Alterations in the Expression of Seed Development Genes Affect Seed Size and Storage Compound Allocation in *Arabidopsis thaliana*

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Abdelhak FATIHI

aus Boudnib, Marokko

Bonn 2012

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

1. Gutachter: Prof. Dr. Peter Dörmann
2. Gutachter: Prof. Dr. Dorothea Bartels
Tag der Promotion: 10-10-2012
Erscheinungsjahr: 2012

For my parents Habiba and Abderahman

It is beyond a doubt that all our knowledge that begins with experience

Immanuel Kant

ERKLÄRUNG:

Hiermit versichere ich, die vorliegende Arbeit selbständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt zu haben. Ich versichere ebenfalls, dass die Arbeit an keiner anderen Hochschule eingereicht wurde.

Bonn, den Mai 2012

Abbreviations:

ABA	Abscisic acid
ADP	Adenosine diphosphate
AGPase	ADP-glucose pyrophosphorylase
ATP	Adenosine triphosphate
bp	base pair
CTAB	cetyl trimethyl ammonium bromide
Col	Columbia
cDNA	complementary DNA
DAF	days after flowering
DAG	diacylglycerol
DAGT	diacylglycerol acyltransferase
DNA	deoxyribonucleic acid
ER	endoplasmic reticulum
EV	empty vector
FA	fatty acid
FAME	fatty acid methyl ester
FIE	fertilization independent endosperm
FIS	Fertilization independent seed
GC	gas chromatography
HEPES	N-2-Hydroxyethylpiperazine-N-2-ethanesulfonic acid
IKU2	haiku2
Inv	invertase
Ler	Landsberg ERECTA
MINI3	miniseed3
PCR	polymerase chain reaction
PgG	polycomp group complex
PHE1	Pheres1
PRC	Polycomb Repressive Complex
PSVs	protein storage vacuoles
PUFAs	polyunsaturated fatty acids
RNA	Ribonucleic acid
SEX1	starch excess1
SHB1	Short Hypocotyl under Blue1
SSE1	Shrunken seed1
SSP	seed storage protein
SUS	sucrose synthase
TAG	triacylglycerol
T-DNA	transfer DNA
UDP	Uridine diphosphate
VLCFAs	very long chain fatty acids
VPEs	vacuolar processing enzymes
WS	Wassileskija
WT	wild type

Table of Contents

1 Introduction 9

1.1 Seed development and storage product accumulation 9

1.1.1 Seed and embryo development of Arabidopsis 10

1.1.1 Regulation of storage reserve biosynthesis in Arabidopsis seeds 13

1.1.2 Seed Storage product accumulation in Arabidopsis 15

1.2 Control of Arabidopsis seed size by developmental genes 24

1.2.1 *FERTILIZATION-INDEPENDENT ENDOSPERM* (*FIE*) 26

1.2.2 *PHERES1 (PHE1)* 26

1.2.3 *HAIKU2 (IKU2)* 27

1.2.4 *MINISEED3 (MINI3)* 27

1.3 The aim of this work 28

2. Materials and methods 29

2.1 Equipment and chemicals 29

2.1.1 Equipment 29

2.1.2 Chemicals 30

2.2 Enzymes and molecular biology kits 32

2.2.1 Enzymes 32

2.2.2 Molecular biology kits 32

2.3 Plant material and growth conditions 33

2.3.1 Seed surface sterilization 33

2.3.2 Plant growth and selection 33

2.3.3 Manual pollination 33

2.3.4 Antibiotic stock concentration for plant selection 34

2.4 Bacterial strains and cloning vectors 35

2.4.1 Bacterial strains 35

2.4.2 Cloning vectors 35

2.5 Molecular techniques 36
2.5.1 Genomic DNA extraction from Arabidopsis leaves 36
2.5.2 PCR screening of Arabidopsis T-DNA mutants 37
2.5.3 DNA agarose gel electrophoresis 38
2.5.4 RNA extraction from Arabidopsis siliques. 40
2.5.5 First-Strand cDNA synthesis and reverse transcription PCR (RT-PCR) 41
2.5.6 Preparation and transformation of chemically competent *E. coli* cells 41
2.5.7 Plasmid DNA isolation from *E. coli* 43
2.5.8 Enzymatic modification of cDNA 43
2.5.9 PCR amplification of the cDNAs and cloning into pGEM®-T Easy vector 44
2.5.10 The Arabidopsis plants transformation using floral dipping 46
2.6 Biochemical and Analytical techniques 47
2.6.1 Fatty acid methyl ester measurements (FAME) 47
2.6.2 Total lipid extraction 47
2.6.3 Total seed protein extraction and quantification 49
2.6.4 Soluble carbohydrate analysis 49
2.6.5 Determination of seed weight 50
2.6.6 Measurement of seed length, width and area 50
2.6.7 SDS-PAGE (Polyacrylamide Gel Electrophoresis) 50
2.7 Light microscopy 52
2.8 Statistics 54
3 Results 55
3.1 Characterization of Arabidopsis seed size candidate genes 55
3.1.1 Isolation of homozygous mutant lines 55
3.1.2 Seed sizes, weight, and morphology 57
3.1.3 Analysis of seed storage reserves 60
3.2 Generation of transgenic Arabidopsis seeds overexpressing seed size candidate genes 65

3.3. Characterization of the seeds of transgenic Arabidopsis overexpression lines 66

3.3.1 Determination of seed size and seed weight 67

3.3.2 Measurements of seed storage products 69

3.4 Characterization of IKU2-OE plants 71

4 Discussion 77

4.1 Characterization of mutant alleles for fie, phe1, mini3 and iku2 77

4.2 Characterization of transgenic Arabidopsis seeds overexpressing seed size candidate genes 80

4.3 Characterization of IKU2-OE plants 82

5 Summary 85

6 References 87

7 Appendix 96

7.1 Synthetic oligonucleotides 96

7.2 Vector Map 98

1 Introduction

1.1 Seed development and storage product accumulation

Seeds provide a rich source of energy and carbon reserves; these reserves include proteins, carbohydrates and oil. Furthermore, vitamins, minerals, and protective phytochemicals are found in the seeds. A large proportion of the human diet and animal livestock is provided by plant seeds from cereals, legumes and oilseed crops. Therefore, the analysis and the understanding of mechanisms controlling seed metabolism and storage product deposition during seed filling are of major economic and agricultural importance. The plant life cycle (Figure 1) comprises the alternation of two generations. The first generation consists of the diploid sporophyte (2n) representing the green plant. The sporophyte undergoes a series of meiotic divisions to produce spores (megaspores and microspores). The second generation is a haploid gametophyte (1n). Megaspores and microspores produce the female gametophyte (megagametophyte), and the male gametophyte (microgametophyte) respectively. The differentiated megagametophyte and microgametophyte produce the female gametes (egg cells) and the male gametes (sperm) respectively (Cove and Knight, 1993).

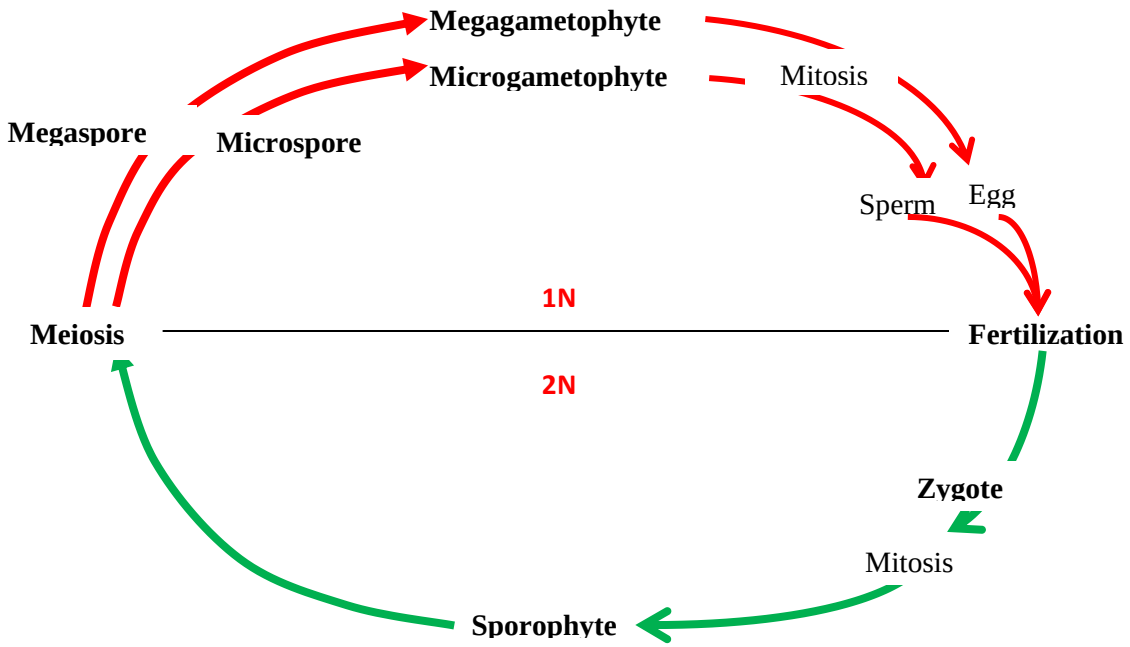


Figure 1. Generalized plant life cycle. The red arrows indicate the haploid gametophyte generation, whereas the green arrows indicate the diploid sporophyte generation.

The angiosperm gametophytes are produced in specialized organs of the flower (Figure 2). These gametophytes are smaller and less complex compared to the sporophytes. The microgametophyte develops within the anther, whereas the megagametophyte is produced by the ovule. Sexual reproduction of the flowering plants requires the delivery of the sperm nuclei via the pollen tube to the embryo sac, where fertilization occurs and the new diploid sporophyte is formed. (Dumas and Mogensen, 1993).

