden, Germany) at the SphI and SalI sites. The vector provides a region encoding for HIS tag sequence to fusion proteins to allow purification by Ni^{2+} affinity chromatography. *AtET*-pQE30 constructs were transformed into *E. coli* strain M15 (Qiagen, Hilden, Germany) carrying its pREP4 plasmid. The expression of fusion proteins is controlled by T5 promoter, which is induced by addition of a suitable amount of IPTG to the bacterial cultures.

Small scale screening of protein expression:

For small-scale expression, starter cultures were routinely set up in 2 ml of LB medium containing both ampicillin (50 mg/l) and kanamycin (50 mg/l) with a single colony picked off of transformants. The cultures were grown at 37°C to the mid logarithmic phase (the OD_{600} reached around 0.5) and could be used immediately or stored overnight at 4°C.

Small cultures were performed by inoculating 10 ml LB medium supplemented with 200 µl of starter cultures and the appropriate ampicillin and kanamycin. The cultures were incubated at 37°C with vigorous shaking at 200 rpm. Once the OD_{600} of bacterial cultures reached around 0.7 protein synthesis was induced by the addition of IPTG into medium at final concentration of 1.0 mM. The cultures were further incubated at different temperatures for protein synthesis. After expression, the bacterial cells were harvested by

centrifugation at 8000 rpm for 5 minutes at 4^{o}C and resuspended in 200 - 300µl of suspension buffer.

The induced bacterial cells were disrupted by ultrasonication for 80 seconds (5 seconds burst and 5 seconds cooling) on ice followed by centrifugation at 12000 rpm for 10 minutes at 4^{o}C. The supernatants containing recombinant proteins were collected and used for further analysis by SDS – PAGE and Western blot.

2.2.4.2. Protein purification:

The recombinant proteins expressed in pQE vector are purified by their His tag, which binds onto nickel-nitriloacetic acid (Ni-NTA) affinity resin (Qiagen, Hilden, Germany) according to the instruction of the manufacture. Although AtET proteins were expressed as soluble form, they could only be purified under denaturing conditions or hybrid conditions. The procedure was the same for both AtET proteins with only minor difference in the volumes of culture depending on the expression level of recombinant proteins. The bacteria were grown at 30^{o}C in 1.5 and 2 litres for AtET1 and AtET2 purifications, respectively. At the end of the induction period, the bacterial cells were harvested by centrifugation at 6000 rpm for 15 minutes.

The bacterial pellet was resuspended in 20 ml of lysis buffer containing 8 M urea or 6 M GuHCl and disrupted by sonication (Vibra cell Sonics & Materials, Schütt) on ice for a total of 8 minutes (5 seconds burst and 5 seconds cooling). Subsequently the sample was centrifuged at 12000 rpm for 30 minutes at 4^{o}C and the supernatant was collected as crude extract for protein purification. The supernatant containing soluble AtET protein was loaded onto a column previously equilibrated from 2 ml of 50% slurry of Ni-NTA. The column was washed with two bed volumes of lysis buffer containing 8M urea, four bed volumes of native washing buffer and finally the bound AtET protein was eluted in 10 ml of elution buffer containing 250 mM imidazole.

The purified AtET protein was dialysed against 1X PBS pH 7.5 buffer in membrane tubing with molecular weight cut-off (MWCO) of 12000-14000 dalton (Spectrum Laboratories, Inc., USA) and further concentrated in polyethylenglycol 6000 (Carl Roth, Germany). Protein concentrations were spectrophotometrically determined according to

Bradford's method (Bradford, 1976) using a standard curve of bovine serum albumin (BSA). The purity of proteins was analyzed by denaturing SDS-PAGE and Western blot.

2.2.5. Western blot analysis

2.2.5.1. Bacterial protein extracts

Bacterial proteins were prepared by mixing one volume of protein solution with one volume of 2X SDS loading buffer and denatured at 95^{o}C for 5 minutes. Proteins were separated on 12% SDS-polyacrylamide gel following by transferring onto nitrocellulose membranes (Schleicher & Schuell, Germany) in transfer buffer for two hours or overnight. The membrane carrying proteins was treated with Roti-Block buffer (Carl Roth, Germany) for an hour at 37^{o}C. Detection of HIS tag fusion proteins was initially performed by incubating the membrane with the primary monoclonal antibody anti-polyhistidine from Sigma. The membrane was then incubated with the second anti-mouse-ALP antibody (Sigma) for an hour at room temperature. The signals were visualized by additions of NBT and BCIP at concentration of 0.33 mg/ml and 0.165 mg/ml, respectively, as substrates for alkaline phosphatase.

2.2.5.2. Plant protein extracts

Leaves of *Arabidopsis* plants were harvested in a 1.5 ml microcentrifuge tube and immediately frozen in liquid nitrogen. Homogenization of the samples in SDS buffer was performed by using an electric drill Eurostar (IKA, Germany) fitted with a micro pestle. Total soluble proteins were collected in the supernatant phase after centrifugation and then determined concentration by Bradford's method. The soluble proteins were separated in a denaturing SDS-polyacrylamide gel and transferred onto nitrocellulose membrane as described above. Unspecific sites on the membrane were blocked by 5% non-fat dry milk in 1X Marvel buffer (blocking buffer) for 2 hours. The membrane was incubated with 9E10 anti c-myc monoclonal antibody (produced at Phytoantibody group, Department of Genetics, IPK, Gatersleben, Germany) diluted 1:50 in blocking buffer for 2 hours at room temperature. Afterward the membrane was washed 3 times for 5 minutes each with 0.5% non-fat dry milk in 1X Marvel buffer before incubation with anti mouse IgG (whole molecule) peroxidase conjugate (Sigma) diluted 1:10000 in 1X Marvel buffer containing 0.5% non-fat dry milk for an hour at room temperature. After three final washes with 1X

Marvel buffer carrying 0.5% non-fat dry milk, 1 time with 1X Marvel buffer and 1 time with PBS for 10 minutes each, the membrane was enhanced chemiluminescence with ECL Western blotting detection kit (Amersham) for 1 minute and exposed to an autoradiografic film (Hyperfilm ECL, Amersham Biosciences).

2.2.6. Electrophoretic mobility shift assay (EMSA)

Oligonucleotides (25 - 28 bp) for binding assays were obtained from Metabion (Martinsried, Germany) as single strands and prepared as a protocol described by Mönke (Mönke *et al.*, 2004). They were generated by heating complementary oligonucleotides in TES buffer for 5 minutes and slowly cooled down to room temperature to allow perfect annealing. The double-stranded DNA oligomers were amplified in pBluescrip SK+ vector and digested with EcoRI and HindIII enzymes. The digested fragments were end-filled with [α-^{32}P]-ATP (3000 Ci/mmol) (Hartmann Analytic, Braunschweig, Germany) using Klenow fragment (Fermentas, Vilnius, Lithuania).

DNA binding reactions were performed in a final volume of 24 µl containing 2x binding buffer, 0.8 - 1.0 µg purified AtET proteins, 1µg bovine serum albumin and 2 µl of labelled oligonucleotide. The mixtures were incubated at room temperature for 30 minutes for binding reactions. Free and bound DNAs were separated on 6% non-denaturing polyacrylamide gels, which were run at a constant voltage of 100V in 0,5X TBE buffer in a cold room (about 8^{o}C). Shifted bands on polyacrylamide gels were exposed to imaging plate (Fuji photo film) for 1 to 2 hours and visualized using Fujifilm FLA 5000 (Fuji, Japan) and Image Reader FLA 5000 program.

2.2.7. Plant culture and genetic manipulation

2.2.7.1. Plant growth and harvest

Arabidopsis seeds were sterilized in 70% (v/v) ethanol for 5 minutes followed by sodium hypochloride solution containing 0.05% triton X-100 for 10 minutes. After three washes with sterile double-distilled water, they were sown on agar-solidized medium consisting of Murashige and Skoog (MS) salts (Duchefa, The Netherlands), 1% sucrose and appropriate vitamins in round Petri dishes. Cultures were maintained in growth room set at 22^{o}C with a 16h photoperiod and light intensity of 70µmolm^{-2}s^{-1}.

2.2.7.2. Extraction of genomic DNA

The rapid genomic DNA extraction from plants was performed according to the protocol described by Edwards (Edwards *et al.*, 1991). Leaf tissues (~ 200 mg) were ground in liquid nitrogen into fine powder and suspended in 800 µl of extraction buffer. The suspension was centrifuged for 10 minutes at full speed in microcentrifuge and extracted with an equal volume of phenol:chloroform. Subsequently, the supernatant was collected into a new tube for precipitation of DNA by supplement of 600 µl of isopropanol. DNA was collected by centrifugation for 10 minutes, washed in 70% ethanol, and allowed to briefly air dry before resuspending in 100 µl of water or TE buffer. The concentration of DNA solution was determined by Nanodrop® ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA).

2.2.7.3. Stable transformation of *Arabidopsis* plants

Transformation of *Arabidopsis thaliana* was performed as a standard protocol described by Clough and Bent (Clough and Bent, 1998). Plants of *A. thaliana* Col and Ws ecotypes were cultivated under short day conditions (8 hours photoperiod) and transferred to long day conditions (16 hours photoperiod) and allowed to grow to the stage of inflorescence. The emerging bolts of plants were clipped to induce growth of multiple secondary bolts for having more unopened floral buds.

Agrobacterium tumefaciens strains pGV2260, pGV3101 carrying the genes of interest were cultured in LB medium supplemented appropriate antibiotics at 28^{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 the 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 hours 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 insertions of transgenes in *Arabidopsis* plants were verified by PCR.

2.2.7.4. Inducible treatment of *Arabidopsis* plants

Arabidopsis seedlings were grown on MS medium for two weeks as described in section 2.2.7.1. For treatments with DEX, seedlings carrying the *AtET*-GR transgenes were either frozen in liquid nitrogen immediately (0 hour DEX) or placed in MS liquid medium containing 30 µM DEX (dissolved in ethanol) for 3, 6, 24 and 48 hours with gentle shaking. Nontransgenic seedlings (wild type plants) were treated similarly with DEX as control. Treated seedlings were collected at different time points and frozen in liquid nitrogen and afterward stored at -80^{o}C.

For analysis of root growth and development, seeds of transgenic *AtET*-GR plants were sown on MS medium in the absence and presence of 10 µM DEX. After two weeks growing on permanent induction, roots were collected and stored for further analyses.

2.7.7.5. Transient expression of AtET-GFP fusions in protoplasts

The mesophyll protoplasts from *Arabidopsis* ecotype Col were prepared in high purity as a protocol previously described (Tiwari *et al.*, 2006). Full lengths of *AtET1* and *AtET2* genes were amplified together with their promoter regions using pfET1-F, pfET1-R and pfET2-F, pfET2-R primers, respectively. The products were introduced into pGKGWG (N9831) and pGBGWG (N9837) vectors (NASC, Nottingham, UK), respectively by gateway cloning technique (Invitrogen, Carlsbad, CA). About 10 µg plasmid carrying *AtET*-GFP fusion was transfected into *Arabidopsis* protoplasts. The vector without ccdB cassette was used as a positive control. After 24h culture at 25^{o}C in dark, the GFP signals were observed by confocal laser scanning microscope (CLSM Meta, Zeiss, Jena, Germany) at Structural Cell Biology groups, IPK, Gatersleben. The Egfp was excited at 488 nm wavelengths by an argon laser and detected between 505 and 520 nm wavelengths.

2.7.7.6. Functional promoter assay

The upstream region (1695 bp) of *AtET2* gene was amplified with proof reading polymerase enzyme using pfET2-F and proET2-R primers. For transcriptional fusion of promoter region to GFP-GUS reporter gene, the resulted product was cloned into pKGWFS7.0 vector (VIB, Ghent University, Belgium) using gateway cloning technique (Invitrogen, Carlsbad, CA).

Histochemical GUS assays for GUS activity in *Arabidopsis* transgenic plants were carried out according to the protocol described by Cheng and co-workers (Cheng *et al.*, 2003). Plant samples were incubated at 37^{o}C for at least 12 hours in GUS buffer added with 1 mM 5-bromo-4-chloro-3-indolyl ß-D-glucuronide (X-gluc). To clear chlorophyll from the green tissues, the stained plant samples were incubated overnight in 70% ethanol at 4^{o}C and then kept in 95% ethanol. GUS staining was observed and recorded by Zeiss SteREO and Axioplan2 microscopes at Gene regulation group. The GFP signal was visualized as indicated in section 2.7.7.5 at Structural Cell Biology group, IPK, Gatersleben.

2.2.8. Quantitative real time PCR experiment

2.2.8.1. RNA extraction

Arabidopsis seedlings under experimental conditions were harvested and immediately frozen in liquid nitrogen. Total RNA was isolated from 100 mg grounded plant material using RNeasy Plant Mini Kit as described in the manufacture's protocol (Qiagen, Hilden, Germany). Isolated RNA in DEPC-treated water (30-50 µl) was digested with RNase-free recombinant DNase I (Roche, Mannhein, Germany) to exclude genomic DNA contamination. Degradation of the RNA was checked by running a denaturing formaldehyde agarose gel (1% v/v) according to Sambrook and Russell (Sambrook and Russell, 2001). The concentration of total RNA was quantified prior to cDNA synthesis by Nanodrop® ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA) at a wavelength of 260 nm. Purity of RNA was assessed by the ratio of the absorbance values at 260 nm and 280 nm, wherein a ratio of about 2.0 was considered a good indication of purity.

2.2.8.2. cDNA synthesis

The first trand cDNA was synthesized by reverse transcription from cleaned total RNA using RevertAid H Minus First Strand cDNA synthesis Kit (Fermentas, Vilnius, Lithuania). 1µg of total RNA and 1µl oligo(dT)$_{18}$ primer were added to each tube to obtain a total volume of 11 µl. Priming was carried out at 70^{o}C for 5 minutes, then rapidly cooled on ice. Thereafter, 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°C for 5 minutes and supplemented 1 µl of Reverse transcriptase (200 units/µl), followed by placing at 42°C for an hour. For inactivation of the enzyme activity, the reaction mixture was heated at 70°C for 10 minutes and finally stored at -20°C for further uses.

2.2.8.3. Real time PCR

cDNAs from the treated *Arabidopsis* and wild type plants were used in the real time PCR experiment. Real time PCR was performed in the ABI 7900 HT (Applied Biosystems, USA), using default thermocycle program for all genes:

50°C for 2 minutes
95°C for 10 minutes
95°C for 15 seconds
60°C for 1 minute } 40 cycles

Individual real time PCR reactions were carried out in a 384-well clear optical reaction plate (Applied Biosystems, USA) with 10 µl final volume per well. Each sample was run in triplicate for each gene to be assayed. The components of a reaction were as follow:

Components	Volume	Working concentration
2X SYBR Green PCR master mix	5 µl	1x
Primer mix	1 µl	50 mM (each primer)
Template	0.1 µl	25 ng
Water	3.9 µl	
Total	10 µl	

Arabidopsis ubiquitin (UBQ10, At4g05320) was used as a reference gene for all real time PCR experiments.

2.2.8.4. Real time PCR data analysis

There are several methods of reporting quantitative gene expression including presentation data as absolute or relative expression levels. Relative gene expression presents the data of the gene of interest relative to some calibrator or internal control genes.

A widely used method to present relative gene expression is the comparative C_T method also referred as $2^{-\Delta\Delta Ct}$ method or ΔΔCt method (Livak and Schmittgen, 2001; Xiong *et al.*, 2006; Schmittgen and Livak, 2008; Yuan *et al.*, 2008).

In this method, a house-keeping gene for instance ubiquitin (UBQ) was used as the endogenous reference gene for target genes. The data were analyzed using an equation

$$2^{-\Delta\Delta Ct} = (C_{t,\text{ target gene}} - C_{t,\text{ hp gene}})_{\text{ time x}} - (C_{t,\text{ target gene}} - C_{t,\text{ hp gene}})_{\text{ time 0}}$$

$$\text{Fold change} = 2^{-\Delta\Delta Ct}$$

Where hp = house-keeping

time x is any time point and time 0 represents the 1x expression of target gene normalized to house-keeping gene.

For the $2^{-\Delta\Delta Ct}$ calculation to be valid, the amplification efficiencies of the target and the endogenous reference must be approximately equal. Amplification efficiencies can be established by completing replicates of the same reactions using a dilution series of cDNA as templates.

2.2.9. Macroarray experiments

Isolattion of mRNAs

Three different *Arabidopsis* plant samples were used in macroarray experiments including homozygous *AtET1*::GR, homozygous *AtET2*::GR and wild type plants. After 2 weeks growing on solid MS medium with and without appropriate antibiotics, *Arabidopsis* seedlings were collected and immediately frozen in liquid nitrogen for isolation of RNAs.

Total RNA was extracted from *Arabidopsis* plants by using RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol for small-scale RNA isolation (see 2.2.8.1). Purification of mRNAs from total RNA preparation was performed using oligo $(dT)_{25}$ Dynabeads (Dynal, Oslo, Norway) as described by the manufacturer. mRNAs carrying their poly$(A)^+$ tails were captured onto oligo $(dT)_{25}$ Dynabeads while nonbinding RNAs were washed away. The mRNA - oligo $(dT)_{25}$ Dynabeads complexes were subjected to cDNA synthesis.

cDNA synthesis and labelling

First-strand cDNAs were synthesized directly on the magnetic beads by reverse-transcriptase PCR (RT-PCR) using oligo $(dT)_{25}$ Dynabeads as primers. A reaction was carried out in 1x RT buffer containing 50 mM Tris-HCl, 50 mM KCl, 10 mM $MgCl_2$, 50 units of ribonuclease inhibitor, 1 mM DTT and 0.5 mM dNTPs in a final volume of 50 µl. The reaction was preheated at 42^{o}C for 2 minutes for primer extension and subsequently added with 1 µl of reverse transcriptase and polymerized at 42^{o}C for an hour. After synthesis of first-strand cDNA, the magnetic beads were washed twice with elution buffer at 95^{o}C for 2 minutes in order to denature poly$(A)^+$ RNA, which was removed from the first-strand cDNA couped to Dynabeads by magnetic separation. cDNA was used as the template in labeling reaction using Klenow polymerase and [α-^{32}P]-dCTP (Hartmann Analytic, Braunschweig, Germany).

Hybridization

REGIA (Regulatory Gene Initiative in *Arabidopsis*) consortium filters containing 1200 transcription factors were pretreated by washing in 0.1x SSC, 0.1% SDS (w/v) for 15 minutes at 65^{o}C and subsequently denatured in 0.4M NaOH, 0.1% SDS (w/v). After washing twice in 6x SSC for 15 minutes each at room temperature, the filters were pre-hybridized in preheated Church buffer at 65^{o}C for at least 2 hours. Labeled cDNA was denatured at 95^{o}C for 3 minutes and added to the roller flasks for hybridization for at least 14 hours. In the following step, the filters were washed twice for 20 minutes each in 2x SSC, 0.1% SDS; 1x SSC, 0.1% SDS and 0.5x SSC, 0.1% SDS respectively. Afterward, the filters were exposed to phospho-imager plates for two or three days. The signals were detected using Fujifilm FLA 5000 (Fuji, Japan) and Image Reader FLA 5000 program.

Data analysis

The signal intensities for all sports on an array filter were determined using the imaging processing software ArrayVision 5.1 (Imaging Research Inc., Ontario, Canada).

2.2.10. Nanobody screening and purification

A phage library was screened for specific recombinant nanobodies against AtET proteins according to a protocol described by Gahrtz and Conrad (Gahrtz and Conrad, 2009). The purified AtET proteins were used as antigens on microtiter plate (Maxisorp, Nunc). At the third round of panning, individual colonies from titration plates were

selected and inoculated in 100 µl of 2x TY medium added 100 µl/ml ampicillin and 1% glucose at 37^{o}C overnight with shaking. Afterwards, the plates were supplemented glycerol to every clone up to a final concentration of 15% for storage at -80^{o}C for further analyses or could be tested soluble nanobodies by ELISA.

The selected clones carrying nanobodies were grown in 2x TY medium containing ampicillin and glucose at indicated above and induce protein syntheses by the addition of 1 mM IPTG. Periplasmic fractions containing nanobodies were prepared by osmotic shock method (Kipriyanov, 2002). The HIS tag-containing nanobodies were purified by immobilized metal affinity chromatography on Ni-NTA agarose (Qiagen, Hilden, Germany).

2.2.11. ELISA

96-well ELISA microtiter plate (Nunc) was coated overnight at room temperature with 100 µl of 10 µg ml^{-1} of the required antigen (AtET1 or AtET2) in PBS. After washing twice with PBS, 200 µl of 3% BSA diluted in PBS were added to the wells for blocking at 25^{o}C for 2 hours. The plate was washed 3 times with PBS followed by adding 50 µl of each bacterial supernatant containing soluble nanobody and incubating for an hour. To remove unbound nanobody, the plate was washed 5 times with PBST and twice with PBS. In the next step, 100 µl of monoclonal anti-c-myc antibody 9E10 diluted in 1% BSA were supplemented per well of the plate and incubated for an hour. The plate was washed 3 times with PBST, twice with PBS and added 100 µl rabbit of anti-mouse IgG-ALP diluted 1:2000. The plate was incubated for an hour and then removed the conjugate by washing as step before. To develop signals, 100 µl of p-nitrophenylphosphate solution (concentration of 1 ng/ml) diluted in 0.1 M diethanolamine-HCl (pH 9.8) were added per well and incubated at 37^{o}C for 15 minutes to an hour. Signals were measured at OD_{405nm} by an ELISA plate reader (SpectraFluor Plus, Tecan) and the obtained data were analyzed by Microsoft Excel.

2.2.12. Screening and verifying for T-DNA insertion mutant lines

The *Arabidopsis* Knock-out Facility (AKF) population at the University of Wisconsin, Madison was used to screen T-DNA insertion into *AtET2* gene in the Wassilewskija (Ws) background. The population lines were transformed with a derivative of the T-DNA vector pD991: 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) incorporated with GET2-F or GET2-R primer.

T-DNA insertion lines of *AtET1* in the Columbia (Col) background were identified from the Salk Institute collection of T-DNA lines transformed with derivative of pBIN19 vector: pROK2 (http://signal.salk.edu/cgi-bin/tdnaexpress) by PCR, using a primer LBa1 (O'Malley *et al.*, 2007) specific for the T-DNA in combination with either a forward primer (GET1-F1, GET1-F2) or reverse primer (GET1-R1, GET1-R2) (see section 2.1.6 for primer sequences).

2.2.13. Alexander and DAPI stainings

For the phenotypic analysis of pollens, anthers from mature flowers just beginning to dehisce were isolated and brushed on a microscope slide. A few small drops of Alexander solution (Alexander, 1969) were added on the slide containing pollens from mutant and wild type plants for staining. Stained pollens were visualized under the Zeiss Axioplan2 microscope to check pollen viability.

Analysis of mature pollens with DAPI was performed as previously described (Park *et al.*, 1998). Pollens were stained with DAPI staining solution and viewed by UV epi-illumination using Zeiss Axioplan2 microscope (Zeiss, Jena, Germany).

2.2.14. Semi-quantitative RT-PCR analysis

Total RNA was isolated from *Arabidopsis* plants using RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the instructions of supplier. After digestion with RNase-free DNase I (Roche, Mannhein, Germany), the treated RNAs were used as template for reverse transcription as described in qPCR experiment (section 2.2.8.2).

Constitutively expressed actin 2 gene (At3g18780) was used as an internal control of the RT reaction efficiency and to confirm the equal amounts and the integrity of RNA used in the RT reactions. The PCRs were carried out with Dream taq polymerase (Fermentas, Vilnius, Lithuania) and two primer pairs ET1-RT-F1, ET1-RT-R1 and ET1-RT-F2, ET1-RT-R2, respectively.

2.2.15. Generation of *AtET1::RNAi* in *et2-1* mutant line

Fragments of 360 and 300 bp corresponding to the C-terminal regions of *AtET1* were amplified using primer pair iET1-F, iET1-R and i2ET1-F, i2ET1-R, respectively. The obtained fragments were first introduce into pDONR207 to verify proper sequences and subsequently cloned into the binary vector pAGRIKOLA by gateway technique (Invitrogen, Carlsbad, CA). Transformation into *et2-1* KO mutant line was done with *Agrobacterium tumefaciens* strain pGV3101 using a standard protocol (Clough and Bent, 1998) described in section 2.2.7.3.

The transgenic plants were selected on soil by spraying herbicide basta solution onto 7-day-old seedlings and repeated four times at two-day intervals. Transgenic plants were readily identified at the end of the basta selection. Untransformed plants remained small and turn yellow rapidly, whereas the resistant plants looked normal and could be transferred to new ports for further analyses.

2.2.16. Generation of a double mutant line

An *et2-1* KO mutant originally isolated in Ws ecotype (Sussman *et al.*, 2000; Ivanov, 2005; Ivanov *et al.*, 2008) was backcrossed repeatedly into Col ecotype to facilitate direct comparison with the other mutants isolated in our laboratory. The *et2-1* mutant was backcrossed 4 times into Col to generate the mutant line *et2-Col*. The homozygous *et2-Col* line was employed in production of a double mutant with *et1-1* mutant line (in Col background). Homozygous lines for the double mutant could be obtained in offspring (T2 or T3) by determination homozygote for both loci using LBa1 in combination with GET1-F1 and GET1-R1, and XR2 with GET2-F and GET2-R primers for *et1-1* and *et2-1* loci, respectively.

2.2.17. *In silico* analyses

Locus identifiers were performed with public microarray databases using web based Genevestigator (https://www.genevestigator.ethz.ch) (Zimmermann *et al.*, 2004) and AtGeneExpress (http://www.weigelworld.org/resources/microarray/AtGenExpress) (Schmid *et al.*, 2005; Kilian *et al.*, 2007). The binding motifs of DNA fragments employed in EMSA were analyzed by using PLACE (A database of plant *cis*-acting regulatory DNA element, http://www.dna.affrc.go.jp/PLACE/) (Higo *et al.*, 1998; Higo *et al.*, 1999; Fauteux and Stromvik, 2009).

3. Results

3.1. Characterization of AtET proteins

3.1.1. Heterologous expression of AtET proteins in *E. coli*

Translation products of the full length reading frames predicted that full length AtET1 and AtET2 proteins consist of 404 and 483 amino acids, respectively, while AtET3 protein comprises only 231 amino acids due to the lack of the repeat region. The global comparison using MAFFT (Multiple Alignment with Fast Fourier Transform, http://align.bmr.kyushu-u.ac.jp/mafft/online/server/) (Katoh *et al.*, 2009) showed an overall 40.2% identity between AtET1 and AtET2 protein sequences, especially high in ET domains (58.3% identity) but less in the GIY-YIG like domains (21.4% identity) (Figure 3). As ET proteins were originally discovered and isolated from seed cDNA expression libraries by South Western hybridization, it has been proposed that they might act as DNA-binding factors. To check these further, pure AtET proteins were required for molecular characterization such as DNA binding studies as well as for the generation of specific antibodies. I therefore introduced the coding sequences of AtET proteins (1209 and 1446 bp for AtET1 and AtET2, respectively) into several expression vectors including the tetA promoter system (pASK-IBA43plus, pASK-IBA45plus, IBA GmbH), and the T7 promoter system (pET-23a, pET-22b, Novagen). However, we failed to get recombinant proteins from these vectors due to unstable synthesis, degradation, as well as toxic effect of AtET proteins to *E. coli* strains. In the last attempt, we cloned the coding sequences into the pQE30 expression vector (Qiagen, Hilden, Germany) between the sites of the restriction enzymes *SphI* and *SalI* (Figure 4B). This commercial expression system allows the production of recombinant proteins in *E. coli* controlled by a phage T5 promoter and two *lac* operator sequences which increase *lac* repressor binding and ensure efficient repression of background expression (Liu *et al.*, 1999; Drees *et al.*, 2004) (Figure 4A). The expressed fusion protein is at the N-terminus tagged with a sequence of 6 histidines (HIS tag), able to bind to a nickel matrix. The tagged protein can be dissociated from binding sites of the matrix by reducing pH or by competition with imidazole, which displaces the tagged protein from the nickel matrix (Kneusel *et al.*, 2000; Drees *et al.*, 2008). Therefore, the purification of recombinant proteins can be carried out under native or denaturing conditions depending on the accumulation of the proteins in soluble or insoluble forms.

The pQE30 plasmids containing either the *AtET1* or the *AtET2* gene were transformed into *E. coli* M15 strain carrying the repressor plasmid pREP4 which produces high level of *lac* repressor for the tight regulation of protein synthesis. *E. coli* cultures were grown to logarithmic phase and immediately used for isopropyl β-D-thiogalactoside (IPTG) mediated induction of foreign protein synthesis.

```
AtET1     M--------------FKRDDYIRTNHDPFFSKWQGFARSMFLRKPISETAELRKTFADYS 46
AtET2     MEFGDGVSFAVVPTVFKREDYKRTKHDTVFSKWQ------------------------- 34
          *              ***:** **:**..*****

AtET1     LISRDLGPKPKILIGANEKENFREGKDLVGRNRVQG----AFQGLYEL-----SHDHGRK 97
AtET2     -----------VLIGSNDWEDFKNGKDGVGRYRVQNLPRKSCPGLYELGVAVIGQEQCRK 83
                     :***:*: *:*::*** *** ***.    :  *****     .::: **

AtET1     ---DDVLVANLGQPESIRSRLRSYSRSFAH----------------------------- 124
AtET2     LEPDIVLASYLGQAESVRSRLQRYGRSGAHLRNVNNLNDCETIESPVKAVTGGLFEDIFS 143
             * **.: ***.**:****: *.** **

AtET1     ----------------------------------------HDLLKQ----------- 130
AtET2     KGGSILYRWAPMGSKREAEATEGMLLSTFDYAWNKGSNGERRQLDLLKKLGDREFMSKRK 203
                                                  ****:

AtET1     -GLSQTILPTTQNKSDNQTEEKKSDSEEEREVSSDAAEKESNS-LPSILRLSRSRPQPVS 188
AtET2     SGISRMLFPFLRNQVGIRIKGEKHVLKEERKLTCDVDEEKSNNFLTSILKLTRSRPQPVS 263
           *:*: ::*  :*: . : : :*   :***:::.*. *::**. *.***:*:********

AtET1     EKHDDIVDESDSASACGVLLEDGTTCTTTPVKGRKRCTEHKGKRLSRVSPGIHIP--CEV 246
AtET2     DRFDE-VDGSCSDIVCGVLLEDGGCCIRSPVKGRKRCIEHKGQRVCRVSPEKQTPPKSEI 322
          ::.*: ** * *  .********  *  :******** ****:*:.****  : *  .*:

AtET1     PTVRE---CEETENICGVILPDMIRCRSKPVSRRKRCEDHKGMRVNAFFFLLNPTERDKA 303
AtET2     FTGQDHHNHKDSDVVCGVILPDMEPCNKRPVPGRKRCEDHKGMRINAFLFLLNQTDREKT 382
           * ::    :::: :********  *..:**. ***********:***:**** *:*:*:

AtET1     VNEDKSKPETST-GMNQEGSGLLCEATTKNGLPCTRSAPEGSKRCWQHKDKTLNHGSSEN 362
AtET2     VKDEKPDPESHTESIEEEALTRFCEATTKNGLPCTRSSPKGSKRCWQHKEKTSSDTSPVY 442
          *:::*..**: * .:::*.   :***************:*:*********:** .. *.

AtET1     VQSATASQVICGFKLYNGSVCEKSPVKGRKRCEEHKGMRITS 404
AtET2     FQPEAAKNVACGVKLGNGLICERSPVKGRKRCEEHKGMRIT- 483
          .*. :*.:* **.** ** :**:******************
```

Figure 3. **Alignment of amino acid sequences of full length AtET1 and AtET2 proteins**.

The deduced amino acids were compared by MAFFT (Multiple Alignment with Fast Fourier Transform) (Katoh *et al.*, 2009). Asterisks indicated identical amino acid residues in two proteins. The sequences were colored in yellow and green representing the GIY-YIG like domain and ET repeat domain, respectively.

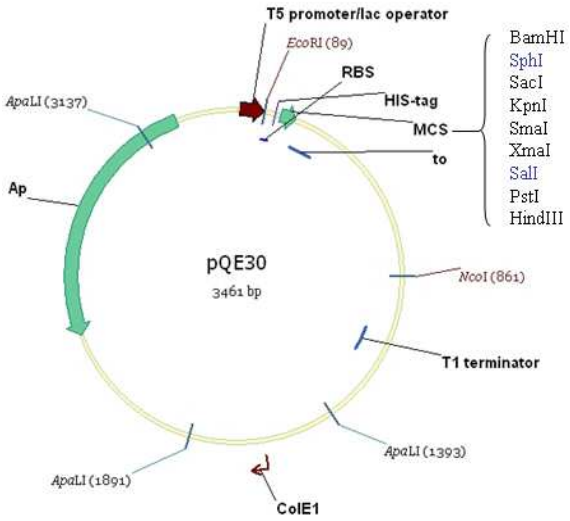


Figure 4A. Schematic drawing of the bacterial expression vector pQE30.

The DNA sequence of interest is transcribed by the *E. coli* RNA polymerase under the control of the T5 promoter (T5) and two transcriptional terminators, to and T1. HIS-tag, N-terminal His-tag sequence; RBS, synthetic ribosomal binding site; MCS, multiple cloning site; Ap, a β-lactamase gene sequence conferring resistance to ampicillin; ColE1, origin of bacterial replication.

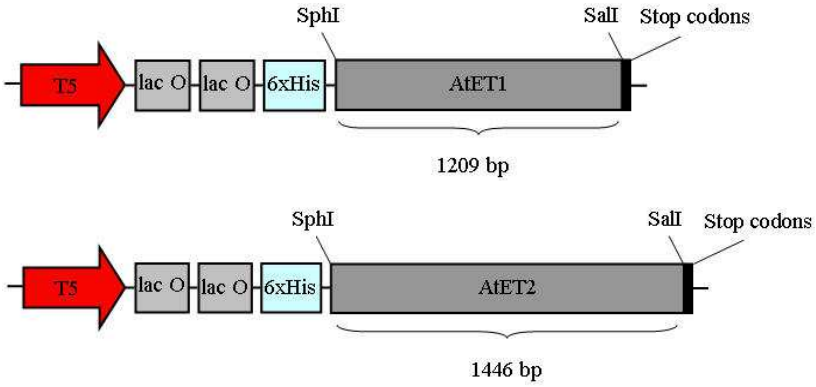


Figure 4B. Schematic representaions of the pQE30-*AtET* constructs.

The coding sequences (1209 and 1446 bp of *AtET1* and *AtET2* genes, respectively) encoding AtET proteins were cloned into pQE30 vector between the *SphI* and *SalI* sites. These constructs were introduced into the host strain *E. coli* M15 carrying the pREP4 repressor plasmid. Expression of fusion proteins was induced by adding IPTG to a final concentration of 1 mM.

A much slower growing rate after induction and the appearance of tiny clumps of bacterial cells suggested that AtET proteins may be toxic to the *E. coli*, even under various growth conditions such as lower growth temperatures and various concentrations of IPTG. Thus, *E. coli* cultures were grown under inducing conditions for 3 hours to avoid degradation of fusion proteins or other negative effects generated by the expressed proteins. Both AtET1 and AtET2 proteins were synthesized and accumulated in the cytoplasm of *E. coli* M15 strain in soluble forms. Unfortunately, these proteins could not be purified by immobilized metal affinity chromatography (IMAC) on nickel-nitrilotriacetic acid (Ni-NTA) agarose under native conditions, most likely because the HIS tag within the folding proteins may not be exposed sufficiently for binding to the Ni-NTA resin (Merits *et al.*, 2000; Debeljak *et al.*, 2006). Consequently, they were not accessible to the Ni-NTA matrix during purification. However, once the recombinant proteins were denatured by denaturants such as guanidine hydrochloride or urea, they were able to bind to the Ni-NTA matrix, providing further evidence that the lack of binding of the HIS tag to the resin depends on conformation (Kneusel *et al.*, 2000). Because under denaturing conditions, most proteins lose their biochemical activities due to disruption of three-dimensional shape, the denatured proteins have to be refolded to their native state (Chaudhuri *et al.*, 1996; Stoker, 2010).

Several small-scale tests showed that purification under hybrid conditions could be used for both AtET proteins. Briefly, purification of the recombinant AtET proteins were initiated under denaturing conditions during lysate and binding steps to Ni-NTA resin (using 6 M GuHCl) and then changed to native conditions during washing and elution steps. Using this protocol for purification, we successfully obtained the full length AtET1 and AtET2 proteins from *E. coli* M15 strain. The purity of the recombinant proteins was tested by SDS-PAGE staining with GelCode Blue or Coomassie blue and monitored by Western blot using anti-HIS tag antibody (Figure 5A and 5B). The sizes of 55 kDa and 65 kDa for fusion proteins AtET1 and AtET2, respectively, were in agreement with earlier predictions from size of the open reading frames of the cloned genes. Protein concentrations were measured by Bradford's method using bovine serum albumin as a standard (Bradford, 1976).

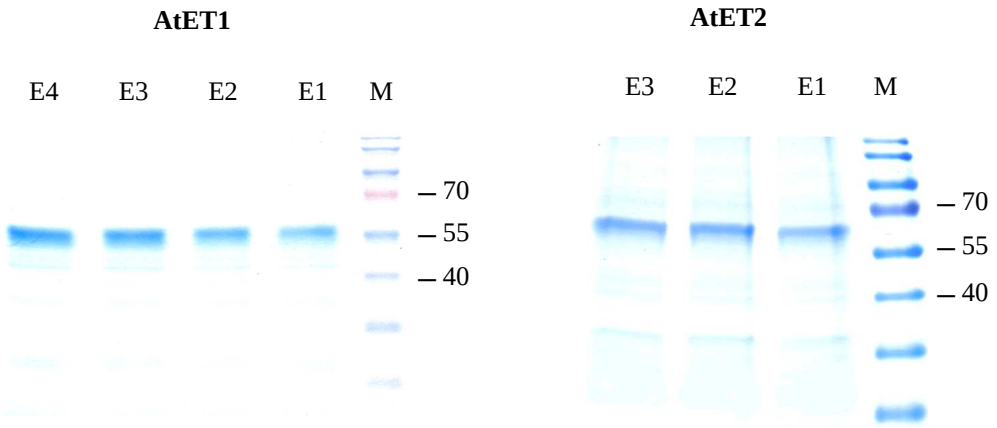


Figure 5A. SDS-PAGE analysis of purified AtET proteins.

The purified proteins were electrophoretically separated on 12% polyacrylamide gels and detected by GelCode blue stain Reagent or Coomassie brillant blue staining. M, Protein marker (Fermentas, Vilnius, Lithuania) and its molecular masses of bands were given in kDa. E, Elution proteins from the Ni-NTA resin.

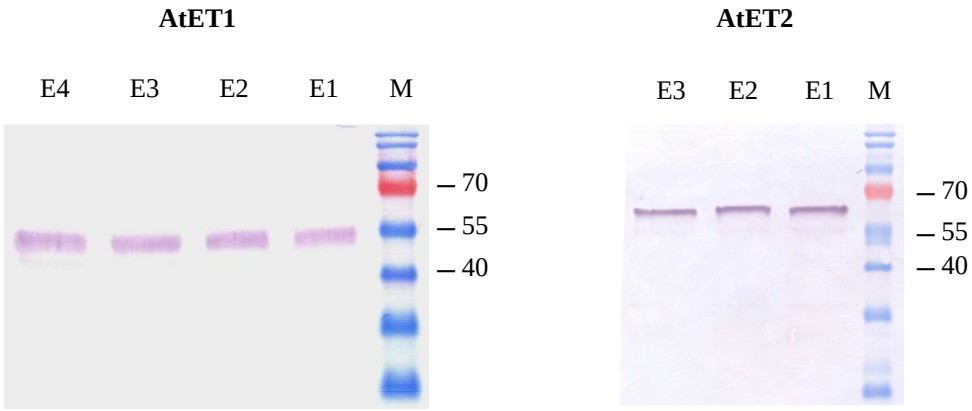


Figure 5B. Western blot analyses of purified AtET proteins with anti-HIS tag antibody. M, Protein marker (Fermentas, Vilnius, Lithuania) and its molecular masses of bands were given in kDa. E, Elution proteins from the Ni-NTA resin. Every lane was loaded with 15 µl of each elution fraction.