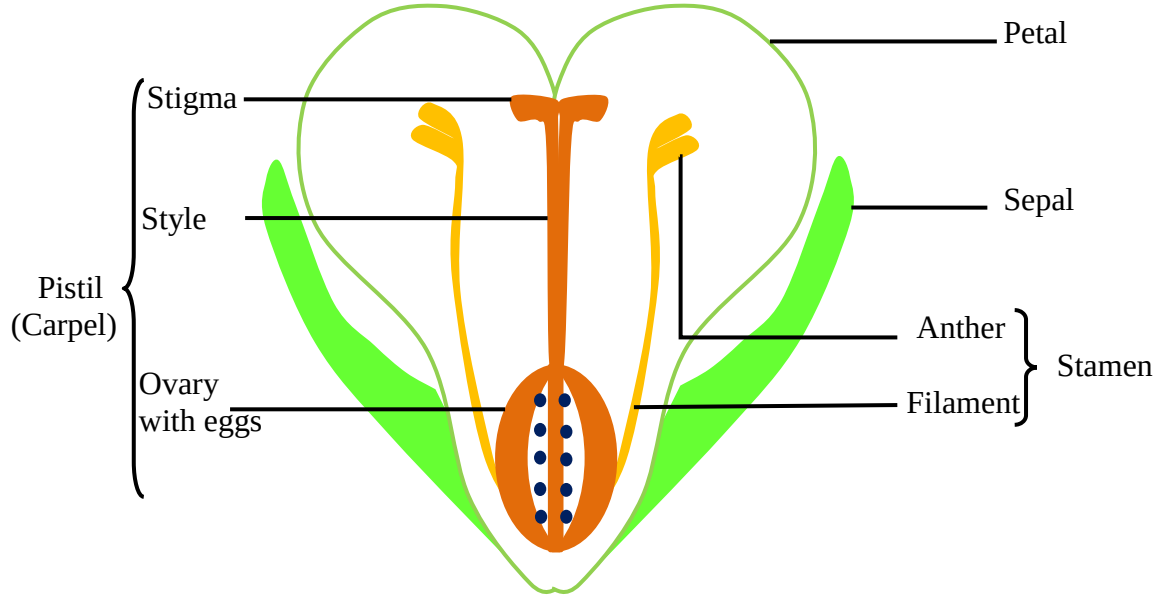


Figure 2. Simplified drawing of an Arabidopsis flower. The flower parts are organized in 4 whorls in the following order: sepals, petals, stamens and pistils. Stamens and pistils represent the reproductive organs; the stamens are the male part where pollen forms, whereas the pistils are the female part where the ovules form.

1.1.1 Seed and embryo development of Arabidopsis

Seed development is a crucial stage in the plant life cycle. The synchronization of the growth of the maternally derived seed coat, the endosperm, and the embryo are indispensable for this process. In flowering plants, seed development is initiated by the double fertilization. During the double fertilization one sperm nucleus unites with the egg cell to form the diploid zygote which then develops into embryo (embryogenesis). The second sperm nucleus fuses with the two polar nuclei to form the triploid endosperm. The endosperm nourishes the embryo during its development. Both, the embryo and the endosperm are protected by the seed coat (testa) which develops from the ovule integuments (Baud *et al.*, 2005). Embryogenesis is the process during which the zygote undergoes several and complex cellular and morphological changes, which lead to the formation of a mature embryo.

Arabidopsis embryogenesis consists of three successive phases: morphogenesis, maturation and desiccation (Figure 3). The morphogenesis phase is characterized by the formation of the embryonic tissue and organ systems, and the specification of the shoot and root apices during the definition of the polar axis of the plant body. The maturation phase is marked by storage reserve accumulation and an increase in seed weight. Finally the desiccation phase is distinguished by the loss of water and the developmental arrest of the embryo (West and Harada, 1993).

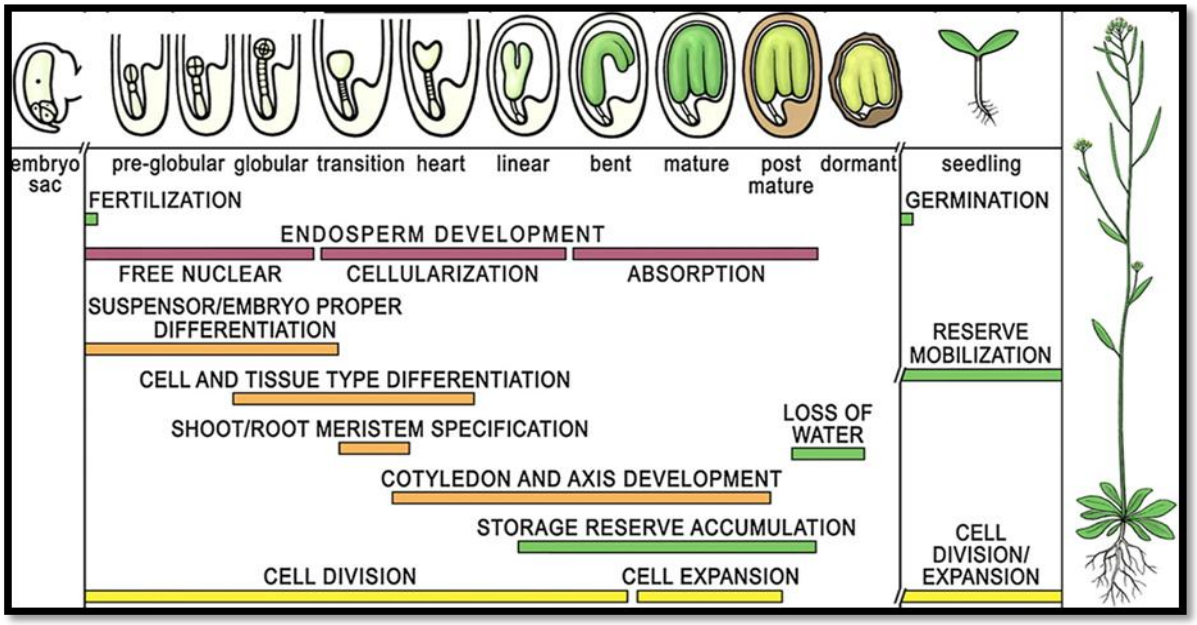


Figure 3. **Schematic representation of Arabidopsis seed development and stages of the life cycle (Garcia *et al.*, 2005).** During Arabidopsis seed development the endosperm and the embryo undergo multiple morphological changes leading to the development of a mature embryo. The late stage of embryo development is characterized by the deposition of storage reserves and loss of water, preparing the seed for desiccation and dormancy.

Arabidopsis seed development which takes approximately 20 days (Figure 4) is characterized by a rapid increase in fresh weight between the 3rd and 7th days after flowering (DAF). The rate of growth then continues to reach a maximum at 15 DAF. Concomitantly, the seed fresh weight decreases dramatically between 16 and 20 DAF corresponding to the loss of water. The water content decreases during seed development to reach a final value of 7% of the mature dry seed weight. Consequently, the seed dry weight increases during this period (Baud *et al.*, 2002). Arabidopsis seed development is under the control of environmental conditions such as light intensity and growth conditions. Moreover, the seed characteristics such as seed dry weight can differ between different ecotypes. For example, one mature desiccated seed of *Arabidopsis thaliana* ecotype Wassilewskija (WS) weighs

approximately 22 µg (Baud *et al.*, 2002), whereas one seed of the *Arabidopsis thaliana* ecotype Col-0 weighs approximately 19 µg (Maisonneuve *et al.*, 2010).

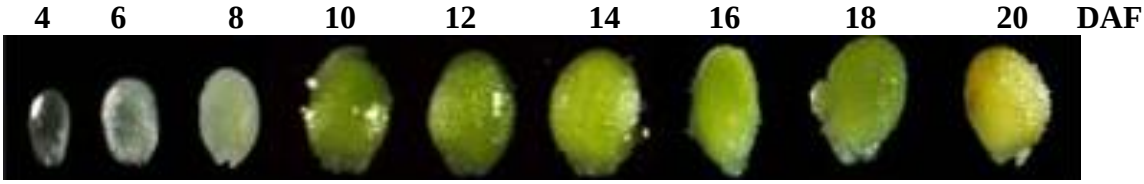


Figure 4. **The seed development of *Arabidopsis thaliana* (ecotype Wassilewskija is characterized by a change in the morphology of the developing seed from 4 to 20 days after flowering (DAF) (Fait *et al.*, 2006).** Arabidopsis seeds reach full maturity after 20 DAF**.** The seed size increases to reach a maximal size at 14 DAF. The preparation of the seed for dormancy and desiccation is associated with the loss of water and decrease in seed fresh weight.

The fully developed embryo (Figure 5) of Arabidopsis seeds consists of a cotyledon and an embryonic axis with shoot and root poles. The cotyledons are rich in storage products such as lipids, proteins and carbohydrates. These storage reserves serve as a source of energy for the germinating seedling before it acquires its photosynthetic activity. Hence, embryogenesis organizes the morphology of the mature plant body and prepares the seed for dormancy and germination (West and Harada, 1993).

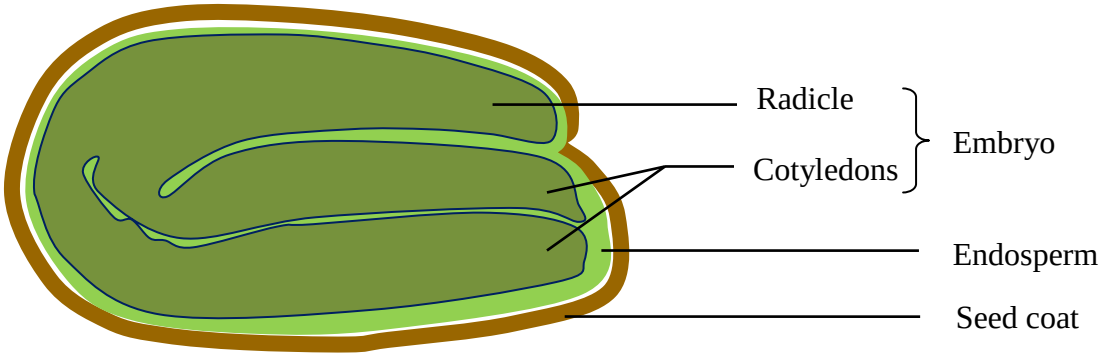


Figure 5. **Drawing of a mature Arabidopsis seed modified from (Muller *et al.*, 2006) .** The mature seed contains a thin layer of endosperm surrounding the embryo. The two tissues are protected by the maternal seed coat.

The rapid growth of the world population, and the boosting demand for new ecological and renewable sources for biofuel and industrial raw materials, resulted in an increase of studies toward the improvement of oilseed yield and seed storage product content. Remarkably, despite the fact that the general biosynthetic pathways of storage products like starch and triacylglycerols (TAGs) are well documented, their regulation and links between the main pathways are still poorly understood (Focks and Benning, 1998; Baud *et al*., 2002).

1.1.1 Regulation of storage reserve biosynthesis in Arabidopsis seeds

A model for the gene regulation of seed storage reserve biosynthesis has been proposed by Hills (Hills, 2004) (Figure 6). This pattern was established based on the in vivo quantification of storage reserves combined with gene expression measurements in developing seeds of legumes. Thus, the use of the ^{13}C-labelling techniques permitted the measurement of carbon incorporation in each single pathway in the biosynthesis of storage reserves in the cytosol and plastid of oilseeds (Hills, 2004). Several studies showed that alterations in embryo and endosperm development affect seed storage product deposition and vice versa. An example how changes in embryo development can affect seed storage product accumulation is the Arabidopsis mutant *trehalose-6-phosphate synthase1* (*tps1*) (Eastmond *et al.*, 2002). The *tps1* mutant is deficient in a gene encoding the enzyme necessary for trehalose biosynthesis. The embryo development of the mutant *tps1* is arrested at the stage of storage reserve accumulation (late torpedo stage). At this stage, *tps1* embryos accumulate more sucrose compared to the wild type, and seeds fail to germinate. Gomez and co-authors (Gomez *et al.*, 2006) reported that approximately 40% of *tps1* seeds were able to germinate after extended stratification and incubation in nutrient medium. These seeds accumulate less protein and TAGs, and more starch compared to wild type seeds. The seeds of Arabidopsis *triacylglycerol1* (*tag1*), a mutant deficient in the diacylglycerol acyltranseferase, accumulate high levels of diacyglycerols (DAGs), and a low content of TAGs. In addition, the maturation of *tag1* mutant seeds is delayed compared to the wild type (Katavic *et al.*, 1995; Zou *et al.*, 1999). Lu and Hills (Lu and Hills, 2002) reported that *tag1* seeds accumulate high amounts of sucrose compared to wild type seeds. Moreover, the *tag1* seedlings are very sensitive to abscisic acid (ABA), glucose and mannitol in the growth medium. In fact, in media containing 0.25 µM ABA, *tag1* seedlings fail to grow, and approximately 70 % of the *tag1* seedlings fail to develop proper cotyledons and leaves, whereas the wild type seedlings were only slightly affected. In addition, the *tag1* seed germination rate is affected by osmotic stress

and other related environmental conditions such as cold and salt stress. The mutation in the *TAG1* gene is assumed to alter the balance of signaling pathways involving ABA and osmotic stress tolerance. Thus, the *tag1* mutant shows that disruption of seed filling can significantly affect the proper development of the embryo. Therefore, the metabolism of storage reserve deposition should not be considered as a single process because it is under the control of regulatory networks that combine the hormonal and developmental signals from the embryo. On the other hand, the embryo development itself is affected by alterations in storage product metabolism (Brocard-Gifford *et al.*, 2003; Hills, 2004).

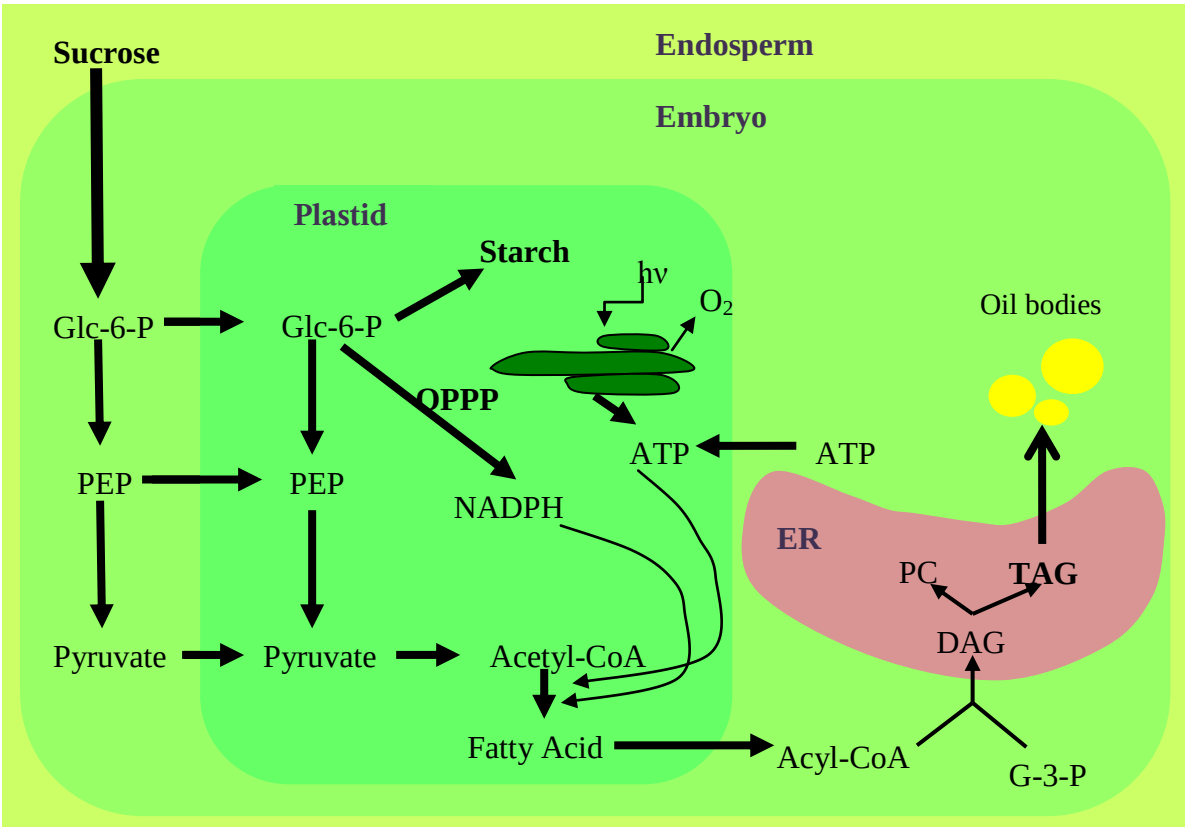


Figure 6. Simplified diagram showing the metabolic pathways of storage product synthesis in a seed modified from (Hills, 2004). During seed filling**,** sucrose is delivered from the mother plant to the developing seed and represents the precursor for the storage product biosynthesis. TAGs are synthesized in the ER, whereas carbohydrates are synthesized in the plastid. Mal-CoA, malonyl-CoA; PEP, phosphoenolpyruvate; OPPP, oxidative pentose phosphate pathway.

The gene expression patterns are clearly different for genes encoding enzymes involved in seed storage compound synthesis (Figure 7) indicating that biosynthetic pathways for these products are under control of different factors. For example, transcripts for starch biosynthesis are expressed at an early stage of seed development which corresponds to the active photosynthesis period. These transcripts are gradually turned off when oil and proteins

start to accumulate corresponding to the seed filling stage (Ruuska *et al.*, 2002). Lin and co-authors (Lin *et al.*, 1999; Lin *et al.*, 2004) reported the isolation of the Arabidopsis *shrunken seed1 (sse1)* mutant. The *SSE1* gene encodes a homolog of the yeast peroxisome biogenesis *PEX16*. Arabidopsis *sse1* mutant seeds accumulate more starch compared to wild type, approximately 1.5 µg/seed. Concomitantly, oil and proteins amounts are reduced to approximately 16 % and 60 % respectively of the wild type levels. Thus, the accumulation of starch results in reduction in carbon flux for the synthesis of proteins and oil in the developing seeds of the *sse1* mutants.

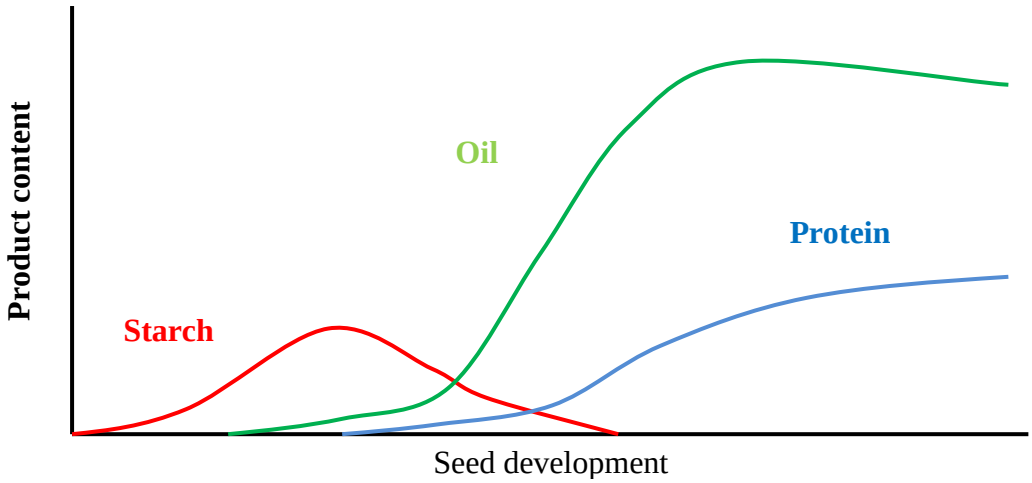


Figure 7. **Accumulation of storage products in Arabidopsis seeds and patterns of expression of genes that are involved in storage product synthesis, adapted from (Hills, 2004).** Genes involved in carbohydrate synthesis are expressed at an early stage of seed development, followed by those involved in oil synthesis, and storage protein genes are expressed at last.

1.1.2 Seed Storage product accumulation in Arabidopsis

Despite the fact that plant seeds are different in morphology and mode of propagation, they have one characteristic in common which is the accumulation of storage reserves necessary to fuel the germination of the seedlings. The main seed storage reserves are oil in the form of triacylglycerols (TAGs), seed storage proteins, and carbohydrates in the form of starch. The proportion of each compound differs with respect to species and ecotypes. In *Arabidopsis thaliana* ecotype WS, TAGs and seed storage proteins represent approximately 40 % of the mature dry seed weight each (Baud *et al.*, 2002), whereas starch is totally absent in the mature dry seeds (Focks and Benning, 1998; Baud *et al.*, 2002).

Starch is made of two glucose polymers, the highly branched amylopectin and relatively unbranched amylase. It is stored as semicrystalline granules inside chloroplasts in photosynthetic organs, and in the amyloplasts in starch storing organs where more than 70% of the starch granule is made of amylopectin. Amylose is mostly made of (1→4) α-D-glucopyranosyl units in a linear linkage. Amylopectin is made primarily of α-D-glucopyranosyl residues connected by (1→4) linkages with approximately 6% of (1→6) linkages at the branch points of the molecule (Buleon *et al.*, 1998). Starch is the most important storage carbohydrate in the seeds. The developing embryos of the plant model Arabidopsis and its close relative oilseed rape (*Brassica napus*) accumulate starch during early stage of development corresponding to the active photosynthesis phase (da Silva *et al.*, 1997; Focks and Benning, 1998; Baud *et al*., 2002). The monosaccharides synthesized during photosynthesis are exported from the chloroplast to the cytosol in the form of triose-phosphates which are the precursors for glucose synthesis during gluconeogenensis. However, a fraction of the glucose remains in the chloroplast and it is stored in the form of transient starch. The developing seeds of Arabidopsis accumulate starch shortly after flowering (3 DAF). The starch content increases rapidly to reach its maximum at the heart stage of embryo development (ca. 1 µg per seed). The content of transient starch decreases markedly after the heart stage, and it is almost absent in the mature dry seeds (Focks and Benning, 1998; Baud *et al*., 2002). In the developing seeds of oilseed rape, the embryo-specific repression of the gene encoding ADP-glucose pyrophosphorylase (AGPase), an enzyme catalyzing the conversion of glucose-1-phosphate to ADP-glucose, the substrate for starch synthesis (Figure 8), results in a reduction in AGPase activity of approximately 50%. In addition, the starch content is reduced up to 70 % in the developing embryos of transgenic seeds at 25 DAF. Furthermore, the amount of glucose-1-phosphate is increased, whereas the amount of ADP-glucose is decreased. Thus, the repression of the AGPase gene led to an inhibition of starch biosynthesis. Moreover, glycolytic enzyme activities, carbon flux through glycolysis, and ATP levels were reduced in these embryos, indicating that glycolysis and respiration were inhibited. The lipid content in the developing transgenic seeds was reduced to approximately 50% of wild type content. There was a slight decrease in total lipid content in the mature transgenic seeds, but no change in fatty acid composition. Furthermore, starch was totally absent in the mature transgenic seeds. These results indicate that starch is not the main carbon source for TAG or sucrose synthesis during embryo maturation, but that its transient accumulation is crucial for normal oil synthesis in the embryo (Vigeolas *et al.*, 2004). The degradation of the starch in the leaves during the dark period provides carbon for

the sink organs and generates energy necessary for cell metabolism. The daily fluctuation of the leaf starch content is indicated by the variation of anabolism and catabolism of starch. Moreover, abiotic conditions such as light intensity affect both the deposition and the mobilization of starch (Zeeman *et al.*, 1998; Zeeman and Ap Rees, 1999; Yu *et al.*, 2001).