The identity of purified AtET proteins were additionally verified by peptide mass fingerprinting using matrix-assisted laser-desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry. After tryptic in-gel digestion, the purified AtET proteins were excised from the gels and used for sample preparations for MALDI-TOF analysis. The peptides of AtET proteins were separated based on the mass-to-charge ratio (m/z). The sequence coverage for different spots varied from 28 to 29.6% and from 20 to 23% for AtET1 and AtET2 fusion proteins, respectively (Figure 6).

```
        10         20         30         40         50         60
MRGSHHHHHH GSACFKRDDY IRTNHDPFFS KWQGFARSMF LRKPISETAE LRKTFADYSL
        70         80         90        100        110        120
ISRDLGPKPK ILIGANEKEN FREGKDLVGR NRVQGAFQGL YELSHDHGRK DDVLVANLGQ
       130        140        150        160        170        180
PESIRSRLRS YSRSFAHHDL LKQGLSQTIL PTTQNKSDNQ TEEKKSDSEE EREVSSDAAE
       190        200        210        220        230        240
KESNSLPSIL RLSRSRPQPV SEKHDDIVDE SDSASACGVL LEDGTTCTTT PVKGRKRCTE
       250        260        270        280        290        300
HKGKRLSRVS PGIHIPCEVP TVRECEETEN ICGVILPDMI RCRSKPVSRR KRCEDHKGMR
       310        320        330        340        350        360
VNAFFFLLNP TERDKAVNED KSKPETSTGM NQEGSGLLCE ATTKNGLPCT RSAPEGSKRC
       370        380        390        400        410        420
WQHKDKTLNH GSSENVQSAT ASQVICGFKL YNGSVCEKSP VKGRKRCEEH KGMRITSVDL

QPSLIS


        10         20         30         40         50         60
MRGSHHHHHH GSACEFGDGV SFAVVPTVFK REDYKRTKHD TVFSKWQVLI GSNDWEDFKN
        70         80         90        100        110        120
GKDGVGRYRV QNLPRKSCPG LYELGVAVIG QEQCRKLEPD IVLASYLGQA ESVRSRLQRY
       130        140        150        160        170        180
GRSGAHLRNV NNLNDCETIE SPVKAVTGGL FEDIFSKGGS ILYRWAPMGS KREAEATEGM
       190        200        210        220        230        240
LLSTFDYAWN KGSNGERRQL DLLKKLGDRE FMSKRKSGIS RMLFPFLRNQ VGIRIKGEKH
       250        260        270        280        290        300
VLKEERKLTC DVDEEKSNNF LTSILKLTRS RPQPVSDRFD EVDGSCSDIV CGVLLEDGGC
       310        320        330        340        350        360
CIRSPVKGRK RCIEHKGQRV CRVSPEKQTP PKSEIFTGQD HHNHKDSDVV CGVILPDMEP
       370        380        390        400        410        420
CNKRPVPGRK RCEDHKGMRI NAFLFLLNQT DREKTVKDEK PDPESHTESI EEEALTRFCE
       430        440        450        460        470        480
ATTKNGLPCT RSSPKGSKRC WQHKEKTSSD TSPVYFQPEA AKNVACGVKL GNGLICERSP
       490        500
VKGRKRCEEH KGMRITVDLQ PSLIS
```

Figure 6. Confirmation of AtET1 (upper panel) and AtET2 (lower panel) proteins expressed in *E. coli* (pQE30 vector). Regions highlighted in red indicate peptides identified by tandem mass spectrometry (MALDI-TOF).

3.1.2. Screening nanobodies from phage libraries

The basic screening procedure of phage libraries against antigens requires multiple rounds of selection to get specifically binding antibodies. Each round of selection comprises a cycle of binding of phage particles to an immobilized antigen, washing away unbound and nonspecifically bound phages, elution of bound phage and propagation of the enriched phage ready for the next round of enrichment (Vaughan *et al.*, 1996; Gao *et al.*, 2002). In the first round of panning, a pool of phages is selected displaying antibody with different specificity and affinity to an antigen. The next round is done to enrich phage particles showing higher antigenic specificity and affinity. After the last panning round single clones were isolated and tested by a suitable binding test to identify clones that produce monoclonal recombinant antibodies.

A phage library displaying 10^8 different nanobodies with randomized CDR3 was screened (Martin Giersberg and Udo Conrad, unpublished, Phytoantibody Group, Department of Genetics, IPK, Gatersleben). Two different purified AtET proteins were used as antigens for screening according to a described protocol (Gahrtz and Conrad, 2009). In each panning round, phages were titrated for pfu (plaque forming unit) in the inputs and outputs to determine the degree of selection. Titration of eluted phages showed a proliferating number of positive phage clones following each panning cycle. The number of positive phage clones specific to AtET1 and AtET2 antigens increased from 6.5×10^5 and 6.1×10^5 to 1.7×10^7 and 5.4×10^7 after the first and third panning rounds, respectively (Table 1). These data suggested that the procedure has been successfully used to enrich AtET-specific phage clones.

Table 1. Enrichment specific phages after each round of panning

Round of panning	Phage eluted against AtET1	Phage eluted against AtET2
1	6.5×10^5	6.1×10^5
2	4.6×10^6	1.5×10^7
3	1.7×10^7	5.4×10^7

After the third round 192 clones of each selection (AtET1 and AtET2) were forced to produce soluble nanobodies and to deliver them into the medium. The supernatants were tested by an indirect ELISA and specific binders were detected (15 against AtET1 and 22 against AtET2). The library was constructed from a nanobody against human TNF alpha and therefore, all 37 selected nanobodies were tested according binding to human TNF alpha and bovine serum albumin (BSA). 4 nanobodies selected against AtET1 and 2 nanobodies against AtET2 were found to be specific. To verify the correct in-frame presences of DNA fragments coding for nanobodies, the plasmid of individual positive clones was isolated and sequenced at IPK using specific primers for pIT2 vector. Sequencing of these clones revealed that there were 3 different functional DNA coding sequences, including a unique sequence for anti-AtET1 nanobody (designated as a.15) and the other two for anti-AtET2 nanobodies (designated as a.18 and a.24).

For further characterization of anti-AtET nanobodies, HB2151 cells harbouring the coding sequences of these three clones were cultured in large scale, followed by induction with 1 mM IPTG overnight at 30^{o}C. Since the HB2151 is a non-suppressor strain, the amber stop codon (TAG) placed at the C-terminal of the coding sequence and prior to the gIII is functional in this host (Hoogenboom, 1991; Suzuki *et al.*, 2005; Marcus *et al.*, 2006; Lin *et al.*, 2008). Therefore, nanobody fragments without gIII protein were synthesized in the presence of IPTG in the cytoplasm as premature proteins and then targeted into periplasm in soluble forms under the control of pelB signal sequence. During the translocation process out of the cytoplasm, the pelB is cleaved by signal peptidase to yield mature proteins - nanobodies (Fabianek *et al.*, 1997; Choi, 2000; Charlton, 2003; Choi and Lee, 2004a; Sørensen and Mortensen, 2005; Pritchard *et al.*, 2006).

Since nanobodies are accumulated in the periplasm of HB2151 cells, we applied the osmotic shock method to release proteins from bacterial cells instead of sonication (Yaagoubi *et al.*, 1994; Kipriyanov, 1998; Kipriyanov, 2002; Mergulhao *et al.*, 2005). Osmotic shock permeabilizes the outer membrane and causes the periplasmic proteins (nanobodies) to leak into the extracellular space, while the inner membrane and cytoplasm retain intact (Wilks and Slonczewski, 2007). In some cases, significant amounts of recombinant proteins were found outside of the bacterial cells due to release of periplasmic components (Jung and Choi, 1997; Choi and Lee, 2004a). By selectively extracting periplasmic material without cell lysis, we could avoid the contamination of bacterial

cytoplasmic proteins. The solution containing nanobodies prepared by osmotic shock according to the protocol described by Kipriyanov (Kipriyanov, 2002) was applied to Ni-NTA superflow affinity column for purification. After the final step of IMAC, nanobodies were found in several fractions and the protein concentrations were determined by the Bradford's method (Bradford, 1976).

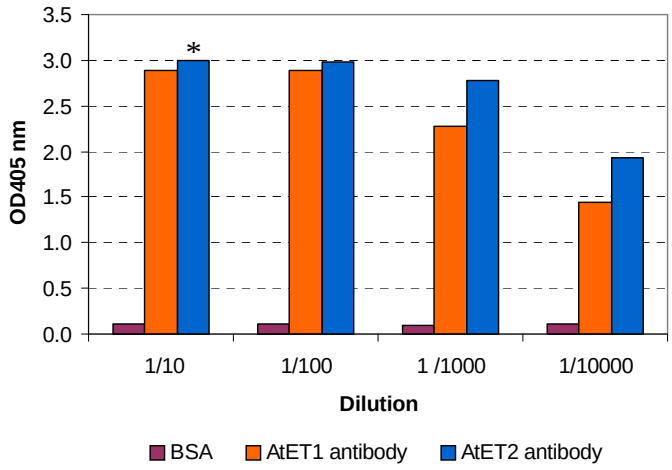


Figure 7. Binding of the nanobodies produced in *E. coli* against AtET antigens detected by ELISA. Several dilutions of nanobodies were used in these assays. BSA was employed as negative control. Every OD value was calculated from three repeated wells of the microtiter plate. The asterik indicated the OD value was higher than 3 and not measurable by an ELISA reader.

The specificity of the nanobodies obtained from screening phage libraries with regard to bind to AtET proteins was analyzed by ELISA. As shown in Figure 7, the high OD_{405} values obtained after 30 minutes incubation demonstrated that these nanobodies could specifically recognize their antigens, AtET1 and AtET2 proteins. The high binding capacity to the AtET antigens also suggested that the phage particles carrying nanobodies against AtET1 and AtET2 proteins were precisely selected. Because every nanobody expressed from pIT2 vector contains both HIS and C-myc tag at the C-terminal to facilitate detection and purification, we monitored nanobodies by SDS-PAGE analysis, and verified again by Western blot analysis with anti-HIS tag antibody. Figure 8 illustrated the pattern of nanobody a.15 against AtET1 protein obtained from superflow affinity column. We also

acquired similar patterns for a.18 and a.24 against AtET2 protein. The size of nanobody (about 17 kDa) was in accordance with a predicted size based on nanobody sequences.

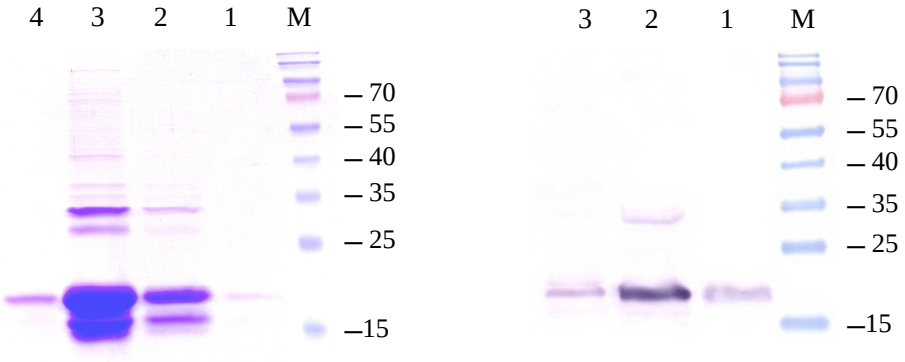


Figure 8. Analysis of the nanobody against AtET1 protein purified from *E. coli* HB2151.

Left: SDS-PAGE analysis of the purified nanobody (a.15). The soluble nanobodies from periplasm were purified by using Ni-NTA agarose. Samples from several fractions were separated on 12% polyacrylamide gel and stained with Coomassie brillant blue. Lane 1, 2, 3, 4: the last washing step, first, second and third elution fractions, respectively.

Right: Western blot analysis of the purified nanobody (a.15) using anti-HIS tag antibody. Lane 1, 2, 3: the first, second and third elution fractions, respectively.

3.1.3. Binding activity of AtET proteins to DNA fragments

Since ET factors contain the C-terminal cysteine repeats (ET repeats) and because they were originally detected and isolated in South Western hybridization, it has been suggested that they might bind to DNA. To characterize the capability of AtET proteins to bind DNA *in vitro*, we studied DNA-protein interactions by electrophoretic mobility shift assays (EMSA). This technique is based on the observation that complexes between DNA-protein and RNA-protein migrate more slowly than free nucleic acid fragments when they are subjected to non-denaturing gel electrophoresis. Since the rate of DNA/RNA migration is shifted or retarded upon protein binding, the technique is also referred to as a gel shift or gel retardation assay (Smith and Delbary-Gossart, 2000).

For the binding experiments, several double-stranded oligonucleotides were prepared from single strands according to the protocol described by Mönke and co-workers (Mönke *et al.*, 2004). The sequences of one strand of these oligonucleotides were presented in material and methods. Recombinant purified HIS tag AtET1 and AtET2 proteins were incubated with radioactively labeled double stranded oligonucleotides (probes). Subsequently, protein-DNA complexes were separated from free probes by polyacrylamide gel electrophoresis under native conditions. When the labeled O.1 probe (see section 2.1.6 in Materials and methods for more detail) was incubated with binding buffer in the absence of AtET protein (AtET1 or AtET2), the mobility of the probe was not retarded and it migrated faster through the gel to the bottom since a protein-DNA complex could not be formed (lane 2, Figure 9A, 9B). In the presence of both AtET protein and labeled probe, two distinct bands were observed in lane 3, 4, and 5 representing a free labeled probe and the DNA-protein complex. We have not seen any additional smear bands formed between shifted bands and the bands of free labeled probes. This observation in combination with clearly shifted bands indicated that AtET proteins formed stable complexes with labeled probes and could be maintained during electrophoresis. Since these probes contained core motifs from several promoters such *napA* (napin *napA*), *USP* (unknown seed protein) or *LeB4* (legumin B4) promoters, it was necessary to determine the binding motifs. Therefore, we prepared other probes (from O.4 to O.7) that each contained only one motif for binding experiments. The data from EMSA revealed that AtET proteins could strongly bind to all new probes as in the cases of O.1, O.2 and O.3.

To further delineate the nucleotide sequences essential for the DNA-protein interactions, three random probes (O.8, O.9 and O.10) were synthesized, which did not carry any known motifs monitoring by PLACE (A database of plant *cis*-acting regulatory DNA element, http://www.dna.affrc.go.jp/PLACE/) (Higo *et al.*, 1998; Higo *et al.*, 1999; Fauteux and Stromvik, 2009). Both AtET1 and AtET2 proteins showed similar binding to these random probes in comparison to the motif-containing probes. These results were in agreement with previously binding experiments based on ELISA reported by Ivanov (Ivanov, 2005). No specific sequence motif had been found for both AtET1 and AtET2 proteins. They bound to all DNA sequences containing the RY element, the gibberellin response element (GARE), and the deleted GARE. Only HRT, an ET factor from *Hordeum vulgare* has been demonstrated to specifically bind to the central sequence of GARE in the

amylase promoter (Raventos *et al.*, 1998). A recent work has reported that BnET, another ET factor from *Brassica napus*, has the capacity to interact with radioactively labeled zinc but no sequence-specific DNA-binding was determined for this factor (Ellerström *et al.*, 2005).

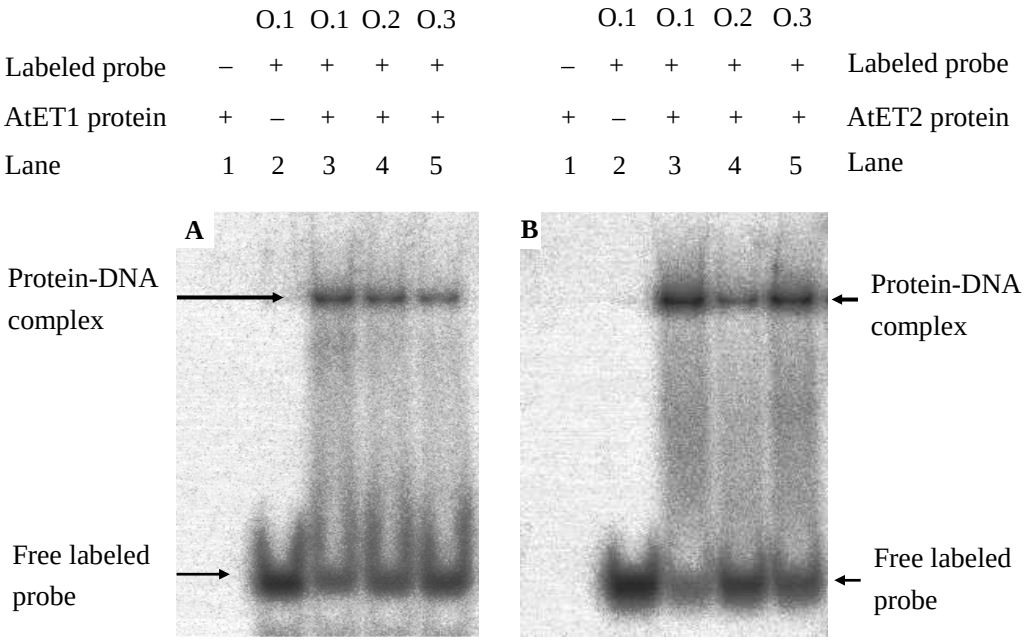


Figure 9. AtET protein-DNA interactions as revealed by EMSA.

Purified AtET proteins were incubated with the ^{32}P-labeled double-stranded oligonucleotides (probes) at room temperature. The complexes were separated from free labeled probes on 4% polyacrylamide gels under native conditions. The upper and lower bands corresponded to protein-DNA complexes and free labeled oligonucleotides, respectively.

As typical EMSA using nuclear extracts, non-specific complexes are usually observed in experiments. To prevent from non-specific complexes formed by non-specific DNA binding proteins, non-specific competitors such as poly dI-dC or Salmon sperm DNA are supplemented into binding reactions. Since we used purified recombinant proteins, the non-specific competitor was not required (Gilmartin and Bowler, 2002). Instead, we used a specific competitor (unlabeled or cold probe) to demonstrate the specificity of the interaction. The addition of unlabeled specific competitor probe O.1 at the concentration of 0.5, 1.0 and 2.5 µg resulted in a partial competition of the binding complexes (Figure 10,

lane 4, 5, 6) indicating that AtET1 protein bound to the provided hot and cold probes. The amount excess of unlabeled probe O.1 was not enough to abolish binding between AtET1 protein and labeled probe. Similar pattern of interaction was also observed when AtET2 protein was incubated with these probes (Figure 10, lane 8, 9, 10).

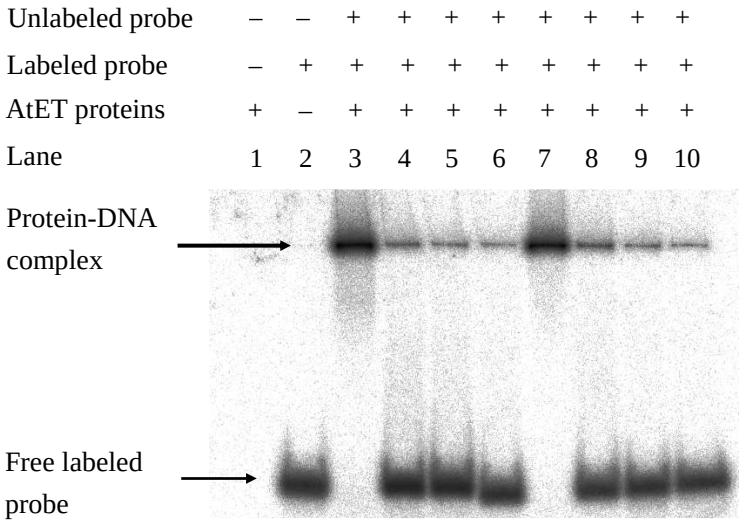


Figure 10. *In vitro* binding activites of AtET proteins in the absence and presence of unlabeled probe. Lane 1 and lane 2 were loaded only with AtET1 protein and labeled probe O.1, respectively. Lanes 3 through 6 contained AtET1 protein, labeled probe O.1 and increasing amount of unlabeled probe O.1 from 0.0, 0.5, 1.0 and 2.5 µg, respectively. Lanes 7 through 10 were similarly prepared, except that AtET1 was substituted by AtET2 protein.

In order to confirm the presence of both AtET proteins in the shifted complexes, the obtained specific nanobodies were added to the EMSA reaction mixtures containing corresponding AtET protein. The specificity of the nanobodies as well as migration positions of specific nanobody-protein complex on Western blot was confirmed by EMSA and vice versa. Figure 11B, lanes 5 showed the position of shifted band between AtET1 protein and labeled probe O.1 in comparison to protein-DNA complexes formed in the presence of specific nanobodies (lanes 6 and 7) demonstrating that nanobodies did not block complex formation.

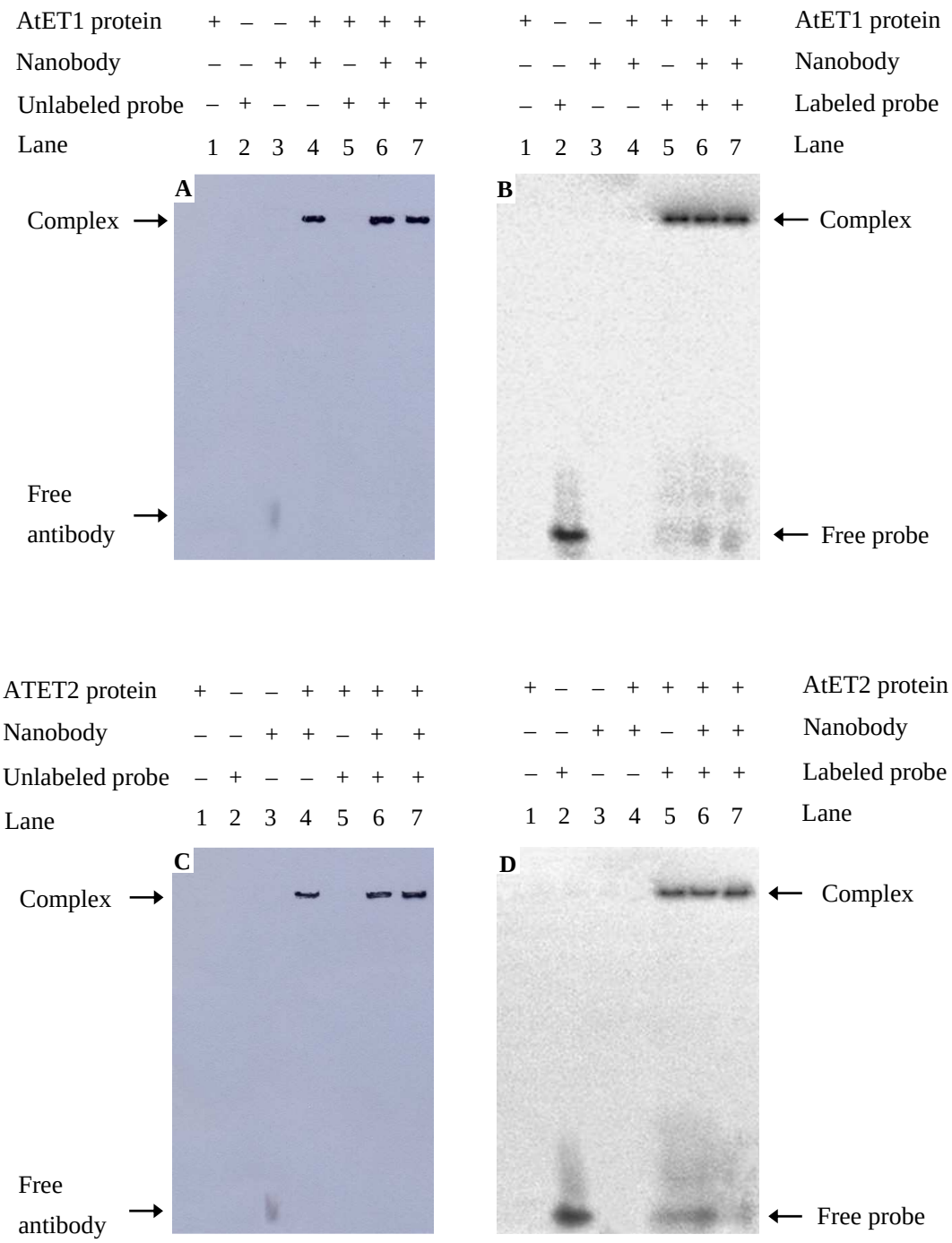


Figure 11. Analysis binding activity of AtET proteins by Western blot (A and C) and EMSA (B and D). Specific nanobodies (a.15 and a.24 against AtET1 and AtET2, respectively) were used in both Western blot and EMSA and added to the indicated reactions. Unlabeled and labeled probe O.1 was in turn employed in Western blot and EMSA. The nanobody bindings did not affect the DNA-protein complexes.

On the Western blot, shifted bands were only observed with specific nanobody independent of unlabeled probe (Figure 11A, lanes 4, 6 and 7). This observation indicated that the AtET1 protein formed complexes in binding reaction of both EMSA and Western blot. In general, the interaction of the nanobody with the protein-DNA complexes would shift the bands to higher apparent molecular weights called supershifts. In our experiments, we have not seen any supershift bands most likely due to the small molecular weights of nanobodies. Similar patterns of interaction between AtET2 protein, labeled and unlabeled probes and its specific nanobody have been observed (Figures 11C and 11D).

3.1.4. A GIY-YIG single strand cutting domain in AtET factors

In addition to the characteristic ET repeats, the AtET factors possess a low level of similarity to the DNA single strand cutting domain present in bacterial UvrC proteins and in GIY-YIG homing nucleases (Derbyshire *et al.*, 1997; Aravind *et al.*, 1999; Verhoeven *et al.*, 2000; Stoddard, 2005). This sequence designated as AtET GIY-YIG like domain is present in the second exon of all AtET proteins, in contrast to the ET repeats detected only in AtET1 and AtET2 sequences, but not in AtET3. The GIY-YIG domains identified in UvrC proteins and in homing nucleases are characterized by the presence of two short motifs “GIY” and “YIG” in the N-terminal part, followed by an arginine residue (R) in the center and a glutamine residue (Q) in the C-terminal part (Kowalski *et al.*, 1999). During the process of *E. coli* nucleotide excision repair (NER), DNA damage recognition and processing are achieved by the action of the UvrA, UvrB and mainly by the UvrC gene products. The N-terminally located GIY-YIG domain is involved in cleavage on the 3’ side, while C-terminal domain makes the 5’ incison to remove the lesion (Dunin-Horkawicz *et al.*, 2006).

Based on the three-dimensional structure of the bacterial GIY-YIG domain (Van Roey *et al.*, 2002), the similarity between the prokaryotic proteins and the plant-specific ET factors is mainly confined to two conserved β-strands and to helix 1. The variable length of the loop between both strands in the plant proteins probably does not disturb the overall structure of the protein. ET factors from barley, rice, poplar and *Arabidopsis* (AtET2, AtET3) all contain extended loops, whereas those from oilseed rape, broad bean and AtET1 have a loop length similar to those found in T4 bacteriophage, *Pseudomonas fluorescens* and *E. coli*. Further sequence similarity between the prokaryotic and plant ET

proteins resides in helix 1. The most highly conserved arginine residue is highlighted (see Figure 40, page 92 in discussion for more detail). It is well established that the replacement of this residue by alanine results in a distortion of activity (Derbyshire *et al.*, 1997; Kowalski *et al.*, 1999; Verhoeven *et al.*, 2000). Therefore, arginine plays an important role for the catalytic function of the DNA single strand cutting domain (Karakas *et al.*, 2007). The sequences that make up GIY-YIG like domains extend from residue 83 to residue 125 (43 amino acids) and from residue 64 to residue 175 (112 amino acids) in AtET1 and AtET2 proteins, respectively (Figure 3). Although these domain sequences vary in length due to the flexible loop, most conserved residues are found here. To demonstrate the functionality of the GIY-YIG-like domain of plant ET factors, the wild-type domain, as well as the arginine to alanine (R>A) replacement allele, was used to replace the corresponding GIY-YIG domain of the *E. coli* UvrC protein (Figure 12).

Plasmids encoding the chimeric proteins were transformed into the UvrC-deficient *E. coli* strain SOLR (Stratagene, La Jolla, CA). Survival rates after ultraviolet (UV) irradiation were determined in four independent experiments. It was shown that the wild-type domain can partially relieve the UV sensitivity of SOLR, whereas the R>A mutation resulted in a reduced survival rate (although still slightly greater than in the presence of an empty vector control) (Figure 13). These data demonstrate that the AtET2 GIY-YIG-like domain can productively cooperate with the C-terminal domain of the bacterial UvrC protein.

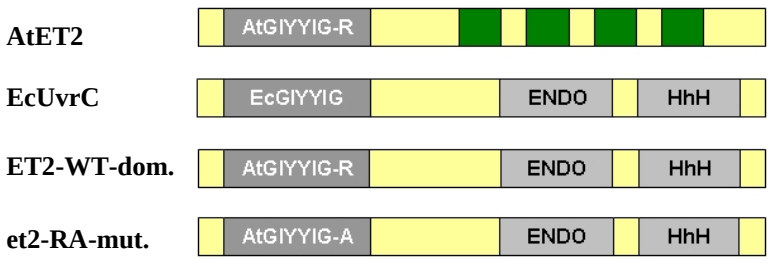


Figure 12. Schematic structure of the domain swapped chimeric proteins.

The four plant ET repeats are shown in green. GIY-YIG represents the N-terminal single strand cutting domain. ENDO and HhH indicate the C-terminal single strand cutting domain ENDO V and the Helix-hairpin-Helix domain, respectively. Domain sizes were not drawn to scale.

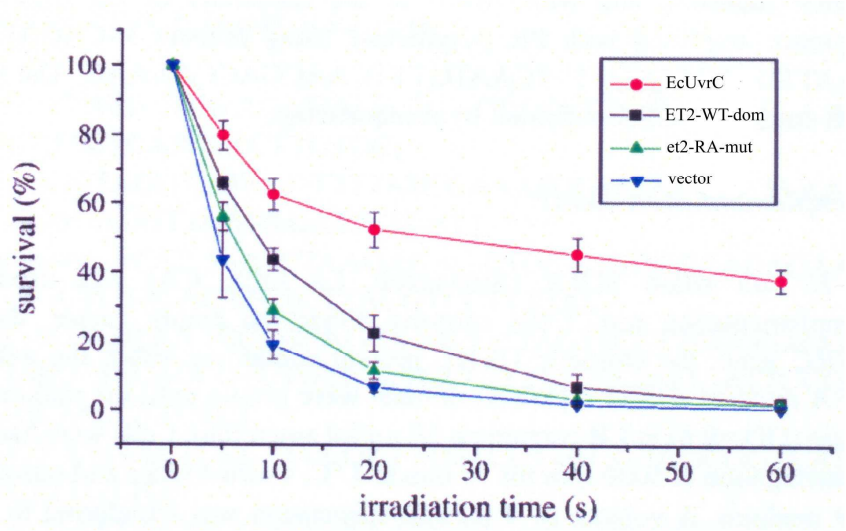


Figure 13. Functional activity assays of the GIY-YIG single strand cutting domain.

UV survival curve of various chimeric ET domain constructs. The UV-deficient *E. coli* strain SOLR was transformed with the authentic *E. coli* wild type UvrC protein (ecuvrc), the *E. coli* UvrC with its N-terminal GIY-YIG domain replaced by the corresponding wild type domain of AtET2 (etwt) and the *E. coli* UvrC protein with its N-terminal GIY-YIG domain replaced by the AtET2 domain carrying the R>A point mutation (etmut). The percent survival rate (%) was given as a function of the irradiation time in seconds. The standard deviation of the mean of four replicates was given.

3.1.5. Subcellular localization of AtET proteins

The data from EMSA experiments showed the binding of AtET proteins to provided oligonucleotides *in vitro*, implicating that they are targeted to the nuclei. Intracellular protein localization can be obtained by in frame fusion of the protein of interest to a marker protein and monitoring the expression of the fusion protein. The green fluorescent protein (GFP), originally isolated from jellyfish (*Aequorea victoria*), is the most commonly used marker protein in localization studies (Chalfie and Euskirchen, 1994; Rizzuto *et al.*, 1995).

To check the proposed function of AtET factors as transcriptional regulators and to investigate in more detail the spatial expression of *AtET* genes, I used an *Arabidopsis* protoplast system for transient expression assays. In these experiments, the full length of *AtET1* (1.60 kb) and *AtET*2 (1.85 kb) genes and their own promoter regions

(approximately 1.3 kb and 1.7 kb, respectively) were amplified from genomic DNA. Cloning into pGKGWG (N9831) and pGBGWG (N9837) vectors (NASC, Nottingham, UK) was performed with the gateway technology developed by Invitrogen (Invitrogen, Carlsbad, CA), resulting in the constructs pGKGWG-*pET1::AtET1* and pGBGWG-*pET2::AtET2*. The pGKGWG and pGBGWG vectors were derived from pGreen backbone and carry different selection marker for transgenic plants (kanamycin and BASTA, respectively) (Zhong et al., 2008) (Figure 14). The vector expressing only GFP was utilized as a transformation control. Constructs containing an inserted AtET gene in frame with GFP were transfected into protoplasts prepared from *Arabidopsis* ecotype Col for transient experiments. The transformed protoplasts were harvested after three days of incubation and used for the detection of GFP fusion proteins.

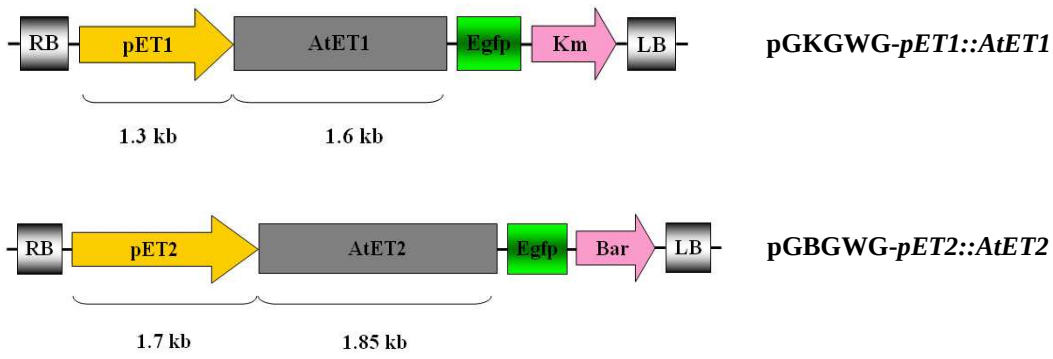


Figure 14. Schematic diagrams of the *AtET-GFP* constructs.

Approximately 2.9 and 3.6 kb of the full length *AtET1* and *AtET2* genomic loci (consisting of the corresponding genes) driven under their native promoter regions were cloned upstream of the GFP encoding sequence. RB, right border of T-DNA; pET1, pET2, *AtET1* and *AtET2* promoters; AtET1, AtET2, full length *AtET1* and *AtET2* genes: Egfp, enhanced green fluorescent protein; Km, kanamycin resistance gene; Bar, BASTA resistance gene; LB, left border of T-DNA.

As shown in Figure 15, visualization by confocal laser scanning microscope (CLSM Meta, Zeiss, Jena, Germany) between 505 and 520 nm emissions demonstrated that transfected protoplasts with a GFP control construct displayed green fluorescence distributed through the entire cytoplasm. In contrast, *Arabidopsis* protoplasts transfected with *AtET1*-GFP and *AtET2*-GFP showed accumulation of GFP only in the nuclei. The intensity of GFP emission

demonstrates the accumulation of GFP fusion proteins in the nuclei of transformed protoplasts. With both AtET-GFP fusion proteins, 15 to 20% of transfected *Arabidopsis* protoplasts showed nuclear accumulation of signals. This observation evidently indicated that AtET1 and AtET2 are nuclear proteins in *Arabidopsis*. In general, GFP seems to need a rather strong promoter to drive sufficient expression for detection and most published examples have used constitutive promoters from viruses such as cytomegalovirus (CMV), SV40 or HIV long term repeat (Tsien, 1998) and CaMV35S promoters (Harper and Stewart JR, 2000).

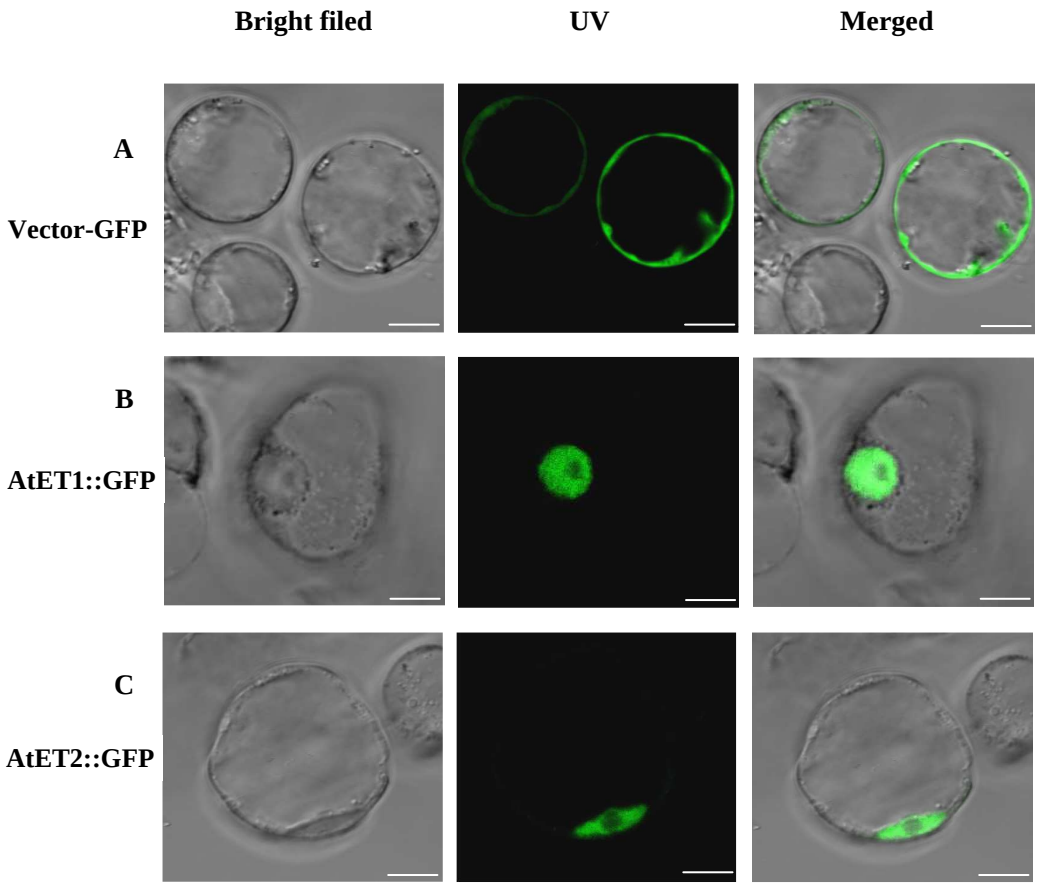


Figure 15. Nuclear localization of the AtET proteins fused to GFP in *Arabidopsis* protoplasts.

Full length sequences of *AtET1* and *AtET2* genes were fused in frame with the GFP encoding sequence and introduced into *Arabidopsis* protoplasts. The expressions of fusion proteins were driven by their native promoters (*AtET1 and AtET2* promoters, respectively). GFP signal was observed by confocal laser scanning microscopy (CLSM Meta, Zeiss, Jena, Germany). GFP alone (A) was used as a control for targeting to cytosol. (B) and (C) corresponded to AtET1 and AtET2 fused to GFP. Bars represented 10µm.

In our experiments the strength of GFP signal in these transient experiments compared to the control verified that AtET proteins were strongly synthesized under their own gene promoters. On the other hand, the observation of GFP fusion protein expression suggested that splicing of introns in the *AtET* genes occurred completely in *Arabidopsis* transient assay. No signals corresponding to the positions of mitochondria or chloroplasts have been visualized. These observations are also consistent with previous reports that AtET proteins (driven by CaMV35S promoter) could be detected in the nuclei of differentiated cells (Ivanov, 2005; Ivanov *et al.*, 2008). The use of AtET-GFP fusion proteins to monitor the accumulation of AtET proteins allowed us to clearly demonstrate the localization of these proteins in the living plant cells *in vitro* as well as to provide another evidence for their function as transcriptional regulators.

3.2. Expression pattern of *AtET* genes

3.2.1. Expression profiling data of *AtET* genes

It is possible to monitor gene expression in the *Arabidopsis* on a genome-wide scale by the data from microarrays. The ATH1 array developed by Affymetrix and The Institute for Genomic Research (TIGR) represents approximately 23750 genes from *Arabidopsis* (Redman *et al.*, 2004). Thousands of arrays based on ATH1 have been performed, of which a significant number are publicly available such as Gene Expression Omnibus (GEO) at NCBI (Edgar *et al.*, 2002), Genevestigator (Zimmermann *et al.*, 2004; Zimmermann *et al.*, 2005; Hruz *et al.*, 2008), AtGenExpress (Expression atlas of *Arabidopsis* development) (Kilian *et al.*, 2007). Taking advantage of publicly accessible ATH1 arrays, we extracted expression data for AtET gene family by using Genevestigator V3 (Zimmermann *et al.*, 2005; Hruz *et al.*, 2008).