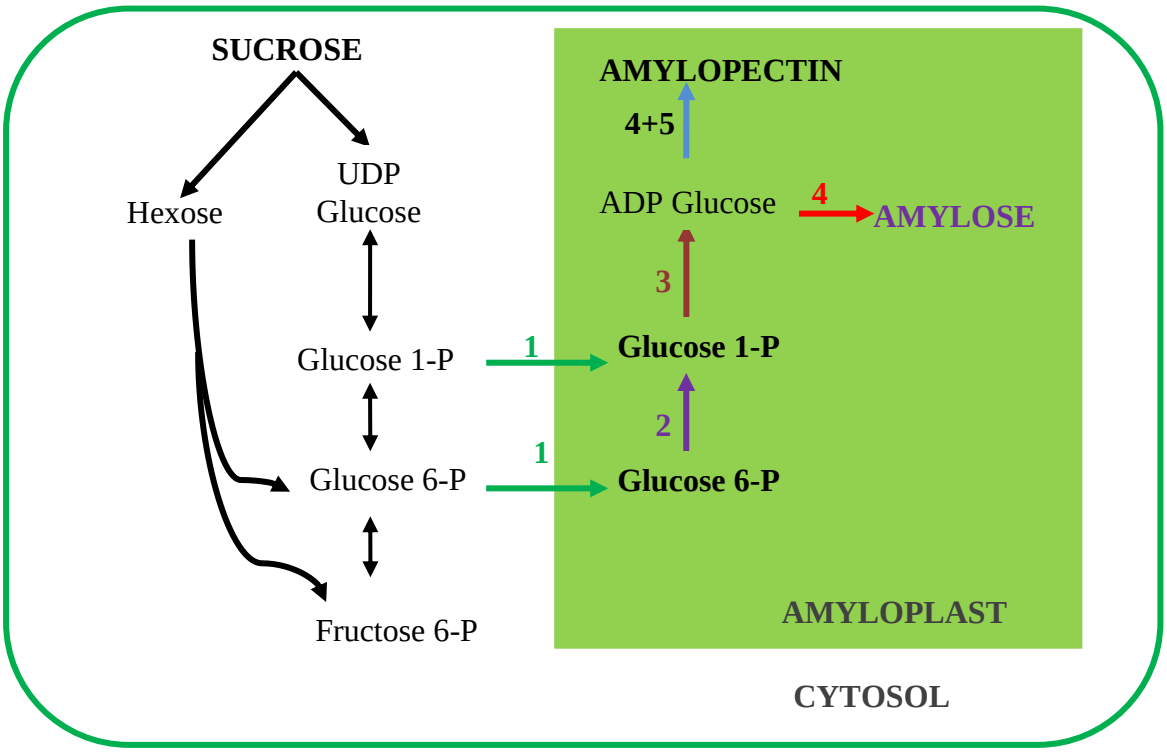


Figure 8. Schematic representation of the pathway of starch synthesis in non photosynthetic starch storing organs, adapted from (Smith *et al.*, 1995). Sucrose is metabolized via glycolysis in the cytosol to produce glucose and fructose. The glucose phosphate crosses the amyloplast envelope. Inside the amyloplast, ADP-glucose is synthesized via ADP-glucose pyrophosphorylase. ADP-glucose is the substrate for the synthesis of the starch polymers amylose and amylopectin catalyzed by different starch synthases and starch-branching enzymes. **1**: hexose-phosphate translocator; **2**: plastidial phosphoglucomutase; **3**: ADP-glucose pyrophosphorylase; **4**: starch synthase; **5**: starch-branching enzyme.

In the developing silique of Arabidopsis, sucrose which is synthesized from glucose and fructose is delivered to the seed coat from the mother plant via the vascular bundle of the funiculus. The sucrose is then cleaved either by sucrose synthase (SUS) or the invertase (Inv). The cleavage of sucrose by SUS produces UDP-glucose. The glucose is imported into the plastid and used for the biosynthesis of transient starch. This transient starch is stored in the seed coat of the developing seed. However, the cleavage of sucrose by Inv produces glucose and fructose which are used by the developing embryo and endosperm as source of energy (Figure 9). In the late stage of Arabidopsis seed development, maternal tissues directly supply

the embryo with sucrose which is cleaved by SUS to produce precursors for the biosynthesis of TAGs and storage proteins. Furthermore, the transient starch is exported from the seed coat to the embryo and later degraded. (Fallahi *et al.*, 2008).

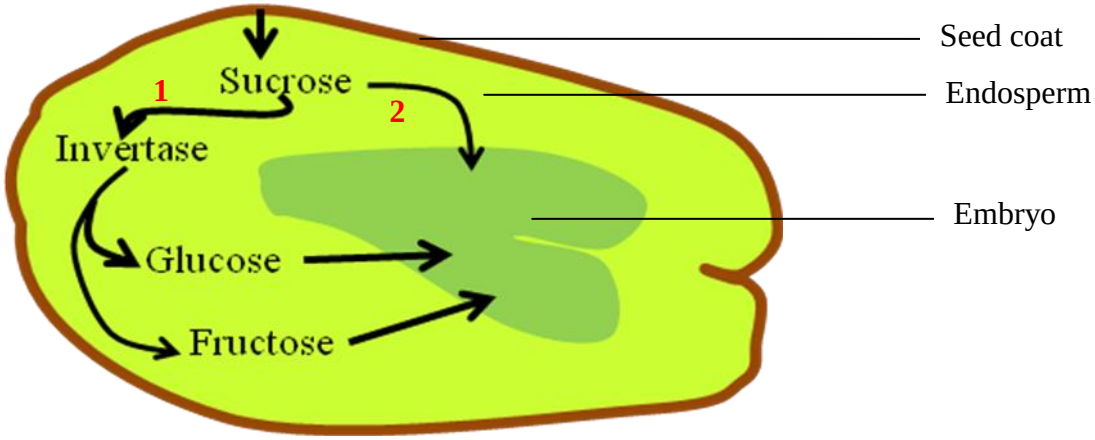


Figure 9. Sugar movement in developing Arabidopsis seed, adapted from (Hill *et al.*, 2003). (1) In the early stage of seed development the sucrose is delivered from mother plant to the endosperm. The sucrose is cleaved by invertase to produce glucose and fructose which are transported to the developing embryo. **(2)** During the late stage of seed development the sucrose is directly delivered to the embryo from the mother plant.

In contrast to sucrose which is stored in the mature dry seed (ca. 500 ng/seed) (Focks and Benning, 1998), developing Arabidopsis seeds accumulate high amounts of glucose and fructose in the early stage of developing. The concentration of these hexoses reaches a maximum of approximately 100 ng/seed at 5 DAF. Glucose and fructose contents decrease rapidly during seed maturation and only traces are present in the mature dry seed. (Focks and Benning, 1998; Baud *et al*., 2002).

In many higher plants, starch is used as the main seed storage reserve. In the mature seeds of cereals and legumes, starch is stored in the endosperm and embryo, respectively. The starch content accounts for 50-80% of the mature dry seed weight (Smith *et al*., 1995). Due to the importance of starch in the human diet and animal feed, in addition to its possible use as a source of renewable energy, the model plant Arabidopsis was extensively used to investigate the enzymes involved in starch biosynthesis. The aim of these studies was to transfer the knowledge gained from Arabidopsis to other crops with starch storing organs and create genetically modified crops with higher starch content.

Three recessive Arabidopsis *starch excess* (*sex*) mutants (*sex1*, *sex4*, and *dbe1*) were isolated. These mutants are characterized by the accumulation of high starch contents in the photosynthetic organs after extended incubation in darkness. Using a forward genetic

approach, Caspar and co-authors (Caspar *et al.*, 1991) isolated the mutant *sex1*. This mutant accumulates a high amount of starch in leaves and seeds. Moreover, the *sex1* mutant accumulates a high amount of starch in siliques, anthers, roots and petals. The authors also reported that the *sex1* mutant growth is affected, suggesting that the proper degradation of the transient starch is indispensible for normal growth. Yu and co-authors (Yu *et al.*, 2001) reported that the *sex1* mutant is defective in a gene encoding the Arabidopsis homolog of the R1 protein from potato. Deficiency in the R1 protein results in an increase in starch content in potato leaves and tubers indicating the involvement of the R1 protein in starch degradation (Lorberth *et al.*, 1998). The *sex4* mutant (Zeeman *et al.*, 1998) has a high starch content similar to the mutant *sex1*. In addition, the *sex4* mutant accumulates two times more starch in the seed coat and anthers compared to wild type. Zeeman and co-authors reported that the *sex4* mutant is deficient in the chloroplastic endo-amylase A2 suggesting that this endo-amylase is indespensable for the degradation of the transient starch. The *dbe1* mutant (Zeeman *et al.*, 1998) is impaired in a gene encoding the chloroplastic debranching enzyme isoform of the isoamylase. The *dbe1* mutant accumulates less than 10% of the starch content in the leaves than the wild type. In addition, this mutant accumulates more soluble branched glucans (phytoglucogen) than the wild type during the day. The accumulated phytoglucogen represents almost 50% of starch content in wild type leaves. This phytoglucogen is later degraded during the night (ca. 90%). The chloroplasts of the *dbe1* mutants are characterized by the presence of small starch granules surrounded by several pockets of materials which are not present in the wild type. The amylopectin structures in the chloroplasts of the mutant were similar to the wild type suggesting that the debranching enzymes are not necessary for starch synthesis in the chloroplasts.