Figure 16 presents Genevestigator data for *AtET* gene expressions for which the data were stable across many microarray experiments. Both *AtET1* and *AtET2* show similar expression pattern in vegetative tissues and organs including stems, leaves and roots as well as in early developmental stages of Col plants (hypocotyls, cotyledons). They are expressed ubiquitously in the plant, but their transcripts are accumulated more in reproductive organs such as pollen (high levels of transcripts in sperm cells) and endosperm. Successively, using data from AtGeneExpress, we could elucidate the matching expression

profiles between *AtET1* and *AtET2* genes during growth and development of *Arabidopsis* (Figure 17).

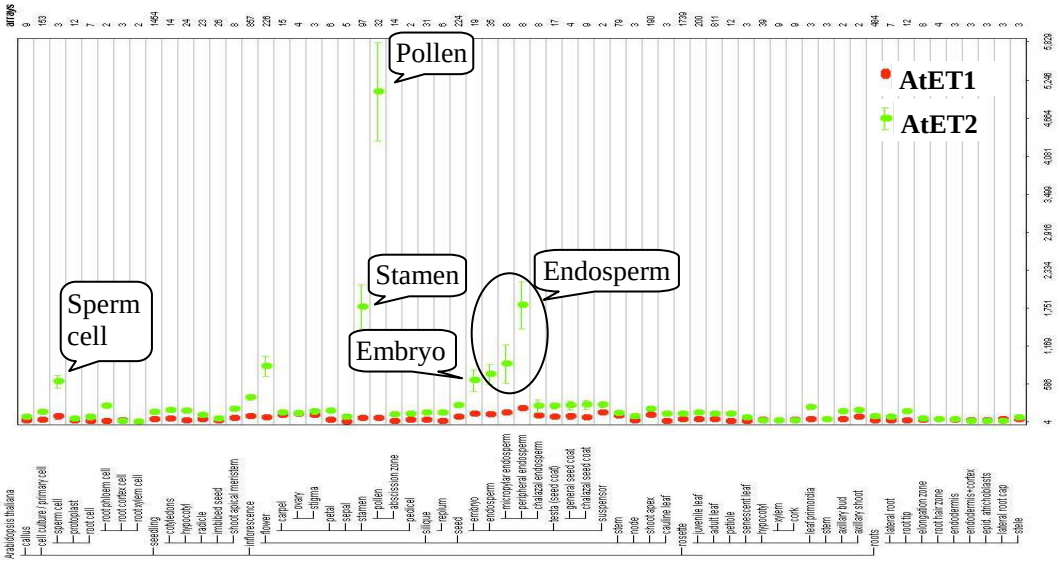


Figure 16. Transcriptional profiles of *AtET* genes from Genvestigator database.

AtET genes show similar expression patterns in most vegetative tissues with the exception of reproductive tissues. *AtET2* is highly expressed in pollen and endosperm tissues. Red and green dots correspond to *AtET1* and *AtET2*, respectively.

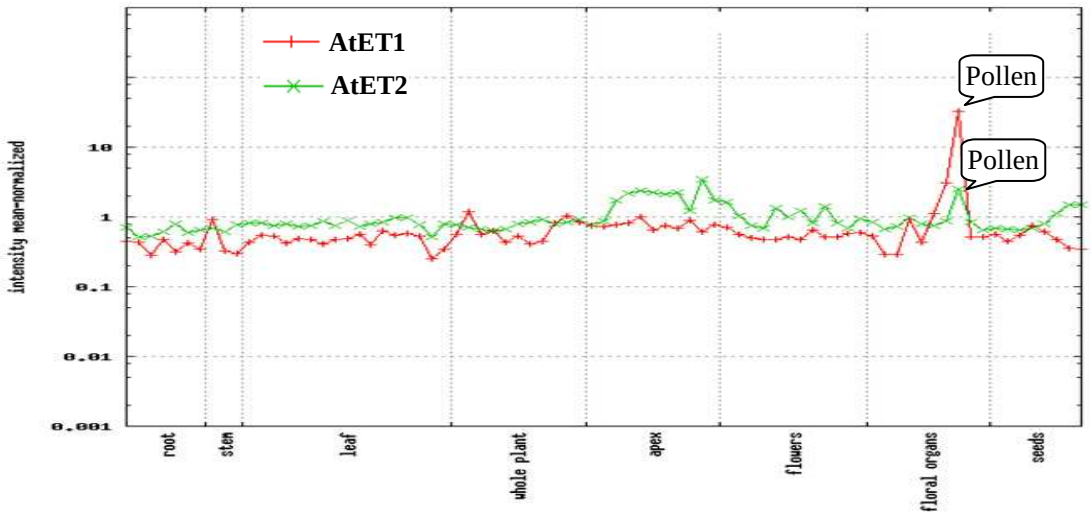


Figure 17. Transcriptional profiles of *AtET* genes analyzed by AtGeneExpress

Red and green lines represent for *AtET1* and *AtET2*, respectively.

3.2.2. Tissue-specific expression pattern of *AtET* genes

For a detailed investigation of spatial and temporal expression of AtET during plant development, I generated *Arabidopsis* reporter lines carrying in frame double reporters β-glucuronidase (GUS) and green fluorescent protein (GFP), driven by the native promoter of *AtET* genes. Promoter region is defined as the nucleotide sequence between *AtET* start codon and the coding area of upstream flanking gene. The reporters were monitored by histochemical GUS activity and by GFP detection in *Arabidopsis* plants at different stages of growth and development.

To this end, 1.7 kb promoter region of *AtET2* gene was amplified from genomic DNA of *Arabidopsis* ecotype Col and fused to the GFP and GUS reporter genes in pKGWFS7.0 vector (VIB Ghent, Belgium) (Figure 18). A suitable clone carrying the promoter sequence was confirmed by sequencing and used for transformation into *Arabidopsis* ecotype Col and Ws. A total of 30 seedlings from each ecotype of transgenic lines showing kanamycin resistance were tested for GUS activity *in situ*. Preliminary observations showed that the active patterns of *AtET2* promoter was similar in both ecotype Col and Ws, therefore we only used trangenic lines of *AtET2* promoter::GFP-GUS (pET2::GFP-GUS) of the Col background for further analyses.

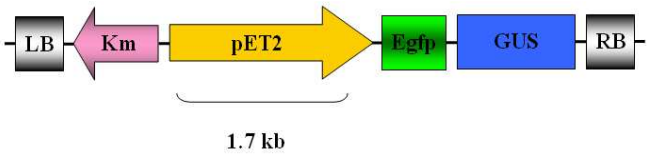


Figure 18. A schematic representation of GFP-GUS expression driven by the *AtET2* promoter.

1.7 kb of *AtET2* promoter region was cloned upstream of coding sequences for GFP-GUS reporter genes. LB, left border of T-DNA; Km, kanamycin resistance gene; pET2, AtET2 promoter; Egfp, enhanced green fluorescent protein; GUS, β-glucuronidase; RB, right border of T-DNA.

The spatial and temporal activity pattern of the pET2::GFP-GUS construct in *Arabidopsis* was analyzed during development from germination to seed stages using at least five independent-GUS positive transgenic lines. Histochemical analysis of developing seedlings (3 and 7 day-old seedlings) showed clear GUS expression in the meristematic region of shoot apices, along the central cylinder of the root and appeared to be increasing intense in the root tips. The other regions of GUS staining were verified in the expanded cotyledons where vascular tissues are initiated (Figure 19, D, E, G and H).

High expression of GUS was found in young leaves, apical meristems while reduced levels of staining were apparent in the first two true leaves as well as in the older leaves of the 15-day-old seedlings (Figure 19, F). In mature *Arabidopsis* plants, the GUS activity was obviously detectable in the veins of rosette (Figure 19, I) and cauline leaves, similar to that observed in true leaves of 15-day-old seedlings. In reproductive organs, GUS expression was observed in the vascular tissues of petals and evident in anthers and in receptacles but not in stigmas at later stages of flower development (Figure 19, J). Further analyses revealed that the *AtET2* promoter was highly active in the mature pollen and in the chalazal region of ovules (Figure 19, K, L, and M). During seed development, GUS staining was well defined in the peduncle and restricted to the vascular system of the siliques (Figure 19, N).

To verify the activity of the pET2::GFP-GUS fusion in developing leaves, stems and in ovules, we examined the GFP signal emitted from these organs by confocal laser scanning microscope (CLSM Meta, Zeiss, Jena, Germany). Transverse and longitudinal sections through the midrib of mature leaves showed GFP signal in the vascular tissues (Figure 19, O and Q). In addition, GFP also was detected in these regions of the longitudinal section through the main stem indicating that the promoter is active in vascular tissues of mature plants (Figure 19, R). In ovules, GFP signal was visualized in the chalazal tissues as observed by GUS staining (Figure 19, M and P).

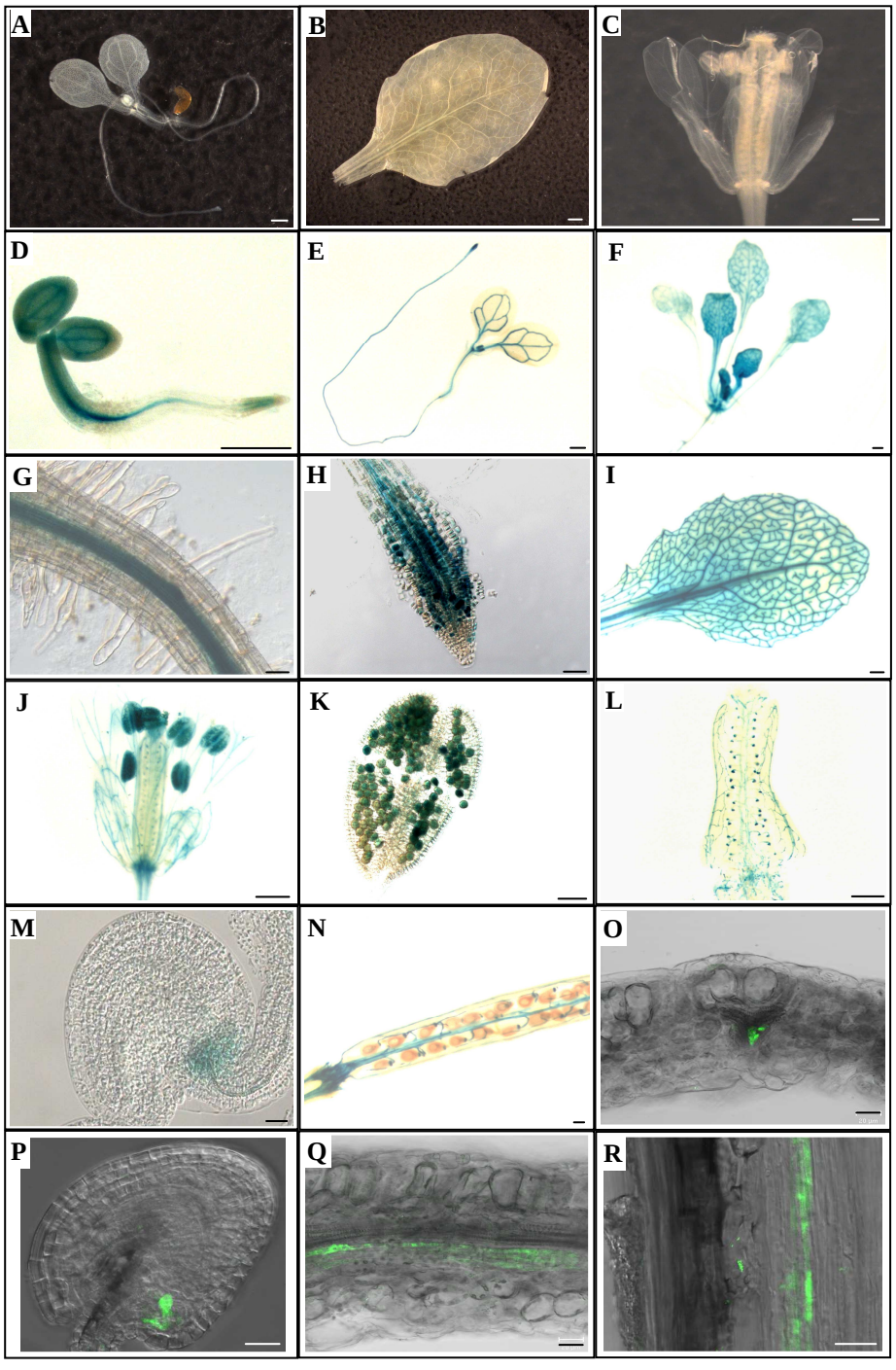

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R

Figure 19. **Pattern activity of the *AtET2* promoter in transgenic *Arabidopsis* plants.**

Histochemical localization of GUS activity in transgenic plants harbouring pET2::GFP-GUS construct. (D), (E), (F), 3, 7 and 15-day-old seedlings, respectively; (G) and (H), root fragment and root tip from seedling in (E); (I), rosette leave; (J), whole mature flower; (K), Close observation of stained pollen; (L), a pistil showing GUS activity in ovules; (M), an isolated ovule (2 days after emasculation); (N), a green silique, in contrast to the seedling (A), leave (B) and flower (C) from wild type Col plants.

GFP detection in transgenic plants by confocal laser scanning microscope. (P), a separated ovule (2 days after emasculation); (O) and (Q), free-hand transverse and longitudinal sections of the leave; (R), longitudinal section of the mature stem.

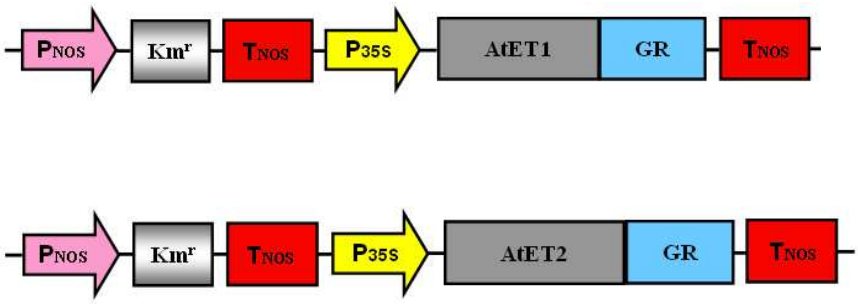
Bars represented 1000 µm (B, and I), 500 µm (A, C, D, E, F, J, L, N), 50 µm (G, H, and K), and 20 µm (M, O, P, Q, and R).

3.3. Regulated expression of AtET genes in *Arabidopsis*

3.3.1. Transgenic lines for overexpression

To control the timing of *AtET* expressions during growth and development of *Arabidopsis* plants, transgenic lines were generated using the glucocorticoid-inducible expression system (Figure 20) as described by Baudry (Baudry *et al.*, 2004). This inducible system is based on the translational fusion of the transcription factor (TF) to the rat glucocorticoid receptor domain (GR). The fusion protein is expressed under the control of a constitutive promoter (35S Cauliflower mosaic virus promoter, CaMV35S) and retained in cytosol by interaction with heat shock proteins (HSPs). After treatment with the synthetic glucocorticoid (dexamethasone, DEX), the fusion protein is released from HSP and translocated to the nucleus.

The coding regions of *AtET1* and *AtET2* genes (1212 and 1449 bp, respectively) were cloned into the p35SR1R2ΔGR vector by gateway technology (Invitrogen, Carlsbad, CA) and subsequently transformed into *Arabidopsis* Col plants. Primary transformants (T_0) were initially identified by their ability to grow on kanamycin-containing medium and then verified by PCR for the genomic integration. Segregation analysis on selective media was done with T_1 plants to obtain homozygous lines in the following generation (T_2 or T_3).



Figue 20. Schematic representation of constructs for the inducible overexpression of AtET factors.

The coding regions of *AtET* genes were fused to the rat GR domain of the p35SR1R2ΔGR vector (Baudry *et al.*, 2004). The expressions of *AtET*::GR were controlled by CaMV35S promoter (P_{35S}). P_{NOS}, nopaline synthase promoter; Km, kanamycin resistance gene; T_{NOS}, nopaline synthase terminator; GR, glucocorticoid receptor.

Several transgenic lines carrying the inducible constructs *AtET1*::GR and *AtET2*::GR were examined for transgene expressions by quantitative real-time PCR (qPCR) using gene-specific primers (qET1-F, qET1-R and qET2-F, qET2-R for *AtET1* and *AtET2*, respectively). Since induction with DEX only gives rise to translocation of *AtET*::GR fusion proteins from cytoplasm to nucleus, the expression of transgenes could be analyzed in plants without induction. Two-week-old seedlings of inducible lines grown on MS media without DEX and wild type Col were prepared for RNA isolation and subsequently for cDNA synthesis. qPCR reactions were run in ABI 7900 HT system using SYBR Green PCR master mix (Applied Biosystems, USA) and performed in triplicate for both biological and technical replicates, resulting in nine data points for each reaction. All data were normalized to expression data of a housekeeping gene UBQ10 (At4g05320) as an endogenous control. The results were analyzed by using ΔΔCt method for comparison (Livak and Schmittgen, 2001; Xiong *et al.*, 2006; Schmittgen and Livak, 2008; Yuan *et al.*, 2008). As shown in Figure 21, all transgenic lines showed increased transcript levels of *AtET1* and *AtET2* compared to wild type Col. Due to the higher expression level of transgenes, *AtET1*::GR 7-7 and *AtET2*::GR 4-1 lines were selected for production of progenies as well as further analyses.

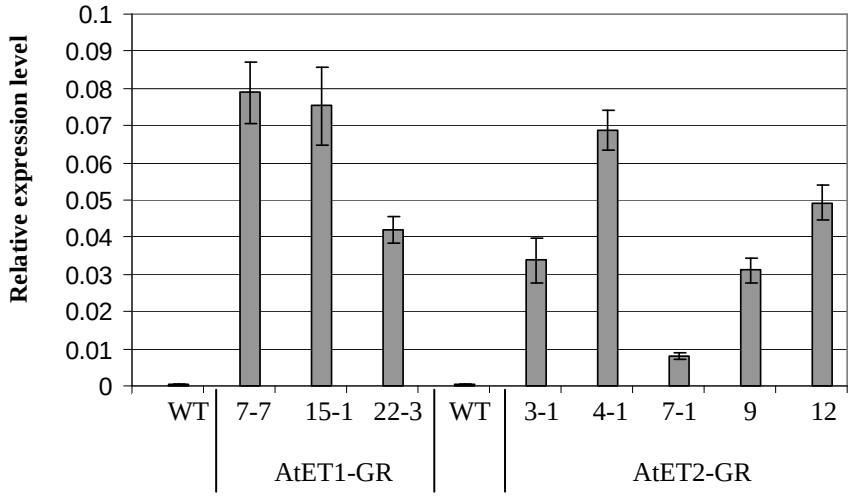


Figure 21. Relative expression levels of AtET::GR constructs in DEX-inducible over-expression lines

Expression of *AtET1* and *AtET2* genes in two-week-old seedlings of several DEX-inducible lines was measured by qPCR compared to the wild type Col samples. Each column showed the mean value of three technical replicates.

3.3.2. Phenotypes of *AtET* expressing plants

Previous data have shown that overexpression of BnET in either *Arabidopsis* or tobacco led to many phenotypic changes including dwarfism, late flowering, reduced germination rate and xylem lignification, and increased anthocyanin content (Ellerström *et al.*, 2005). To analyze phenotypes of inducible overexpression plants, the seeds of homozygous lines (*AtET1*::GR 7-7-2 and *AtET2*::GR 4-1-3) were sown on MS medium in the absence and presence of 10 µM of DEX. The morphological characteristics were compared to those of wild type Col at different developmental stages. However, no significant changes in germination rate were observed between wild type Col and *AtET*::GR lines (both 7-7-2 and 4-1-3 lines) under non-induced conditions (without DEX) and induced conditions (with 10 µM of DEX). The wild type Col and *AtET*::GR seedlings growing under non-induced and induced conditions displayed similar development of cotyledons and first true leaves within 7 days after imbibition, except for the beginning of root development. The lines overexpressing *AtET1* or *AtET2* showed short main roots

under induced conditions in comparison to the wild type Col and corresponding seedlings under non-induced conditions (Figure 22). Nevertheless, they exhibited a root length identical to the wild type Col and non-induced seedlings within two weeks of continued induction. In an attempt to investigate the influence of AtET overexpression on root growth and development, we applied array hybridization for detection of induced transcription factors (TFs).

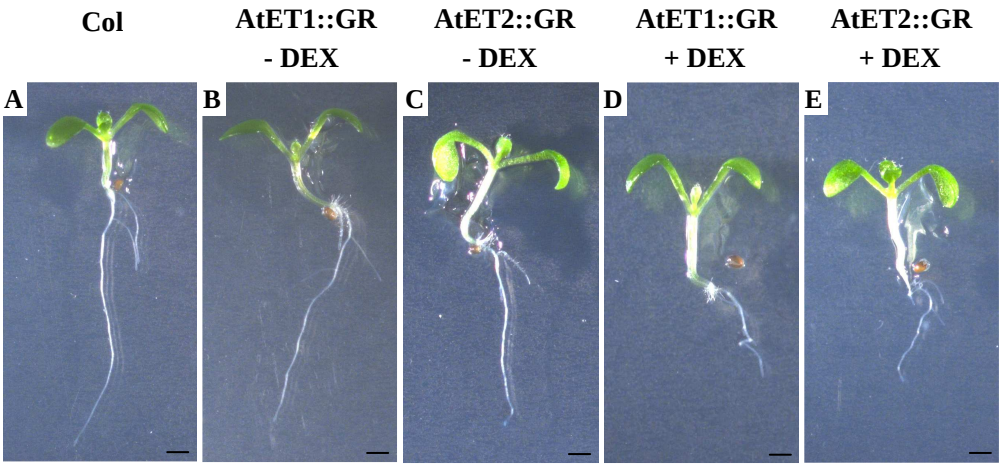


Figure 22. Phenotypes of induced *AtET*::GR seedlings.

Seeds of wild type Col, *AtET1*::GR 7-7-2 and *AtET2*::GR 4-1-3 were sown on MS medium in the absence and presence of 10 µM of DEX. 7-day-old transgenic seedlings carrying *AtET1*::GR (B) and *AtET2*::GR (C) looked like wild type (A). Induced seedlings containing *AtET1*::GR (D) and *AtET2*::GR (E) exhibited growth retardation of roots after 7 days. Bars represented 1000 µm.

The seeds from inducible overexpression line (*AtET2*::GR 4-1-3) were placed on vertical plates containing MS medium to facilitate root growth. Samples were prepared from roots of 7-day-old seedlings and used for macroarray with REGIA filter carrying around 1200 transcription factors. A list of TFs detected by macroarray experiments was annotated according to TAIR (Garcia-Hernandez *et al.*, 2002; Rhee *et al.*, 2003) and given in Table 2. Up-regulated TFs might be involved in several development processes since ethylene response factor (ERF), WRKY18 and zinc finger proteins were induced (Laity *et al.*, 2001; Xu *et al.*, 2006; Li *et al.*, 2009). One of them, the NAC-LIKE transcription factor (At1g69490) is up-regulated in the quiescent center and might be included in lateral root

formation (Guo and Gan, 2006). In sum, these TFs could potentially function downstream of AtET2 either directly or indirectly during root development. Further analyses will be required to test this hypothesis.

Table 2: TF genes up- or down-regulation in AtET2::GR seedling treatment

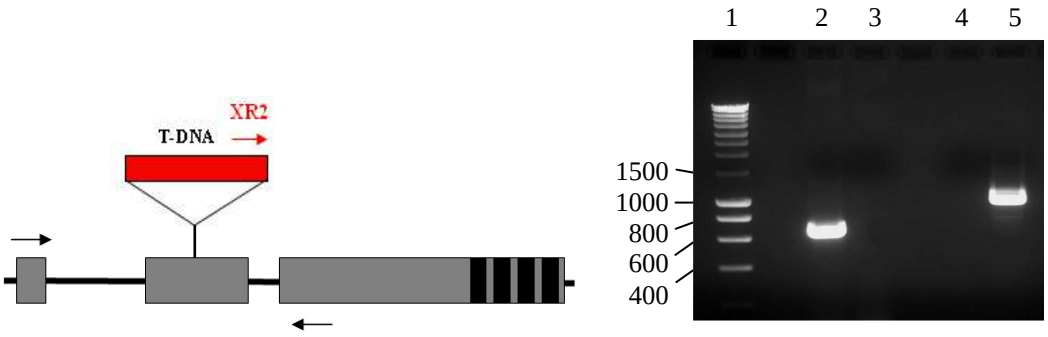
ATG number	Annotation	Induction factor	
		1st filter	2nd filter
At1g28310	zinc finger domain-containing protein	3.62	4.31
At1g69490	NAC-LIKE transcription factor	6.66	4.73
At2g18670	zinc finger protein	3.96	3.51
At4g18170	WRKY18 transcription factor	3.03	3.59
At4g39070	Zinc finger (B-box type) family protein	9.38	3.31
At5g61590	ERF (ethylene response factor)	3.01	6.51
		Repression factor	
		1st filter	2nd filter
At2g44940	AP2 domain-containing protein	9.38	12.10
At2g46510	bHLH18 transcription factor	10.57	6.34
At3g25730	AP2 domain-containing transcription factor	6.60	6.84
At5g37260	a MYB family transcription factor	5.53	12.12
At5g44080	bZIP13 transcription factor family protein	4.36	5.78

3.4. Molecular characterization of T-DNA insertions in *AtET* genes

3.4.1. Verification of T-DNA insertion in line *et2* after backcrosses

To identify loss-of function alleles of *AtET2* (At5g56780), we screened accessible T-DNA mutant collections on the background of both Col and Ws accessions. Searching from the GABI-Kat database (http://www.gabi-kat.de) revealed the availability of several mutant lines (GABI_559E11, GABI_606A02, GABI_586H02) in the Col background with insertions in the exons of the *AtET2* gene. However, these lines did not pass the quality control in the T2 generation at GABI upon our request, thus they were unrecovered for

further work. Previously, a single line containing a T-DNA insertion in *AtET2* gene was selected by pool screening from the collection of the *Arabidopsis* Knock-out Facility (AKF), University of Wisconsin (Ivanov, 2005). This line was transformed with a derivative of the T-DNA vector pD991 into Wassilewskija2 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 to facilitate direct comparison with other mutants, all in the Col background. Therefore, the line *et2-1* was backcrossed at least four times before use for further analyses and renamed to *et2-Col*.



	518
AtET2 2nd exon:	TGTCCGGGTCTGTATGAGCT
T-DNA RB in *et2-1*:	TATTCGGGCCTAACTTTGGTGTGTCCGGGTCTGTATGAGCT

.

Figure 23. Genomic organization of the *AtET2* gene showing the T-DNA insertion in the *et2-1* allele.

A. Upper left: Structure of *AtET2* gene with T-DNA insertion site. Exons were indicated in grey boxes, and primers for genotyping in small arrows.

B. Upper right: Confirmation of homozygous knock-out line after backcrosses into Col. XR2 and gene-specific primers were used for line *et2-Col* (lane 2 and 3), and Col (lane 4 and 5), respectively, and would produce 963 bp of wild type allele and around 700 bp of T-DNA flanking region. The lack of wild type allele product demonstrated that line *et2-Col* was homozygous for the T-DNA insertion.
SmartLadder DNA (Eurogentec, Belgium) given in bp was used as a marker.

C. Lower: T-DNA RB (right boder) written in red was inserted at position 518 relative to the start codon.

Homozygous plants for the insertion in *et2-Col* were screened and confirmed through PCR analysis using T-DNA right border primer (XR2) in combination with either a forward or a reverse gene-specific primers (GET2-F or GET2-R, respectively). The wild type Col serving as control should generate a PCR product (963 bp) with the gene-specific primers but not with the T-DNA primer XR2. 30 plants were genotyped in order to select an *et2-Col* homozygous line. As shown in Figure 23B, only the indicated line *et2-Col* produced a band of approximately 700 bp with mutant primers, suggesting that it is homozygous for T-DNA insertion and can be used for further analysis. The result obtained from sequencing of the right border and the T-DNA flanking region revealed that T-DNA was inserted in the second exon at the position of nucleotide 518 relative to the start codon.

3.4.2. Isolation of insertional mutant lines for *AtET1*

Arabidopsis lines containing T-DNA insertions in *AtET1* gene (At4g26170) were identified and obtained from the SIGnAL T-DNA collection (Salk Institute Genomic Analysis Laboratory). According to the sequence data found in the database (http://signal.salk.edu/cgi-bin/tdnaexpress), the insertions in these lines were predicted in the exon (SALK_000422) or the introns (SALK_026258, 094357, 146916). We selected these four mutant lines for further analysis and renamed them *et1-1*, *et1-2*, *et1-3* and *et1-4*, respectively (Figure 24).

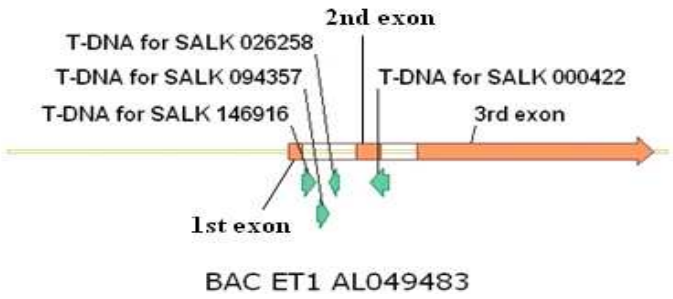


Figure 24. Positions of the T-DNA insertion in potential mutant lines for *AtET1* gene in the SIGnAL T-DNA collection. Four SALK lines including 000422, 026258, 094357 and 146916 were selected for primarily mutant analyses and designated as line *et1-1*, *et1-2*, *et1-3* and *et1-4*, respectively.

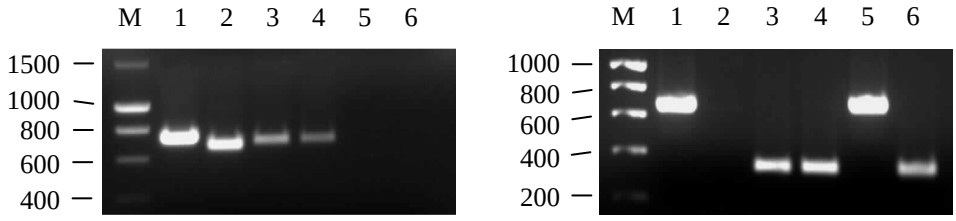


Figure 25. Genotyping of *et1* lines containing T-DNA insertions in the *AtET1* gene.

In both pictures, PCR products amplified from *et1-1*, *et1-2*, *et1-3* and *et1-4* lines were loaded in lane 1, 2, 3, 4, respectively. Lane 5 and 6 contained products generated from genomic DNA of Col. SmartLadder DNA (Eurogentec, Belgium) given in bp was used as a marker.

Left picture: Left border primer (LBa1) was used in combination with gene-specific primer GET1-F1 for *et1-1* and *et1-2* or GET1-R1 for *et1-3* and *et1-4*.

Right picture: Two pairs of gene-specific primers (GET1-F1 in association with GET1-R2 or GET1-R1) were used to amplify wild type alleles and would produce products of 681 and 343 bp. The absence of wild type allele in *et1-2* (lane 2) demonstrated that this line was homozygous for T-DNA insertions.

About 30 *Arabidopsis* plants for each mutant line were genotyped in the first generation to verify the insertions and determine whether the line was heterozygous or homozygous. To this end, the genomic DNAs isolated from these plants were used for PCR with T-DNA left border primer (LBa1) and gene-specific primers (GET1-F1, GET1-R1, GET1-R2) depending on the position of T-DNA in *AtET1* sequence. The sizes of amplified PCR products were estimated by gel electrophoresis and were around 790, 700, 770 and 750 bp for *et1-1, et1-2, et1-3,* and *et1-4*, respectively, as seen in the Figure 25 (left picture). The wild type Col was used as control. Homozygous knock-out lines were detected by PCR and were expected to have a band with T-DNA primers and not with the gene-specific primers. The results indicated that only line *et1-2* was homozygous for T-DNA insertion in the *AtET1* gene whereas the other three lines (*et1-1, et1-3* and *et1-4*) were heterozygous and therefore require the screening of the following generations for homozygosity.

3.4.3. Analysis of T-DNA integration sites in *et1*

To precisely identify the integration sites of T-DNA insertions in the *AtET1* locus, the flanking sequences of T-DNA were amplified with LBa1 and gene-specific primers and cloned into pCR®II vector (Invitrogen, Carlsbad, CA). In each case, two independent clones for a particular PCR product were selected to exclude PCR or sequencing error and entirely sequenced at Plant Genome Resources Center, IPK, Gatersleben. The sequencing of genomic DNA flanking regions revealed three lines which harbored T-DNA insertions in the first intron for line *et1-2*, *et1-3*, and *et1-4*, respectively, downstream of the start codon of *AtET1*.

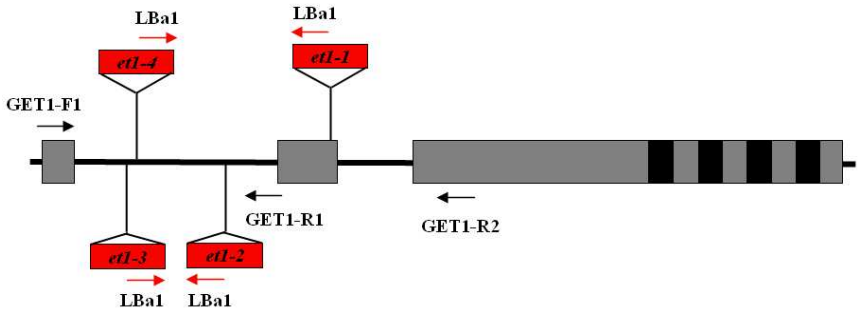


Figure 26. Scheme to illustrate the T-DNA insertion sites in *AtET1*.

Introns and exons were shown in lines and grey boxes, respectively. The coordinates of the T-DNA insertions in the coding region were indicated with respect to the transcription start site. The T-DNA inserts were not drawn to scale. Primers for genotyping were indicated in small arrows.

In addition, comparison of the T-DNA flanking sequence to the corresponding *AtET1* Col wild type loci showed that the integration of T-DNA into line *et1-1* (SALK_000422) induced a deletion of 23 bp of the second exon. The removal of these 23 bp shifts the *et1-4* insertion site into the second exon at nucleotide 382 from start codon of *AtET1* and generates a good candidate for a knock-out mutant. A more detailed characterization of T-DNA insertions in the *AtET1* gene is shown in Figure 26.

3.4.4. Transcription analysis of *et1*

To determine whether the T-DNA insertions influenced the expression of *AtET1* gene, we employed RT-PCR analysis. Since all knock-out lines carried T-DNA insertions close

to the 5' end of the *AtET1* gene, the *AtET1* transcript gene can be detected with several different gene-specific primers within the coding region and downstream of the point of insertions. The product of actin 2 gene (At3g18780) was used to quantify the amount of templates in the PCR reaction. All three mutant lines displayed reduced transcript level of *AtET1* in comparison to Col (Figure 27). Due to pleiotropic phenotypes such as sterility even in heterozygote line *et1-4* could not be used for any further analysis. As judged from the transcript level, line *et1-1* could be a null mutation and therefore, is currently employed for the creation of a double mutant with the line carrying the T-DNA in *AtET2* gene (*et2*-Col).

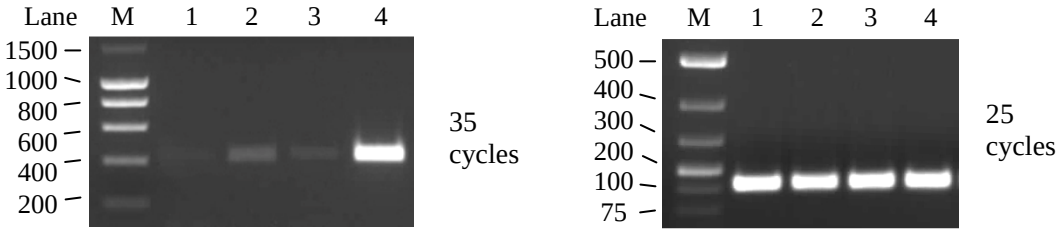


Figure 27. Analysis of transcript levels of *et1* mutant lines

Left: Amplification of AtET1 transcript by ET1-RT-F2 and ET1-RT-R2 primers. Lane 1, 2, 3, 4 were loaded with products from *et1-1*, *et1-2*, *et1-3*, and Col. SmartLadder DNA (Eurogentec, Belgium) given in bp was used as a marker (lane M).

Right: Expression of actin 2 gene (At3g18780) was used as loading control for the corresponding lanes in left panel. GeneRuler^{TM} 1kb DNA Ladder Plus (Fermentas, Vilnius, Lithuania) given in bp was used as a marker (lane M).

3.4.5. Phenotypic analysis of *et1*

The phenotype of homozygous *et1-1* knock-out mutant was inspected and compared to wild type Col under standard growth conditions. Development and growth of this knock-out plant appeared normal, and throughout the rosette stage mutants were indistinguishable from wild type Col plants. Similarly, mutant plants exhibited normal floral sizes, leave shape and numbers as well as branching. Flowering time, silique sizes and seed morphology of mutant and Col plants did not display appreciable differences.

Remarkably, previous data showed that the null mutation in *AtET*2 gene (*et2-1*) strongly increased germination of immature seeds compared to the wild type Ws (Ivanov, 2005). For this reason, immature seeds from green siliques were isolated and placed on MS medium containing basic components. As summarized in Figure 28, the germination rate of immature seeds of the *et1-1* line was higher in comparison to the wild type Col. Thus, this result is in agreement with observations reported by Rumen Ivanov (Ivanov, 2005), although germination rate was not as high as in the *et2-1* mutant.

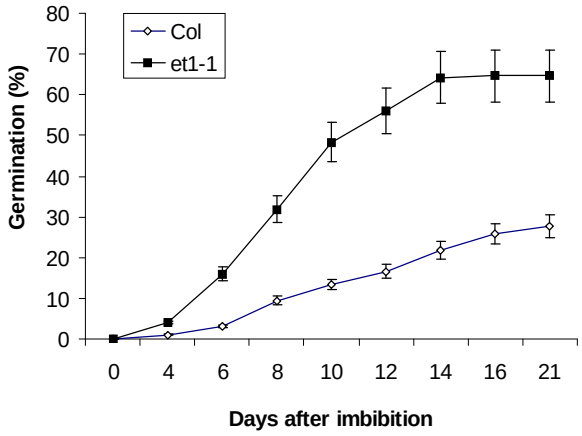


Figure 28. Germination rates of immature seeds from *et1-1* mutant line.

Isolated immature seeds were sown on MS medium supplemented basic vitamins. Germination rate was calculated for three weeks and compared to wild type Col.

3.4.6. Creation of double knock-out mutant

Phenotypic analyses demonstrated that single knock-out mutants in the *AtET1* (*et1-1* line) and *AtET*2 (*et2-1* line) genes showed similar phenotype in delaying germination of immature seeds. In addition, *et2-1* mutant line exhibited significant reduction in lignin content of leaves and stems compared to wild type (Ivanov *et al.*, 2008). Besides these characters single knock-out mutants lack other discernible morphological phenotypes. It is reasonable to speculate that phenotype in a given single mutant might be hidden by overlapping function of *AtET1* and/or *AtET*2 genes (see discussion in pages 99-100 for more detail). Moreover, both these genes also display similar expression patterns during growth and development of plants. To analyze if genetic redundancy could mask essential functions of individual genes, I generated double mutants impaired in both *AtET1* and *AtET*2 genes. Although all mutant lines exhibited reduced expression of *AtET1* gene, only

et1-1 (SALK_000422) line showed nearly complete loss of *AtET1* transcript. Both *et1-1* and *et2-Col* mutants (see 3.4.1 and 3.4.2) were shown to contain T-DNA insertions in the second exon of *AtET1* and *AtET*2 genes, respectively. Thus far, I crossed *et1-1* x *et2-Col* to produce F1 progeny heterozygous for both T-DNA insertions. Current experiments aim to identify either a double homozygous mutant line or a line which is homozygous for one allele and heterozygous for the other in case of gametophytic or zygotic lethality.