Mature Arabidopsis seeds accumulate oil in the form of TAGs which constitute approximately 94% of the seed storage oil (Figure 10). The TAGs are stored in the embryo and make up to 40% of the mature dry seed weight (North *et al.*, 2010). Oil content in Arabidopsis seeds is variable between ecotypes, and it is under the control of environmental conditions such light and growth conditions (Focks and Benning, 1998; Baud *et al*., 2002; O'Neill *et al.*, 2003; Li *et al.*, 2006). The analysis of seed oil content in 360 *Arabidopsis thaliana* ecotypes showed that the average oil content represents approximately 38% of the seed dry weight, with most accessions having an average oil content between 33% and 43% (O'Neill *et al*., 2003). Interestingly, the set of FA found in the seed TAG is conserved in most ecotypes of *Arabidopsis thaliana*, despite the fact that they are grown in different geographic locations, and under different environment conditions (Millar and Kunst, 1999).

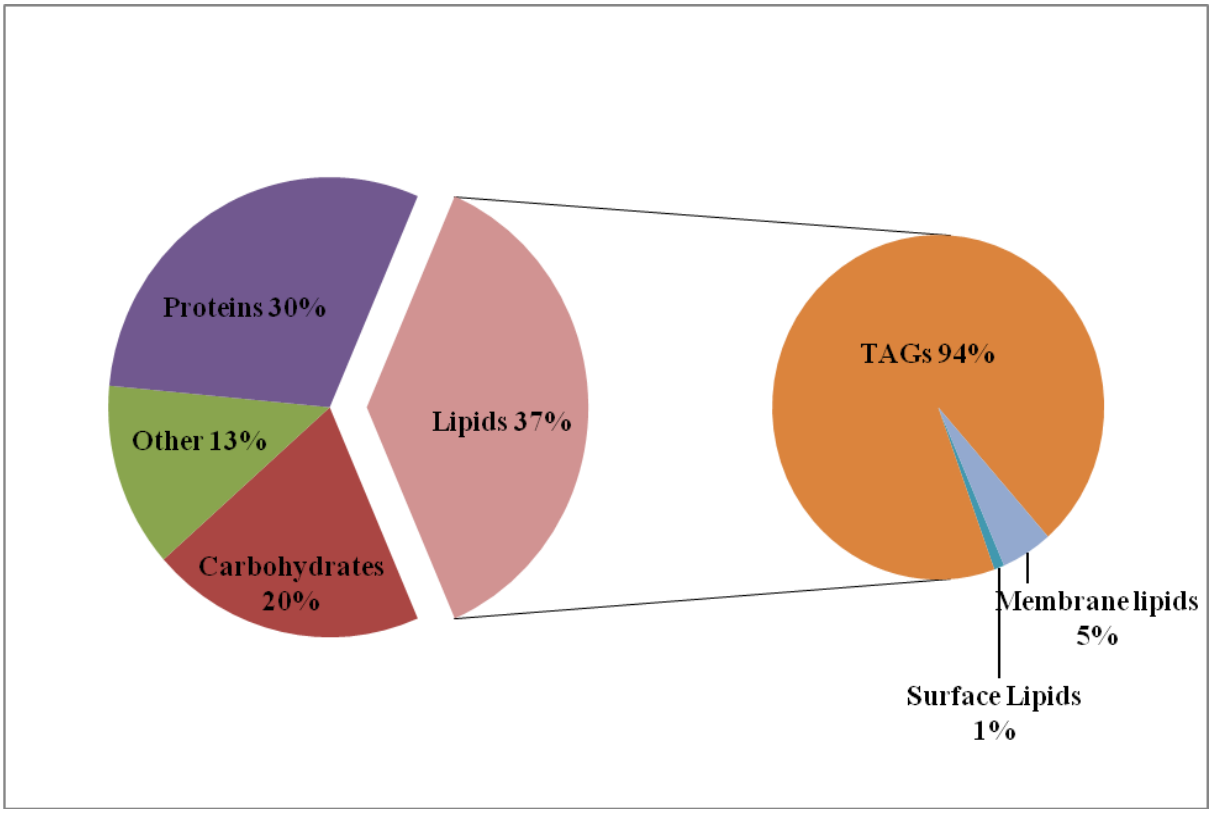


Figure 10. Relative distribution of lipids and other storage compounds in mature Arabidopsis seeds adapted from (Ohlrogge and Browse, 1995; Li *et al.*, 2006; Molina *et al.*, 2006; Beisson *et al.*, 2007)

TAGs are neutral hydrophobic triesters of glycerol with fatty acids esterified at each of the three hydroxyl groups of the glycerol backbone (Figure 11). The highly reduced carbons of the fatty acids attached to the glycerol backbone of the TAGs provide more energy during oxidation than the oxidation of proteins or carbohydrates. Thus, the TAGs are the main storage compounds in the seeds of many plants and in animal tissues (Alvarez and Steinbuchel, 2002; Baud and Lepiniec, 2008). In developing oilseeds, the fatty acids are first synthesized in the plastid stroma, and later exported to the cytoplasm in the form of fatty acyl-CoAs. The assembly of the fatty acyl-CoA with the glycerol-3-phosphate which takes place in the ER results in the formation of the TAGs (Ohlrogge and Browse, 1995). The TAG biosynthesis pathway is known as the Kennedy pathway. The first step of TAG synthesis involves two successive acylations with acyl-CoA at the sn-1 and sn-2 position of glycerol-3-phosphate to produce phosphatidic acid (PA). The first acylation is catalyzed by glycerol-3-phosphate acyltransferase, while the second acylation is catalyzed by lyso-phosphatidic acyltransferase. The PA is hydrolysed by phosphatidate phosphatase to produce diacylglycerol (DAG). The sn-3 position of DAG is acylated by diacylglycerol acyltransferase (DAGT) to produce TAG. DAGT is a unique enzyme specific to TAG

synthesis (Voelker and Kinney, 2001; Chapman and Ohlrogge, 2012). In the mature seeds of most plants, TAGs are stored in special spherical organs called oil bodies (oleosomes). Each oleosome comprises a TAG matrix surrounded by a monolayer of phospholipids enclosed in a layer of structural proteins called oleosins (Huang, 1996). The size of the mature oilseed oleosomes is directly correlated with their TAG and oleosin contents (Ross *et al.*, 1993; Tzen *et al.*, 1993), and the seed oil content is reflected by the number of oleosins. Arabidopsis and oilseed rape seeds with high oil contents have an increase of approximately 20% in oleosin number compared to wild types (Murphy and Cummins, 1989; Siloto *et al.*, 2006). The suppression of the Arabidopsis gene *OLEO1* encoding the most abundant oleosin in seeds resulted in the accumulation of large oil bodies in the embryos of transgenic seeds. In addition, the germination of these seeds was delayed, and they accumulated less oil and more proteins compared to wild type, whereas the starch and sucrose contents were similar to wild type (Siloto *et al.*, 2006).

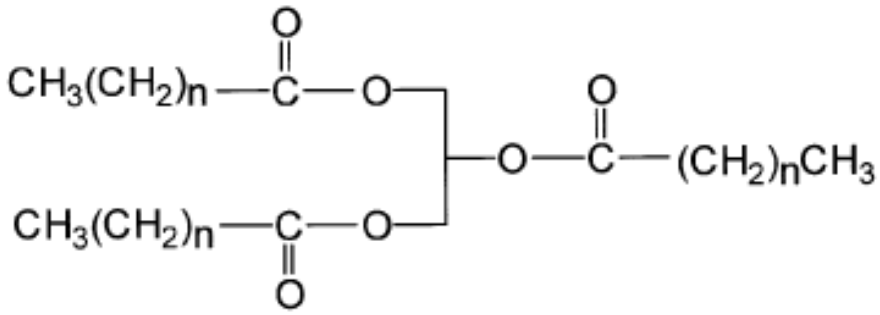


Figure11. Simplified schematic of TAG.

The biosynthesis of fatty acids is an indispensable metabolic pathway in all plant cells. The complete repression of genes involved in fatty acids biosynthesis results in lethality. Thus, no mutants with deficiency in fatty acids biosynthesis have been reported (Ohlrogge and Browse, 1995). In the developing seed of Arabidopsis, TAG biosynthesis begins approximately 6 DAF (early heart stage of embryo development). The amount of TAGs increases rapidly to reach its maximum at 18 DAF, this increase is accompanied with an increase in embryo size and seed dry weight (Focks and Benning, 1998; Baud *et al.*, 2002; Gomez *et al.*, 2006). The Arabidopsis seed oil is composed of 8 major fatty acids with high amounts of saturated fatty acids, polyunsaturated fatty acids (PUFAs) and very long chain fatty acids (VLCFAs) (Figure12). Palmitic acid (16:0) and stearic acid (18:0) are saturated fatty acids, while oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3) are

unsaturated. The VLCFAs in Arabidopsis seed are mainly arachidic acid (20:0), 11-eicosenoic acid (20:1), 11,14-eicosadienoic acid (20:2), and erucic acid (22:1) (O'Neill *et al*., 2003). The cotyledons of mature Arabidopsis embryos contain approximately 87% of total seed oil content, whereas seed coat and endosperm contain around 13% (Li *et al*., 2006). The mature Arabidopsis seed endosperm contains high amounts of 16:0, 18:2, 18:3, 18:1 and 20:1 which were presumed to be stored in the form of TAG. The storage of TAGs in Arabidopsis endosperm suggests a possible role of this tissue in fuelling seed germination and seedling development (Penfield *et al.*, 2004; Li *et al*., 2006).

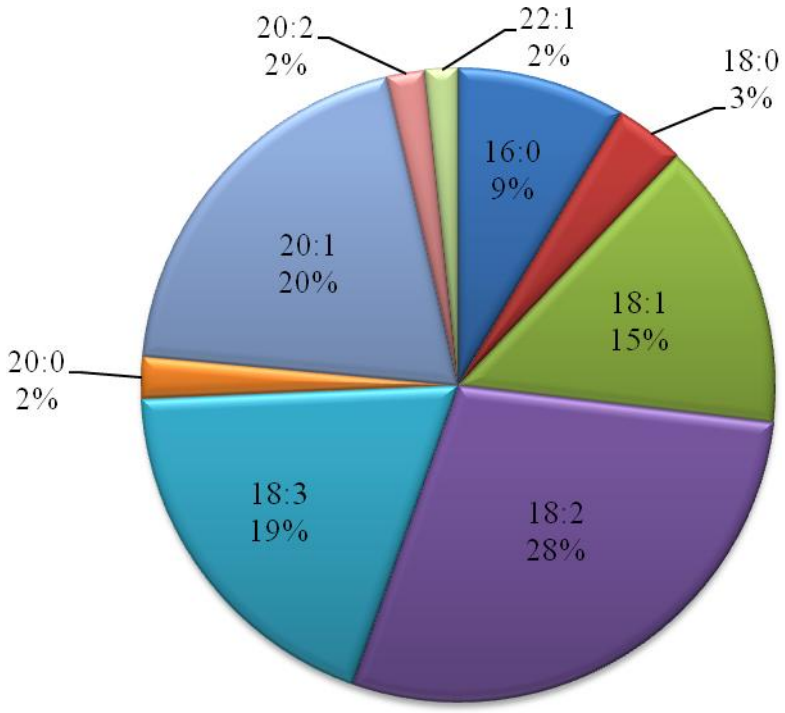


Figure12. Fatty acid composition of seed oil in mature dry seed of *Arabidopsis thaliana* Columbia-0, modified from (Li *et al*., 2006).

The key genes controlling fatty acid and storage oil accumulation in Arabidopsis seeds are poorly investigated, despite the fact that the enzymes involved in the FA and TAG biosynthesis are well described. Previous studies have shown that the overexpression of single genes encoding enzymes controlling FA or TAG biosynthesis results in an alteration of oil content. The overexpression of rapeseed lysophosphatidic acid acyltransferase isozymes in Arabidopsis seeds induces an increase in seed oil content of approximately 13% compared to wild type. Moreover, the dry weight of the transgenic Arabidopsis seeds was increased by approximately 6% (Maisonneuve *et al*., 2010). Jako and co-authors (Jako *et al.*, 2001) reported that the overexpression of DAGT in Arabidopsis seeds induces an increase in seed

oil content, seed weight, and seed yield compared to wild type. In contrast, the overexpression of 3-ketoacyl-acyl-carrier protein synthase III (KAS III) from *Cuphea hookeriana* in rapeseed, using a seed specific promoter, results in the reduction in total oil content (Dehesh *et al.*, 2001). Several key genes controlling the TAG biosynthesis were isolated during the screening for Arabidopsis mutants with alterations in seed storage oil deposition. Katavic and co-workers (Katavic *et al*., 1995) isolated the Arabidopsis *tag1* mutant which is deficient in the gene encoding the DGAT. The developing seeds of the *tag1* mutants accumulate less oil at a reduced rate associated with a decrease in VLCFAs biosynthesis. Moreover, the seed development and maturation of *tag1* were delayed. Focks and Benning (Focks and Benning, 1998) isolated a mutant causing wrinkledness of the seeds. Homozygous *wrinkled1* (*wri1*) seeds accumulate TAGs at a low rate, and the total oil content is reduced by approximately 80%. The TAG fatty acid composition of *wri1* is characterized by a reduction in oleic acid (18:1) and linoleic acid (18:2). Concomitantly, the levels of erucic acid (22:1) and lionlenic acid (18:3) are increased. The *WRI1* gene encodes a transcription factor of the APETALA2/ethylene-responsive element-binding protein (AP2/EREB) family. This transcription factor is involved in the control of carbon metabolism during embryo development and maturation, and it is expressed mainly in the seeds (Cernac and Benning, 2004). Recently, it has been shown that the overexpression of the maize (*Zea mays) WRI1,* using a seed specific promoter, induces an increase in seed oil content and reduction in starch content by 46% and 60%, respectively. The increase of oil content in transgenic maize plants had no effect on important agronomic traits such as germination rate, seed yield and protein content (Shen *et al.*, 2010). The *shrunken seed1 (sse1)/pex16* mutant is another mutant impaired in seed oil content. The mature dry seeds of *sse1* are characterized by an increase in starch content and a concomitant reduction in oil content (Lin *et al*., 1999). Lin and co-authors (Lin *et al*., 2004) reported that the reduced oil content and the low rate of fatty acid synthesis which led to an increase in starch content in the *sse1* mature seeds is caused by an alteration in peroxisome biogenesis.

Seed storage proteins are the main reserve source of nitrogen and sulfur in the mature seeds. In developing seeds, the endosperm is the main tissue accumulating storage proteins in monocotyledoneous plants, whereas in the dicotyledoneous plants they are stored in the embryo (Hughes and Galau, 1989). The seed storage proteins are first synthesized as precursors in the ER, and later exported and stored in their mature form (protein bodies) in special vacuoles called protein storage vacuoles (PSVs) (Herman and Larkins, 1999). The

conversion of the precursors of seed storage proteins into their mature forms is achieved by the seed specific vacuolar processing enzymes (VPE) (Haranishimura *et al.*, 1995). Mature seeds of Arabidopsis mutants deficient in one of the gene encoding the vacuolar processing enzyme (*βVPE*) accumulate high amounts of precursors of seed storage proteins in addition to their mature forms (Shimada *et al.*, 2003). The seed storage proteins are variable in their structures, yet they have three common characteristics. They are all synthesized in specific tissues at defined stage of development. All are stored in the mature seeds as protein bodies. Moreover, these proteins display high polymorphisms both in the same genotypes and in genotypes from the same species (Shewry *et al.*, 1995). Mature Arabidopsis seeds accumulate storage proteins in the form of 12S globulin and 2S albumin (Heath *et al*., 1986). In developing Arabidopsis seeds, storage proteins start to accumulate at the early stage (5 DAF). The amount of storage proteins increases during seed maturation to reach a final concentration of 9 µg in mature dry seed (Focks and Benning, 1998; Baud *et al*., 2002). Shimada and co-authors (Shimada *et al.*, 2003; Shimada *et al.*, 2006) isolated two Arabidopsis mutants *atvsr1* and *maigo1* (*mag1/vsp29*) impaired in seed storage protein accumulation. In fact, the two mutants missort the seed storage proteins by secreting them from the cells into the extracellular space in the mature seeds. As a result, the mutant seeds are characterized by the small size of their PSVs. In addition the two mutants accumulate high amounts of precursors of storage proteins in their mature seeds.

1.2 Control of Arabidopsis seed size by developmental genes

Several Arabidopsis seed developmental genes were isolated. Mutations in these genes result in alterations in embryo and/or endosperm development which as a consequence may affect seed storage compound deposition and seed size. At least two groups of genes were described on the basis of their effect and action on plant seed development, the polycomp group complex (PgG), and the SHB1 co-activator group.

In Drosophila *(Drosophila melanogaster*), the PcG proteins suppress homeotic gene expression by changing the chromatin structure (Simon and Tamkun, 2002). Therefore, the PcG complex has been defined as a group of genes whose single mutations result in phenotypes similar to those of Drosophila *Polycomb* (*Pc*) mutants (transformation of the second and third thoracic segments into the eighth abdominal segment), or which can enhance the phenotypes of Drosophila *Pc* mutant alleles (Grimaud *et al.*, 2006). In Drosophila and mammals, the repression of PcG target genes is the results of the coupled

action of two large multi-protein complexes known as Polycomb Repressive Complex 1 (PRC1) and PRC2 (Lund and van Lohuizen, 2004). The screening for mutations affecting different aspects of plant development such as seed and flower development, and the vernalization response, led to the identification of a number of PcG genes of the PCR2 complex. The amino acids of the proteins encoded by these genes show high sequence similarity to animal PRC2 proteins (Breiling *et al.*, 2007). In Arabidopsis, endosperm development is under the control of 5 known PcG proteins of the PCR2 complex. These proteins include FERTILIZATION INDEPENDENT SEED 2 (FIS2), FERTILIZATION INDEPENDENT ENDOSPERM (FIE/FIS3), MEDEA (MEA/FIS1), MULTICOPY SUPRESSOR OF IRA (MSI1), and SWINGER (SWN) (Pien and Grossniklaus, 2007). These PcG proteins control gene expression in the female gametophyte during endosperm development through imprinting (Wang *et al.*, 2006; Hennig and Derkacheva, 2009). Before fertilization, the PcG complex suppresses endosperm development and plays an important role in regulating seed development after fertilization. Thus, mutations in genes encoding proteins of the PcG complex result in initiation of endosperm over-proliferation without cellularization in the absence of fertilization. After fertilization, the mutant embryo development is arrested and the embryos exhibit different defects in cell proliferation and morphogenesis (Leroy *et al.*, 2007; Pien and Grossniklaus, 2007). The gene expression of *PHERES1 (PHE1);* a gene encoding a MADS-box protein, is regulated via imprinting by *FIS* and *MEA* (Köhler *et al.*, 2003a; Villar *et al.*, 2009). Moreover, the seed abortion phenotype of the *mea* mutant is rescued by reducing the expression levels of *PHE1*.

In addition to the PgG proteins, *SHORT HYPOCOTYL UNDER BLUE 1 (SHB1)* (Zhou *et al*., 2009), *IKU2* (Garcia *et al.*, 2003), and *MINI3* (Luo *et al.*, 2005) are major regulators of Arabidopsis seed development. The gain-of-function overexpression mutation *shb1-D* induces increases in seed size, whereas in *shb1* mutant, a loss-of-function allele, the seed size is reduced. At four days after pollination, the developing seed of *shb1-D* mutant are characterized by the presence of large embryo sacs and endosperm cellularization is delayed compared to wild type. *shb1-D* mature seeds have more and larger cells compared to wild type. The authors have shown that *SHB1* is binding with the promoter region of *MINI3* and *IKU2* in vivo though the SHB1 protein does not contain a signature DNA binding motif. Thus, the *SHB1* activity appears to be necessary for the correct expression of *MINI3* and *IKU2*. *SHB1*, *MINI3* and *IKU2* may function in the same signaling pathway controlling embryo and endosperm development.

To investigate the role of Arabidopsis seed developmental genes in storage deposition, four candidate genes where chosen in this study. *FIE* and *PHE1* are members of the PgG complex, whereas *MINI3* and *IKU2* belong to the SHB1 co-activator group respectively.

1.2.1 *FERTILIZATION-INDEPENDENT ENDOSPERM* (*FIE*)

Ohad and co-workers (Ohad *et al*., 1996) isolated an Arabidopsis mutant capable of starting endosperm development without fertilization. The mutant was designated *fie* which stands for fertilization independent endosperm. Female gametophytes carrying the *fie* mutation undergo embryo abortion at the heart stage even if they were pollinated by wild type pollen. In the absence of fertilization, *FIE/fie* heterozygous plants produce elongated siliques containing seed-like structures which are characterized by the absence of embryo development and the presence of an endosperm surrounded by the seed coat. Thus, the *FIE* gene is indispensable to prevent early seed development in the absence of fertilization. Ohad and co-authors (Ohad *et al.*, 1999) reported that the *FIE* gene encodes a protein associated to the family of WD Polycomb group proteins that control pattern formation during embryogenesis in insects and mammals. Moreover, *FIE* is expressed in both reproductive and vegetative tissues before and after fertilization. It has been reported that pollination of maternal *fie* mutants with hypomethylated pollen bearing the wild type *FIE* allele is able to rescue the endosperm development in these mutants. This result indicates that endosperm development does require neither fertilization nor embryogenesis and it represents a new opportunity to produce apomictic crops in which endosperm is the main storage organ in the mature seeds (Vinkenoog *et al.*, 2000).