3.5. Down regulation of *AtET* during plant development

3.5.1. Generation of *AtET1::*RNAi in *et2* mutant line

In addition to the generation of a double mutant, I attempted to down-regulate the expression of *AtET* through a combination between a gene knock-out and a gene knock-down. In this approach, *AtET1* gene was silenced through selective posttranscriptional degradation (known as RNA interference, RNAi), while *AtET2* gene was completely disrupted by a T-DNA insertion mutant (*et2-1* mutant). To this end, silencing constructs were generated using the gateway-compatible pAGRIKOLA vector from the AGRIKOLA (Arabidopsis Genomic RNAi Knock-out Line Analysis) resource in which gene-specific sequence tags (GSTs) for silencing are separated by an intron spacer. This spacer consisted of two head-to-head introns that enabled splicing of the encoded transcript regardless of its orientation (Hilson *et al.*, 2004). We employed two DNA fragments for silencing constructs including 360 bp (named as i_1*ET1*) and 300 bp (named as i_2*ET1*) corresponding to the third exon of the *AtET1* gene. These fragments were inserted in sense and antisense orientation into destination vector pAGRIKOLA yielding the final plasmids pA_i_1ET1 and pA_i_2ET1, respectively (Figure 29).

The expression of both *RNAi* constructs was controlled by the constitutive CaMV35S promoter. Sequence verified constructs were introduced into *A. tumefaciens* strain pGV3101 and subsequently transformed into *et2-1* mutant line. A total of 48 independent plant lines were selected for both *RNAi* constructs by their ability to growth under BASTA treatment and additionally examined for the presence of trangenes using PCR (Figure 30). Transgenic plants carrying pA_i_1ET1 and pA_i_2ET1 were designated as i_1*ET1 et2-1* and i_2*ET1 et2-1* plants, respectively. A minimum of 5 individual plants per construct were selected to provide a collection of *RNAi* lines for further analyses.

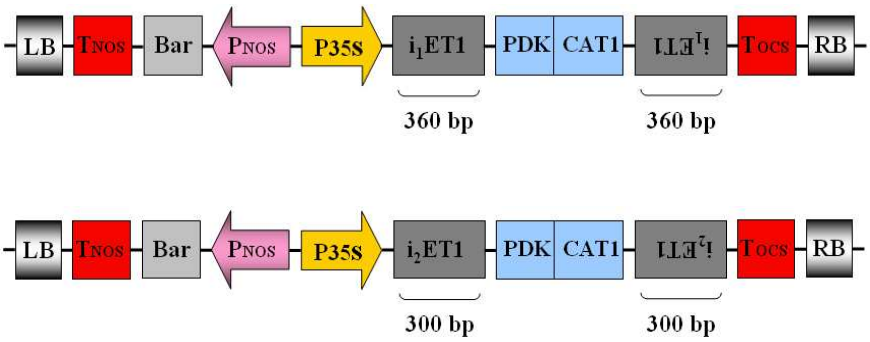


Figure 29. A schematic drawing of two RNAi-mediated silencing constructs

The hairpin cassettes were produced using partial cDNA fragments of *AtET1* gene (360 and 300 bp) that were introduced into the region flanking the PDK and CAT introns. LB, left boder; T_{NOS}, nopaline synthase terminator; Bar, BASTA resistance gene; P_{NOS}, nopaline synthase promoter; P_{35S}, CaMV35S promoter; PDK, 2nd intron of the *Flaveria* PDK gene; CAT, intron of the castor bean CAT gene; T_{OCS}, octopine synthase terminator; RB, right boder.

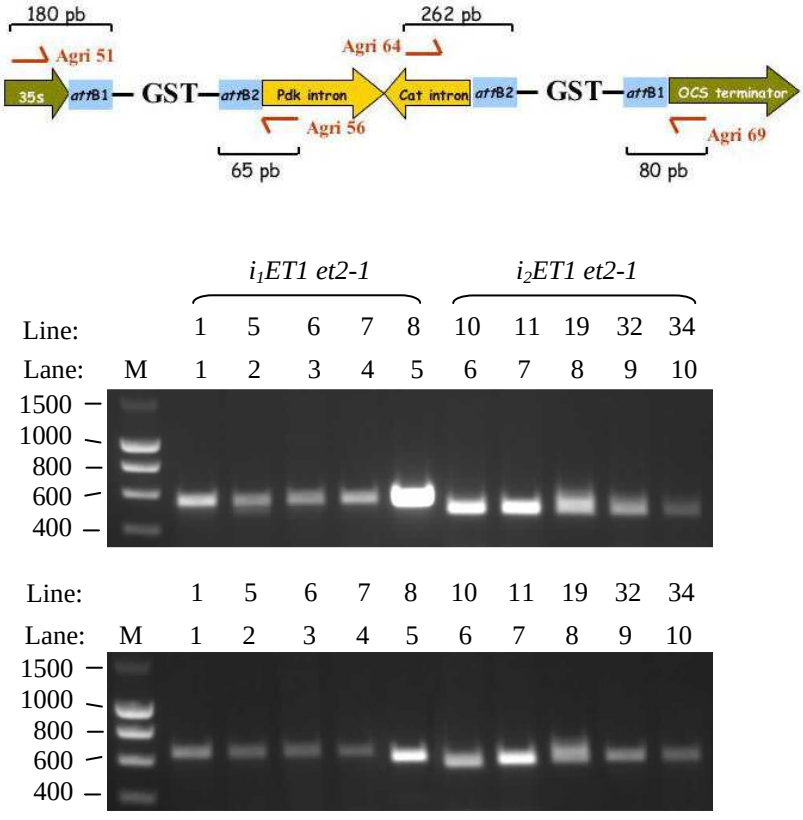


Figure 30. Verification of specific GSTs present in *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants.

Five independent transformed lines per construct were confirmed to contain GSTs using pAGRIKOLA-specific primer pairs (Agri 51/56 and Agri 64/69).

Upper: Structure of recombined hairpin cassette with the inverted GST repeats and a scheme of primer annealing sites.
Middle: Agri 51/56 produced 605 and 545 bp for *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1*, respectively.
Lower: Agri 64/69 yielded 702 and 642 bp for *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1*, respectively.

SmartLadder DNA (Eurogentec, Belgium) given in bp was used as a marker.

3.5.2. Reduced expression level of *AtET1* in *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants

In order to determine that the repression of *AtET1* gene expression was due to the overexpression of RNAi constructs, the *AtET1* transcript was monitored in the *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants. Because RNAi produced a series of *Arabidopsis* transgenic lines showing different levels of down-regulation of the *AtET1* gene, we tested the expression in 5 transgenic lines per construct using total RNA prepared from rosette leaves of each plant line. The expression of *AtET1* gene in *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plant lines was analyzed by conventional RT-PCR using ET1-RT-F2 and ET1-RT-R2 primers. This primer pair has been found to be specific for the *AtET1* sequence and amplified a 416 bp fragment from the third exon. Amplification of the actin 2 gene (At3g18780) product was used as a constitutive control to show that equal amounts of RNA had been used in the experiments. Primer ACT2-F and ACT2-R generated 134 and 220 bp from cDNA and genomic DNA of actin 2 gene, respectively.

As shown in Figure 31A, two lines (*i$_2$ET1#10 et2-1* and *i$_2$ET1#32 et2-1*) did not display any reduction in transcript level of *AtET1* gene compared to products from *et2-1* mutant and Ws (lanes 11 and 12). Other lines exhibited a substantial decrease (*i$_1$ET1#1 et2-1*, *i$_1$ET1#5 et2-1*, *i$_2$ET1#19 et2-1*, and *i$_2$ET1#34 et2-1*) or a nearly complete loss (*i$_1$ET1#6 et2-1*, *i$_1$ET1#7 et2-1*, *i$_1$ET1#8 et2-1*, and *i$_2$ET1#11 et2-1*) of the transcripts in comparison to *et2-1* mutant and wild type Ws. All these lines were used for further analyses of phenotypic alterations.

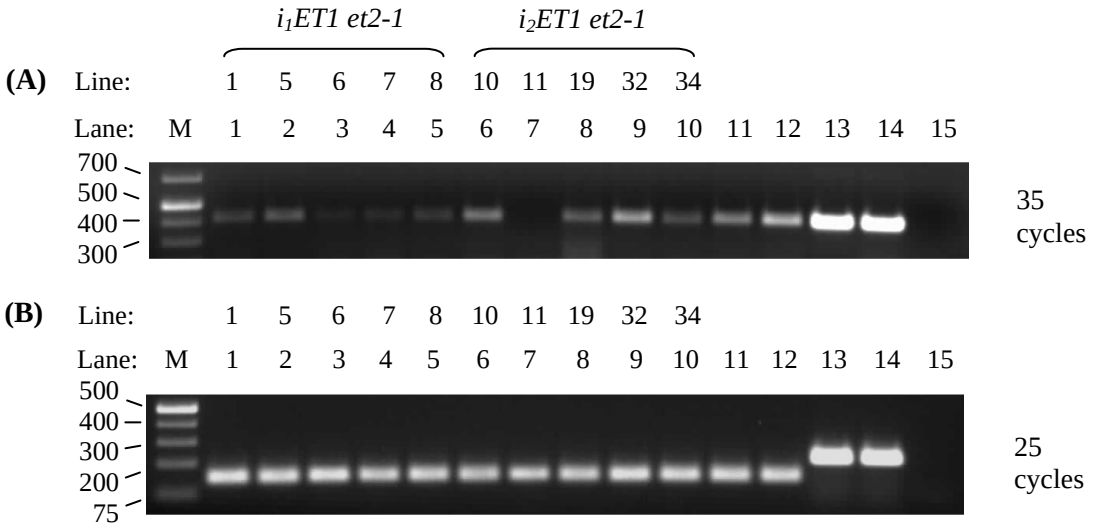


Figure 31. RT-PCR analysis of *AtET1* gene silencing.

A. The analysis of *AtET1* transcript levels in 10 independent lines (lanes 1 to 10). Lanes 11 and 12, PCR products from cDNA of *et2-1* mutant and wild type Ws. Lanes 13 and 14, PCR products from genomic DNA of *et2-1* mutant and wild type Ws. Lane 15, negative control.

B. Expression of actin 2 gene (At3g18780) was used as a loading control for the corresponding lanes in (A) panel.

GeneRuler^{TM} 1kb DNA Ladder Plus (Fermentas, Vilnius, Lithuania) given in bp was used as a marker (lane M).

3.5.3. Phenotypes of *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants

In general RNAi plants can exhibit a series of phenotypic changes that are proportional to the silencing of target gene expression (Chuang and Meyerowitz, 2000). Since *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants were produced by a combination between knock-down of *AtET1* gene expression and a mutation in *AtET2* gene (*et2-1* line), the phenotypes of these plants are expected to exhibit significant differences in comparison to wild type Ws and *et2-1* mutant plants. However, morphological analyses of *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plant lines carrying both constructs (pA-i$_1$ET1 and pA-i$_2$ET1) did not reveal any obvious phenotypic changes in rosette leave shape and number, but were overall clearly smaller than wild type Ws and *et2-1* mutant plants (Figure 32).

A B C D

Figure 32. Comparison of morphological phenotypes of *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants. Three-week-old seedlings were grown on soil. A, wild type Ws; B, *et2-1* mutant plant; C, *i$_2$ET1#19 et2-1* line; D, *i$_1$ET1#6 et2-1* line.

As *AtET* genes are markedly expressed for instance in reproductive organs (Figure 16 and 17), we examined whether loss of function in both *AtET* genes has any effects on pollen development. To test for pollen viability, mature pollen from wild type Ws, *et2-1* mutant, *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants were colored with Alexander staining (Alexander, 1969). The viable pollen are usually stained dark purple or brightly red with this dye, whereas nonviable pollen can be recognized by their smaller size and flattened shape as well as a pale red or a light purple stain. As shown in Figure 33, *et2-1* mutant plants exhibited uniform pollen grains similar to the wild type Ws, whereas *i$_2$ET1#10 et2-1*, *i$_2$ET1#19 et2-1*, and *i$_1$ET1#6 et2-1* plants showed two distinct types of pollen including normal and aberrant phenotypes in comparison to wild type Ws (Figure 33, C, D and E). Approximately 10, 30, and 50% aberrant pollen were observed in *i$_2$ET1#10 et2-1*, *i$_2$ET1#19 et2-1*, and *i$_1$ET1#6 et2-1* plants, respectively, compared to 1% aberrant pollens in wild type Ws and in *et2-1* mutant plants (Figure 33). Therefore, *AtET* genes are critical for pollen viability.

To additionally investigate developmental stages and nuclear composition of the pollen, we stained dehiscing anthers with DAPI (4, 6-diamidino-2-phenylindole) solution and visualized nuclei under UV illumination. The wild type Ws pollens showed clearly a diffusely stained vegetative nucleus and two condensed sperm nuclei (tricellular pollen). A similar pattern of staining was visualized in the pollen of *et2-1* mutant, indicating that nuclear divisions occurred during pollen development. In contrast, DAPI staining could not detect any fluorescence from aberrant pollens tested by Alexander staining. This

supplemental observation confirmed that aberrant pollen were most likely inviable and could not function in fertilization. Besides the normal pollen similar to mature pollen from Ws and *et2-1* mutant plants, pollen from *i*$_2$*ET1#19 et2-1*, and *i*$_1$*ET1#6 et2-1* plants exhibited a diffuse vegetative nucleus and only one compact generative-like nucleus (bicellular pollen), indicating a failure to progress through the second mitotic division (Figure 34, G and H).

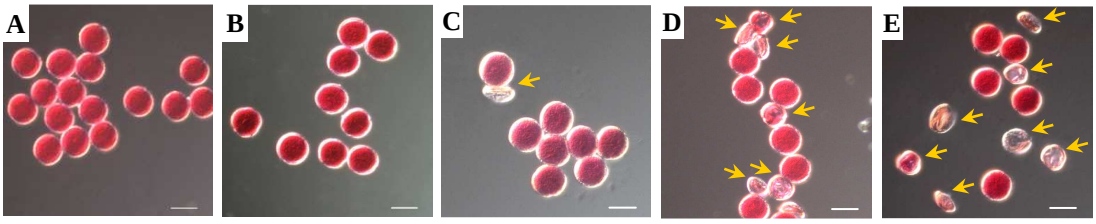


Figure 33. Pollen phenotypes in representative *i*$_1$*ET1 et2-1* and *i*$_2$*ET1 et2-1* plants.

Alexander staining was performed with pollen grains from wild type Ws (A), *et2-1* mutant (B), *i*$_2$*ET1#10 et2-1* (C), *i*$_2$*ET1#19 et2-1* (D), and *i*$_1$*ET1#6 et2-1* (E) plants. Yellow arrows indicated the aberrant pollen grains. Bars 20 µm.

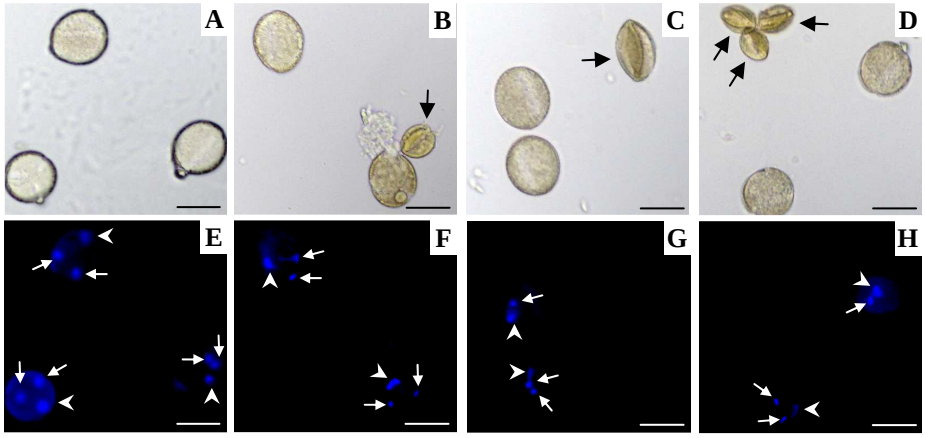


Figure 34. The nuclear constitution of mature pollens from *i*$_1$*ET1 et2-1* and *i*$_2$*ET1 et2-1* plants. Phase contrast images of pollen grains from wild type Ws (A), *i*$_2$*ET1#10 et2-1*, *i*$_2$*ET1#19 et2-1*, and *i*$_1$*ET1#6 et2-1* plants (B, C and D, respectively). The same samples were used for fluorescence image detection (E, F, G, H). The two brightly stained small nuclei are sperm nuclei (indicated by white arrows) and the large and more diffuse nucleus is the vegetative cell nucleus (indicated by white arrow heads). All aberrant pollen (indicated by black arrows) did not contain any nuclei. Bars 20 µm.

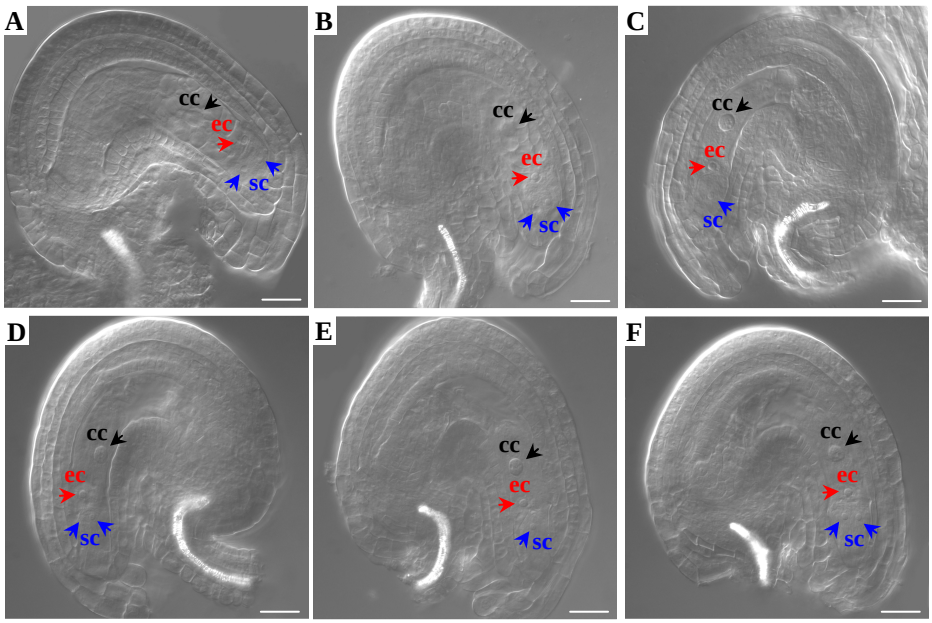


Figure 35. Female gametophytic phenotypes in *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants.

Microscopy images of isolated ovules (2 days after emasculation) from *i$_2$ET1#10 et2-1* (C), *i$_2$ET1#19 et2-1* (D), and *i$_1$ET1#6 et2-1* (E and F) plants showed no alteration in gametophytic phenotype relative to Ws (A), *et2-1* mutant (B) plants. Cc, central cell, ec, egg cell; sc, synergid cell. Bars represented 20 µm.

To further examine female gametophytic phenotypes, we visualized isolated ovules from *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants under microscope and compared to Ws, *et2-1* mutant. As shown in Figure 35, the female gametophytes developed identically to Ws, *et2-1* mutant, *i$_2$ET1#10 et2-1*, *i$_2$ET1#19 et2-1* and *i$_1$ET1#6 et2-1* plants. They exhibited normal central cells, egg cells as well as synergid cells. This observation suggested that female gametophyte development was not affected in *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants.

Dissection of green siliques from *i$_1$ET1#5 et2-1*, *i$_1$ET1#6 et2-1*, *i$_1$ET1#7 et2-1* *i$_1$ET1#8 et2-1* and *i$_2$ET1#19 et2-1* plant lines revealed that they contained healthy developing seeds together with infertile ovules and aborted seeds, whereas siliques from wild type Ws and *et2-1* mutant plants consisted of uniformly green developing seeds (Figure 37, H and I). The *i$_2$ET1#10 et2-1* and *i$_2$ET1#32 et2-1* plant lines had only few infertile ovules as well as aborted seeds. Initially, aborted seeds were yellow or transparent, while siblings were green. Later in development, the aborted seeds were dark red or dark brown in appearance and shrunken (Figure 37, J to O). Depending on morphological

phenotypes, the *i₁ET1 et2-1* and *i₂ET1 et2-1* plant lines were grouped into normal, moderate and strong phenotype levels corresponding to phenotype class A, B, and C. *i₂ET1#10 et2-1* and *i₂ET1#32 et2-1* lines appeared normal and were grouped into phenotype A. *i₁ET1#1 et2-1*, *i₁ET1#5 et2-1*, *i₂ET1#19 et2-1*, and *i₂ET1#34 et2-1* were clustered into phenotype B, while *i₁ET1#6 et2-1*, *i₁ET1#7 et2-1*, *i₁ET1#8 et2-1*, and *i₂ET1#11 et2-1* exhibited the most severe alterations and were therefore grouped into phenotype C (Table 3).

Table 3. Phenotypic analyses of *i₁ET1 et2-1* and *i₂ET1 et2-1* plants

Genetics	Line	Phenotype		Healthy seeds	Aborted seeds	Infertile ovules	Aborted embryos
		level	class				
i₁ET1 et2-1/et2-1	1	Moderate	B	33 ± 5	14 ± 6	6 ± 2	heart
i₁ET1 et2-1/et2-1	5	Moderate	B	27 ± 6	10 ± 4	9 ± 4	heart
i₁ET1 et2-1/et2-1	6	Strong	C	16 ± 5	22 ± 8	11 ± 5	globular-heart
i₁ET1 et2-1/et2-1	7	Strong	C	14 ± 7	25 ± 9	10 ± 4	globular
i₁ET1 et2-1/et2-1	8	Strong	C	8 ± 7	28 ± 12	8 ± 4	globular
i₂ET1 et2-1/et2-1	10	Normal	A	52 ± 6	3 ± 2	7 ± 3	no
i₂ET1 et2-1/et2-1	11	Strong	C	18 ± 8	20 ± 8	10 ± 3	globular-heart
i₂ET1 et2-1/et2-1	19	Moderate	B	32 ± 6	15 ± 7	5 ± 2	heart
i₂ET1 et2-1/et2-1	32	Normal	A	55 ± 7	4 ± 2	5 ± 3	no
i₂ET1 et2-1/et2-1	34	Moderate	B	38 ± 6	16 ± 5	6 ± 2	heart
ET1/ET1 ET2/ET2	Ws	Normal	A	59 ± 4	1	3 ± 2	no
ET1/ET1 et2-1/et2-1	*et2-1*	Normal	A	56 ± 5	1	5 ± 3	no

To determine the effect of loss of function of both *AtET* genes on earlier stages of seed formation, we isolated seeds from representative *i₁ET1 et2-1* and *i₂ET1 et2-1* plants for clearing and observed the morphology of developing embryos in the seeds by light microscopy Zeiss Axioplan2. No significant morphological differences were detected between *et2-1* mutant and wild type Ws at corresponding stages of development as shown in Figure 38 (A to J). Analysis of seeds from *i₁ET1 et2-1* and *i₂ET1 et2-1* plants revealed clearly two types, normal and aberrant seeds. The normal seeds developed similar

morphologies to wild type and *et2-1* mutant seeds. On the contrary, aberrant seeds were often smaller in size compared to the wild type Ws and *et2-1* mutant at the same stages of development. Their embryo development was arrested at globular (Figure 38, K to N) or globular-heart transition stage (Figure 38, O to Q) and mainly found in phenotype class C, approximately 60% of embryos (Figure 36). Other abnormal embryos were blocked at heart stage and predominantly appeared in phenotype class B, around 35% (Figure 36 and Figure 38, R to T). Thus the seeds carrying these aberrant embryos were aborted and died before maturity as seen in Figure 37 (L, N and O). This observation indicated that phenotypes of abnormal embryos were most likely related to the transcript levels of *AtET* genes in *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants.

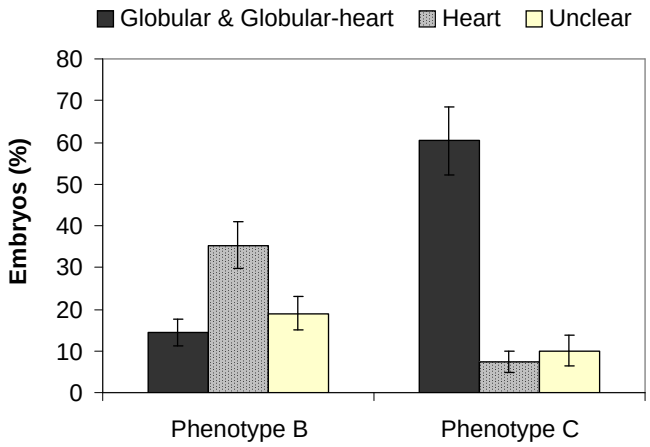


Figure 36. Embryo development in *i$_1$ET1 et2-1* and *i$_2$ET1et2-1* plants.

Dissected siliques revealed an increase in aberrant embryos arrested at globular and globular-heart transition stages in phenotype C and arrested at heart stage in phenotype B plants. N = 15 siliques were analyzed for each line.

The results from combination between silenced *AtET1* and *et2-1* mutant strongly suggested that eliminating or substantially reducing the two *AtET* genes (both *AtET1* and *AtET2* genes) was seriously detrimental to plants and these gene products were essential for normal growth and development of *Arabidopsis* plants. To completely define if one or both parents were affected by loss of function of both *AtET* genes, it was necessary to carry out reciprocal crosses by pollinating the wild type Ws, *et2-1* mutant females with pollen from RNAi plants and *vice versa*. These crosses have been completed and will be analyzed soon.

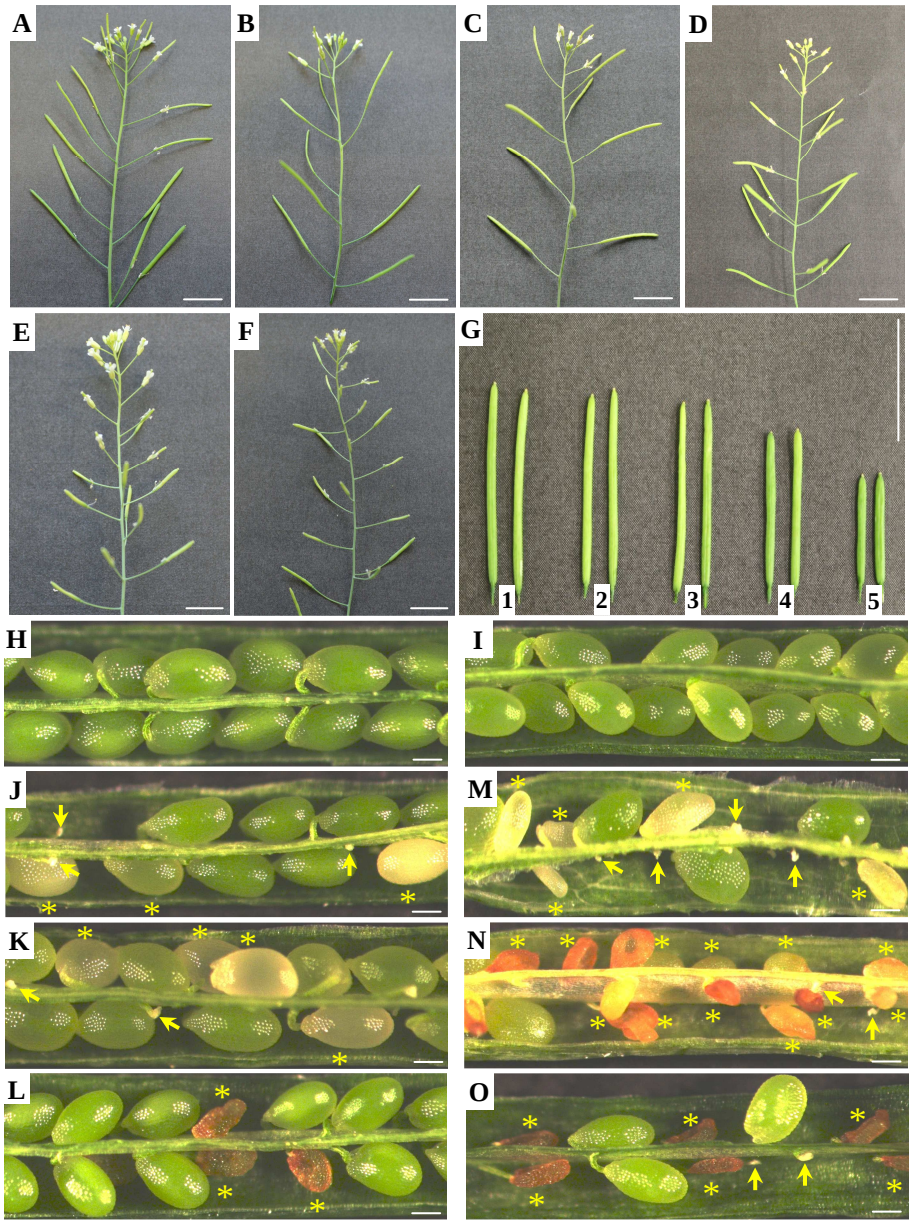


Figure 37. Seed abortion in *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants.

Inflorescences from *i$_2$ET1#10 et2-1* (C), *i$_2$ET1#19 et2-1* (D), *i$_1$ET1#6 et2-1* (E), and *i$_1$ET1#8 et2-1* (F) plants relative to the wild type Ws (A), *et2-1* mutant (B) plants. Mature green siliques from wild type Ws, *et2-1* mutant, *i$_2$ET1#10 et2-1*, *i$_2$ET1#19 et2-1* and *i$_1$ET1#6 et2-1* (G1, G2, G3, G4 and G5, respectively). Bars represented 1 .0 cm.

Dissected siliques from wild type Ws (H) and *et2-1* mutant (I) plants were compared to siliques from *i$_1$ET1#5 et2-1* (J and K), *i$_2$ET1#19 et2-1* (L), *i$_1$ET1#7 et2-1* (M), *i$_1$ET1#6 et2-1* (N) and *i$_1$ET1#8 et2-1* (O) plants. The *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* lines contained infertile ovules (indicated by yellow arrows) and aborted seeds (pointed by yellow asterisks). Bars represented 200 µm.

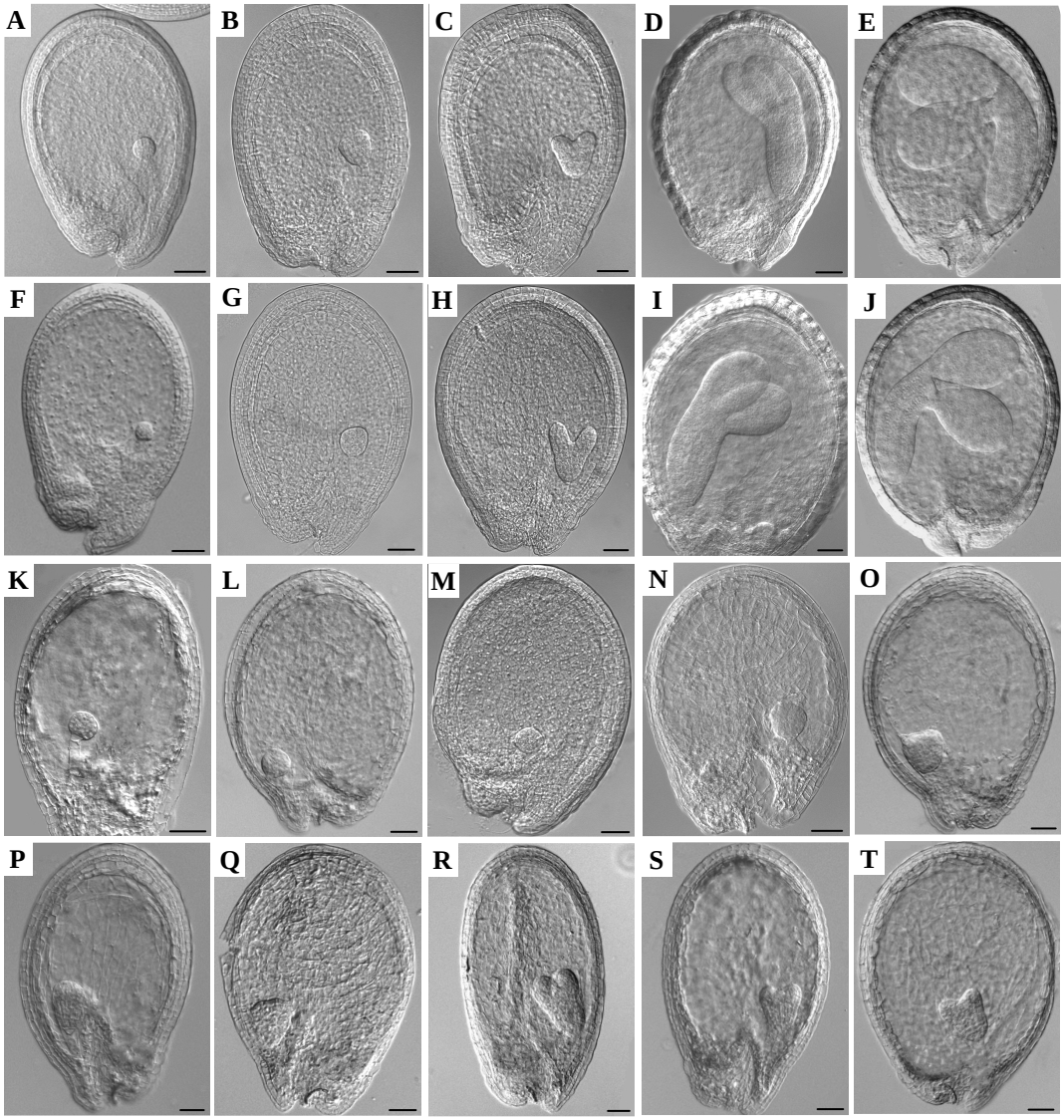


Figure 38. Aberrant embryo development in *i$_1$ET1 et2-1 and i$_2$ET1 et2-1 plants*.

Developing seeds from Ws, *et2-1* mutant, *i$_1$ET1 et2-1* and *i$_2$ET1 et2-1* plants were cleared and examined with light microscopy Zeiss Axioplan2.

A to E, globular, globular-heart transition, heart, torpedo and walking stick stages of Ws, respectively. Pictures F to J showed embryo development in *et2-1* mutant plants corresponding to stages of Ws. Aberrant seeds containing embryos arrested at globular stage (K to N), globular-heart transition stage (O to Q) were observed in *i$_1$ET1#6 et2-1*, *i$_1$ET1#8 et2-1* plants or at heart stage (R to T) in *i$_1$ET1#5 et2-1* and *i$_2$ET1#19 et2-1* plants. Bars represented 50 µm.

4. Discussion

Differentiation and development requires a precise control of gene expression. Gene regulation at the transcriptional level implies the interaction between transcription factors and the corresponding target genes. Transcription factors have been classified on the basis of their highly conserved DNA-binding domains. The DNA-binding domain of the HRT (HORDEUM REPRESSOR OF TRANSCRIPTION) class of transcription regulators consists of several highly conserved repeats with typical cysteine patterns. Founding members of the strictly plant specific HRT class have been originally isolated from broad bean (*Vicia faba*) and barley (*Hordeum vulgare*). For a more detailed functional analysis we have studied a corresponding gene family, designated as EFFECTOR OF TRANSCRIPTION (ET), in the model plant *Arabidopsis*. The ET gene family consists of three members. Both *AtET1* and *AtET2* genes encode full length proteins, whereas *AtET3* encodes a truncated version lacking the C-terminal ET repeats. All what is currently known about HRT/ET has been published in three papers (Raventos *et al.*, 1998; Ellerström *et al.*, 2005; Ivanov *et al.*, 2008). The published data suggest important developmental functions of HRT/ET during seed and xylem development, most likely connected to gibberellin-mediated processes.

Here, we describe experiments which provide essential tools for a further functional characterization of this family of plant regulators, including a) the synthesis of recombinant proteins for DNA binding studies, generation of specific nanobodies, and identification of DNA-binding properties of AtET proteins *in vitro*, b) domain swapping experiments to analyze their single strand cutting activity and c) the generation and characterization of loss of function mutants supporting an important function of ETs in seed development.

4.1. Expression of recombinant AtET proteins in *E. coli*

The study of biological functions of the AtET proteins has been hampered by the difficulties in obtaining sufficient amounts of pure proteins for the generation of specific nanobodies as well as for DNA binding studies. Obviously, AtET proteins are difficult to be synthesized as recombinant proteins in *E. coli* most likely due to their cysteine-rich domains. In total, the four repeats contain twelve conserved cysteine residues at the C-

terminal (Ellerström *et al.*, 2005; Ivanov *et al.*, 2008). Thiol (sulfhydryl) groups of cysteines are involved in the formation of disulfide bonds, crucial for protein stability, enzyme catalysis and redox status (Leichert and Jakob, 2004; Meyer and Hell, 2005; Kaur and Bachhawat, 2007; Hansen *et al.*, 2009; Alvarez *et al.*, 2010). Moreover, it is well known that cysteines can form aberrant intra- or intermolecular disulfide bridges (Meyer and Hell, 2005), which can cause mis-folding and degradation (Benita *et al.*, 2006). Thus, attempts to produce cysteine-rich recombinant proteins in *E. coli* often failed because of low stability and the presence of many thiol groups toxic to the host cells (Xiong and Ru, 1997; Huang *et al.*, 2009a).

Several trials to synthesize AtET proteins in various expression vectors containing *tetA* or *T7* promoters supported the view that these proteins seem to be toxic to *E. coli* strains even when grown under low temperature. Although *T7* promoter controlled expression systems facilitate high levels of protein synthesis, they have been reported to be leaky (Giacalone *et al.*, 2006). The *tetA* systems have been successfully applied for tightly regulated bacterial expression of heterologous proteins. In several cases, the basal level of *tetA* promoter activity is present even in the absence of inducer (Grkovic *et al.*, 2002). Thus, leaky expression can be a problem when recombinant proteins have detrimental effects on host cells and suppression or reduction of basal expression may be employed to reduce toxicity (Schumann and Ferreira, 2004; Bongers *et al.*, 2005). These negative effects of recombinant proteins on host cells can be overcome by using vectors which possess a tight expression control prior to induction (Giacalone *et al.*, 2006). We therefore attempted to synthesize AtET proteins in *E. coli* employing the vector pQE30 (Qiagen, Hilden, Germany). This vector contains two *lac* repressor binding sites to ensure efficient repression of the *T5* promoter by high levels of *lac* repressor protein in *E. coli* before induction (Liu *et al.*, 1999; Scheidegger *et al.*, 2001; Drees *et al.*, 2004; Blevins *et al.*, 2007). Using the pQE30 vector, we were finally able to get low level expression of AtET proteins in *E. coli* strain M15. Although proteins were accumulated in soluble forms in *E. coli* cytoplasm, they could only be purified under denaturing or hybrid conditions. Probably, the protein did not bind sufficiently to the Ni-NTA resin due to non-suitable protein conformation (Kneusel *et al.*, 2000; Debeljak *et al.*, 2006). Using large cultures and applying the above described hybrid conditions for purification we could obtain sufficient amounts of AtET proteins for further analyses such as DNA-binding properties and specific nanobody production.

4.2. DNA-binding properties of AtET proteins

Since ET factors were originally detected and isolated by South Western approaches as DNA-binding proteins, it is assumed that they might act as transcription factors. Therefore, efforts were directed on the identification of DNA motifs recognized by ET factors. Using EMSA, it was shown that AtET factors do bind to DNA, however, the results indicate that there is no obvious sequence specificity detectable. This finding is in agreement with previous data obtained for AtET and BnET factors. In both cases, binding assays did not exhibit any specific sequence in the DNA fragments used as probes (Ellerström *et al.*, 2005; Ivanov, 2005). Only HRT, an ET factor from *Hordeum vulgare*, has been reported to interact with a central sequence of the GA response element (Raventos *et al.*, 1998). Currently, we interpret this as a difference between ET factors from monocots and dicots.

Thus, AtET factors most likely function as non-sequence specific DNA-binding proteins, since they bind to DNA fragments containing seed-specific *napA*, *USP* or *LeB4* promoters as well as random probes. The synthetic random oligonucleotides used as controls did not match with known transcriptional regulatory motifs in plants monitored by PLACE (A database of plant *cis*-acting regulatory DNA element, www.**dna**.affrc.go.jp/**PLACE**/) (Higo *et al.*, 1998; Higo *et al.*, 1999; Fauteux and Stromvik, 2009). A possible explanation for the described binding properties of ET could be the requirement of a yet unknown interacting factor. Previous yeast two-hybrid experiments have identified a bHLH transcription factor as well as a SWAP domain containing nuclear factor as putative interaction partners of AtET2 (R. Ivanov and H. Bäumlein, data unpublished). Clearly, this requires further confirmation by a more detailed biochemical analysis. It is also conceivable that non-specific DNA-binding domains might assist a protein sliding mechanism along the DNA (Gowers and Halford, 2003; Gao and Skolnick, 2009; Huang *et al.*, 2009b). Thus, non-specific interaction could be an important intermediate step in the process of sequence-specific recognition and binding (Kalodimos *et al.*, 2004). The observed detection of the AtET-DNA complex by specific nanobodies further supports the DNA binding properties of ET factors. The observation that nanobodies did not block the interaction between AtET and the probe oligonucleotides suggests that the nanobodies do recognize protein regions outside of the DNA-binding ET

domain. Obviously, the generated nanobodies provide an excellent experimental tool for further molecular analysis of ET factors.

4.3. Putative functions of GIY-YIG domain in ET factors

Since ET factors have been located in the nucleus, where they can act as regulators of other genes, we searched for the presence of recognizable protein domains. In addition to the characteristic ET repeats, this search identified a low level of similarity to the DNA single strand cutting domain present in bacterial UvrC protein 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). The protein 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 nucleotide sequence between both single strand cuts is removed and replaced by the correctly repaired strand. The two single strand cuts are processed by two structurally and functionally distinct domains of the UvrC protein. A C-terminal domain consisting of an Endonuclease V (ENDOV) and a Helix-hairpin-Helix (HhH) subdomains is required for the cut 8 bp 5´ of the lesion, whereas the N-terminal GIY-YIG domain inserts the cut 4 bp 3`of the lesion (Figure 39) (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). In addition to the UvrC protein, the GIY-YIG domain (see Figure 40 for more detail) is also present in homing nucleases, which are encoded within mobile group I, group II and archaea introns as well as in inteins, intervening sequences which are spliced and excised post-translationally (Stoddard, 2005; Dunin-Horkawicz *et al.*, 2006; Dassa *et al.*, 2009). Therefore, the domain has been re-designated URI for UvrC and intron-encoded endonucleases (Aravind *et al.*, 1999; Singh *et al.*, 2002).

The sequence similarity between plant ET factors and the URI domain proteins is restricted to the single strand cutting GIY-YIG domain. This suggests that a UvrC-like ancestral domain has been recruited by ET proteins and attached to the DNA-binding ET repeats. A corresponding domain shuffling event would be consistent with the exon-intron structure of ET genes, with the GIY-YIG domain represented by a separate second exon. 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. This showed that the plant domain productively interacts with the C-terminal UvrC ENDOV/HhH domain. Since the single amino acid exchange (R>A) results in the near complete loss of this activity, the single strand cutting activity of the GIY-YIG domain must clearly be required for its molecular function. Thus, we suggest that the nicking activity of the plant ET factor GIY-YIG domain may be involved in the catalysis of changes in higher order DNA structure, such as, for example, nucleosome sliding (Choi *et al.*, 2002; Xiao *et al.*, 2003). The presence of the transient break in DNA within nucleosome is needed to initiate gene expression (Haince *et al.*, 2006). Alternatively, it may contribute to the relaxation of supercoiled chromatin domains, which are implicated in the control of differentiation and development. The importance of the relief of torsional tension in DNA to the triggering of transcriptional activation has been recently described (Ju *et al.*, 2006).

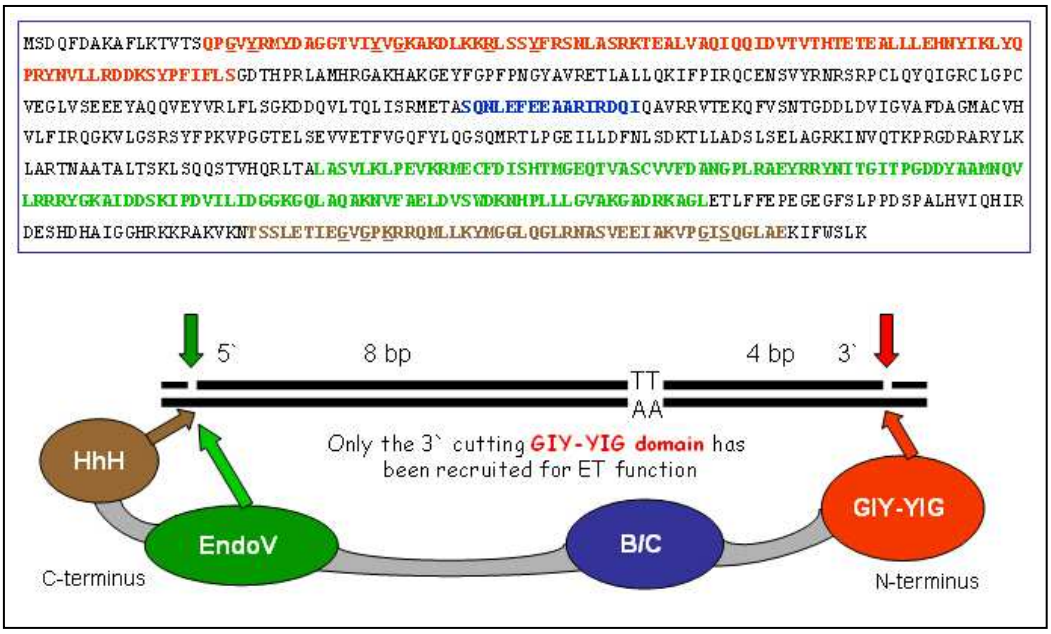


Figure 39. Schematic representation of working model of UvrC protein.

Upper: The amino acid sequence of UvrC protein.

Lower: The 3' incision of the damaged DNA takes place at 4 bp downstream of the lesion. The 5' incision then follows at 8 bp upstream of the lesion. HhH, Helix-hairpin-Helix motif; EndoV, Endonuclease V domain; B/C, UvrBC domain; GIY-YIG domain.

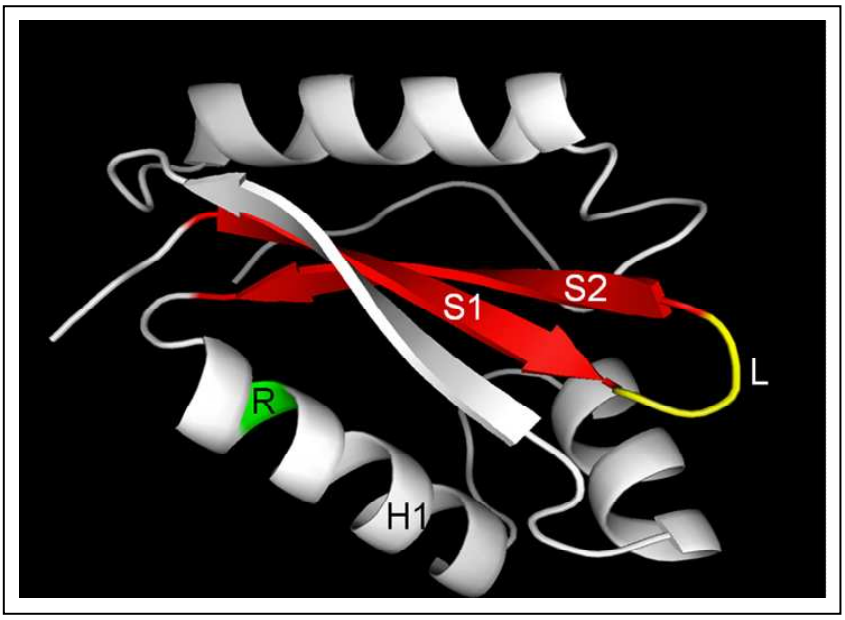


Figure 40. Three dimension model of GIY-YIG domain of phage T4.

Strands 1 and 2 were shown in red (S1 and S2). The interstrand loop (L) was given in yellow and extended in most ET proteins with the exception of AtET1. The position of helix 1 and catalytic arginine were marked as H1 and R, respectively (Ivanov, 2005).

Remarkably, the HhH domain of the UvrC protein can be considered as the ancestor protein domain for another class of plant regulators like *Arabidopsis* DEMETER (DME) and REPRESSOR OF SILENCING 1 (ROS1) (Choi *et al.*, 2002; Gong *et al.*, 2002; Xiao *et al.*, 2003; Choi *et al.*, 2004b; Morales-Ruiz *et al.*, 2006). DME can introduce single strand nicks in MEA promoter and activate maternal MEA allele expression (Choi *et al.*, 2002; Xiao *et al.*, 2003; Choi *et al.*, 2004b). A possible conclusion 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 repeats. The resulting protein factor is not involved in the repair process but acts as a gene regulator. The regulatory mechanism - in part analogous to the function of DME and ROS1 - includes 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.

4.4. Loss of function of *AtET* genes

AtET1 and AtET2 are two closely related proteins in AtET family and their overall sequences show 40.2% identity, especially high in ET repeats (domains, 58.3% identity) but less homology in GIY-YIG like domains (21.4% identity) (Figure 3). Analyses of public databases (Genevestigator and AtGeneExpress) revealed that *AtET1* and *AtET2* genes have somewhat similar expression patterns, suggesting a functional redundancy among these family members. Both these genes are expressed at low levels in vegetative organs such as leaves, roots, stems, and shoot apexes and higher levels in sexual organs (Figure 16 and Figure 17).

Single knock-out mutants of *AtET* genes (*et1-1* and *et2-1*) exhibit minor phenotypic differences in comparison to wild types (ecotypes Col and Ws). Besides an increased germination rate of immature seeds (in both *et1-1* and *et2-1* mutants) and a reduced lignin content in *et2-1* mutant plants (Ivanov, 2005; Ivanov *et al.*, 2008), neither *et1-1* nor *et2-1* single mutants have severe effects, most like due to genetic redundancy in the *AtET* family. It is conceivable that mutations in one of two genes would still allow function of the other but mutations in both genes would completely block the expression and result in much more severe effects (Hobbie and Estelle, 1995). Alternatively, both single mutants might be conditional mutants, with phenotypic changes to become obvious only under certain conditions (Meissner *et al.*, 1999; Bouche and Bouchez, 2001; Pawloski *et al.*, 2006). To further analyze the proposed functional redundancy we are constructing a double mutant by crossing between *et1-1* and *et2-1* mutant lines. In parallel, we employed an RNAi approach as an alternative to silence both *AtET1* and *AtET2* genes. Since the common functional domains between these genes exhibit 21.4 – 58.3% similarity at the nucleotide level, we could not knock-down both of them by a single RNAi construct. We therefore undertook a combinational approach between RNAi and single mutant to silence *AtET1* in the *et2-1* mutant for dissecting their function. Both RNAi constructs deployed to silence *AtET1* did not reveal an obvious vegetative phenotype. However, the pollen viability and development were impaired as detected by Alexander staining and DAPI staining. Therefore, fertilization rates might have been reduced as a result of lower level of viable pollen produced by these plant lines. Besides pollen grains with three nuclei (tricellular pollens) as in wild type, we observed several bicellular pollen, containing one diffuse DAPI stained vegetative nucleus and a more compact generative-like nucleus in *i$_1$ET1 et2-*

1 and *i_2ET1 et2-1* plants displaying phenotypes B and C (Figure 34, G and H). This observation suggests that the first microspore division is normal, whereas the second division of the generative cell is disrupted generating bicellular pollen (Durbarry *et al.*, 2005; Iwakawa *et al.*, 2006; Gusti *et al.*, 2009). The result is consistent with the above mentioned predominant expression of *AtET* genes in pollen (Figure 16, 17 and Figure 19, J, K). Furthermore, the existence of aberrant pollen without any nuclei and bicellular pollen indicating loss of function in both *AtET* genes may affect earlier and/or later stages of pollen development. Since ovule development may have ceased prior to fertilization, we examined mature female gametophytes from *i_1ET1 et2-1* and *i_2ET1 et2-1* plants. Contrary to the defective pollen development, we did not detect any phenotypic change in these organs with the normal development of synergid, central and egg cells at maturity. These results imply that AtET genes are essential only for the male gametophyte development, revealing a sex-specific function of ETs.

The expression of RNAi in *et2-1* mutant line (*i_1ET1 et2-1 and i_2ET1 et2-1* plants) resulted in a range of embryo phenotypes. Most embryos arrested at the globular stage and at the transition between globular and heart stage (Table 3, phenotype C and Figure 38, K to Q), while others continued development and arrested at the heart stage (Table 3, phenotype B and Figure 38, R to T). The seeds containing aborted embryos appeared yellow transparent and became dark red or dark brown and shrunken at maturity (Figure 37, J to O), possibly because the absence of a mature embryo caused the seed coat to collapse (Figueroa *et al.*, 2005). Moreover, the presence of aborted embryos showed that fertilization occurred and that embryogenesis was interrupted after fertilization (Kunieda *et al.*, 2008). It was therefore possible that the embryos were morphologically disrupted and the seeds were aborted due to lack of normal endosperm (Koehler *et al.*, 2003; Iwakawa *et al.*, 2006; Johnston *et al.*, 2007; Gusti *et al.*, 2009). Although *Arabidopsis* endosperm does not store the reserves of the seed, it most probably controls the flux of nutrients delivered by the vascular tissue of the mother to the embryo and protects the embryo from physical and osmotic stresses (Garcia *et al.*, 2003). Since successive divisions and cellularisation are two important events of endosperm development occurring early after fertilization, it is necessary to analyze endosperm in young seeds. However, it was difficult to distinguish between abnormal and normal seed when they both were young and white (Li and Thomas, 1998; Kristof *et al.*, 2008). Therefore, additional experiments should be performed for examination of endosperm development, for instance by introducing endosperm-specific

markers such as KS22 (Ungru *et al.*, 2008) or KS177 (Sorensen *et al.*, 2001; Garcia *et al.*, 2003; Ingouff *et al.*, 2005).

Analyses of *i$_1$ET1 et2-1* and *i2ET1 et2-1* plants revealed a number of aberrant developmental phenotypes that correlated with the relative reduction in transcript levels of *AtET1*, while *AtET2* transcript was completely lost. In particular, the plant lines with more pronounced reduction in *AtET1* transcripts (lines 6, 7, 8, 11, displaying phenotype class C) showed defects in early embryogenic stages, whereas less severe reduction led to proceed to heart stage of embryogenesis (lines 1, 5, 19, 34, exhibiting phenotype class B). Thus, only embryos weakly affected by loss of function may have been able to develop into mature seeds, while severely influenced embryos may form shrunken seeds. Since these defective seeds have not been observed in either of the single *et1-1* or *et2-2* mutant, it is reasonable that the defective seeds carrying aborted embryos are related to the loss of function in both *AtET* genes. Loss of their expression severely affects male gametophyte development and embryogenesis suggesting that they are crucial for regular development of embryos and survival of plants or they might control multiple gene networks affecting plant development. In plant species, epigenetic modifications of chromatin, which regulates transcription, have been proposed to play an important role in modulation of many developmental pathways (Berger and Gaudin, 2003). Taken together with putative function of GIY YIG domain, we speculate that AtET factors are likely to participate as components of chromatin-remodelling complexes and play a role in transcriptional regulation through nicking activity.

The lack of phenotypic alterations in individual mutants is probably because of functional redundancy in *AtET* family. Redundant functions are often due to different members of a family having overlapping developmental roles. It has previously been reported among the member of several plant multigene families. Gene members can exhibit functional redundancy, depending on the extent of divergence of function by changes in their coding sequences and/ or expression patterns. An interesting example of partial redundancy within gene families has been reported for the *Arabidopsis CAULIFLOWER* and *APETALA1* genes in which double mutants have a dramatic cauliflower-like floral meristem defect, while *cauliflower* single mutants have a wild type phenotype and *apetala1* single mutants have a milder floral-defective phenotype (Budziszewski *et al.*, 2001). Since 65 % of *Arabidopsis* genes are considered to be

members of gene families, redundant functions might be expected in this species (Budziszewski *et al.*, 2001; Sappl *et al.*, 2004). Moreover, 37.4 % of *Arabidopsis* gene families have more than 5 members in comparison with 12.1 % in *Drosophila melanogaster* and 24.0 % in *Caenorhabditis elegans* (Mladek *et al.*, 2003; Shopinski *et al.*, 2006) which creates redundant function and explains why many gene knock-out do not exhibit any visible phenotype (Willmann, 2001).

In other cases, phenotypic analyses might be difficult or impossible due to lethal mutants in homozygous state or redundant function of members in large multigene families. For example, there are 31 combinations of single, double, and multiple knock-outs for a family of five genes (Bouche and Bouchez, 2001). Using RNAi approach or in combination with mutants may be suitable choice for phenotypic investigation in these cases. For instance, in *Arabidopsis* the CUC1 (CUP-SHAPE-COTYLEDON1), CUC2, and CUC3 regulate embryonic shoot meristem formation and boundary specification of lateral organs including cotyledons (Aida *et al.*, 1999; Hibara *et al.*, 2006). Overexpression of MIR164A and MIR164B targeting mRNAs of CUC1 and CUC2 in the cuc3-2 mutant caused an almost complete lack in axillary bud formation, demonstrating a redundant function of CUC1, CUC2 and CUC3 (Hibara *et al.*, 2006; Nikovics *et al.*, 2006; Raman *et al.*, 2008).

Conclusions and outlook

As members of ET family, AtET factors are characterized by several highly conserved sequences with typical cysteine patterns. With this characteristic, AtET proteins are difficult to be synthesized as recombinant proteins in *E. coli*. By using a tightly regulated expression vector and modified purification, both AtET proteins were obtained successfully from *E. coli*.

Analyses of protein-DNA interaction by EMSA demonstrated that purified AtET proteins bind to DNA fragments containing seed-specific *napA*, *USP* or *LeB4* promoters as well as random sequence probes. Thus, AtET factors most likely function as non-sequence specific DNA-binding proteins. It is possible that binding property is restricted to ET domains (ET repeats). This observation was supported by the results from Western blot since specific nanobodies against AtET proteins did not block binding between proteins

and DNA fragments. In addition to ET repeats, AtET factors consist of a DNA single strand cutting domain (GIY-YIG-like domain) similar to that of bacterial UvrC protein and GIY-YIG homing nucleases, which successfully replaced GIY-YIG domain of the UvrC protein in *E. coli*. This result suggested that AtET factors with their inherent DNA nicking activity may be involved the catalysis of changes in higher order DNA structure, such as, for example, nucleosome sliding. It is highly interesting to note that the plant-specific AtET factors have recruited GIY-YIG domain and attached to the DNA-binding ET repeats in order to control DNA nicking.

It is conceivable that AtET factors with the critical biochemical activities such as DNA nicking and subsequent epigenetic control analogous to the imprinted DNA glycosylase DME would play essential role during important developmental stages such as plant reproduction. Unfortunately, knocking-out each of the *AtET* genes (*et1-1* and *et2-1* mutants) did not exhibit any obvious morphological phenotypes, suggesting that their function may be masked by functional redundancy. I therefore initiated a double mutant by crossing between *et1-1* and *et2-1*. However, I have not yet obtained homozygous mutant lines for both loci and these experiments are currently being continued. Interestingly, a combination between knock-down for *AtET1* (RNAi) and knock-out for *AtET2* (*et2-1* mutant) revealed novel phenotypes. Aberrant pollen development and subsequent disruption of embryogenesis in these mutant plants led to embryo abortion. It is unclear if there is endosperm and/or suspensor failure. Since we lack the data of endosperm and suspensor development in these plants, additional experiments such as detailed morphological marker analysis using CLSM will be needed to distinguish between these possibilities.

Taken together, the data suggest that AtET factors are crucial for male gametophyte development as well as regulation of embryogenesis. Making a parallel to the DNA glycosylase DME involved in excision of methylated nucleotides and thereby establishment of active transcription marks, the DNA nicking activity of AtETs together with their expression in the seeds suggest a possibility of a comparable function of ET in imprinting. They might involve as components of chromatin-remodelling complexes and modulate gene transcription through their nicking activity.

5. Summary

EFFECTOR OF TRANSCRIPTION (ET) factors are structurally characterized by highly conserved C-terminal cysteine containing, zinc and DNA binding repeats, strictly confined to the plant kingdom as evidence from their occurrence in both higher plants [for instance, BnET of *Brassica napus* and HRT (*Hordeum* Repressor of Transcription) of *Hordeum vulgare*] and lower plants such as *Physcomitrella patens*. ET homologues were initially isolated as novel transcription factors expressed abundantly in seeds of *H. vulgare* and *B. napus*. Subsequent work in our lab has established that AtET2 in *Arabidopsis* is involved during seed germination and xylem differentiation presumptively through the modulation of a plant hormone gibberellin.

In the present study, I have focused on the three following objectives: (i) technology development such as AtET protein production in bacteria and nanobody production for further functional work; (ii) biochemical analysis such as the DNA binding activity of AtET and an *in vivo* demonstration of a conserved DNA single strand cutting functional domain in AtET; and (iii) functional analysis of the AtET family based on T-DNA mutants and RNAi constructs. The major aspect of the third objective was to establish by genetic analysis that the functionally redundant ET genes have pleiotropic role during plant development as implied from their ubiquitous gene expression, particularly during plant reproduction.

The full length of AtET proteins (approximate 55 and 65 kDa for AtET1 and AtET2 proteins, respectively) were synthesized and purified from bacteria as recombinant proteins. A total of the four ET repeats within AtET1/AtET2 contain twelve conserved cysteine residues at the C-terminal. Inherently, AtET proteins are difficult to be synthesized as recombinant proteins in *E. coli* due to their cysteine-rich domains, likely because cysteines could form aberrant intra- or intermolecular disulfide bridges leading to cause mis-folding and degradation. By using the tightly regulated expression vector and modified purification, we obtained sufficient amount purified AtET proteins for the nanobody production using phage library screening. Both the purified AtET proteins and corresponding nanobodies serve as established resources for further experiments.

Protein-DNA interaction performed by electrophoretic mobility shift assay (EMSA) demonstrated that AtET factors bind not only to DNA fragments containing seed-specific promoters but also random probes. Thus, they most likely function as non-sequence specific DNA binding proteins. It is possible that additional interacting factors might be required to direct specific DNA binding. The observation that specific nanobodies against AtET proteins did not block the interaction between AtET factors and the oligonucleotide probes suggests that the nanobodies do recognize protein regions outside of the DNA-binding ET domain.

Transfection of full length *AtET1* and *AtET2* genes under their native promoters in *Arabidopsis* protoplasts in transient assays indicated the nuclear localization of the corresponding AtET proteins. The accumulation of these proteins allowed us to clearly demonstrate the localization in the living plant cells *in vitro*. Both the DNA binding activity and nuclear localization of AtET1 and AtET2 provide clear evidence for their function as transcriptional regulators.

Another interesting characteristic feature of AtET factors is DNA single strand cutting domain (GIY-YIG-like domain) similar to that of bacterial UvrC protein and GIY-YIG homing nuclease. This domain is located in the second exon in all AtET members separated from ET domain by an intron. By domain swapping experiments, I could demonstrate that the AtET2 GIY-YIG-like domain can productively cooperate with the C-terminal domain of the bacterial UvrC protein. Since the single amino acid exchange (R>A) results in the near complete loss of this activity, the single strand cutting activity of the GIY-YIG domain must clearly be required for its molecular function. This domain might be involved in the catalysis of changes in higher order DNA structure, such as, for example, nucleosome sliding or may contribute to the relaxation of supercoiled chromatin domains.

Functional redundancy between AtET1 and AtET2 was expected due to (a) their homologous protein sequence and conserved domains; and (b) their overlapping gene expression reported in publicly available indexed microarrays. Thus, temporal and spatial expression pattern of *AtET2* alone was examined by marker analysis in transgenic lines carrying a double reporter (GUS, β-glucuronidase; GFP, green fluorescent protein) transcriptionally fused to a 1.7 kb putative AtET2 promoter (pET2). The reporters were detected either by histochemical (GUS) analysis or detection of GFP. GUS staining was

well defined in meristem and vascular tissues of seedlings such as shoot apexes, root tips, vein of expanded cotyledons, and central cylinder of roots. In mature plants, GUS activity was detected in the veins of rosette and cauline leaves, similar to that observed in seedlings. In reproductive organs, GUS was apparently detectable in pollen and in the chalazal region of the ovules but no expression was found in the embryo sacs. These observations were further confirmed by GFP detection as well. Together, while pET2 promoter expression is ubiquitous through plant developmental stages, it is exclusively expressed in the male gametophyte (pollen), not in the female (embryo sac) implying a sex-specific expression of ET.

Analyses of individual T-DNA insertion mutants of AtET gene family did not show obvious morphological phenotypes indicating the possibility of functional redundancy between these family members. An exception to this rule was a subtle precocious seed germination phenotype in *et1* mutant in analogy to the previously reported data for *et2*. I therefore constructed a double mutant between *et1-1* and *et2-1* by genetic crossing. As an alternative to this ongoing work, a combination between knock-down for *AtET1* (RNAi) and knock-out for *AtET2* (*et2-1* mutant) was performed to dissect their function. The obtained data indicate that loss of function in both *AtET* genes can cause pollen developmental defect, but has no alteration on female gametophyte development. In addition, down expression of these genes disrupts embryogenesis and consequently lead to embryo abortion at various stages, suggesting that AtET factors are essential for male development as well as normal embryogenesis in *Arabidopsis*. These mutant phenotypes in reproductive tissues are consistent with the observed expression patterns of the AtET2 promoter.

Overall, the data reported in this work support the notion that both AtET1 and AtET2 are novel and redundant DNA binding plant transcription factors with DNA nicking activity that function pleiotropically during plant development. Equally important is the discovery that these factors are critical for male gametophyte development as well as regulation of embryogenesis.

6. Zusammenfassung

EFFECTOR OF TRANSCRIPTION (ET) Faktoren sind strukturell durch C-terminale, hoch-konservierte, DNA- und Zink- bindende repetitive Protein-Domänen mit einem definierten Cystein-Muster charakterisiert. Sie wurden ursprünglich als putative Transkriptionsregulatoren während der Samenentwicklung von *Vicia faba*, *Brassica napus* und *Hordeum vulgare* beschrieben. ET-Faktoren werden ausschließlich in Pflanzen gefunden und das Vorkommen von ET-ähnlichen Genen im Genom von *Physcomitrella patens* belegt ihre phylogenetische Konservierung. Nachfolgende Arbeiten an *Arabidopsis thaliana* zeigten, dass ET-Faktoren -wahrscheinlich durch die Modulation des Phytohormons Gibberellin- an der Regulation der Samenkeimung sowie der Xylemdifferenzierung beteiligt sind.

Drei Entwicklungen stehen im Mittelpunkt der vorliegenden Arbeiten: a) die Schaffung methodischer Voraussetzungen für die weitere Analyse von ET-Faktoren durch Synthese in Bakterien und die Erzeugung spezifischer Antikörper (*nanobodies*), b) die biochemische Analyse der DNA-Bindungsaktivität sowie die *in vivo* Untersuchung zur Funktion einer DNA-Einzelstrang schneidenden Proteindomäne und c) die genetische Charakterisierung der ET-Genfamilie auf der Basis von T-DNA-Mutanten und RNAi-Konstruktionen. Letzteres zeigt, dass die funktionell redundanten ET-Gene eine pleiotrope Rolle bei der Kontrolle pflanzlicher Entwicklung, insbesondere während der Reproduktion spielen.

Vollständige ET-Proteine (AtET1, 55 kD und AtET2, 65 kD) wurden in *E. coli* synthetisiert und als rekombinante Proteine isoliert. Die geringe Syntheseeffizienz wird auf die hohe Zahl konservierter Cystein-Reste in den vier ET-Domänen zurückgeführt. Wahrscheinlich führt die Ausbildung aberranter intra- und intermolekularer Disulfidbrücken zu Fehlfaltung und Abbau der Proteine. Durch die Anwendung eines streng regulierten Expressionsvektors sowie eines modifizierten Verfahrens zur Proteinreinigung wurden ausreichende Mengen an ET-Protein für die Erzeugung spezifischer Antikörper (*nanobodies*) aus *phage display*-Bibliotheken erhalten. Sowohl die gereinigten ET-Proteine als auch die erzeugten Antikörper sind wertvolle Werkzeuge für weitere Experimente.

Untersuchungen zur Interaktion isolierter ET-Proteine mit DNA durch *electromobility shift assays* (EMSA) zeigen, dass ET-Faktoren sowohl an samenspezifische Genpromotoren, aber auch Sequenz-unabhängig an DNA binden. Möglicherweise werden weitere interagierende Proteine für eine Sequenzmotiv-spezifische DNA-Bindung benötigt. Die Beobachtung, dass die erzeugten Antikörper die Interaktion von ET-Faktoren und DNA nicht blockieren, spricht dafür, dass die Antikörper Proteinregionen außerhalb der DNA-Bindungsdomäne erkennen.

Transiente Expression von ET-Genen unter Kontrolle der nativen Promotoren in *Arabidopsis* Protoplasten zeigt, dass ET-Proteine *in vivo* im Zellkern vorliegen. Kernlokalisation sowie DNA-Bindung legen eine Funktion von ET-Faktoren als Genregulatoren nahe.

Eine interessante Besonderheit von ET-Faktoren ist das Vorkommen einer DNA-Einzelstrang-schneidenden Protein-Domäne (GIY-YIG-Domäne). Ähnliche Domänen werden in bakteriellen UVRC-Proteinen sowie in GIY-YIG-Nukleasen (*homing nucleases*) gefunden. In allen bekannten ET-Genen wird diese Domäne allein vom zweiten Exon kodiert, während die repetitiven ET-Domänen von einem separaten Exon kodiert werden. Versuche mit ausgetauschten Protein-Domänen (*domain swapping*) belegen, dass die GIY-YIG-Domäne der ET-Faktoren funktionell mit der C-terminalen Domäne bakterieller UVRC-Proteine kooperieren kann. Ein Aminosäureaustausch im aktiven Zentrum (R>A) führt zum Verlust der kooperativen Aktivität. Dies zeigt, dass die DNA-Einzelstrang-schneidende Aktivität der ET-Faktoren für ihre molekulare Funktion benötigt wird. Die Domäne könnte z. B. Veränderungen höherer DNA-Strukturen katalysieren, wie sie bei Veränderungen von Nukleosomen (*nucleosome sliding*) oder bei der Dekondensation von Chromatin (*supercoiled chromatin*) beschrieben wurden.

Für die detaillierte Analyse des Expressionsmusters von *AtET2* wurden transgene *Arabidopsis*-Linien erzeugt, die ein doppeltes Reportergen (GUS, β-glucuronidase; GFP, green fluorescent protein) unter der Kontrolle eines 1.7 kb *AtET2*-Genpromotors exprimieren.

Die Reporter wurden histochemisch bzw. fluoreszenzanalytisch nachgewiesen. Beide Reporter werden in Meristemen und vaskulärem Gewebe von Keimlingen und adulten

Pflanzen wie z. B. Spross- und Wurzelapex, Gefäßen von Kotyledonen sowie im Zentralzylinder der Wurzel nachgewiesen. In reproduktiven Organen wird GUS-Aktivität in Pollen sowie in der chalazalen Region der Ovulen, aber nicht im Embryosack, gefunden. GFP-Detektion bestätigen diese Beobachtungen. Die Promotoraktivität von *AtET2* im männlichen, nicht aber im weiblichen Gametophyten weist auf eine sex-spezifische Expression von ET-Genen hin.

Individueller T-DNA-Insertionsmutanten der *AtET*-Genfamilie zeigen keine auffälligen phänotypischen Veränderungen. Zusammen mit den oben beschriebenen Vorkommen konservierter Proteindomänen sowie dem ähnlichen Genexpressionsmuster legen diese Befunde die funktionelle Redundanz der Mitglieder der Genfamilie nahe. Ausgenommen davon ist der schwach ausgeprägte Phänotyp bezüglich der vorzeitigen Keimung von *et1* und *et2*-Mutanten. Entsprechende Doppelmutanten werden gegenwärtig erzeugt und untersucht. Als eine Alternative zur Ausschaltung beider Gene wurde ein *AtET1*-spezifisches RNAi-Konstrukt in die *et2*-Mutante transformiert. Bisherige Daten zeigen, dass der Verlust beider *ET*-Genfunktionen zu Störungen der Pollenentwicklung, nicht aber der Entwicklung des weiblichen Gametophyten führt. Außerdem führt der Verlust beider *ET*-Genfunktionen zu einer Unterbrechung der Embryogenese und dem Absterben des Embryos.

Die beschriebenen Ergebnisse charakterisieren ET-Faktoren als neue, redundant funktionierende, DNA-bindende Transkriptionsfaktoren mit DNA-Einzelstrang-schneidender Aktivität und pleiotropen Funktionen während der pflanzlichen Entwicklung. Die vorläufige Mutantenanalyse belegt darüber hinaus wichtige Funktionen während der Entwicklung des männlichen Gametophyten sowie der frühen Embryogenese.

7. References

Aida, M., T. Ishida and M Tasaka (1999). Shoot apical meristem and cotyledon formation during *Arabidopsis* embryogenesis: interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. *Development* 126: 1563-1570.

Alexander, M. P. (1969). Differential staining of aborted and nonaborted pollen. *Stain Technol* 44(3): 117-122.

Alvarez, C., L. Calo, L. C. Romero, I. Garcia and C. Gotor (2010). An O-acetylserine(thiol)lyase homolog with L-cysteine desulfhydrase activity regulates cysteine homeostasis in *Arabidopsis*. *Plant Physiol* 152(2): 656-669.

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(5): 1223-1242.

Baud, S., J-P. Boutin, M. Miquel, L. Lepiniec and C. Rochat (2002). An integrated overview of seed development in *Arabidopsis thaliana* ecotype Ws. *Plant Physiol. Biochem.* 40: 151-160.

Baudry, A., M. A. Heim, B. Dubreucq, M. Caboche, B. Weisshaar and L. Lepiniec (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *The Plant Journal* 39: 366-380.

Bäumlein, H., S. Mise'ra, H. Luerssen, K. Koelle, C. Horstmann, U. Wobus and A. J. Mueller (1994). The FUS3 gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *The Plant Journal* 6: 379-387.

Benita, Y., M. J. Wise, M. C. Lok, I. Humphery-Smith and R. S. Oosting (2006). Analysis of high throughput protein expression in *Escherichia coli*. *Mol Cell Proteomics* 5(9): 1567-1580.

Bentsink, L., J. Hanson, C. J. Hanhart, H. Blankestijn-de Vries, C. Coltrane, P. Keizer, M. El-Lithy, C. Alonso-Blanco, M. Teresa de Andrése, M. Reymond, F. van Eeuwijk, S. Smeekens and M. Koornneef (2010). Natural variation for seed dormancy in *Arabidopsis* is regulated by additive genetic and molecular pathways. *PNAS* 107: 4264-4269.

Berger, F. (1999). Endosperm development. *Curr Opin Plant Biol* 2(1): 28-32.

Berger, F. and V. Gaudin (2003). Chromatin dynamics and *Arabidopsis* development. *Chromosome Res* 11(3): 277-304.

Berleth, T. (1998). Experimental approaches to *Arabidopsis* embryogenesis. *Plant Physiol Biochem* 36: 69-82.

Bertauche, N., J. Leung and J. Giraudat (1996). Protein phosphatase activity of abscisic acid insensitive 1 (ABI1) protein from *Arabidopsis thaliana*. *Eur J Biochem* 241(1): 193-200.

Bewley, J. D. (1997). Seed Germination and Dormancy. *Plant Cell* 9(7): 1055-1066.

Bies-Etheve, N., P. Gaubier-Comella, A. Debures, E. Lasserre, E. Jobet, M. Raynal, R. Cooke and M. Delseny (2008). Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. *Plant Mol Biol* 67(1-2): 107-124.

Blevins, J. S., A. T. Revel, A. H. Smith, G. N. Bachlani and M. V. Norgard (2007). Adaptation of a luciferase gene reporter and lac expression system to Borrelia burgdorferi. *Appl Environ Microbiol* 73(5): 1501-1513.

Bongers, R. S., J. W. Veening, M. Van Wieringen, O. P. Kuipers and M. Kleerebezem (2005). Development and characterization of a subtilin-regulated expression system in *Bacillus subtilis*: strict control of gene expression by addition of subtilin. *Appl Environ Microbiol* 71(12): 8818-8824.

Bouche, N. and D. Bouchez (2001). *Arabidopsis* gene knockout: phenotypes wanted. *Curr Opin Plant Biol* 4(2): 111-117.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248-254.

Brady, S. M., S. F. Sarkar, D. Bonetta and P. McCourt (2003). The ABSCISIC ACID INTENSITIVE3 (ABI3) gene is modulated by farnesylation and is involved in auxin signaling and lateral root development in *Arabidopsis*. *The Plant Journal* 34: 67-75.

Braybrook, S. A., S. L. Stone, S. Park, A. Q. Bui, B. H. Le, R. L. Fischer, R. B. Goldberg and J. J. Harada (2006). Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *PNAS* 103: 3468-3473.

Brocard, I. M., T. J. Lynch and R. R. Finkelstein (2002). Regulation and role of the *Arabidopsis* Abscisic Acid-Insensitive 5 gene in abscisic acid, sugar, and stress response. *Plant physiology* 129: 1533-1543.

Budziszewski, G. J., S. P. Lewis, L. W. Glover, J. Reineke, G. Jones, L. S. Ziemnik, J. Lonowski, B. Nyfeler, G. Aux, Q. Zhou, J. McElver, D. A. Patton, R. Martienssen, U. Grossniklaus, H. Ma, M. Law and J. Z. Levin (2001). *Arabidopsis* genes essential for seedling viability: isolation of insertional mutants and molecular cloning. *Genetics* 159(4): 1765-1778.

Capron, A., S. Chatfield, N. Provart and T. Berleth. Eds. (2009). Embryogenesis: Pattern formation from a single cell. The Arabidopsis book.

Chalfie, M. and G. Euskirchen (1994). Green fluorescent protein as a marker for gene expression. *Science*.

Charlton, K. A.(2003). *Expression and isolation of recombinant antibody fragments in E. coli*. Antibody Engineering: Methods and Protocols. K. Lo Benny**:** 245-254.

Chaudhury, A. M., A. Koltunow, T. Payne, M. Luo, M. R. Tucker, E. S. Dennis and W. J. Peacock (2001). Control of early seed development. *Annu Rev Cell Dev Biol* 17: 677-699.

Cheng, N. H., J. K. Pittman, B. J. Barkla, T. Shigaki and K. D. Hirschi (2003). The *Arabidopsis* cax1 mutant exhibits impaired ion homeostasis, development, and hormonal responses and reveals interplay among vacuolar transporters. *The Plant Cell* 15: 347-364.

Choi, J. H. (2000). Efficient secretory production of alkaline phosphatase by high cell density culture of recombinant *Escherichia coli* using the *Bacillus* sp. endoxylanase signal sequence. *Appl Microbiol Biotechnol* 53: 16.

Choi, J. H. and S. Y. Lee (2004a). Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* 64: 11.

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 (2004b). An invariant aspartic acid in the DNA glycosylase domain of DEMETER is necessary for transcriptional activation of the imprinted MEDEA gene. *PNAS* 101: 7481-7486.

Chuang, C. F. and E. M. Meyerowitz (2000). Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 97(9): 4985-4990.

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.

Curaba, J., t. Moritz, R. Blervaque, F. Parcy, V. Raz, M. Herzog and G. Vachon (2004). AtGA3ox2, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in *Arabidopsis*. *Plant Physiol* 136: 3660-3669.

Dassa, B., N. London, B. L. Stoddard, O. Schueler-Furman and S. Pietrokovski (2009). Fractured genes: a novel genomic arrangement involving new split inteins and a new homing endonuclease family. *Nucleic Acids Res* 37(8): 2560-2573.

Dean Rider, S., Jr., M. R. Hemm, H. A. Hostetler, H. C. Li, C. Chapple and J. Ogas (2004). Metabolic profiling of the *Arabidopsis* pkl mutant reveals selective derepression of embryonic traits. *Planta* 219(3): 489-499.

Dean Rider, S., Jr., J. T. Henderson, R. E. Jerome, H. J. Edenberg, J. Romero-Severson and J. Ogas (2003). Coordinate repression of regulators of embryonic identity by PICKLE during germination in *Arabidopsis*. *Plant J* 35(1): 33-43.

Debeljak, N., L. Feldman, K. L. Davis, R. Komel and A. J. Sytkowski (2006). Variability in the immunodetection of His-tagged recombinant proteins. *Anal Biochem* 359(2): 216-223.

Deblaere, R., B. Bytebier, H. De Greve, F. Deboeck, J. Schell, M. Van Montagu and J. Leemans (1985). Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. *Nucleic Acids Res* 13: 4777-4788.

Derbyshire, Victoria, Joseph C. Kowalski, John T. Dansereau, Charles R. Hauer and Marlene Belfort (1997). Two-domain structure of the td intron-encoded endonuclease I-Tev I correlates with the two-domain configuration of the homing site. *J. Mol. Biol.* 265: 494-506.

Donohue, K. (2009). Completing the cycle: maternal effects as the missing link in plant life histories. *Phil. Trans. R. Soc* 364: 1059-1074.