1.2.2 *PHERES1 (PHE1)*

The *PHERES* (*PHE1*) gene encodes a type I MADS-box protein. In developing seeds of Arabidopsis wild type, *PHE1* is expressed shortly after fertilization (3 DAP), and only in the female gametophyte. The expression level of *PHE1* reaches its maximum in the pre-globular stage of embryo development and decreases rapidly afterwards. In Arabidopsis wild type plants, *PHE1* expression is restricted to the embryo and chalazal region of the endosperm of developing seeds (Köhler *et al.*, 2003a). The *PHE1* gene expression is under the control of the Fertilization Independent Seed Polycomb Group (*FIS-PcG* complex) and the Ubiquitine-Specific Protease gene *AtUBP26*. In developing seeds of both mutants *FIS-*

PcG group and the *Atubp26*, the expression of the *PHE1* gene is up-regulated compared to wild type (Köhler *et al.*, 2003a; Luo *et al.*, 2008). The *PHE1* gene is the only known Arabidopsis imprinted gene to be parentally expressed and maternally repressed (Köhler *et al.*, 2005).

1.2.3 *HAIKU2 (IKU2)*

Garcia and co-workers (Garcia *et al*., 2003) isolated the mutant *haiku 2(iku2)* which is characterized by a reduced seed size. The mutation in the *IKU2* gene is sporophytic recessive, and results in an early arrest of endosperm growth. As consequences, the mutant embryos are characterized by an early cellularization and restrain of cell proliferation. In addition, the seed coat cell elongation of *iku2* is reduced compared to wild type. Seeds of homozygous *iku2* are reduced by approximately 25% and 14% in length and width respectively. Thus, *iku2* mature seeds are more spherical comparing to the oblong wild type seeds. Despite their small size, *iku2* seeds germinate normally, and develop into plants morphologically indistinguishable from wild type plants. Luo and co-authors (Luo *et al*., 2005) reported that *IKU2* encodes a leucine-rich repeat transmembrane protein kinase and that *IKU2* is strictly expressed in the endosperm of developing seeds at the early stage of development.

1.2.4 *MINISEED3 (MINI3)*

Luo and collaborators (Luo *et al*., 2005) isolated the Arabidopsis mutant *mini3* with reduced seed size. This gene encodes a WRKY transcription factor. Seeds of homozygous *mini3* plants have a similar phenotype as the previously described seed mutant *iku2*. Thus, the *mini3* mutation is sporophytic recessive, and plants homozygous for *mini3* produce seeds markedly smaller than wild type. Contrary to *IKU2* which is only expressed in the endosperm of developing seeds, *MINI3* is expressed in the endosperm and the embryo shortly after fertilization. *IKU2* expression is reduced in the mutant *mini3*, whereas *MINI3* expression is similar to wild type in the mutant *iku2*, suggesting a possible regulation of *IKU2* by *MINI3*.

The two genes *IKU2* and *MINI3* are under control of the gene *SHORT HYPOCOTYL UNDER BLUE1 (SHB1)* (Zhou *et al*., 2009). In the *short hypocotyl under blue 1-Dominant (shb1-D*) mutant, a gain of function mutant of *SHB1*, the expression of *MINI3* and *IKU2* is increased, whereas the expression of the two genes is reduced in the loss of function mutant of *SHB1*. The enhanced expression of *IKU2* and *MINI3* in the *shb1-D* background results in an increase in seed weight and seed yield per plant compared to wild type.

1.3 The aim of this work

Enhancing seed size and therefore increasing seed storage compound content is of major economic and agricultural importance. Thus, the analysis of genes controlling embryo and/or endosperm development represents a new opportunity to reach this goal. In this study, the role of four Arabidopsis candidate genes in seed storage product deposition was investigated. (i) Homozygous mutants of the candidate genes *FIE* (At3g20740), *PHE1* (At1g65330), *MINI3* (At1g55600) and *IKU2* (At3g19700) were analyzed by measuring seed storage reserve deposition in mature seeds compared to the wild types background. (ii) cDNAs encoding the candidate genes were expressed in *Arabidopsis thaliana Col-0* using the embryo specific glycinin promoter, and seed storage product deposition in the mature seeds of transgenic lines was investigated. The aim of this study was to increase seed size and storage compound content by seed specific overexpression of these candidate genes in Arabidopsis. The final outlook is to apply our findings to generate transgenic oilseed crops, like rapeseed, with increased seed size and seed storage compound content.

2. Materials and methods

2.1 Equipment and chemicals

2.1.1 Equipment

Analytic balance XS205	Labomedic
Gas Chromatograph (GC) 6890 Series with Flame Ionization Detector (FID)	Agilent
Centrifuge 5424	Eppendorf
Centrifuge 5810R	Eppendorf
Centrifuge Biofuge 13 Heraeus	SEPATECH
Stereomicroscope SZX 16	Olympus
Light microscope BH-2	Olympus
Microscope digital camera DP72	Olympus
Fluorescence illuminator X-Cite®120Q	EXFO
Ultracentrifuge Optima L 90K	Beckman Coulter
Mixer Mill 400	Retsch
Microtome RM 2155	Leica Microsystems
Spectrophotometer Specord 205	Analytik Jena
Spectrophotometer Nano Drop 1000	Peqlab
Software for microscopy CellSens Standard	Olympus
BH2-RFL-T2 power source	Olympus
Water bath	Köttermann Lab Devices
Vortex Certomat®MV	Braun
pH Meter pH Level 1	InoLab WTW
Sterile bench Model 1.8	Holten LaminAir
Thermocycler T Professional	Biometra
Thermocycler T Personal	Biometra
Gel documentation system E-BOX 100/26M	Peqlab

UV-Transilluminator DP-001 T1A	Vilber Lourmat
Micro Pulser Electroporator	Bio-Rad
Heating Mixing Block MB-102	BIOER
Speed vacuum SPD121P	Thermo Savant
Refrigerator vapor trap RVT 4104	Thermo Savant

2.1.2 Chemicals

Acetic acid	AppliChem
Acetone	AppliChem
30 % Acrylamide	AppliChem
Adenosine triphosphat (ATP)	Sigma
Agarose	Peqlab
Ammonium acetate	Sigma
Amonium persulfate (APS)	Bio-Rad
Bacto peptone	Duchefa Biochemie
Bromophenol blue	Serva
Boric acid	Fisher Scientific
Chloroform	Merck
Confidor WG 70	Rolfs
Calcium chloride	Merck
Diethyl pyrocarbonate (DEPC)	Sigma
Dithiothreitol (DTT)	AppliChem
99% Ethanol	AppliChem
100% Ethanol	Merck
Sodium chloride	Duchefa Biochemie
Sodium hypochlorite solution	Hedinger
Sorbitol	Duchefa Biochemie
Tris-(hydroxymethyl)-aminomethane, Trometamol	Duchefa Biochemie
Murashige and Skoog medium	Duchefa Biochemie
Phyto agar	Duchefa Biochemie
Proplant	Rolfs
Ethidium bromide	Serva

Isopropanol	AppliChem
Tryptone	AppliChem
Yeast extract	Duchefa Biochemie
Ethylendiaminetetraacetic acid (EDTA)	Roth
Lauroylsarcosine-sodium salt	Serva
Hexadecyltrimethylammonium bromide(CTAB)	Sigma
Phenol/Chloroform/isoamylalcohol	Sigma
β-mercaptoethanol	Sigma-Aldrich
Lithium chloride	Sigma
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Duchefa Biochemie
5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal)	AppliChem
Glucose	AppliChem
4-(2-Hydroxyethyl)-1-piperazine-1-ethansulfonic acid (HEPES)	AppliChem
Potassium chloride	Merck
Magnesium chloride hexahydrate	AppliChem
Manganese (II) chloride tetrahydrate	AppliChem
Magnesium sulfate, anhydrous	Sigma
Triton X-100	Sigma
Sucrose	Duchefa Biochemie
Glycerol	Duchefa Biochemie
Silwet Gold	Rolfs
3N Methanolic HCl	Sulpelco
Methanol	J.T. Baker
N-Hexane	Merck
Pentadecanoic acid	Sigma
Formic acid	Sigma
Imidazole	Sigma
Nicotinamide adenine dinucleotide phosphate ($NADP^+$)	Sigma
Sodium dodecyl sulfate (SDS)	AppliChem
Tetramethylethylenediamine (TEMED)	Bio-Rad
Glycine	AppliChem
Technovit 7100	Heraeus Kulzer
Technovit 3040	Heraeus Kulzer

Toluidine blue	Sigma-Aldrich
Paraformaldehyde	Sigma
Glutaraldehyde	Serva

2.2 Enzymes and molecular biology kits

2.2.1 Enzymes

T_4 DNA ligase	New England biolabs
DCS Pol-DNA polymerase	DNA Cloning Service
Taq DNA polymerase	Fermentas
Pfu DNA polymerase	Fermentas
Paq 5000 hotstart PCR master mix	Agilent
MluI	Fermentas
Xhol	Fermentas
EcoRI	Fermentas
PstI	Fermentas
Glucose-6-P dehydrogenase	Sigma
Hexokinase	Sigma
Phosphogluco-isomerase	Sigma
Invertase	Sigma

2.2.2 Molecular biology kits

High-Speed Plasmid Mini kit	DNA cloning service
SuperScript® III First Strand Kit	Invitrogen
NucleoSpin Plasmid preps	Machery-Nagel
Gel/PCR DNA fragment extraction kit	DNA cloning service
pGEM-T easy vector system	Promega
Cloned JET PCR cloning kit	Fermentas
DC Protein Assay kit	Bio-Rad
RevertAid™ Premium First Strand cDNA Synthesis Kit	Fermentas

2.3 Plant material and growth conditions

2.3.1 Seed surface sterilization

Arabidopsis seeds were surface sterilized in a 1.5 ml microcentrifuge tube by immersion in 99 % (v/v) ethanol for 10 seconds and air-dried for 10 min. The seeds were incubated for 10 min with gentle shaking in 2.4% hypochlorite solution. The seeds were washed five times with sterilized, double deionized water (ddH_2O), air-dried under the laminar flow hood and kept sterile at 4°C.