Drees, J., J. Smith, F. Schafer and K. Steinert (2004). High-throughput expression and purification of 6xHis-tagged proteins in a 96-well format. *Methods Mol Med* 94: 179-190.

Dunin-Horkawicz, Stanislaw, Marcin Feder and Janusz M Bujnicki (2006). Phylogenomic analysis of the GIY-YIG nuclease superfamily *BMC Genomics* 7: 1-19.

Durbarry, A., I. Vizir and D. Twell (2005). Male germ line development in *Arabidopsis*. duo pollen mutants reveal gametophytic regulators of generative cell cycle progression. *Plant Physiol* 137: 297-307.

Edgar, R., M. Domrachev and A. E. Lash (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1): 207-210.

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(6): 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 develomental processes. *Plant Molecular Biology* 59: 663-681.

Fabianek, Renata A., Martina Huber-Wunderlich, Rudi Glockshuber, Peter Künzler, Hauke Hennecke and Linda Thöny-Meyer (1997). Characterization of the *Bradyrhizobium japonicum* CycY protein, a membrane-anchored periplasmic thioredoxin that may play a role as a reductant in the biogenesis of c-type cytochromes. *The journal of biological chemistry* 272: 4467-4473.

Fallahi, H., G. N. Scofield, M. R. Badger, W. S. Chow, R. T. Furbank and Y. L. Ruan (2008). Localization of sucrose synthase in developing seed and siliques of *Arabidopsis thaliana* reveals diverse roles for SUS during development. *J Exp Bot* 59(12): 3283-3295.

Fauteux, F. and M. V. Stromvik (2009). Seed storage protein gene promoters contain conserved DNA motifs in *Brassicaceae, Fabaceae* and *Poaceae*. *BMC Plant Biol* 9: 126.

Figueroa, P., G. Gusmaroli, G. Serino, J. Habashi, L. Ma, Y. Shen, S. Feng, M. Bostick, J. Callis, H. Hellmann and X. D. Deng (2005). *Arabidopsis* has two redundant cullin3 proteins that are essential for embryo development and that interact with RBX1 and BTB proteins to form multisubunit E3 ubiquitin ligase complexes in vivo. *The Plant Cell* 17: 1180-1195.

Finkelstein, R. R. (1994). Mutations at two new *Arabidopsis* ABA response loci are similar to the abi3 mutations. *The Plant Journal* 5: 765-771.

Finkelstein, R. R., S. S.L. Gampala and C. D. Rock (2002). Abscisic acid signaling in sedds and seedlings. *The Plant Cell*: 15-45.

Finkelstein, R. R. and T. J. Lynch (2000). The *Arabidopsis* abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. *Plant Cell* 12(4): 599-609.

Finkelstein, R. R. and C. R. Somerville (1990). Three classes of absiscic acid (ABA) insensitive mutations in *Arabidopsis* define genes that control overlapping subsets of ABA responses. *Plant Physiol* 94: 1172-1179.

Finkelstein, R. R., M. L. Wang, T. J. Lynch, S. Rao and H. M. Goodman (1998). The *Arabidopsis* abscisic acid response locus ABI4 encodes an APETALA 2 domain protein. *Plant Cell* 10(6): 1043-1054.

Finkelstein, R., W. Reeves, T. Ariizumi and C. Steber (2008). Molecular aspects of seed dormancy. *Annu Rev Plant Biol* 59: 387-415.

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(1): 91-101.

Foley, M. E. (2001). Seed dormancy: an update on terminology, physiological genetics, and quantitative trait loci regulating germinability. *Weed Science* 49: 305-317.

Franco-Zorrilla, J. M., P. Cubas, J. A. Jarillo, B. Ferna´ndez-Calvı´n, J. Salinas and J. M. Martı´nez-Zapater (2002). AtREM1, a member of a new family of B3 domain-containing genes, is preferentially expressed in reproductive meristems. *Plant Physiol* 128: 418-427.

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

Gahrtz, M. and U. Conrad (2009). *Immunomodulation of plant function by in vitro selected single-chain Fv intrabodies*, Humana Press.

Gao, M. and J. Skolnick (2009). From nonspecific DNA-protein encounter complexes to the prediction of DNA-protein interactions. *PLoS Comput Biol* 5(3): e1000341.

Garcia, D., V. Saingery, P. Chambrier, U. Mayer, G. Jurgens and F. Berger (2003). *Arabidopsis* haiku mutants reveal new controls of seed size by endosperm. *Plant Physiol* 131(4): 1661-1670.

Gazzarrini, S., Y. Tsuchiya, S. Lumba, M. Okamoto and P. McCourt (2004). The transcription factor FUSCA3 controls developmental timing in *Arabidopsis* through the hormones gibberellin and abscisic acid. *Developmental Cell* 7: 373-385.

Giacalone, M. J., A. M. Gentile, B. T. Lovitt, N. L. Berkley, C. W. Gunderson and M. W. Surber (2006). Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *BioTechniques* 40: 355-363.

Gilmartin, P. M. and C. Bowler. Eds. (2002). Molecular plant biology, OXFORD University Press.

Giraudat, J., B. M. Hauge, C. Valon, J. Smalle, F. Parcy and H. M. Goodman (1992). Isolation of the *Arabidopsis* ABI3 gene by positional cloning. *Plant Cell* 4(10): 1251-1261.

Goldberg, R. B., G. de Paiva and R. Yadegari (1994). Plant Embryogenesis: Zygote to Seed. *Science* 266(5185): 605-614.

Gong, Z., T. Morales-Ruiz, R. R. Ariza, T. Rolda´ n-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.

Gowers, D. M. and S. E. Halford (2003). Protein motion from non-specific to specific DNA by three-dimensional routes aided by supercoiling. *EMBO J* 22(6): 1410-1418.

Grelet, J., A. Benamar, E. Teyssier, M. H. Avelange-Macherel, D. Grunwald and D. Macherel (2005). Identification in pea seed mitochondria of a late-embryogenesis abundant protein able to protect enzymes from drying. *Plant Physiol* 137(1): 157-167.

Grkovic, S., M. H. Brown and R. A. Skurray (2002). Regulation of bacterial drug export systems. *Microbiol Mol Biol Rev* 66(4): 671-701, table of contents.

Guo, Y. and S. Gan (2006). AtNAP, a NAC family transcription factor, has an important role in leaf senescence. *The Plant Journal* 46: 601-612.

Gusti, A., N. Baumberger, M. Nowack, S. Pusch, H. Eisler, T. Potuschak, L. D. Veylder, A. Schnittger and P. Genschik (2009). The *Arabidopsis thaliana* F-Box protein FBL17 Is essential for progression through the second mitosis during pollen development. *PLoS One* 4.

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

Hansen, R. E., D. Roth and J. R. Winther (2009). Quantifying the global cellular thiol-disulfide status. *Proc Natl Acad Sci U S A* 106(2): 422-427.

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

Harper, B. and N. Stewart JR (2000). Patterns of green fluorescent protein expression in transgenic plants. *Plant Molecular Biology Reporter* 18: 141a-141i.

Henderson, J. T., H. C. Li, S. Dean Rider, A. P. Mordhorst, J. Romero-Severson, J. C. Cheng, J. Robey, Z. R. Sung, S. C. de Vries and J. Ogas (2004). PICKLE acts throughout the plant to repress expression of embryonic traits and may play a role in gibberellin-dependent responses. *Plant Physiol* 134: 995-105.

Herman, E. M. . Ed. (1995). Cell and molecular biology of seed oil bodies Seed development and Germination, Marcel Dekker Inc.

Hibara, K., M. R. Karim, S. Takada, K. Taoka, M. Furutani, M. Aida and M. Tasaka (2006). *Arabidopsis* CUP-SHAPED COTYLEDON3 regulates postembryonic shoot meristem and organ boundary formation. *Plant Cell* 18(11): 2946-2957.

Higo, K., Y. Ugawa, M. Iwamoto and H. Higo (1998). PLACE: a database of plant *cis*-acting regulatory DNA elements. *Nucleic Acids Res* 26(1): 358-359.

Higo, K., Y. Ugawa, M. Iwamoto and T. Korenaga (1999). Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* 27(1): 297-300.

Hilson, P., Joke Allemeersch, Thomas Altmann, Se´bastien Aubourg, Alexandra Avon, Jim Beynon, Rishikesh P. Bhalerao, Fre´de´rique Bitton, Michel Caboche, Bernard Cannoot, Vasil Chardakov, Ce´cile Cognet-Holliger, Vincent Colot, Mark Crowe, Caroline Darimont, Steffen Durinck, Holger Eickhoff, Ande´ol Falcon de Longevialle, Edward E. Farmer, Murray Grant, Martin T.R. Kuiper, Hans Lehrach, Ce´line Le´on, Antonio Leyva, Joakim Lundeberg, Claire Lurin, Yves Moreau, Wilfried Nietfeld, Javier Paz-Ares, Philippe Reymond, Pierre Rouze´, Goran Sandberg, Maria Dolores Segura, Carine Serizet, Alexandra Tabrett, Ludivine Taconnat, Vincent Thareau, Paul Van Hummelen, Steven Vercruysse, Marnik Vuylsteke, Magdalena Weingartner, Peter J. Weisbeek, Valtteri Wirta, Floyd R.A. Wittink, Marc Zabeau and Ian Small2 (2004). Versatile gene-specific sequence tags for *Arabidopsis f*unctional genomics: Transcript profiling and reverse genetics applications. *Genome Research* 14: 2176-2189.

Hirner, B., W. N. Fischer, D. Rentsch, M. Kwart and W. B. Frommer (1998). Developmental control of H+/amino acid permease gene expression during seed development of *Arabidopsis*. *Plant J* 14(5): 535-544.

Hobbie, L. and M. Estelle (1995). The axr4 auxin-resistant mutants of *Arabidopsis thaliana* define a gene important for root gravitropism and lateral root initiation. *The Plant Journal* 7: 211-220.

Hou, A., K. Liu, N. Catawatcharakul, X. Tang, V. Nguyen, W. A. Keller, E. W. Tsang and Y. Cui (2005). Two naturally occurring deletion mutants of 12S seed storage proteins in *Arabidopsis thaliana*. *Planta* 222(3): 512-520.

Hruz, T., O. Laule, G. Szabo, F. Wessendorp, S. Bleuler, L. Oertle, P. Widmayer, W. Gruissem and P. Zimmermann (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics* 2008: 420747.

Hsieh, K. and A. H. C. Huang (2004). Endoplastic reticulum, oleosins, and oils in seeds and tapetum cells. *Plant Physiol* 136: 3427-3434.

Huang, Y. F., C. C. Huang, Y. C. Liu, Y. J. Oyang and C. K. Huang (2009b). DNA-binding residues and binding mode prediction with binding-mechanism concerned models. *BMC Genomics* 10 Suppl 3: S23.

Huang, Y., Z. Su, Y. Li, Q. Zhang, L. Cui, Y. Su, C. Ding, M. Zhang, C. Feng, Y. Tan, W. Feng, X. Li and L. Cai (2009a). Expression and Purification of glutathione transferase-small ubiquitin-related modifier-metallothionein fusion protein and its neuronal and hepatic protection against D-galactose-induced oxidative damage in mouse model. *J Pharmacol Exp Ther* 329(2): 469-478.

Hundertmark, M. and D. K. Hincha (2008). LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* 9: 1-22.

Ingouff, M., J. Haseloff and F. Berger (2005). Polycomb group genes control developmental timing of endosperm. *Plant J* 42(5): 663-674.

Ivanov, R (2005). Molecular characterization of the EFFECTOR OF TRANSCRIPITION (ET) gene family in *Arabidopsis* and its role in plant development. Halle, Germany, Martin-Luther-Universität Halle-Wittenberg. Doctor**:** 149.

Ivanov, R, T Jens, C. Andreas, S. Anna, L. H. Diep, H.-P. Mock, C. Bernhard, T. Annegret and B. Helmut (2008). EFFECTOR OF TRANSCRIPTION2 is involved in xylem differentiation and includes a functional DNA single strand cutting domain. *Developmental Biology* 313: 93-106.

Iwakawa, H., A. Shinmyo and M. Sekine (2006). *Arabidopsis* CDKA;1, a cdc2 homologue, controls proliferation of generative cells in male gametogenesis. *Plant J* 45(5): 819-831.

Johnston, A. J., P. Meier, J. Gheyselinck, S. E. J. Wuest, M. Federer, E. Schlagenhauf, J. D. Becker and U. Grossniklaus (2007). Genetic subtraction profiling identifies genes essential for *Arabidopsis* reproduction and reveals interaction between the female gametophyte and the maternal sporophyte. *Genome Biology* 8.

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.

Jung, W-J. and S-J. Choi (1997). Secretion of human angiogenin into periplasm and culture medium with Eukaryotic signal sequence by *Escherichia coli*. *J. Biochem. Mol. Biol.* 30: 80-84.

Jurgens, G. (2001). Apical-basal pattern formation in *Arabidopsis* embryogenesis. *EMBO J* 20(14): 3609-3616.

Kagaya, Y., K. Ohmiya and T. Hattori (1999). RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res* 27: 470-478.

Kagaya, Y., R. Toyoshima, R. Okuda, H. Usui, A. Yamamoto and T. Hattori (2005). LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. *Plant Cell Physiol* 46(3): 399-406.

Kalodimos, C. G., N. Biris, A. M. J. J. Bonvin, M. M. Levandoski, M. Guennuegues, R. Boelens and R. Kaptein (2004). Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* 305: 386-389.

Kang, J-Y., H-I Choi, M-Y. Im and S. Y. Kim (2002). *Arabidopsis* basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. *The Plant Cell* 14: 343-357.

Karakas, E., J. J. Truglio, D. Croteau, B. Rhau, L. Wang, B. Van Houten and C. Kisker (2007). Structure of the C-terminal half of UvrC reveals an RNase H endonuclease domain with an Argonaute-like catalytic triad. *EMBO J* 26(2): 613-622.

Katoh, K., G. Asimenos and H. Toh.(2009). *Multiple alignment of DNA sequences with MAFFT*. Bioinformatics for DNA sequence analysis. David Posada, Humana Press. 537**:** 39-64.

Kaur, J. and A. K. Bachhawat (2007). Yct1p, a novel, high-affinity, cysteine-specific transporter from the yeast *Saccharomyces cerevisiae*. *Genetics* 176(2): 877-890.

Keith, K., M. Kraml, N. G. Dengler and P. McCourt (1994). fusca3: A heterochronic mutation affecting late embryo development in *Arabidopsis*. *The Plant Cell* 6(5): 589-600.

Kilian, J., D. Whitehead, J. Horak, D. Wanke, S. Weinl, O. Batistic, C. D'Angelo, E. Bornberg-Bauer, J. Kudla and K. Harter (2007). The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J* 50(2): 347-363.

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

King, S. P., J. E. Lunn and R. T. Furbank (1997). Carbohydrate content and enzyme metabolism in developing Canola siliques. *Plant Physiol* 114(1): 153-160.

Kipriyanov, S. M.(1998). *Purification, characterization, and biotinylation of single-chain antibodies*. Molecular diagnosis of infectious diseases. U. Reischl, Humana Press. 13**:** 615-622.

Kipriyanov, S. M. .(2002). *High-Level periplasmic expression and purification of ScFvs*. Antibody phage display: Methods and protocols. P. M. O'Brien and R. Aitken, Humana Press. 178**:** 333-341.

Kneusel, Richard E., Joanne Crowe, Melanie Wulbeck and Joachim Ribbe.(2000). *Procedures for the analysis and purification of His-tagged proteins*. The nucleic acid protocols - Handbook. Ralph Rapley, Humana Press**:** 921-934.

Koag, M. C., R. D. Fenton, S. Wilkens and T. J. Close (2003). The binding of maize DHN1 to lipid vesicles. Gain of structure and lipid specificity. *Plant Physiol* 131(1): 309-316.

Koehler, C., L. Hennig, R. Bouveret, J. Gheyselinck, U. Grossniklaus and W. Gruissem (2003). *Arabidopsis* MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *EMBO J* 22: 4804-4814.

Koornneef, M., G. Reuling and C. M. Karssen (1984). The isolation and characterisation of abscisic acid-insensitive mutants of *Arabidopsis thaliana*. *Plant Physiol* 90: 463-469.

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-TevI: coincidence of computational and molecular findings. *Nucleic Acids Res* 27(10): 2115-2125.

Krebbers, E., L. Herdies, A. De Clercq, J. Seurinck, J. Leemans, J. Van Damme, M. Segura, G. Gheysen, M. Van Montagu and J. Vandekerckhove (1988). Determination of the processing sites of an *Arabidopsis* 2S Albumin and characterization of the complete gene family. *Plant Physiol* 87(4): 859-866.

Kristof, J. A., J. L. Coppersmith, K. Hong, P. P. Liu, T. M. Homrichhausen, J. Sun, R. C. Martin and H. Nonogaki (2008). An *Arabidopsis thaliana* embryo arrest mutant exhibiting germination potential. *Seed Science Research* 18: 55-65.

Kroj, T., G. Savino, C. Valon, J. Giraudat and F. Parcy (2003). Regulation of storage protein gene expression in *Arabidopsis*. *Development* 130(24): 6065-6073.

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

Kunieda, T., N. Mitsuda, M. Ohme-Takagi, S. Takeda, M. Aida, M. Tasaka, M. Kondo, M. Nishimura and I. Hara-Nishimura (2008). NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in *Arabidopsis*. *Plant Cell* 20(10): 2631-2642.

Kwong, R. W., A. Q. Bui, H. Lee, L. W. Kwong, R. L. Fischer, R. B. Goldberg and J. J. Harada (2003). LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. *The Plant Cell* 15(1): 5-18.

Lara, P., L. Onate-Sanchez, Z. Abraham, C. Ferrandiz, I. Diaz, P. Carbonero and J. Vicente-Carbajosa (2003). Synergistic activation of seed storage protein gene expression in *Arabidopsis* by ABI3 and two bZIPs related to OPAQUE2. *J Biol Chem* 278(23): 21003-21011.

Laux, T., T. Wurschum and H. Breuninger (2004). Genetic regulation of embryonic pattern formation. *Plant Cell* 16 Suppl: S190-202.

Lee, H., R. L. Fischer, R. B. Goldberg and J. J. Harada (2003). *Arabidopsis* LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. *PNAS* 100: 2152-2156.

Leichert, L. I. and U. Jakob (2004). Protein thiol modifications visualized *in vivo*. *PLoS Biol* 2(11): e333.

Leung, J., S. Merlot and J. Giraudat (1997). The *Arabidopsis* ABSCISIC ACID-INSENSlTIVE2 (AB12) and ABI1 genes encode homologous Protein Phosphatases 2C involved in Abscisic Acid signal transduction. *The Plant Cell* 9: 759-771.

Li, H. C., K. Chuang, J. T. Henderson, S. D. Rider, Jr., Y. Bai, H. Zhang, M. Fountain, J. Gerber and J. Ogas (2005). PICKLE acts during germination to repress expression of embryonic traits. *Plant J* 44(6): 1010-1022.

Li, Z. and T. L. Thomas (1998). PEI1, an embryo-specific zinc finger protein gene required for heart-stage embryo formation in *Arabidopsis*. *Plant Cell* 10(3): 383-398.

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(25): 17688-17692.

Liu, X. Q., S. Zhang, X. M. Pan and C. C. Wang (1999). A novel method for increasing production of mature proteins in the periplasm of *Escherichia coli*. *Protein Science* 8: 2085-2089.

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

Lopez-Molina, L. and N. H. Chua (2000). A null mutation in a bZIP factor confers ABA-insensitivity in *Arabidopsis thaliana*. *Plant Cell Physiol* 41(5): 541-547.

Lopez-Molina, L., S. Mongrand and N-H. Chua (2001). A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in *Arabidopsis*. *PNAS* 98: 4782-4787.

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.

Luerssen, H., V. Kirik and S. Mise'ra (1998). The FUSCA3 gene of *Arabidopsis thaliana* encodes a product with partial homology to Vp1/ABI3-like regulatory proteins. *Plant Journal* 15: 755-764.

Mansfield, S. G. and L. G. Briarty. Eds. (1991). Early embryogenesis in *Arabidopsis thaliana*. II. The developing embryo, Can.J.Brot.

Mayer, U., R. A. Torres Ruiz, T. Berleth, S. Misera and G. Jürgens (1991). Mutations affecting body organization in the *Arabidopsis* embryo. *Nature* 353: 402-407.

McCarty, D. R., T. Hattori, C. B. Carson, V. Vasil, M. Lazaz and I. K. Vasil (1991). The Viviparous-1 developmental gene of maizes encodes a novel transcriptional activator. *Cell* 66: 895-905.

Meinke, D. W. (1995). Molecular genetics of plant embryogenesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46: 369-394.

Meinke, D. W., L. H. Franzmann, T. C. Nickle and E. C. Yeung (1994). Leafy cotyledon mutants of *Arabidopsis*. *Plant Cell* 6(8): 1049-1064.

Meissner, R. C., H. Jin, E. Cominelli, M. Denekamp, A. Fuertes, R. Greco, H. D. Kranz, S. Penfield, K. Petroni, A. Urzainqui, C. Martin, J. Paz-Ares, S. Smeekens, C. Tonelli, B. Weisshaar, E. Baumann, V. Klimyuk, S. Marillonnet, K. Patel, E. Speulman, A. F. Tissier, D. Bouchez, J. J. Jones, A. Pereira, E. Wisman and et al. (1999). Function search in a large transcription factor gene family in *Arabidopsis*: assessing the potential of reverse genetics to identify insertional mutations in R2R3 MYB genes. *Plant Cell* 11(10): 1827-1840.

Mergulhao, F. J., D. K. Summers and G. A. Monteiro (2005). Recombinant protein secretion in *Escherichia coli*. *Biotechnol Adv* 23(3): 177-202.

Meyer, A. J. and R. Hell (2005). Glutathione homeostasis and redox-regulation by sulfhydryl groups. *Photosynth Res* 86(3): 435-457.

Meyer, K., M. P. Leube and E. Grill (1994). A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. *Science* 264(5164): 1452-1455.

Mladek, C., K. Guger and M. T. Hauser (2003). Identification and characterization of the ARIADNE gene family in *Arabidopsis*. A group of putative E3 ligases. *Plant Physiol* 131(1): 27-40.

Mönke, G., L. Altschmied, A. Tewes, W. Reidt, H. P. Mock, H. Bäumlein and U. Conrad (2004). Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. *Planta* 219: 158-166.

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 54-incision. *Nucleic Acids Res* 26: 462-468.

Morales-Ruiz, T., A. P. Ortega-Galisteo, M. I. Ponferrada-Marin, M. I. Martinez-Macias, R. R. Ariza and T. Roldan-Arjona (2006). DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci U S A* 103(18): 6853-6858.

Murphy, D. J. (1993). Structure, function and biogenesis of storage lipid bodies and oleosins in plants. *Prog Lipid Res* 32(3): 247-280.

Murphy, D. J. and J. Vance (1999). Mechanisms of lipid-body formation. *Trends Biochem Sci* 24(3): 109-115.

Nikovics, K., T. Blein, A. Peaucelle, T. Ishida, H. Morin, M. Aida and P. Laufs (2006). The Balance between the MIR164A and CUC2 Genes Controls Leaf Margin Serration in *Arabidopsis*. *The Plant Cell* 18: 2929-2945.

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(11): 2910-2917.

Ogas, J., S. Kaufmann, J. Henderson and C. Somerville (1999). PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in *Arabidopsis*. *PNAS* 96: 13839-13844.

Ooms, J., K. M. Leon-Kloosterziel, D. Bartels, M. Koornneef and C. M. Karssen (1993). Acquisition of desiccation tolerance and longevity in seeds of *Arabidopsis thaliana Plant Physiol* 102(4): 1185-1191.

Otegui, M. S., R. Herder, J. Schulze, R. Jung and L. A. Staehelin (2006). The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. *Plant Cell* 18(10): 2567-2581.

Pang, P. P., R. E. Pruitt and E. M. Meyerowitz (1988). Molecular cloning, genomic organization, expression and evolution of 12S seed storage protein genes of *Arabidopsis thaliana*. *Plant Molecular Biology* 11: 805-820.

Parcy, F., C. Valon, A. Kohara, S. Misera and J. Giraudat (1997). The ABSCISIC ACID-INSENSITIVE3, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of *Arabidopsis* seed development. *Plant Cell* 9(8): 1265-1277.

Parcy, F., C. Valon, M. Rayna, P. Gaubier-Comella, M. Delseny and J. Girauda (1994). Regulation of gene expression programs during *Arabidopsis* seed development: Roles of the AB13 locus and of endogenous Abscisic Acid. *The Plant Cell* 6: 1567-1582.

Park, S. and J. J. Harada (2008). *Arabidopsis* embryogenesis. Plant Embryogenesis. Mari'a F. Sua'rez and Peter V. Bozhkov, Humana Press. 427.

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(19): 3789-3799.

Pawloski, L. C., M. K. Kandasamy and R. B. Meagher (2006). The late pollen actins are essential for normal male and female development in *Arabidopsis*. *Plant Mol Biol* 62(6): 881-896.

Pritchard, M. P., L. McLaughlin and T. Friedberg.(2006). *Establishment of functional human cytochrome P450 monooxygenase system in Escherichia coli*. Cytochrome P450 protocols. Ian R. Phillips and Elizabeth A. Shephard**:** 19-31.

Raghavan, V. (2002). Induction of vivipary in *Arabidopsis* by silique culture: implications for seed dormancy and germination. *American Journal of Botany* 89: 766-776.

Raman, S., T. Greb, A. Peaucelle, T. Blein, P. Laufs and K. Theres (2008). Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in *Arabidopsis thaliana*. *Plant J* 55(1): 65-76.

Raventos, 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(36): 23313-23320.

Raz, V., J. H. Bergervoet and M. Koornneef (2001). Sequential steps for developmental arrest in *Arabidopsis* seeds. *Development* 128(2): 243-252.

Redman, J. C., B. J. Haas, G. Tanimoto and C. D. Town (2004). Development and evaluation of an *Arabidopsis* whole genome Affymetrix probe array. *Plant J* 38(3): 545-561.

Reyes, J., M. J. Rodrigo, J. M. Colmenero-Flores , J. V. Gil, A. Garay-Arroyo , F. Campos, F. Salamini, D. Bartels and A. A. Covarrubias (2005). Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects *in vitro*. *Plant, Cell and Enviroment* 28: 709-718.

Rizzuto, R., M. Brini, P. Pizzo, M. Murgia and T. Pozzan (1995). Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells. *Current Biology* 5: 635-642.

Rodriguez, P. L., G. Benning and E. Grill (1998). ABI2, a second protein phosphatase 2C involved in abscisic acid signal transduction in *Arabidopsis*. *FEBS Lett* 421(3): 185-190.

Rohde, A., M. Van Montagu and W. Boerjan (1999). The ABSCISIC ACID-INSENSITIVE 3 (ABI3) gene is expressed during vegetative quiescence processes in *Arabidopsis*. *Plant, Cell and Environment* 22: 261-270.

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

Santos-Mendoza, M., B. Dubreucq, M. Miquel, M. Caboche and L. Lepiniec (2005). LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in *Arabidopsis* leaves. *FEBS Letters* 579: 4666-4670.

Sappl, P. G., L. Onate-Sanchez, K. B. Singh and A. H. Millar (2004). Proteomic analysis of glutathione S -transferases of *Arabidopsis thaliana* reveals differential salicylic acid-induced expression of the plant-specific phi and tau classes. *Plant Mol Biol* 54(2): 205-219.

Scheidegger, P., W. Weiglhofer, S. Suarez, S. Console, J. Waltenberger, M. S. Pepper, R. Jaussi and K. Ballmer-Hofer (2001). Signalling properties of an HIV-encoded angiogenic peptide mimicking vascular endothelial growth factor activity. *Biochem J* 353(Pt 3): 569-578.

Schmid, M., T. S. Davison, S. R. Henz, U. J. Pape, M. Demar, M. Vingron, B. Scholkopf, D. Weigel and J. U. Lohmann (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat Genet* 37(5): 501-506.

Schmittgen, T. D. and K. J. Livak (2008). Analyzing real-time PCR data by the comparative C_T method. *Nature protocols* 3: 1101-1108.

Schumann, W. and L. C. S. Ferreira (2004). Production of recombinant proteins in *Escherichia coli*. *Genetics and Molecular Biology* 27: 442-453.

Shopinski, K. L., M. J. Iqbal, J. L. Shultz, D. Jayaraman and D. A. Lightfoot (2006). Development of a pooled probe method for locating small gene families in a physical map of soybean using stress related paralogues and a BAC minimum tile path. *Plant Methods* 2: 1746-1811.

Siloto, R. M. P., K. Findlay, A. Lopez-Villalobos, E. C. Yeung, C. L. Nykiforuk and M. M. Moloney (2006). The accumulation of oleosomes determines the size of seed oilbodies in *Arabidopsis*. *The Plant Cell* 18: 1961-1974.

Singh, S., G. E. Folkers, A. M. Bonvin, R. Boelens, R. Wechselberger, A. Niztayev and R. Kaptein (2002). Solution structure and DNA-binding properties of the C-terminal domain of UvrC from *E.coli*. *EMBO J* 21(22): 6257-6266.

Skriver, K. and J. Mundy (1990). Gene expression in response to abscisic acid and osmotic stress. *Plant Cell* 2(6): 503-512.

Smith, F. S. and S. Delbary-Gossart.(2000). *Electrophoretic mobility shift assay (EMSA)*. Colorectal cancer: Methods and protocols. S. M. Powell, Humana. 50**:** 249-257.

Sørensen, Hans Peter and Kim Kusk Mortensen (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology* 115 113–128.

Sorensen, M. B., A. M. Chaudhury, H. Robert, E. Bancharel and F. Berger (2001). Polycomb group genes control pattern formation in plant seed. *Current Biology* 11(4): 277-281.

Souter, M. and K. Lindsey (2000). Polarity and signalling in plant embryogenesis. *J Exp Bot* 51(347): 971-983.

Stoddard, B. L. (2005). Homing endonuclease structure and function. *Q Rev Biophys* 38(1): 49-95.

Stone, S. L., S. A. Braybrook, S. L. Paula, L. W. Kwong, J. Meuser, J. Pelletier, T. F. Hsieh, R. L. Fischer, R. B. Goldberg and J. J. Harada (2008). *Arabidopsis* LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proc Natl Acad Sci U S A* 105(8): 3151-3156.

Stone, S. L., L. W. Kwong, K. M. Yee, J. Pelletier, L. Lepiniec, R. L. Fischer, R. B. Goldberg and J. J. Harada (2001). LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *PNAS* 98: 11806-11811.

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.

Suzuki, M., C. Y. Kao and D. R. McCarty (1997). The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. *The Plant Cell* 9: 799-807.

Suzuki, M., H. H-Y. Wang and D. R. McCarty (2007). Repression of the LEAFY COTYLEDON 1/B3 Regulatory Network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. *Plant Physiol* 143: 902-911.

Ting, J. T. L., K. Lee, C. Ratnayake, K. A. Platt, R. A. Balsamo and A. H. C. Huang (1996). Oleosin genes in Maize kernels having diverse oil contents are constitutively expressed independent of oil contents. *Planta* 199: 158-165.

Tiwari, S., S. Wang, G. Hagen and T. J. Guilfoyle. Eds. (2006). Transfection assays with protoplasts containing integrated reporter genes. *Arabidopsis* protocols, Humana Press.

To, A., C. Valon, G. Savino, J. Guilleminot, M. Devic, J. Giraudat and F. Parcy (2006). A network of local and redundant gene regulation governs *Arabidopsis* seed maturation. *Plant Cell* 18(7): 1642-1651.

Tsien, R. Y. (1998). The green fluorescent protein. *Annu Rev Biochem* 67: 509-544.

Ulmasov, T., J. Murfett, G. Hagen and T. J. Guilfoyle (1997). Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* 9: 1963-1971.

Ungru, A., M. K. Nowack, M. Reymond, R. Shirzadi, M. Kumar, S. Biewers, P. E. Grini and A. Schnittger (2008). Natural variation in the degree of autonomous endosperm formation

reveals independence and constraints of embryo growth during seed development in *Arabidopsis thaliana*. *Genetics* 179: 829-841.

van der Klei, H., J. V. Damme, P. Casteels and E. Krebbers (1993). A fifth 2s albumin isoform is present in *Arabidopsis thaliana*. *Plant Physiol* 101: 1415-1416.

Van Roey, P., L. Meehan, J. C. Kowalski, M. Belfort and V. Derbyshire (2002). Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-TevI. *Nat Struct Biol* 9(11): 806-811.

Verdier, J. and R. D. Thompson (2008). Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. *Plant Cell Physiol* 49(9): 1263-1271.

Verhoeven, Esther E. A., Marian van Kesteren, Geri F. Moolenaar, Rob Visse and Nora Goosen (2000). Catalytic sites for 3' and 5' incision of *Escherichia coli n*ucleotide excision repair are both located in UvrC. *The journal of biological chemistry* 275: 5120–5123.

Vicient, C. M., N. Bies-Etheve and M. Delseny (2000). Changes in gene expression in the leafy cotyledon1 (lec1) and fusca3 (fus3) mutants of *Arabidopsis thaliana* L. *J Exp Bot* 51(347): 995-1003.

Vigeolas, H., J. T. van Dongen , P. Waldeck, D. Huhn and P. Geigenberger (2003). Lipid storage metabolism is limited by the prevailing low oxygen concentraions within developing seeds of oilseed rape. *Plant Physiol* 133: 2048-2060.

Voelker, T. and A. J. Kinney (2001). Variations in the biosynthesis of seed-storage lipids. *Annu Rev Plant Physiol Plant Mol Biol* 52: 335-361.

Wan, L., A. R. Ross, J. Yang, D. D. Hegedus and A. R. Kermode (2007). Phosphorylation of the 12 S globulin cruciferin in wild-type and abi1-1 mutant *Arabidopsis thaliana* (thale cress) seeds. *Biochem J* 404(2): 247-256.

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

West, M. and J. J. Harada (1993). Embryogenesis in Higher Plants: An Overview. *Plant Cell* 5(10): 1361-1369.

West, M., K. M. Yee, J. Danao, J. L. Zimmerman, R. L. Fischer, R. B. Goldberg and J. J. Harada (1994). LEAFY COTYLEDON1 is an essential regulator of late embryogenesis and cotyledon ldentity in *Arabidopsis*. *The Plant Cell* 6: 1731-1745.

Western, T. L., D. J. Skinner and G. W. Haughn (2000). Differentiation of mucilage secretory cells of the *Arabidopsis* seed coat. *Plant Physiol* 122(2): 345-356.

Willmann, M. R. (2001). *Arabidopsis* enters the post-sequencing era. *Trends Plant Sci* 6(2): 51.

Wobus, U. and H. Weber (1999). Seed maturation: genetic programmes and control signals. *Curr Opin Plant Biol* 2(1): 33-38.

Xiao, B., Y. Huang, N. Tang and L. Xiong (2007). Over-expression of a LEA gene in rice improves drought resistance under the field conditions. *Theor Appl Genet* 115(1): 35-46.

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.

Xiong, J., Z. H. Zhu, J. S. Liu, Y. Wang and H. S. Wu (2006). The expression of toll-like receptor 2, 4 of livers in mice with systemic inflammatory response syndrome. *Hepatobiliary Pancreat Dis Int* 5(1): 143-146.

Xiong, Y. and B. Ru (1997). Purification and characteristics of recombinant mouse metallothionein-I from *Escherichia coli*. *J Biochem* 121(6): 1102-1106.

Yaagoubi, A. E., M. Kohiyama and G. Richarme (1994). Localization of DnaK (chaperone 70) from *Escherichia coli* in an osmotic-shock-sensitive compartment of the cytoplasm. *J Bacteriol* 176(22): 7074-7078.

Yuan, J. S., D. Wang and C. N. Stewart, Jr. (2008). Statistical methods for efficiency adjusted real-time PCR quantification. *Biotechnol J* 3(1): 112-123.

Zhang, H., S. D. Rider, Jr., J. T. Henderson, M. Fountain, K. Chuang, V. Kandachar, A. Simons, H. J. Edenberg, J. Romero-Severson, W. M. Muir and J. Ogas (2008). The CHD3 remodeler PICKLE promotes trimethylation of histone H3 lysine 27. *J Biol Chem* 283(33): 22637-22648.

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(23): 3100-3112.

Zhong, S., Z. Lin, R. G. Fray and D. Grierson (2008). Improved plant transformation vectors for fluorescent protein tagging. *Transgenic Res* 17(5): 985-989.

Zimmermann, P., L. Hennig and W. Gruissem (2005). Gene-expression analysis and network discovery using Genevestigator. *Trends in Plant Science* 10: 407-409.

Zimmermann, P., M. Hirsch-Hoffmann, L. Hennig and W. Gruissem (2004). GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. *Plant Physiol* 136(1): 2621-2632.

Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 313 (2008) 93–106
DEVELOPMENTAL BIOLOGY
www.elsevier.com/developmentalbiology

EFFECTOR OF TRANSCRIPTION2 is involved in xylem differentiation and includes a functional DNA single strand cutting domain

Rumen Ivanov, Jens Tiedemann, Andreas Czihal, Anna Schallau, Le Hong Diep, Hans-Peter Mock, Bernhard Claus, Annegret Tewes, Helmut Bäumlein *

Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, D-06466 Gatersleben, Germany

Received for publication 18 June 2007; revised 26 September 2007; accepted 26 September 2007
Available online 11 October 2007
Dedicated to the 65th birthday of Prof. Dr. Ulrich Wobus.

Abstract

EFFECTORS OF TRANSCRIPTION2 (ET) are plant-specific regulatory proteins characterized by the presence of two to five C-terminal DNA- and Zn-binding repeats, and a highly conserved cysteine pattern. We describe the structural characterization of the three member *Arabidopsis thaliana ET* gene family and reveal some allelic sequence polymorphisms. A mutation analysis showed that *AtET2* affects the expression of various *KNAT* genes involved in the maintenance of the undifferentiated state of cambial meristem cells. It also plays a role in the regulation of *GA5* (gibberellin 3-beta-dioxygenase) and the cell-cycle-related *GASA4*. A correlation was established between *AtET2* expression and the cellular differentiation state. AtET–GFP fusion proteins shuttle between the cytoplasm and nucleus, with the *AtET2* product prevented from entering the nucleus in non-differentiating cells. Within the nucleus, AtET2 probably acts *via* a single strand cutting domain. A more general regulatory role for ET factors is proposed, governing cell differentiation in cambial meristems, a crucial process for the development of plant vascular tissues.

Keywords: Gene regulation; Xylem differentiation; Transcription factors; Single strand cutting; GIY-YIG domain

Introduction

Plant growth is determined by the action of a small number of cells present at the shoot and root apical meristems. The vascular cambium is a secondary meristem, derived from the shoot apex. Although these meristems differ in function, a growing body of evidence suggests that their regulation shares many common principles and related genes (Groover, 2005). One of the products of cambial activity is the xylem, which develops towards the centre of the stem. The differentiation of xylem cells from the cambium is characteristically accompanied by a gradual accumulation of lignin, which therefore serves as a useful indicator of the progression of xylem cell differentiation. This process is controlled by the activity of several factors, including the phytohormone gibberellin (GA) and transcription factors of the *KNOTTED1*-like homeobox *KNAT* family (Hake et al., 2004; Scofield and Murray, 2006).

The plant hormone gibberellin (GA) is essential for the differentiation of the vascular tissues. Experiments in poplar, hybrid aspen and tobacco (Israelsson et al., 2003; Biemelt et al., 2004) have demonstrated that transgenic plants which ectopically express the biosynthetic gibberellin 3 beta dioxygenase encoding gene (*GA5*) exhibit significantly increased levels of xylem lignification. On the contrary, depletion of active GA by the ectopic expression of a gene which encodes the GA degrading enzyme *GA2-oxidase* inhibits lignin accumulation in tobacco (Biemelt et al., 2004). Furthermore, expression profiling in hybrid aspen showed an induction of GA-regulated genes in the early stages of cell differentiation near the cambial meristem (Hertzberg et al., 2001a).

Members of the *KNAT* gene family act as major regulators of several GA-mediated functions by inhibiting both the biosynthesis of and the meristematic response to GA. The *KNAT* genes *BREVIPEDICELLUS* (*BP*) and *SHOOTMERISTEMLESS* (*STM*) act redundantly to repress the transcription of *Arabidopsis thaliana GA5* (gibberellin 3-beta-dioxygenase) (Hay et al., 2002). In addition, the tobacco protein NTH15 represses the expression of

* Corresponding author. Fax: +49 39482 5500.
E-mail address: baumlein@ipk-gatersleben.de (H. Bäumlein).

0012-1606/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2007.09.061

the GA20 oxidase gene *NTC12* by interacting directly with an element in its promoter (Sakamoto et al., 2001a). *KNAT* gene expression in the cambium is essential for the control of xylem differentiation and lignin formation (Smith and Hake, 2003; Brown et al., 2005; Ehlting et al., 2005). In particular, *BP* activity prevents cambium-derived cells from differentiating into lignified xylem tissue (Mele et al., 2003). Other class I *KNAT* genes such as *At STM* and *KNAT2* and poplar *KNAP2* have similar activity (Israelsson et al., 2003; Ko and Han, 2004; Schrader et al., 2004a; Groover, 2005; Demura and Fukudo, 2007).

In addition to *KNAT* genes, the previously characterized members of the EFFECTOR OF TRANSCRIPTION (ET) family including the barley protein HORDEUM REPRESSOR OF TRANSCRIPTION (HRT) are also involved in GA-mediated processes of xylem differentiation (Raventos et al., 1998; Ellerstrom et al., 2005). Originally, ET factors have been isolated as DNA-binding proteins by south-western screens from *Hordeum vulgare*, *Brassica napus* and *Vicia faba*. They represent strictly plant-specific proteins characterized by one (*Physcomitrella patens*), two (*V. faba*), three (*H. vulgare*), four (*B. napus*, *A. thaliana*) or five (*Populus trichocarpa*) highly conserved cysteine-containing structural domains with a $CX_{8-9}CX_{10}CX_2H$ consensus sequence, designated as ET repeats. These repeats have been shown to bind zinc in a stoichiometric ratio of close to 1:1, although the cysteine pattern differs greatly from classical zinc finger motifs. The HRT protein interacts with gibberellin response elements of various promoters, whereas *B. napus* ET – although it is able to strongly interact with DNA – does not show a clear sequence specificity as it was shown by an ELISA-based binding assay (Mönke et al., 2004 and Mönke et al., unpublished). Barley HRT is targeted to the nucleus and functional tests in plant cells indicated that HRT can regulate the activity of certain GA-responsive promoters, including two α-amylase gene promoters. Northern hybridizations indicate that *HRT* transcripts accumulate to low levels in various tissues and a role for HRT in mediating developmental and phytohormones-responsive gene expression have been discussed (Raventos et al., 1998). Recently, we have described a dicot ET factor from *B. napus* (BnET) providing evidence for its role in gibberellin signaling modulation and cell differentiation. BnET is also targeted to the nucleus and its ectopic expression in either *Arabidopsis* or tobacco causes a pleiotropic phenotype including dwarfism due to shorter internodes and late flowering, reduced germination rate, increased anthocyanin content and reduced xylem lignification as a marker for terminal cell differentiation. Transient expression in protoplasts and transcript analysis support the notion that this is most likely due to a transcriptional repression of GA-controlled genes. In contrast to other GA-deficient mutants, the shorter internodes were due to fewer but not smaller cells, suggesting a function of BnET in GA-mediated cell division control (Ellerstrom et al., 2005).

In this paper, we present the initial characterization of the ET family in *Arabidopsis*. A T-DNA insertion in the *AtET2* gene leads to defects in xylem differentiation as detected by distortions of lignification. Array hybridization and RT-PCR analysis demonstrate altered expression of several GA-related genes and members of the *KNAT* family. Two of the three *AtET* genes are specifically up-regulated in differentiating cells and their regulation involves post-transcriptional control of their nuclear localization, preventing the AtET1 and AtET2 factors from entering the nucleus in non-differentiating cells. The molecular function of ET proteins as regulators of transcription most likely involves the activity of a functional single strand cutting domain. The data suggest a novel function of ET factors in the regulation of cell differentiation in cambial meristems.

Materials and methods

Molecular cloning

PCR, restriction digestion and DNA ligation were performed according to standard protocols (Molecular Cloning Third Edition, eds. Sambrook J. and Russel D., CSH Laboratory Press).

CAPS marker for the mutated AtET1 allele

A 1244-bp genomic fragment spanning the mutation site and an *Eco*RI cleavage site was amplified from the *AtET1* sequence, using primers 5′-ATGTTCAAGAGAGACGACTACA and 5′-ATCCTCGCATCGTTTTCTCC. The amplicon was digested with *Eco*RI (Amersham) and sized by agarose gel electrophoresis. The wild-type allele produced a 1082-bp fragment, whereas the frame-shifted allele produced a 897-bp one.

Plant transformation

The Ws-2 ecotype was transformed by vacuum infiltration as described (Bechthold et al., 1993).

ProAtET2-driven GUS expression

A 1.5-kbp upstream region of *AtET2* was placed ahead of the GUS reporter gene in the plasmid pMDC162 (Curtis and Grossniklaus, 2003), using GATEWAY cloning technology (Invitrogen). GUS activity was assayed in 2- to 4-week-old homozygous T3 plants following standard histological procedures, and the signal was visualized with a “Axioplan 2 imaging mot” (Zeiss, Jena, Germany) light microscope or a “StereoLumar V12” (Zeiss, Jena, Germany) binocular microscope.

In situ hybridization

Segments from the basal 5 mm of stem were fixed for 3 h, following vacuum infiltration with 2% paraformaldehyde, 0.2% glutardialdehyde, 0.01% Triton X100 in 0.5 M cacodylic acid buffer, pH 6.8. After two washing steps of 30 min each in the same buffer, the specimens were dehydrated in a graded ethanol series (10%, 30%, 50%, 70%, 90%, twice 96%, 100%) for at least 30 min per step. All buffers were prepared with ddH_20 treated with 0.05% DEPC. The embedding method was adapted from Tiedemann et al. (2000), using reduced incubation times. Samples were taken from plants subjected either to short days (8 h light) until bolting, or maintained under long days (16 h light). Hybridization was with *in vitro* transcribed riboprobes as described previously (Tiedemann et al., 2001) with the following modifications: hybridization conditions were 16 h at 50 °C, and the sections were washed (2×1 h) in 50% formamide in 0.5× SSC. Following RNase-A digestion (20 μg/ml, 5 mM NaCl, 1 mM EDTA, 10 mM Tris–HCl, pH 7.5), the sections were subjected to an additional stringent wash (50% formamide in 0.1× SSC, 50 °C) for a further 30 min. The primers used for the gene-specific probe synthesis were

T7 promoterAtET1 (antisense probe): 5′-AAACGACGGCCAGTGAATTG TAATACGACTCACTATAGGGCGAGTGACAACCAAACCGAAGAG;

T3 promoterAtET1 (sense probe): 5′-AAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGGTTATATTCTCAGTTTCTTCACATTG;
T7 promoter *AtET2* (antisense probe): 5′-AAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGGTTGGTATCAGAATAAAAGGA;
T3 promoter *AtET2* (sense probe): 5′-AAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGGTCACAACATCAGAGTCTTTATG.

Gene-specific regions of the primers are underlined. Immunological DIG detection was performed according to the manufacturer's instructions (Roche, Mannheim, Germany).

T-DNA insertion line

The T-DNA insertion line *et2-1* was isolated from the *Arabidopsis* knock-out facility collection, following an established pool screening strategy (Sussman et al., 2000). The primer pairs used for the detection of the wild-type *AtET2* allele were 5′-ATGGAATTCGGCGACGGCG and 5′-GGTGATTCTCATTCCCTTATG, and those for the T-DNA insertion allele were 5′-TGGGAAAACCTGGCGTTACCCAACTTAAT and 5′-TGCTCTTCACATCTCTTACGTCCTTTTAC.

Lignin measurement

Total lignin content was determined following thioacidolysis, using a published procedure (Campbell and Ellis, 1992). Four-week-old plants were pooled into five pools including five plants each, and 200 mg of stems and rosette leaves was extracted. Three independent measurements (technical replicates) were performed from each sample. Recovery rates for each individual experiment were determined by analyzing parallel samples with appropriate amounts of authentic lignin.

Hypocotyl growth induction

Single plants were grown in soil, and hypocotyl growth was stimulated by repeated clipping of the bolting stem over 5 weeks. At the end of this period, fresh hand sections were prepared. Lignin auto-fluorescence was visualized under an Axiovert135 fluorescent microscope (Zeiss, Jena) using an excitation wavelength of 325 nm with detection through a 420-nm long pass emission filter.

Array hybridization

Plants were grown for 10 days in liquid half-strength MS medium (Duchefa) with 10 g/l sucrose (Sigma) and 0.5 g/l MES buffer (Duchefa), pH 5.6. Poly A RNA was isolated using the Dynabeads mRNA Direct Kit (Dynal Biotech), according to the manufacturer's instructions. First strand cDNA was synthesized directly on the beads using AMV Reverse Transcriptase (Promega). The probes were labeled with ^{33}P-dCTP via random priming using the Megaprime DNA labeling Kit (Amersham) and hybridized to the REGIA 1200 *At* transcription factors filters (Paz-Ares and REGIA-Consortium, 2002). Data analysis was performed with Array VisionTM software (Imaging research Inc., Brock University, Ontario, Canada). Two independent experiments were performed and only expression differences of more than three-fold were retained.

RT-PCR analysis

Total plant RNA was isolated from 10-day-old seedlings, leaves, stems, flowers, siliques or dry seeds using the Total RNA Isolation Reagent (Biomol). Single stranded cDNA was synthesized using the Revert Aid First strand cDNA Synthesis Kit (Fermentas). Amplicons were separated by 1.5% agarose gel electrophoresis and were transferred to Hybond+ membranes (Amersham). cDNA probes were labeled with [^{32}P]dCTP by the means of the RediprimeTM II Random Prime Labeling Kit (Amersham) and hybridized to the membranes. Signal detection was achieved with a Phosphoimager (Fujifilm). The RT-PCR primer pairs were

AtET1: 5′-ATGTTCAAGAGAGACGACTACATTGC and 5′-AAGATGTCATTCTCATCCCCTTGTGC;
AtET2: 5′-CTATATCATCGGTTTTATCGAAATGGAATT and 5′-AAGTGATGCAGAGGTTAGGTGATTCTCATT;
compromised AtET2: 5′-ATCTAAGAGAGAAGCTGAGGCAACAGAAG and 5′-TGCTCTTCACATCTCTTACGTCCTTTTAC;
STM: 5′-AGAGTGGTTCCAACAGCA and 5′-TTAGTTCCTTGGGGAGGA;
KNAT1/BP1: 5′-CACCGTCTGTCTCTGCCTCCTCTA and 5′-ATTCCGCCAACGCTACCTTCTCT;
KNAT2: 5′-CGAACTCGCTACCGCTTTGTCCT and 5′-TCGCGGTCATTGCTTCTTTGTTG;
KNAT3: 5′-CCGGCGGTGGAGAAAACAA and 5′-TCCCCCATCGAACATATTAGCATC;
KNAT6: 5′-CTCCGCCGGTGAAAATCGTGT and 5′-GGTTCCGTAGCTGCATCTCAATCT;
FIL: 5′-ATGTCTATGTCGTCTATGTCC and 5′-TTAATAAGGAGTCACACCAACG;
GA5: 5′-ATGGCCGTAAGTTTCGTAAC and 5′-TTAGATGGGTTTGGTGAGCC;
GASA4: 5′-ATGGCTAAGTCATATGGAGC and 5′-TCAAGGGCATTTTGGTCCAC;
AtEf-1Bα: 5′-AGGAGAGGGAGGCTGCTAAG and 5′-AATCTTGTTGAAAGCGACAATG.

Protoplast transformation and transient assay

Transient expression experiments were performed as described elsewhere (Ellerstrom et al., 2005). *AtET2* and *GASA4* promoters were cloned into pGUS1 to drive the expression of the GUS reporter gene. Transformed protoplasts were grown in K3 medium containing 0.9 μM 6-benzylaminopurine and 0.1 μM 1-naphthalene acetic acid. For the localization of expression, the *AtET1*, *AtET2* and *AtET3* coding regions were used to generate a translational fusion to *EGFP*, driven by the CaMV35S promoter in pFF19g (ProCaMV35s-MCS-EGFP-ter) (Hofius et al., 2004). The empty pFF19g was used as control. The protoplast suspension was incubated in K3 medium adjusted for either non-differentiating (4.5 μM 6-benzylaminopurine, 10 μM 1-naphthalene acetic acid, 4.5 μM 2,4-dichlorophenoxyacetic acid) or differentiating (0.9 μM 6-benzylaminopurine, 0.1 μM 1-naphthalene acetic acid) conditions. EGFP signals were measured *in vivo* with a confocal laser scanning microscope, using an argon laser for induction at 488 nm and detection at 521 nm (CLSM Meta, Zeiss, Jena). The identity of the EGFP signal was verified by measuring the signal wavelength in λ-stack mode.

In vitro mutagenesis of AtET2

The coding sequence of *AtET2* was cloned into pCR2.1 (Invitrogen) and PCR mutagenized using Pfu polymerase (Stratagene) and the mutagenized primer sequence 5′-GAAAGTGTTAGGTCTGCACTTCAGCGTTATGG. The position of the mutation is underlined. The introduction of the mutation was confirmed by resequencing.

Domain swapping

The wild-type UVRC promoter and gene were amplified from *Escherichia coli* DH5α using primers 5′-GCTGATGTCAAAATCATCATG and 5′-TCAATGTTTCAACGACCAGAAG and were cloned into pCR2.1. To express the chimeric UVRC protein containing the AtET2 GIY-YIG domain, the *E. coli* UVRC promoter was amplified using 5′-GCTGATGTCAAAATCATCATG and 5′-CCCGGGCTTGATAATGTCTCCGCA. The amplicon was inserted into pCR2.1 and the resulting plasmid was linearized by *Sma*I restriction. The coding sequence for the AtET2 GIY-YIG domain, either with or without the R>A mutation, was amplified by Pfu polymerase with primers 5′-ATGTCTTGTCCGGGTCTGTATGAG and 5′-GATATCGTTAAGGTTGTTAACAT to ensure a blunt ended product. The amplicon was inserted into the linearized vector downstream of the UVRC promoter. The resulting construct was linearized by

*Eco*RV digestion and was ligated to the remainder of the UVRC coding sequence amplified with Pfu polymerase using primers 5′-CAGATCCAGC-AAATTGATGTA and 5′-TCAATGTTTCAACGACCAGAAG. The fidelity of both constructs was confirmed by resequencing.

Complementation assay

E. coli strain SOLR (Stratagene, La Jolla, CA) was used for the complementation test. Cells carrying either the empty vector, the *E. coli* UVRC gene, the chimeric UVRC protein containing either the wild-type or R>A mutated AtET2 GIY-YIG domain were grown until the mid-exponential phase (OD=0.6) in LB containing 50 μg/ml ampicillin. Cells were harvested by centrifugation at 5000 rpm for 10 min at 4 °C, washed twice and resuspended in M9 medium. A volume of 4 ml cell suspension was transferred to a 70-mm diameter Petri dish, producing a <2-mm-deep liquid layer. Irradiation was applied with a 4-W UV lamp (254 nm) from a distance of 90 cm in a dark room for 5, 10, 20, 40 or 60 s. At each time point, 100 μl cell suspension were diluted in M9 medium, and survival rate after irradiation was compared to that of the non-irradiated sample.

Results

Ecotype-specific polymorphisms in the AtET gene family

The *At* genome contains three ET-like sequences (Fig. 1). Of these, *AtET1* (At4g26170) shares the highest level of sequence homology to *BnET*. *AtET2* (At5g56780) and *AtET3* (At5g56770) are located adjacent to one another, so they probably represent the result of a duplication event. *AtET2* is an intact coding sequence, whereas *AtET3* lacks the coding region for the Zn- and DNA-binding C-terminal cysteine repeats (Fig. 1). Resequencing revealed that the Wassilewskaja-2 (Ws-2) and Columbia-0 (Col-0) ecotypes are polymorphic with respect to *AtET1*. The Ws-2 allele is functional, but the coding region in Col-0 is interrupted by stop codons, suggesting that it is probably a pseudogene. An adenine deletion at position 114 of the Col-0 cDNA sequence has generated a reading frame shift. Since this deletion creates an *Eco*RI site, it was possible to design a CAPS marker to discriminate between the Col-0 and Ws-2 alleles. The chosen amplicon was of length 1244 bp and contains an additional *Eco*RI site at position 162, which serves as an internal digestion control. *Eco*RI restriction thus generated both a 1082-bp and a 162-bp fragment from the Ws-2 allele, and a profile of 897 bp, 185 bp and 162 bp from the Col-0 allele. Of 85 *At* accessions screened in this way, only "Limeport" carries the Col-0 allele.

Despite the lacking C-terminal repeats, some *AtET3* transcript can be detected by RT-PCR (data not shown). Further resequencing revealed a mis-annotation of the exon–intron gene structure (At5g56770), as well as a 155-bp shorter coding region in the Ws-2 allele, resulting from a four base pair duplication at position 602, which creates a frameshift followed immediately by a stop codon. *AtET2* is intact in both Col-0 and Ws-2.

A GIY-YIG single strand cutting domain in ET factors

Since ET factors are located in the nucleus (see below), where they act as regulators of other genes, we searched for the presence of recognizable protein domains. In addition to the characteristic ET repeats, this identified a low level of similarity to the DNA single strand cutting domain present in bacterial UVRC proteins and in GIY-YIG homing nucleases (Derbyhire et al., 1997; Aravind et al., 1999, Verhoeven et al., 2000; Stoddard, 2005). The AtET GIY-YIG like domain is present in the second exon. Based on the derived three-dimensional structure of the bacterial GIY-YIG domain (Van Roey et al., 2002), the similarity

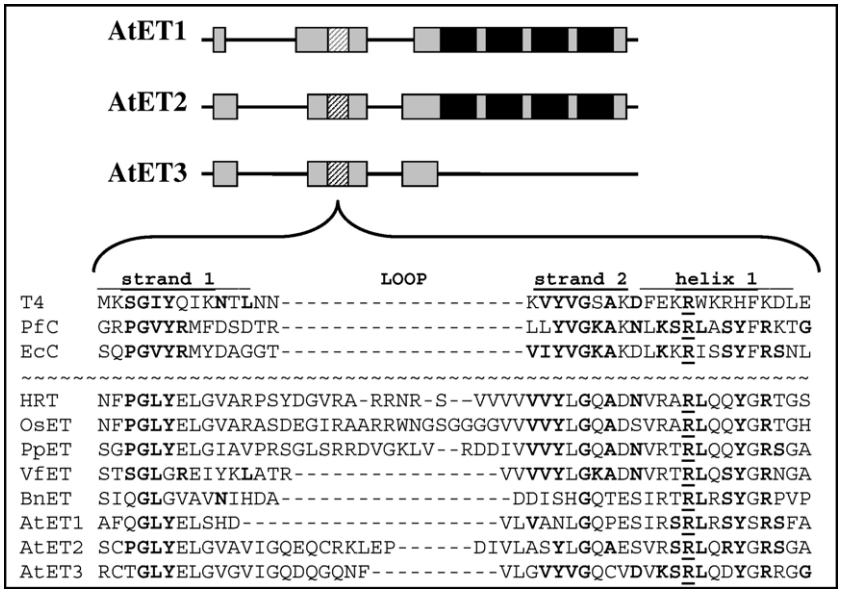


Fig. 1. Structure of the *AtET* gene family. Schematic representation of the three ET genes in ecotype Ws-2: exons are indicated by grey boxes and the ET repeat regions are shown in black. The GIY-YIG-like single strand cutting domain in the second intron is hatched. Domain sequences of prokaryotic and plant origin are shown (T4: phage T4; PfC: UVRC protein of *Pseudomonas fluorescens*; EcC: UVRC protein of *E. coli*; HRT: hordeum repressor of transcription, a barley ET factor; OsET, a rice ET factor; PpET: a poplar ET factor; VfET a broad bean ET factor; BnET: an oilseed rape ET factor; AtET1, 2, 3, AtET factors). The structural features strand, loop and helix are indicated. Conserved amino acids residues are highlighted in bold, and the arginine residue in helix1 is underlined.

between the prokaryotic proteins and the plant-specific ET factors is mainly confined to two conserved β-strands and helix 1 (Fig. 1). The variable length of the loop between both strands in the plant proteins probably does not disturb the overall structure of the protein. ET factors from barley, rice, poplar and *At* (AtET2, AtET3) all contain extended loops, whereas those from oilseed rape, broad bean and AtET1 have a loop length similar to those in T4 bacteriophage, *Pseudomonas fluorescens* and *E. coli*. Further sequence similarity between the prokaryotic and plant ET proteins resides in helix 1. The most conserved arginine residue is highlighted (Fig. 1). It is well established that the replacement of this residue by alanine results in a distortion of activity (Derbyhire et al., 1997; Kowalski et al., 1999; Verhoeven et al., 2000). To demonstrate the functionality of the GIY-YIG-like domain of plant ET factors, the wild-type domain, as well as the arginine to alanine (R>A) replacement allele, was used to replace the corresponding GIY-YIG domain of the *E. coli* UVRC protein (Fig. 2B). Plasmids encoding the chimeric proteins were transformed into the UVRC-deficient *E. coli* strain SOLR. Survival rates after UV irradiation were determined in four independent experiments. It was clear that the wild-type domain can partially relieve the UV sensitivity of SOLR, whereas the R>A mutation resulted in a reduced survival rate (although still slightly greater than in the presence of an empty vector control) (Fig. 2A). These data demonstrate that the AtET2 GIY-YIG-like domain can productively cooperate with the C-terminal domain of the bacterial UVRC protein.

To further confirm the importance of the GIY-YIG domain of the AtET2 factor also in plant cells, transient expression has been performed using the promoter of a NAM transcription factor gene (At4g28500). This gene promoter was chosen since the corresponding gene is down-regulated in the *et2-1* mutant described below. The transient co-expression of *AtET2* in *At* protoplasts resulted in an increased activity of the *NAM* promoter-GUS construct (Fig. 2C). In contrast, the R>A mutation in the AtET2 protein caused a near complete loss of this effect, as the promoter activity remained at a similar level to that in the control (empty vector). Thus, the GIY-YIG domain must be involved in the function of the AtET2 protein.

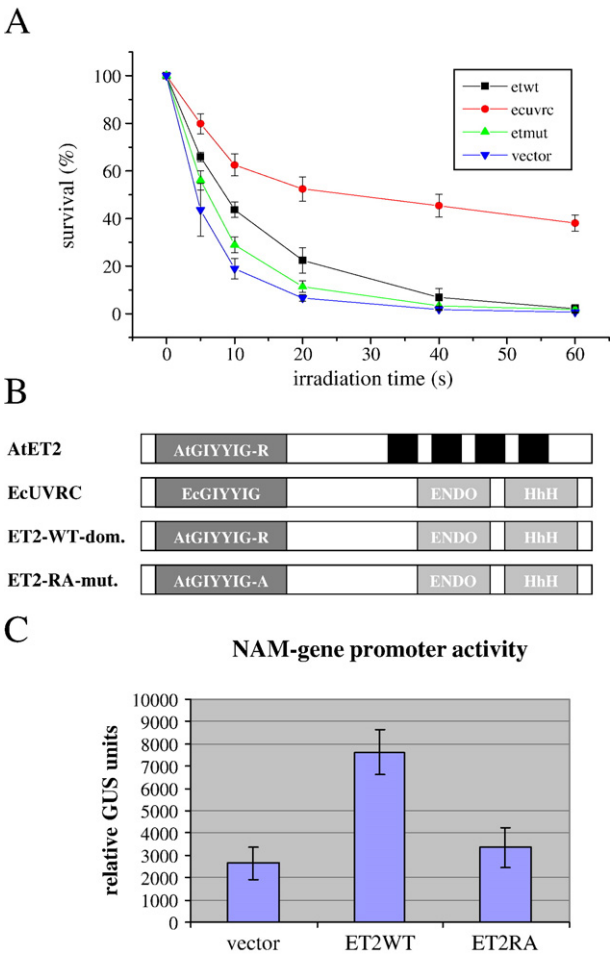


Fig. 2. Functional activity assays of the GIY-YIG single strand cutting domain. (A) UV survival curve of various chimeric ET-domain-constructs. The UV-deficient *E. coli* strain SOLR was transformed with the authentic *E. coli* wild-type UVRC protein (ecuvrc), the *E. coli* UVRC with its N-terminal GIY-YIG domain replaced by the corresponding wild-type domain of AtET2 (etwt) and the *E. coli* UVRC protein with its N-terminal GIY-YIG domain replaced by the AtET2 domain carrying the R>A point mutation (etmut). The survival rate in % is given as a function of the irradiation time in seconds. The standard deviation of the mean of four replicates is given. (B) Schematic structure of the domain swapped chimeric proteins. The four plant ET repeats are shown in black. GIY-YIG represents the N-terminal single strand domain cutting domain and ENDO and HhH, respectively, the C-terminal single strand cutting domain ENDO V and the Helix–hairpin–Helix domain. Domain sizes are not drawn to scale. (C) Transient co-expression of the AtET2 wild-type factor (ET2WT) and the AtET2 mutant factor (ET2RA) containing the R>A mutation with a *NAM* gene promoter-GUS reporter construct. The empty vector was used as a negative control. Four batches of protoplasts and plasmid preparations have been used. The bars represent the standard deviation.

Expression of AtET1 and AtET2 in vascular tissues

Due to low native expression levels, RT-PCR was used to assess the expression patterns of *AtET1* and *AtET2*. Both genes were expressed ubiquitously in plant organs (Fig. 3A). Surprisingly, the *AtEt2* transcript was undetectable in the cauline leaves. During seed development, *AtET2* was more strongly expressed in the early stages and was down-regulated in mature seeds, whereas *AtET1* was predominantly expressed in mature seeds. This inverse expression pattern in early and late seed development precisely reflects previously reported patterns (de Folter et al., 2004). An analysis of *At* plants transformed with a *ProAtET2-GUS-nosT* construct provided an explanation for the ubiquitous expression of *AtET2*. The promoter activity was detectable within the vascular tissues of stems, hypocotyls, leaves and flowers in homozygous T3 plants (Figs. 3B, 4B, C and 5A). In mature rosette leaves, expression was detectable in the whole vascular bundle region, apart from the fibre caps. In cross sections of shoots from plants at the time of flowering, expression was present in the xylem parenchyma as well as in the phloem and cambium (Fig. 4C). This pattern of expression was completely reproducible across over 20 independent transgenic individuals.

To further evaluate the reporter gene-based data, *AtET1* and *AtET2* transcripts were localized by *in situ* hybridization. Both transcripts were detected within the xylem parenchyma cells in the vascular bundles (Figs. 4D–H). In shoots – grown under

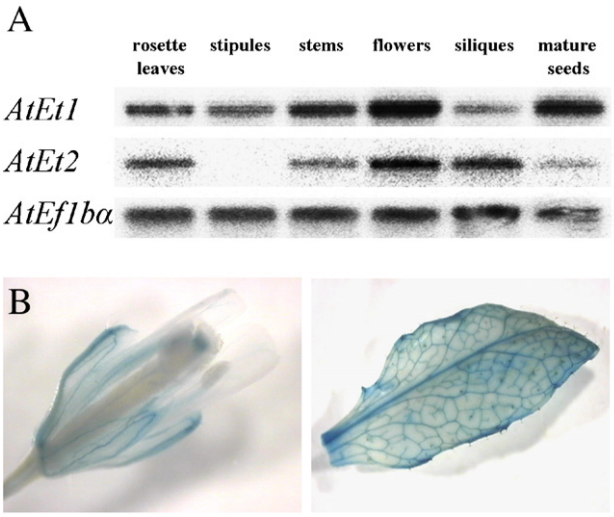


Fig. 3. Expression of *AtET1* and *AtET2*. (A) RT-PCR analysis of transcripts from various plant organs shows the ubiquitous expression of both genes. *AtEt2* mRNA was not detected in stipules. Expression is normalized according to the constitutively expressed *Ef1bα*. (B) GUS staining of a *ProAtET2-GUS-nosT* transgenic line. AtET2 expression is detected mainly in the vascular tissues of flowers (left) and leaves (right). The pattern is reproducible in >20 independent lines.

short day conditions – showing pronounced secondary growth, both transcripts were present in living xylem parenchyma cells, but a strong signal was also present in the dormant cambial zone (Figs. 4D, G). In addition, transcripts were also detectable in the remnant cytoplasm of differentiated fibre cells. *AtET1* and *AtET2* transcripts were also found in protoxylem element parenchymal cells. Although the expression pattern is essentially similar in plants showing little secondary growth, signal intensity was somewhat lower (Figs. 4E, H). These *in situ* hybridization data confirmed that the *ProAtET2-GUS-nosT* lines faithfully represent the expression pattern of *AtET2* and showed that *AtET1* and *AtET2* are turned on during the differentiation of the fibre elements.

Lignin content is reduced in ET2 mutant plants

An initial functional analysis of the *AtET* genes was effected by a study of *et2-1*, a Ws-2 T-DNA-insertion allele of *AtET2* selected from the *Arabidopsis* knock-out facility (AKF) collection (Sussman et al., 2000). The insertion event interrupts the second exon of the gene at nucleotide position 518 and is in the homozygous state, as demonstrated by both Southern hybridization and PCR. The absence of transcript was verified by RT-PCR. The *et2-1* mutant showed no obvious differences from wild type with respect to internode number, height, flowering time and leaf morphology. Since lignification is accepted as a reliable marker for the differentiation of xylem tissue (Mele et al., 2003), the lignin content of leaves and stems of mature plants was compared. The *et2-1* mutant line contained about 30% less lignin than did the wild type, both in the leaves and in the stems (Fig. 4A). This level of reduction corresponds well to changes in *AtET1* and *AtET2* expression levels in the vascular bundles and specifically in the xylem.

Furthermore, we have taken advantage of the potential of *Arabidopsis* for secondary growth as it has been described by Zhao et al. (2000). As in stems and leaves, the *ProAtET2-GUS-nosT* construct was active in the hypocotyl xylem (Fig. 5A). Repeated clipping of the bolting stem generated an increase in the diameter of the central cylinder of the hypocotyl, as a consequence of a prolonged xylem differentiation process induced by a delay in flowering time. After a week induction period, sections of Ws-2 and *et2-1* hypocotyls were compared by fluorescence microscopy (Fig. 5B). Wild-type plants reacted with a significant increase in the diameter of the central cylinder and of the hypocotyl as a whole (Figs. 5C, D). In contrast, the corresponding changes in *et2-1* were only modest (Figs. 5E, F). A quantification of these histological data is shown in Fig. 5B. Taken together, the data suggest that the lack of a functional AtET2 product perturbs normal cambial function and lignification.

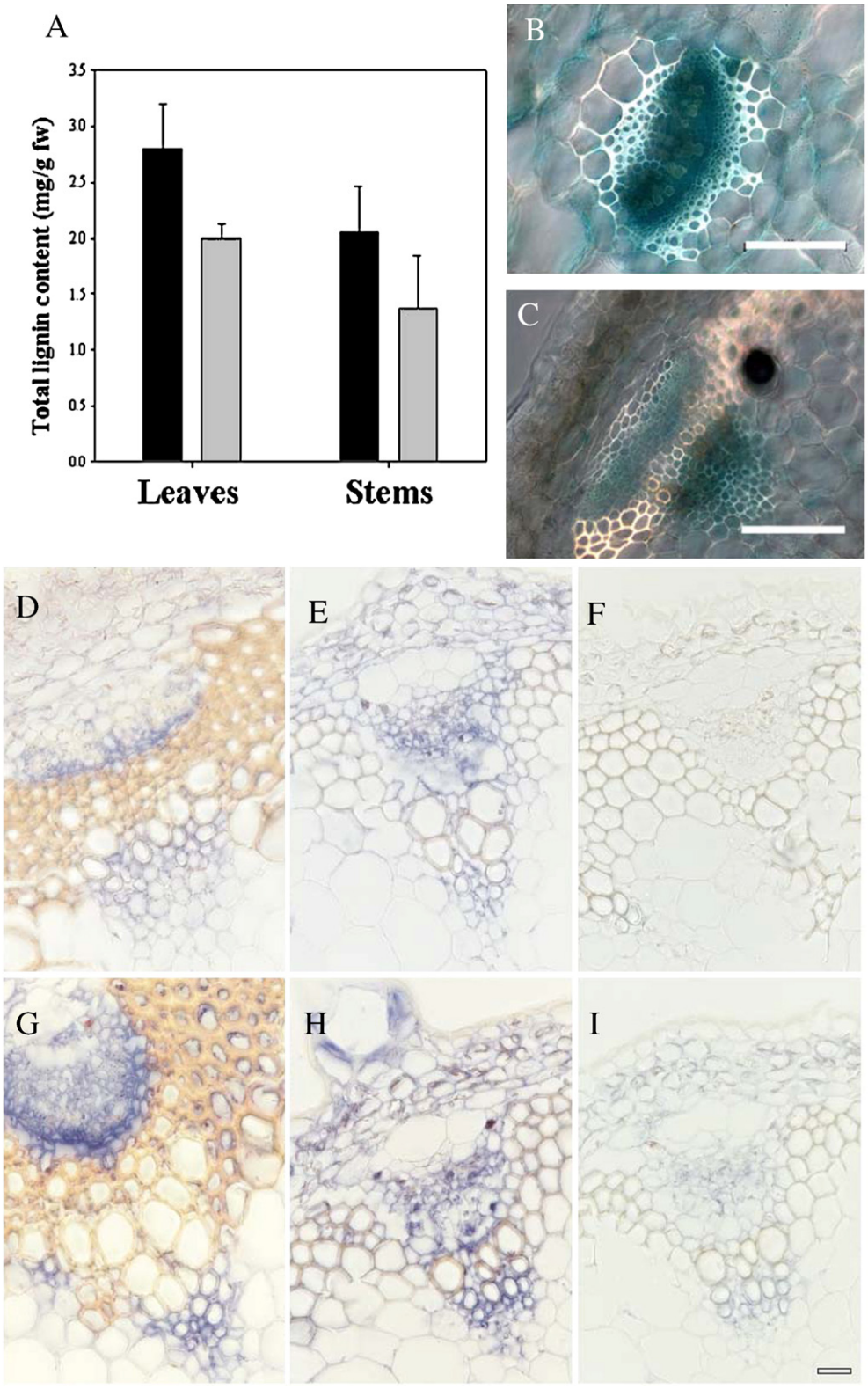
The et2-1 mutation affects the expression of meristem identity genes

A combined filter array hybridization and RT-PCR experiment was performed to elucidate the role of *AtET2* at the molecular level. The 1200 transcription factor REGIA consortium macro array was able to identify factors showing a differential pattern of expression between the *et2-1* mutant and the wild type (Table 1). Several meristem identity genes were up-regulated in the mutant, including *KNAT6* (class I) and *KNAT3* (class II). As not all KNAT gene family members are represented on the array, RT-PCR assays were applied for the members not represented on the array, including *BP1* (known to act as an inhibitor of lignification in the cambium of both *At* and poplar), *STM* and *KNAT2*. *BP1*, *KNAT6* and *KNAT3* were all up-regulated in the *et2-1* mutant, whereas *STM* and *KNAT2* were down-regulated. Interestingly, the *YABBY* gene *filamentous flowers* (*FIL*), a negative regulator of *KNAT* class I genes, was also up-regulated in *et2-1* mutant plants (Fig. 6A), as was the putative GA response inhibitor *lateral root primordia* (*LRP*), a member of the *SHI* family (Table 1).

Fig. 4. Lignification and *AtET2* expression in *et2-1* mutant plants. (A) Reduction in the lignin content of leaves and stems of *et2-1* (grey columns) compared to Ws-2 wild-type (black columns) plants. Four-week-old plants were pooled (five pools containing five plants each). In each pool, lignin concentration was determined by three independent measurements (technical replicates). The error bars represent the standard deviation of all measurements. (B, C) GUS staining of a 4-week-old plant showing the expression of *AtET2* in the xylem of (B) leaves (cross section of a leaf and the central vein) and (C) stems (cross section of the middle of the second internode above the rosette leaves). Bar length=100 μm. (D–I) *In situ* hybridization with *AtET1* probes (D–F) and *AtET2* probes (G–I). Sense controls are shown in panels F and I. Shoots of short (D, G) and long (E, H, F, I) day grown plants. Both ET factor mRNAs were detected in the cambial cells, the xylem parenchyma and the phloem region. Remnants of cytoplasm within fiber cells also show some hybridization signal (compare upper right area of panel G). Artefactual probe retention occurs in the protoxylem element cell wall. Bar length=20 μm.

AtET2 is required to suppress the GA response

BnET over-expression in tobacco down-regulates GA-responsive genes and triggers a feedback response in the GA biosynthesis pathway (Ellerstrom et al., 2005). Several of the genes up-regulated in *et2-1* are directly or indirectly modulated by GA. We therefore investigated a possible role of *AtET2* in the GA response, using a transient expression assay in *At* protoplasts and the GA-responsive *GASA4* promoter (Herzog et al., 1995). As previously shown for *BnET* (Ellerstrom et al.,



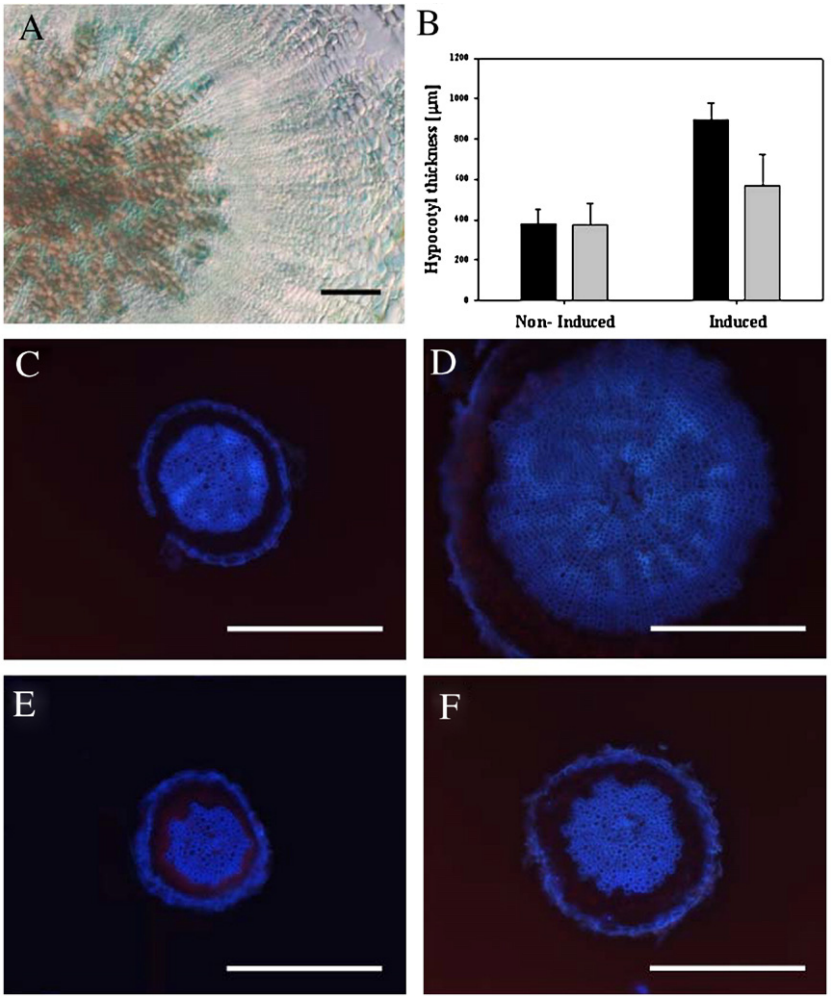


Fig. 5. Xylem differentiation in hypocotyls. (A) GUS (driven by *ProAtET2*) staining of a hypocotyl section from a 2-week-old plant, showing activity in the hypocotyl xylem tissue. Bar length = 30 µm. (B) Hypocotyl growth induced by repeated clipping of the bolting shoot results in an increase in diameter over wild-type hypocotyls (black columns), an effect which is much less pronounced in the *et2-1* mutant (grey columns). Twenty plants each of wild-type and mutant have been analyzed and the standard deviation is given. (C) Ws-2 wild-type hypocotyl in non-clipped 5-week-old plants. (D) Ws-2 wild-type hypocotyl of 5-week-old clipped plants. (E) *et2-1* mutant hypocotyl in non-clipped 5-week-old plants. (F) *et2-1* mutant hypocotyl of 5-week-old clipped plants. Bar length in panels C–F = 400 µm.

2005), the co-expression of *AtET2* driven by a constitutive CaMV35S promoter down-regulates the ProGASA4GUS construct in terms of its inducibility by GA (Fig. 6B). *GASA4* transcript was detectable by RT-PCR in the *et2-1* mutant (Fig. 6C). Similarly, *GA5* was more strongly induced in the *et2-1* mutant compared to wild type (Fig. 6D). These data support the notion that *AtET2* is required for the suppression of the GA response.

Transcriptional regulation of AtET2 includes a feedback mechanism

Feedback regulation is a common means of transcriptional control. The influence of *AtET2* expression on the activity of its own promoter was analyzed in a transient expression assay, where the activity of the *ProAtET2GUS* construct was shown to be reduced by the co-expression of *AtET2* (Fig. 6E). An *in vivo* verification was sought by determining the transcript level of the *et2-1* T-DNA-insertion allele. The 5′ end of the *et2-1 AtET2* transcript was tracked by RT-PCR. The lack of a functional AtET2 product resulted in an increased level of the truncated transcript (Fig. 6F), consistent with the transient assay data. Although differences in transcript stability cannot be excluded, the data are more consistent with the action of an auto-regulatory negative feedback mechanism regulating *AtET2* expression.

AtET2 expression depends on cell fate

The decision between maintenance in the meristematic state and cell differentiation can be modulated *in vitro* by the application of phytohormones (Valente et al., 1998; Grafi, 2004). We therefore established an *At* protoplast system, which can be triggered into a non-differentiating or a differentiating cell population by two different hormone regimes. Under high levels of auxin and cytokinin, non-differentiated cells are spherical and well separated from one another (Figs. 7A, C, E). Lowering the hormone concentrations drives the cell population

Table 1
Genes up-regulated in the *et2-1* mutant, as determined from a hybridization experiment with a transcription factor array filter

AGI number	Name	Function	Induction factor
At1g23380	KNAT6	Meristem identity, gibberellin response	33
At1g35540	ARF Protein	Auxin response	10
At2g40740	WRKY55		17
At3g25710	AtbHLH 32		42
At2g45190	Fil (Filamentous-Flowers)	Meristem identity, regulation of KNAT	38
At3g15030	TCP4	Cell division, leaf morphogenesis	42
At4g22070	WRKY31		19
At5g08330	TCP family bHLH protein	Auxin-induced protein	15
At5g12330	LRP1	Gibberellin response	14
At5g25220	KNAT3	Meristem identity, gibberellin response	30
At5g53980	Homeobox-leucine zipper protein		25

Only factors induced at >3-fold in two replicate hybridizations are included.

towards reorganization, enlargement and clumping. Finally, cell wall lignification is initiated, with the formation of xylem elements being taken as an indicator of an advanced level of differentiation (Figs. 7B, D, F). In addition, the expression of the two meristem identity genes *STM* and *BP1* was used as a marker. Both genes were highly expressed in the non-differentiating cell population, and both were down-regulated in the differentiating one (Fig. 7G). Semi-quantitative determination of *AtET1* and *AtET2* transcript levels showed that both were more abundant in the differentiating than in the non-differentiating population (Fig. 7G). *AtET1* and *AtET2* exhibited the same overall expression pattern, although the difference between the populations was more pronounced for the former.

AtET–GFP fusions exhibit differentiation dependent sub-cellular localization

Although the expression of *AtET1* and *AtET2* correlated with differentiation, low levels of both transcripts were nevertheless detectable in the non-differentiating cell population, suggesting a further possibly post-transcriptional control process to regulated *AtET1* and *AtET2* functions. The sub-cellular localization of all three ET factors was determined using C-terminal EGFP fusions driven by the CaMV 35S promoter and expressed in the protoplast system described above. Wave-length scanning was used to verify that the signal was from EGFP, rather than being an artefact of autofluorescence. None of the three fusion proteins was present in the nuclei of cells cultured under non-differentiating conditions (Figs. 7H–J);

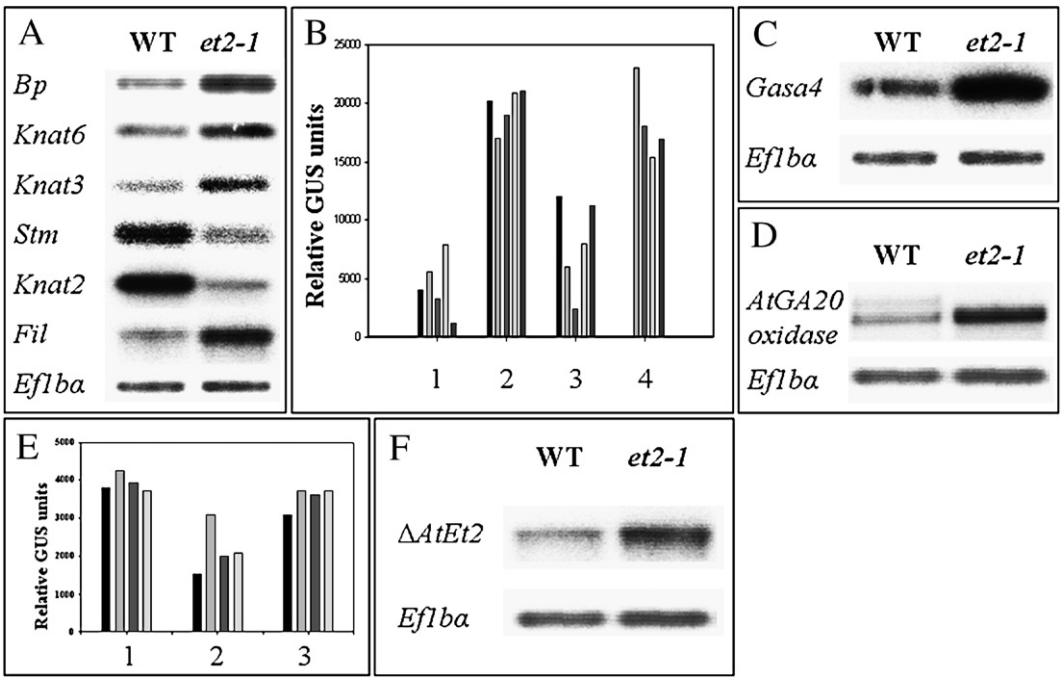


Fig. 6. *AtET2*-mediated regulation of *KNAT* genes and the GA response. (A) RT-PCR analysis illustrates up-regulation of *BP1*, *KNAT6*, *KNAT3* and *FIL* transcripts in the *et2-1* mutant and down-regulation of *STM* and *KNAT2* transcripts. (B) Repression of GA-induced promoter activity in a transient *At* expression system. The test construct consisted of the *GASA4* promoter fused to GUS and terminated by the nopaline synthase terminator. *AtET2* was expressed under the control of the CaMV35S promoter and terminated with the nopaline synthase terminator. Results from five independent experiments are shown. (1) ProGASA4-GUS-NOS; (2) ProGASA4-GUS-NOS+GA3; (3) ProGASA4-GUS-NOS+GA3+ProCaMV35S-AtET2-NOS; (4) ProGASA4-GUS-NOS+GA3+ProCaMV35S-NOS. (C) Over-accumulation of *GASA4* transcript in 10-day-old *et2-1* mutant seedlings. (D) Enhanced expression of *GA5* in leaves in 4-week-old *et2-1* mutant plants. (E) Transient expression of *AtET2* in protoplasts shows reduced activity of GUS driven by ProAtET2: (1) ProAtET2-GUS-NOS; (2) ProAtET2-GUS-NOS+ProCaMV35S-AtET2-NOS; (3) ProAtET2+ProCaMV35S-NOS. (F) Increased level of truncated transcript derived from the T-DNA insertion allele of *AtET2* indicates negative feedback regulation. Expression levels in panels A, C, D and F are normalized with respect to the constitutively expressed *Ef1Bα*.

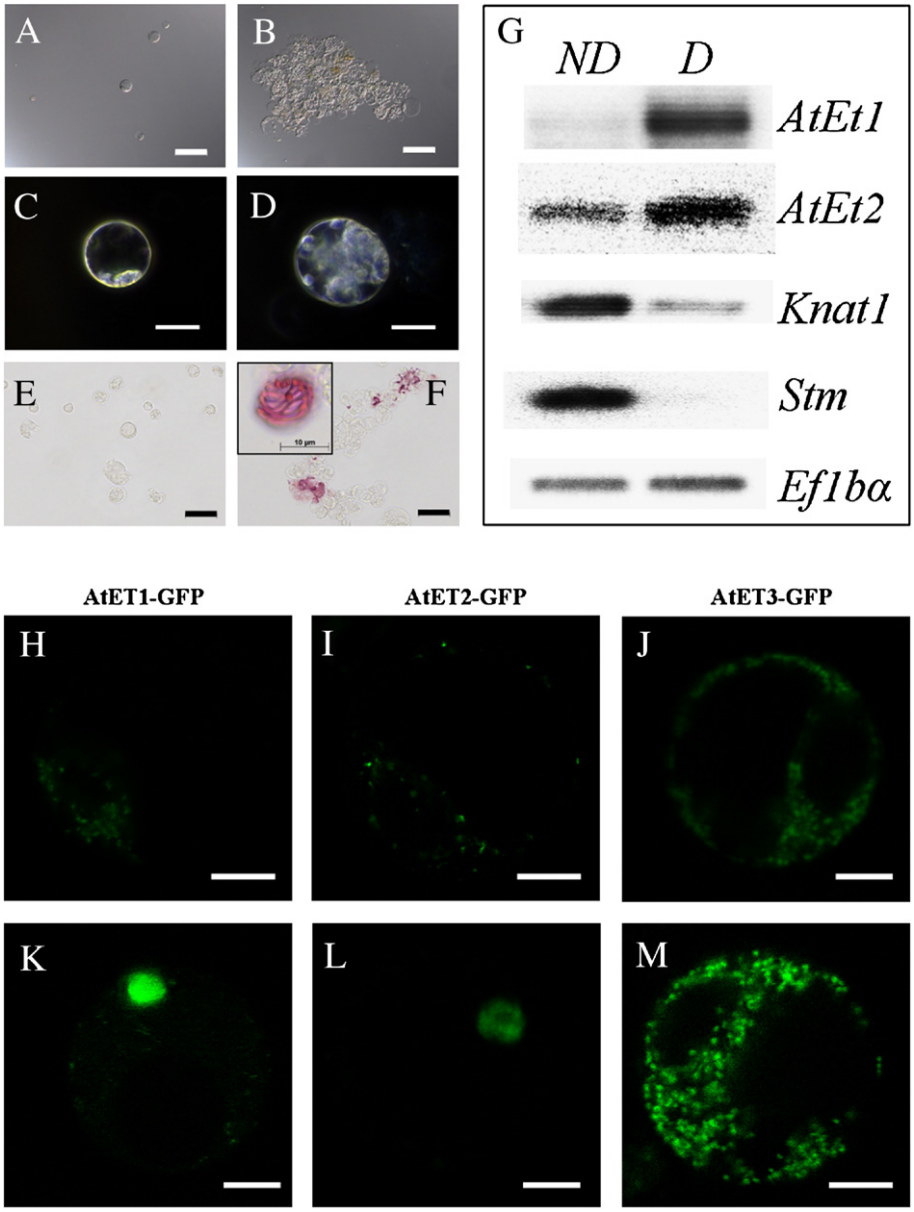


Fig. 7. Transcriptional regulation of *AtET1* and *AtET2* in protoplasts, cultured under non-differentiating (A, C, E) or differentiating (B, D, F) conditions. (A) Separated single cells (non-differentiating conditions). Bar length = 100 μm. (B) Clustered cells (differentiating conditions). Bar length = 100 μm. (C, D) Single cell cultured under non-differentiating (C) and differentiating (D) conditions. Bar length = 20 μm. (E, F) Lignification of cells under non-differentiating (E) and differentiating (F) conditions. Bar length = 50 μm. The insert shows a magnified cell with lignin incrustation, resembling a differentiating xylem element. Bar length = 10 μm. (G) Increased abundance of *AtET1* and *AtET2 transcripts in differentiating cells. The expression of KNAT1* and *STM* is used as a marker for meristem identity (ND, non-differentiating cells; D, differentiating cells). Expression is normalized against the constitutively expressed *EF1Bα*. (H–M) Subcellular localization of AtET fusion proteins. All three AtET proteins were translationally fused to EGFP and expressed in protoplasts. Bar length = 5 μm. (H–J) Under non-differentiating conditions, all three fusion proteins are located in the cytoplasm and none in the nucleus. In differentiating cells AtET1–EGFP (K) and AtET2-EGFP (L) fusion proteins are present in the nucleus. (M) No nuclear translocation was observed for the AtET3–EGFP fusion.

instead, the signal was dispersed throughout the cytoplasm and did not co-localize with either the plastids or the mitochondria (data not shown). In contrast, the *AtET1–EGFP* and *AtET2–EGFP* fusions were expressed in the nuclei of cells grown under conditions triggering cell differentiation (Figs. 7K–M). Although some EGFP signal was still detectable in the cytoplasm under these conditions, it was rather weak. The shift into the nucleus did not occur for the *AtET3–EGFP* fusion (Fig. 7M). We conclude that, even though *AtET1* and *AtET2* are somewhat expressed in non-differentiating cells, the gene products are prevented from entering the nucleus and are therefore inactive as transcriptional regulators.

Discussion

The heterologous ectopic expression of *B. napus BnET* in tobacco and *At* induces alterations in the programming of cell differentiation (Ellerstrom et al., 2005). The molecular basis of ET function and its interaction with established regulatory pathways has been described here through a

detailed analysis of the small *AtET* gene family, in which three members revealed a level of allelic variation at the sequence level. A characteristic feature of the family is the presence of the highly conserved repetitive motifs, and it is these that allow for the clear discrimination between ET and other gene products with regular cysteine patterns. The absence of any orthologs in non-plant genomes suggests that the ET proteins encode plant-specific process(es). The allelic differences identified probably do have an impact on the functionality of *AtET1*. This gene is structurally intact in Ws-2 and many other ecotypes but is represented by a frame-shifted pseudogene in Col-0 and Limeport. As a result, there must be at least some partial functional redundancy between *AtET1* and *AtET2*. *AtET3* appears to be a truncated duplication of *AtET2* and encodes a product that lacks the characteristic ET repeats. We have no data at present to determine whether the *AtET3* product is non-functional or whether it acts as a dominant-negative factor. *AtET2* is the only functional *ET* gene in Col-0.

Xylem differentiation and secondary growth are reduced in et2-1

The differentiation of xylem from the cambial meristem involves a series of overlapping processes, including secondary cell wall formation and lignification. The inactivation of *AtET2* in the *et2-1* mutant caused an overall reduction in lignification and compromised the capacity of the hypocotyl to produce xylem tissue. Thus, it seems probable that *AtET2* is necessary for the orderly differentiation of xylem elements and fibre cells. The reduced level of lignification is therefore probably a secondary (although specific) effect of a delay to or a decrease in the differentiation capability of cambial derivatives. Since *et2-1* plants do not exhibit macroscopical changes in plant architecture, the lignification phenotype appears highly localized and cannot be attributed to pleiotropy. This interpretation is also consistent with experimental data which show that the constitutive expression of *BnET* prevents the de-differentiation process in tobacco leaf cells (Ellerstrom et al., 2005). *In situ* hybridization experiments clearly showed that the expression of *AtET1* and *AtET2* was strongly enhanced in the cambial zone. At this stage, the vascular tissues are already present, and the remaining cambium ceases to function as a meristem. In perennial species such as *Populus tremula*, but not in *At*, this state is reversible (Schrader et al., 2004b). As ET factors likely act as the trigger for these differentiation processes, we suggest that ET function is not restricted to the induction of xylem differentiation but also suppresses cambial meristematic activity, and in particular its capacity for cell division. The ectopic expression in tobacco of *BnET* resulted in the suppression of cell division and the promotion of cell differentiation (Ellerstrom et al., 2005). An analogous role for *AtET2* is suggested by the differential expression profiles of the *et2-1* mutant and wild type. Two *TCP* genes with a high homology to rice *PCF1* and *PCF2* are strongly up-regulated in the mutant. The products of these two genes act as positive regulators of the replication factor PCNA (Kosugi and Ohashi, 1997), which is consistent with the involvement of *AtET2* in cell cycle repression.

Among the genes showing altered expression patterns in the *et2-1* mutant are several *KNAT* family members (including *BP1*), which are involved in cambial function and xylem differentiation. BP1 is also involved in the regulation of internode patterning in the florescence (Smith and Hake, 2003). The over-expression of *BP1* in *At* results in a decreased level of lignification, whereas its loss of function leads to the over-accumulation of lignin (Mele et al., 2003). The enhanced expression of *BP1* in the *et2-1* mutant may therefore provide an explanation for the reduction in lignin level. Overall, it is likely therefore that AtET2 is required to suppress *BP1* activity in cambium-derived cells in order to allow their differentiation into lignified xylem cells. The process might also involve *BP1*-related genes such as *KNAT6* and *KNAT3*, both of which are as yet functionally poorly characterized, but which are up-regulated in the *et2-1* mutant. No change in expression between wild-type and *et2-1* mutant is detected for *KNAT7*, identified in transcript profiling experiments and shown to be involved in fiber differentiation (Ehlting et al., 2005; Brown et al., 2005). The two class I KNAT genes *STM* and *KNAT2* behave differently, as their expression was significantly down-regulated in the *et2-1* mutant. Since at the same time *FIL* was induced, and the loss of *FIL* function resulted in the up-regulation of *BP1*, *KNAT2* and *STM* (Kumaran et al., 2002), we hypothesize that *FIL* may act in an *AtET2*-dependent manner to suppress *BP1* and in an *AtET2*-independent manner to down-regulate *STM* and *KNAT2*. The same biased mode of regulation has been demonstrated recently for the chromatin remodeling factor *FIE*, which acts in a complex with *CLF* to regulate *BP1* and *KNAT6*, but requires a different interacting partner to interact with STM and *KNAT2* (Katz et al., 2004; Guyomarc'h et al., 2005).

The differential regulation of *KNAT* genes may explain the partial similarity in phenotype between the *BnET* over-expressing tobacco and the *At et2-1* mutant. Both show decreased lignin accumulation, suggesting a function for the ET factors in xylem differentiation. The inactivation of *ET* prevented the blocking of *BP1* expression and resulted in a prolongation of the meristematic state. Its over-expression may trigger an alternative *KNOX* pathway, which also results in a decrease in lignin formation. Our hypothesis rests on the assumptions (as yet unverified) that both *ET* genes have the same function, and that the differential regulation of the tobacco *KNOX* genes is in principle similar to what occurs in *At*.

The effect of ET on the differentiation of xylem cells has some long-term implications for application in the area of biomass production. High lignin content presents a major limitation for the efficient fermentation of plant fibers into biofuel (Himmel et al., 2007), and thus it is conceivable that the controlled expression of *ET* genes could allow for the modification of the lignin content in biomass crops such as poplar and willow.

ET acts as a specific regulator of GA-mediated processes

An important function of *KNAT* genes lies in their role as negative regulators of GA-mediated processes (Sakamoto et al., 2001a,b; Hay et al., 2002). GA is required for the early stages of cell differentiation at both the shoot apical and the cambial meristem but is otherwise known to inhibit organogenesis (Ezura and Harberd, 1995; Hertzberg et al., 2001b; Israelsson et al., 2003; Hay et al., 2004). At least two ET proteins (HRT and BnET) have been demonstrated to modulate GA responses either *in vitro* and *in vivo* (Raventos et al., 1998; Ellerstrom et al., 2005). We have shown, via both transient expression and transcript profiling, that AtET2 acts as a negative regulator of the GA induced *GASA4*, which is known to be involved in the control of cell division (Aubert et al., 1998). This is consistent AtET2 functioning both as an inhibitor of cell division and GA response. A further level of complexity relates to the induced expression in the *et2-1* mutant of *GA5*, which is feedback regulated by GA, and acts as an important checkpoint between GA biosynthesis and response (Olszewski et al., 2002). Since *GA5* is not ectopically expressed in the mutant (data not shown), the induced *GASA4* expression cannot be a secondary effect of enhanced GA biosynthesis but rather appears to represent an independent regulatory event. Therefore, *AtET2* must be involved in the control of both GA biosynthesis and the GA response, as we have suggested elsewhere (Ellerstrom et al., 2005).

Transcriptional and post-transcriptional regulation of AtET1 and AtET2

As regulators of other transcription factors, the expression of *ET* requires precise temporal and spatial control. We have shown that the expression of *AtET* includes negative auto-regulation as well as regulation at the transcriptional and post-transcriptional levels. Two independent approaches – a transient assay and the analysis of the loss-of-function *AtET2* mutation – demonstrated that the AtET2 product interacts with its own gene promoter. It remains unclear, however, whether this is due to direct auto-regulation or is the outcome of a feedback signaling loop. Nevertheless, this observation provides some clues as to why *AtET2* is expressed at such a low level. The transcription of *AtET1* and *AtET2* is mainly restricted to differentiating cells, but some low-level expression continues in non-differentiating cells, indicating that additional levels of regulation must be required for the biased function of ET factors in non-differentiating and differentiating cells. It is intriguing that both *AtET1* and *AtET2* expression can toggle between the cytoplasm and nucleus, depending on cell fate. The AtET1 and AtET2 gene products in the meristematic cells are kept inactive by being prevented entry into the nucleus. A similar shuttling in response to exogenous signals has been noted for a range of animal (STAT and SMADs), yeast (Aft1) and plant (PHOR1) proteins (Darnell, 1997; Heldin et al., 1997; Yamaguchi-Iwai et al., 2002; Amador et al., 2001). For the moment, the nature of the trigger is unclear, as is whether the shuttling process is a cause or a consequence of the differentiation process.

An interesting structural detail of the three AtET proteins is that none possesses an obvious nuclear localization signal (NLS), so their entry into the nucleus probably requires interaction with an additional factor carrying such a signal. This feature appears to be unique for the *AtET* family, since other monocot and dicot *ET* sequences do have a *bona fide* NLS (Raventos et al., 1998; Ellerstrom et al., 2005). It is possible, of course, that they carry an NLS sequence, which is not recognized by current motif detection software. The *AtET3* product, however, does not enter the nucleus under differentiation conditions. It completely lacks the characteristic ET repeats, and so is probably a truncated version of *AtET2*. As a result, it is uncertain whether *AtET3* is even a functional gene. The lack of Zn- and DNA-binding repeats may completely compromise its functionality, or it may act as a dominant-negative regulator of the level of protein–protein interaction.

Overall, the data point to ET factors functioning as novel regulators of cell differentiation required for xylem differentiation in the cambial meristem, a crucial process in the development of vascular plants.

ET-mediated gene regulation may include the insertion of DNA single strand nicks

The bacterial UVRC protein is required for DNA excision repair (Friedberg et al., 1995). The protein is targeted to UV-induced DNA lesions, where it introduces a single strand cut 8-bp 5′ and another 4-bp 3′ of the lesion. The two cuts are processed by two structurally and functionally distinct domains, the former involving the C-terminal ENDO V and Helix–hairpin–Helix (HhH) and the latter involving the N-terminal GIY-YIG (Lin and Sancar, 1992; Friedberg et al., 1995; Derbyhire et al., 1997; Kowalski et al., 1999; Aravind et al., 1999; Verhoeven et al., 2000; Van Roey et al., 2002; Stoddard, 2005). The GIY-YIG domain is also present in the so-called "homing nucleases", which are encoded within mobile group I, group II and archaea introns, as well as in inteins (intervening sequences which are spliced and excised post-translationally; Stoddard, 2005). As a result, the domain has been re-designated URI (UVRC and intron-encoded endonucleases; Aravind et al., 1999). The sequence similarity between plant ET factors and the prokaryotic UVRC proteins is only in the single strand cutting GIY-YIG domain, which suggests that a UVRC-like ancestral domain may have been recruited by ET proteins and attached to the DNA-binding ET repeats. The suggested domain shuffling event is consistent with the exon–intron structure of ET genes, with the GIY-YIG domain represented by a separate second exon. The HhH domain has been identified in a number of plant regulatory proteins, such as DME and ROS1 (Choi et al., 2002, 2004; Morales-Ruiz et al., 2006).

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. This showed that the plant domain productively interacts with the C-terminal UVRC ENDOV/HhH domain. The functional importance of the AtET2 GIY-YIG domain is

also involved in stimulating the activity of a *NAM* gene promoter, a putative first target promoter of AtET2. Since the single amino acid exchange (R > A) results in the near complete loss of this activity, the single strand cutting activity of the GIY-YIG domain must clearly be required for its molecular function. Thus, we suggest that the nicking activity of the plant ET factor GIY-YIG domain may be involved in the catalysis of changes in higher order DNA structure, such as, for example, nucleosome sliding (Langst and Becker, 2001). Alternatively, it may contribute to the relaxation of supercoiled chromatin domains, which are implicated in the control of differentiation and development. The importance of the relief of torsional tension in DNA to the triggering of transcriptional activation has been recently described (Ju et al., 2006).

Our conclusion 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 repeats. The resulting protein factor is not involved in the repair process but acts as a gene regulator. The regulatory mechanism – in part analogous to the function of DME and ROS1 – includes the insertion of nicks, with an impact on higher order structures of chromatin packed DNA required for differentiation processes. *In planta* approaches are needed to test this hypothesis further.

Acknowledgments

The authors thank Sabine Skiebe for the preparation of the protoplasts and Elke Liemann for plant transformation. We thank Heike Schmuths and Ralf Horres for providing DNA of various *At* accessions and acknowledge the help of Andrei D. Shutov with amino acid alignments. Tzvetina Brumbarova is acknowledged for her general support, discussions and critical reading of the manuscript. We thank Tim Sharbel and www.smartenglish.co.uk for linguistic advice. This research was supported by grants from the IPK and the Deutsche Forschungsgemeinschaft (Ba1235/7-1).

References

Amador, V., Monte, E., Garcia-Martinez, J.L., Prat, S., 2001. Gibberellins signal nuclear import of PHOR1, a photoperiod-responsive protein with homology to *Drosophila* armadillo. Cell 106, 343–354.

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

Aubert, D., Chevillard, M., Dorne, A.M., Arlaud, G., Herzog, M., 1998. Expression patterns of GASA genes in *Arabidopsis thaliana*: the GASA4 gene is up-regulated by gibberellins in meristematic regions. Plant Mol. Biol. 36, 871–883.

Bechthold, N., Elli, J., Pelletier, G., 1993. *In planta Agrobacterium*-mediated transfer by infiltration of adult *Arabidopsis thaliana* plants. CR Acad. Sci. Paris 316, 1194–1199.

Biemelt, S., Tschiersch, H., Sonnewald, U., 2004. Impact of altered gibberellin metabolism on biomass accumulation, lignin biosynthesis, and photosynthesis in transgenic tobacco plants. Plant Physiol. 135, 254–265.

Brown, D.M., Zeef, L.A.H., Ellis, J., Goodacre, R., Turner, S.R., 2005. Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17, 2281–2295.

Campbell, M., Ellis, B., 1992. Fungal elicitor mediated responses in pine cell cultures: induction of phenylpropanoid metabolism. Planta 486, 409–417.

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

Choi, Y., Harada, J.J., Goldberg, R.B., Fischer, R.L., 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. U. S. A. 101, 7481–7486.

Curtis, M.D., Grossniklaus, U., 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol. 133, 462–469.

Darnell Jr., J.E., 1997. STATs and gene regulation. Science 277, 1630–1635.

de Folter, S., Busscher, J., Colombo, L., Losa, A., Angenent, G.C., 2004. Transcript profiling of transcription factor genes during silique development in *Arabidopsis*. Plant Mol. Biol. 56, 351–366.

Demura, T., Fukudo, H., 2007. Transcriptional regulation in wood formation. Trends Plant Sci. 12, 64–70.

Derbyhire, V., Kowalski, J.C., Dansereau, J.T., Hauer, C.R., Belfort, M., 1997. Two-domain structure of the *td* intron-encoded endonuclease I–*Tev*I correlates with the two-domain configuration of the homing site. J. Mol. Biol. 265, 494–506.

Ellerstrom, M., Reidt, W., Ivanov, R., Tiedemann, J., Melzer, M., Tewes, A., Moritz, T., Mock, H.P., Sitbon, F., Rask, L., Baumlein, H., 2005. Ectopic expression of effector of transcription perturbs gibberellin-mediated plant developmental processes. Plant Mol. Biol. 59, 663–681.

Ehlting, J., Mattheus, N., Aeschliman, D.S., Li, E., Hamberger, B., Cullis, I.F., Zhuan, J., Kandea, M., Mansfield, S.D., Samuels, L., Ritland, K., Ellis, B.E., Bohlmann, J., Douglas, C.J., 2005. Global transcript profiling of primary stems from *Arabidopsis thaliana* identifies candidate genes fro missing links in lignin biosynthesis and transcription regulators of fiber differentiation. Plant J. 42, 618–640.

Ezura, H., Harberd, N.P., 1995. Endogenous gibberellin levels influence in-vitro shoot regeneration in *Arabidopsis thaliana* (L.) Heynh. Planta 197, 301–305.

Friedberg, E.C., Walker, G.C., Siede, W., 1995. Nucleotide excision repair in prokaryotes. DNA Repair and Mutagenesis. ASM Press, Washington, DC, pp. 191–232.

Grafi, G., 2004. How cells dedifferentiate: a lesson from plants. Dev. Biol. 268, 1–6.

Groover, A.T., 2005. What genes make a tree a tree? Trends Plant Sci. 10, 210–214.

Guyomarc'h, S., Bertrand, C., Delarue, M., Zhou, D.X., 2005. Regulation of meristem activity by chromatin remodelling. Trends Plant Sci. 10, 332–338.

Hake, S., Smith, H.M., Holtan, H., Magnani, E., Mele, G., Ramirez, J., 2004. The role of *KNOX* genes in plant development. Annu. Rev. Cell Dev. Biol. 20, 125–151.

Hay, A., Kaur, H., Phillips, A., Hedden, P., Hake, S., Tsiantis, M., 2002. The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. Curr. Biol. 12, 1557–1565.

Hay, A., Craft, J., Tsiantis, M., 2004. Plant hormones and homeoboxes: bridging the gap? BioEssays 26, 395–404.

Heldin, C.H., Miyazono, K., ten Dijke, P., 1997. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature 390, 465–471.

Hertzberg, M., Sievertzon, M., Aspeborg, H., Nilsson, P., Sandberg, G., Lundeberg, J., 2001a. cDNA microarray analysis of small plant tissue samples using a cDNA tag target amplification protocol. Plant J. 25, 585–591.

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlen, M., Teeri, T.T., Lundeberg, J., Sundberg, B., Nilsson, P., Sandberg, G., 2001b. A transcriptional roadmap to wood formation. Proc. Natl. Acad. Sci. U. S. A. 98, 14732–14737.

Herzog, M., Dorne, A.M., Grellet, F., 1995. GASA, a gibberellin-regulated gene family from *Arabidopsis thaliana* related to the tomato GAST1 gene. Plant Mol. Biol. 27, 743–752.

Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady,

J.W., Foust, T.D., 2007. Biomass recalcitrance: engineering plants and enzymes for biofuel production. Science 315, 804–807.

Hofius, D., Hajirezaei, M.R., Geiger, M., Tschiersch, H., Melzer, M., Sonnewald, U., 2004. RNAi-mediated tocopherol deficiency impairs photoassimilate export in transgenic potato plants. Plant Physiol. 135, 1256–1268.

Israelsson, M., Eriksson, M.E., Hertzberg, M., Aspeborg, H., Nilsson, P., Moritz, T., 2003. Changes in gene expression in the wood-forming tissue of transgenic hybrid aspen with increased secondary growth. Plant Mol. Biol. 52, 893–903.

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

Katz, A., Oliva, M., Mosquna, A., Hakim, O., Ohad, N., 2004. FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. Plant J. 37, 707–719.

Ko, J.H., Han, K.H., 2004. *Arabidopsis* whole-transcriptome profiling defines the features of coordinated regulations that occur during secondary growth. Plant Mol. Biol. 55, 433–453.

Kosugi, S., Ohashi, Y., 1997. PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. Plant Cell 9, 1607–1619.

Kowalski, J.C., Belfort, M., Stapleton, M.A., Holpert, M., Dansereau, J.T., Pietrokovski, S., Baxter, S.M., Derbyhire, V., 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.

Kumaran, M.K., Bowman, J.L., Sundaresan, V., 2002. YABBY polarity genes mediate the repression of KNOX homeobox genes in *Arabidopsis*. Plant Cell 14, 2761–2770.

Langst, G., Becker, B., 2001. ISWI induces nucleosome sliding on nicked DNA. Mol. Cell 8, 1085–1092.

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

Mele, G., Ori, N., Sato, Y., Hake, S., 2003. The *knotted1*-like homeobox gene *BREVIPEDICELLUS* regulates cell differentiation by modulating metabolic pathways. Genes Dev. 17, 2088–2093.

Mönke, G., Altschmied, L., Reidt, W., Mock, H.P., Bäumlein, H., Conrad, U., 2004. Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. Planta 219, 158–166.

Morales-Ruiz, T., Ortega-Galisteo, A.P., Ponferrada-Marin, M.I., Martinez-Macias, M.I., Ariza, R.R., Rolan-Arjona, T., 2006. *DEMETER* and *REPRESSOR OF SILENCING1* encode 5-methylcytosine DNA glycosylases. Proc. Natl. Acad. Sci. U. S. A. 103, 6853–6858.

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

Paz-Ares, J., REGIA-Consortium, 2002. REGIA, an EU project on functional genomics of transcription factors from *Arabidopsis thaliana*. Comp. Funct. Genom 3, 102–108.

Raventos, D., Skriver, K., Schlein, M., Karnahl, K., Rogers, S.W., Rogers, J.C., Mundy, J., 1998. HRT, a novel zinc finger, transcriptional repressor from barley. J. Biol. Chem. 273, 23313–23320.

Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahori, S., Matsuoka, M., 2001a. KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. Genes Dev. 15, 581–590.

Sakamoto, T., Kobayashi, M., Itoh, H., Tagiri, A., Kayano, T., Tanaka, H., Iwahori, S., Matsuoka, M., 2001b. Expression of a gibberellin 2-oxidase gene around the shoot apex is related to phase transition in rice. Plant Physiol. 125, 1508–1516.

Schrader, J., Nilsson, J., Mellerowicz, E., Berglund, A., Nilsson, P., Hertzberg, M., Sandberg, G., 2004a. A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. Plant Cell 16, 2278–2292.

Schrader, J., Moyle, R., Bhalerao, R., Hertzberg, M., Lundeberg, J., Nilsson, P., Bhalerao, R.P., 2004b. Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome. Plant J. 40, 173–187.

Scofield, S., Murray, J.A.H., 2006. KNOX gene function in plant stem cell niches. Plant Mol. Biol. 60, 929–946.

Smith, H.M.S., Hake, S., 2003. The interaction of two homeobox genes, *BREVIPEDICELLUS* and *PENNYWISE*, regulates internode patterning in the *Arabidopsis* inflorescence. Plant Cell 15, 1717–1727.

Stoddard, B.L., 2005. Homing endonuclease structure and function. Q. Rev. Biophys. 38, 49–95.

Sussman, M.R., Amasino, R.M., Young, J.C., Krysan, P.J., Austin-Phillips, S., 2000. The *Arabidopsis* knockout facility at the University of Wisconsin-Madison. Plant Physiol. 124, 1465–1467.

Tiedemann, J., Neubohn, B., Muntz, K., 2000. Different functions of vicilin and legumin are reflected in the histopattern of globulin mobilisation during germination of vetch (*Vicia sativa* L.). Planta 211, 1–12.

Tiedemann, J., Schlereth, A., Muntz, K., 2001. Differential tissue-specific expression of cysteine proteinases forms the basis for the fine-tuned mobilization of storage globulin during and after germination in legume seeds. Planta 212, 728–738.

Van Roey, P., Meehan, L., Kowalski, J.C., Belfort, M., Derbyshire, V., 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.A., van Kesteren, M., Moolenaar, G.F., Visse, R., Goosen, N., 2000. Catalytic sites for the 3′ and 5′ incision of *Escherichia coli* nucleotide excision repair are both located in UvrC. J. Biol. Chem. 275, 5120–5123.

Valente, P., Tao, W., Verbelen, J.-P., 1998. Auxins and cytokinins control DNA endoreduplication and deduplication in single cells of tobacco. Plant Sci. 134, 207–215.

Yamaguchi-Iwai, Y., Ueta, R., Fukunaka, A., Sasaki, R., 2002. Subcellular localization of Aft1 transcription factor responds to iron status in *Saccharomyces cerevisiae*. J. Biol. Chem. 277, 18914–18918.

Zhao, C., Johnson, B.J., Kositsup, B., Beers, E.P., 2000. Exploiting secondary growth in *Arabidopsis*. Construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. Plant Physiol. 123, 1185–1196.

9. Acknowledgements

First and foremost, my deepest appreciation goes 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. I would have not accomplished my PhD. program without his advice, encouragement, and constant support throughout my graduate study.

I am sincerely grateful to the mentor PD Dr. habil. Udo Conrad, Head of Phytoantibody group, IPK, Gatersleben for his helpful guidance, explanation, and discussions along this research and critical reading manuscript.

I am very thankful to Dr. Amal Joseph Johnston for the help in mutant analyses and for valuable discussion and critical reading manuscript.

I would like to thank former and current directors of IPK Gatersleben, Prof. Dr. Ulrich Wobus and Prof. Dr. Andreas Graner for creating the scientifically stimulating atmosphere during the whole work time.

I wish to thank all my colleagues in Gene Regulation group for various supports, Elke Liemann for taking care plants, Andreas Czihal, Annet Bushing, Sabine Skiebe, Monika Gottowik for excellent technical assistance. Special thanks to Dr. David Koszegi, Dr. Astrid Junker for the help at beginning time. My gratitude goes to Maria Mildner, Dr. Corina Gryczka, Anna Schallau for the help in different ways.

During my study in IPK, I have been fortunate to work with colleagues and people from many different countries. I thank those colleagues and friends, including Dr. Andrea Matros for help in MALDI-TOF, Dr. Twan Rutten for preparation of microscopy pictures, Dr. Gudrun Mönke for her helpful advice in protein expression and purification, Dr. Dmitri Demidov and Dr. David Riewe for kindly providing the vector and bacterial strain for RNA interference. I would like to thank Dr. Martin Giersberg and Silke Krause for phage library screening experiments.

I would like to extend my gratitude to Hanoi University of Science, Vietnam National University, Hanoi for allowing me away such a long time. My sincere thanks go to Prof. Dr. Nong Van Hai for introducing me to IPK.

My stay and PhD. work was supported by Ministry of Education and Training of Vietnam (MOET), partly from IPK and German Academic Exchange Service (DAAD). The IPK's grant helped me to prolong my stay at IPK to finish all laboratory work. My acknowledgements go to MOET, DAAD and IPK for all financial supports.

Last, but not least, I would like to thank my parents, my two sisters for their love, support and endless endurance, which has been my invaluable source of strength throughout my study.

Curriculum vitae

Personal data

Full name	:	Le, Hong Diep
Gender	:	Male
Date of birth	:	20th April 1971
Place of birth	:	Hanam, Vietnam
Marital status	:	Unmarried
Nationality	:	Vietnamese
E-mail	:	le@ipk-gatersleben.de , diep.le-hong@student.uni-halle.de

Education and employment

October 2006-present:	PhD student at Gene Regulation group, Department of Molecular Genetics, Leibniz- Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany.
2000 - 2006:	Assistant lecturer and researcher at Department of Plant Physiology and Biochemistry, Faculty of Biology, Vietnam National University, Hanoi.
1996 - 2000:	Master student at Faculty of Biology, Vietnam National University, Hanoi. Title of thesis “Comparative study of microsatellites derived from Vietnamese upland rice (*Oriza sativa* L.) accessions”.
1991 - 1995:	Student at Faculty of Biology, Vietnam National University, Hanoi. Title of thesis: “Micropropagation of sugar cane (*Saccharum officinarum* L.) ROC10 accession”.
1985 - 1988:	High school in Hanam, Vietnam.

Publication

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

Posters

Diep L. H., Mönke G., Vorwieger A., Matros A., Mock H-P., Conrad U., and Bäumlein H. (2008) Characterization of EFFECTOR OF TRANSCRIPTION (ET) in *Arabidopsis thaliana*. The fourth IPK Plant Student Science Conference, Gatersleben.

Diep L. H., Mönke G., Junker A., Matros A., Mock H-P., Conrad U., and Bäumlein H. (2009) Molecular characterization of EFFECTOR OF TRANSCRIPTION (ET) in *Arabidopsis thaliana*. The fifth IPK Plant Student Science Conference, Halle.

Diep L. H., Mönke G., Junker A., Matros A., Mock H-P., Conrad U., and Bäumlein H. (2009) Molecular characterization of EFFECTOR OF TRANSCRIPTION (ET) in *Arabidopsis thaliana*. IPK Day, Gatersleben.

Diep L. H., Johnston A.J., Czihal A., Ivanov R., Rutten T., Bäumlein H. (2010) Reproductive functions of EFFECTOR OF TRANSCRIPTION (ET) in *Arabidopsis thaliana*. IPK Day, Gatersleben.