2.3.2 Plant growth and selection

Arabidopsis surface-sterilized seeds were sown on Petri dishes with Murashige and Skoog (MS) medium (1 X Murashige and Skoog salts, 1 % sucrose, 0.05 % MES-KOH, pH 5.7; 1% Phyto agar was then added to the medium and autoclaved for 20 min at 120°C.) To select transgenic lines with T-DNA insertion, selective antibiotics were added to the cool medium under the Laminar Flow Hood. The plates were sealed by wrapping parafilm around them and kept in the dark at 4°C for three days for stratification, then transferred to a growth cupboard (Rumed, Rubarth Apparate) with 16 h light/8 h dark cycle at 20 °C. After 2 weeks the plants were transferred into plastic pots with soil (Topf Erde, vermiculite 2:1). The pots were placed into plastic trays. Before planting, the soil was watered with a solution of 1.5 ml/l Proplant, 1.5 ml/l boric acid, and 20 mg/ml Condifor**.** The plants were kept in the growth chamber (temperature 20°C, air humidity 55%, fluorescent light, cool white/warm white (1:1), light intensity: ca 150 µE in shelves, light period: 16 h light), watered regularly and treated in the same way for all trays. When the siliques shattered, the plants were covered with paper bags and watered for one or two more weeks. After stopping the watering, the plants were left to complete dryness. Seeds were harvested, dried under vacuum for 48, and stored at 4°C.

2.3.3 Manual pollination

For efficient manual pollination, crossing of *Arabidopsis* plants should be done in the morning. Before starting the manual pollination, mature siliques, open flowers, and open buds are removed from the inflorescence of the pollen recipient plant. This is done using a magnifying glass or binocular microscope with scissors and fine forceps. The inflorescence is gently fixed using a sticky tape, and then sepals, petals, and anthers are gently removed from the closed flower buds. The stigmas of emasculated flowers were manually pollinated by flowers from pollen donator plants. Pollination is achieved by gentle touching of stigmas

from emasculated flower buds of the pollen recipient plant, with stamen of mature flowers from pollen donor plant. The pollinated inflorescences are marked and covered with plastic foil for approximately 3 to 4 days. The mature siliques from manually pollinated flowers are harvested prior to opening and kept at room temperature for further maturation.

2.3.4 Antibiotic stock concentration for plant selection

The following antibiotics kanamycin monosulfate, hygromycin B, and sulfadiazine were dissolved in sterile ddH_2O to final concentrations of 50 mg/ml, 20 mg/ml, and 7.5 mg/ml, respectively, and stored at - 20°C. The Table 1 shows the mutant lines for the candidate genes used in this study and antibiotics used for the selection of homozygous lines.

Table.1: Arabidopsis mutants alleles used in this study:

Mutant	Mutant line	Ecotype	Antibiotic resistance	Origin
mini3-2	Salk_050364	*Col-0*	Kanamycin	SALK Institute
mini3-3	Salk_056336	*Col-0*	Kanamycin	SALK Institute
mini3-1	EMS mutant	*Ler*	Non	Luo *et al*. (2005)
phe1-6	RATM11-2982-1_H	*No-0*	Hygromycin	SALK Institute
phe1-7	GT_5_6830	*Ler*	Kanamycin	SALK Institute
phe1-1	ET189	*Ler*	Kanamycin	SALK Institute
iku2-4	Salk_073260	*Col-0*	Kanamycin	SALK Institute
iku2-5	Salk_110598	*Col-0*	Kanamycin	SALK Institute
iku2-3	EMS mutant	*Ler*	Non	Luo *et al*. (2005)
fie-13	GK-362D08	*Col-0*	Sulfadiazine	SALK Institute
fie-14	GK-532F01	*Col-0*	Sulfadiazine	SALK Institute
fie-11	*fie-11*(N8172)	*C24*	Non (γ radiation mutant)	SALK Institute

2.4 Bacterial strains and cloning vectors

2.4.1 Bacterial strains

Two *Escherichia coli (E. coli)* strains were used for plasmid transformation:

1. Xl10 Gold®Ultracompetent cells (Stratagene) {Tet^{r}Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*) 173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F´ *proAB lacI*^{q}ZΔ*M15* Tn*10* (Tet^r) Amy Cam^r]. (Genes listed signify mutant alleles. Genes on the F´ episome, however, are wild-type unless indicated otherwise}.

2. ElectroSHOX competent cells (Bioline){F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ* ΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara* Δ139 Δ(*ara, leu*)7697 *gal*U *gal*K λ- *rps*L (Str^R) *nup*G *λ*}.

The two *E.coli* strains were grown in LB medium containing a final concentration of 100 μg/ml ampicillin or 50 μg/ml kanamycin from a stock solution of 100 mg/ml and 50 mg/ml, respectively.

Two *Agrobacterium tumefaciens* (*A. tumefaciens*) strains were used for the floral dipping of *Arabidopsis* plants, namely pGV2260 and LB4404. The two strains were grown in YEP medium with final concentration of 50 μg/ml kanamycin and 60 μg/ml rifampicin from a stock solution of 50 mg/ml and 60 mg/ml, respectively. Rifampicin is dissolved in dimethyl sulfoxide (DMSO).

LB		**YEP**	
Tryptone	10.0 g/l	Yest extract	10 g/l
Yeast extract	5.0 g/l	Bacto peptone	10 g/l
NaCl	5.0 g/l	NaCl	5 g/l
pH 7.2 (with NaOH)		pH 7.0 (with NaOH)	

LB and YEP liquid media were autoclaved for 20 min at 120 °C. LB and YEP agar plates were made by adding 1.5% (w/v) of agar to the liquid media after adjusting the pH and autoclaved.

2.4.2 Cloning vectors

Plasmids used for subcloning of PCR products include the pGEM® -T Easy vector (Promega), pUC19 (Stratagene), and the pJET1.2 vector (Fermentas). The three cloning vectors confer ampicillin resistance in bacteria. For *Arabidopsis* plant transformation, the binary vector pBinGlyRed1 (Prof. Edgar Cahoon, Department of Biochemistry, University of Nebraska-Lincoln, Nebraska, USA) was used. pBinGlyRed1 confers kanamycin resistance in

bacteria and Basta resistance in plants. Furthermore, it carries a DsRed marker gene (*Discosoma sp*. red fluorescent protein).

2.5 Molecular techniques

2.5.1 Genomic DNA extraction from Arabidopsis leaves

Fresh leaves were harvested and transferred to 1.5 ml microcentrifuge tubes containing two 3-mm stainless steel beads, and immediately frozen in liquid nitrogen. The frozen leaves were ground to a fine powder using a mixer mill (Retsch). 1ml of CTAB buffer was added to the tubes followed by incubation at 65°C with vigorous shaking for 10 min. 400µl chloroform was added and the samples shaken for complete homogenization and centrifuged for 5 min at 3000 rpm. After transferring the aqueous phase into fresh microcentrifuge tubes, 700 µl of isopropanol was added and the solution was mixed by inverting the tubes 5 to 6 times and placed on ice for 10 min to precipitate the DNA. The precipitated DNA was pelleted by centrifugation for 5 min at 10 000g, washed with 70 % (v/v) ice-cold ethanol, dried, and dissolved by gentle pipetting in 50 µl ddH_2O. 0.5 µl of RNaseA (10 µg/ml) was added to each sample followed by incubation at room temperatures for 15 min. The quantity and quality of the extracted DNA was analysed using the spectrophotometer (NanoDrop). The concentration was calculated by measuring the DNA absorbance at the wave length of 260 nm (A_{260}). An absorbance of 1.0 unit at 260 nm (A_{260}) = 50 µg/ml DNA.

DNA concentration (µg/ml) = A_{260} x dilution factor x 50 µg/ml

The purity of the extracted DNA was controlled by calculating the (A_{260}/A_{280}) ratio. A high purity DNA sample has a (A_{260}/A_{280}) ratio between 1.7 and 2. The extracted DNA is stored at -20°C.

CTAB buffer 500 ml:

140 mM sorbitol	12.8 g
220 mM Tris-HCl, pH 8	110 ml of 1 M stock solution
22 mM EDTA	22 ml of 0.5 M
800 mM NaCl	80 ml of 5 M or 23.2 g
1% Sarkosyl	5 g N-lauroylsarcosine
0.8% CTAB	4 g

2.5.2 PCR screening of Arabidopsis T-DNA mutants

The screening for homozygous plants for each T-DNA insertion line was done by PCR of genomic DNA from each line. PCR was performed using combinations of gene specific forward and reverse primers for each locus, and T-DNA left border specific primer (LB) as shown in Table 3. The left border specific primers were:
For Salk lines LB, bn78; Riken lines LB, bn130; GABI kat LB, bn142.

PCR reaction (25 µl):

2.5 µl 10 X PCR buffer
0.75 µl 50 mM $MgCl_2$
0.5 µl 10 mM dNTP-mix (10 mM of each)
2.5 µl forward primer 10 pmol/µl
2.5 µl reverse primer 10 pmol/µl
10-500 ng template DNA
0.5 µl DNA polymerase (5 units/µl)

The annealing (T_a) temperature used for the PCR depends on the melting temperature (T_m) of the chosen primers; generally an annealing temperature is 5**°C** bellow the lowest T_m of the pair of primers used (Michael A. Innis, 1990). The melting temperature can be calculated using the following formula (Rychlik and Rhoads, 1989) : T_m = 64.9°C + 41°C x (number of G's and C's in the primer – 16.4)/N. N: is the length of the primer.

Table 2: PCR program

Temperature	Time	Cycle step	Cycle number
94°C	3 min	Initial Denaturation	1
94°C	30 sec	Denaturation	
T_m-5°C	30 sec	Annealing	35
72°C	1 min/kb (Taq), 2min/kb (pfu)	Extension	
72°C	10 min	Final Extension	1
4°C	hold	Hold storage	-

2.5.3 DNA agarose gel electrophoresis

The products of the PCR amplification, restriction enzyme digestion, and the extracted DNA were analysed by gel electrophoresis using 0.8% agarose gels with 1 X TBE (Tris/borate/EDTA) buffer. To prepare a 0.8% gel, 3.2 g agarose was added to 400 ml 1 X TBE buffer and boiled in a microwave oven. After cooling to about 55°C, ethidium bromide (EtBr) was added to a final concentration of 5 μg/ml. The gel was immediately poured into a chamber with a comb. After the solidification of the gel, the comb was removed and the chamber with gel was placed into a tank filled with 1 x TBE buffer covering the top of the gel**.** The DNA samples were mixed with 6 x gel loading buffer (DNA), and loaded with size standard (1 kb ladder, Fermentas) into the wells. The gel was run under an electric voltage of 100-120 V until the separation of size standard bands was achieved. DNA bands were visualised using an UV transilluminator.

10 X TBE buffer (1 l):

108g	890 mM Tris base
55g	890 mM boric acid
40ml	0.5 M Na_2-EDTA (pH 8.0)

Ethidium Bromide solution: 1% ethidium bromide is dissolved in H_2O (1 mg/ml) and stored in the darkness at 4°C.

6 x Gel loading buffer (DNA) 10 ml:

25mg	0.25 % bromphenol blue
3.5g	30 % glycerol
25mg	0.25 % xylene Cyanole

The loading buffer was stored at 4°C

0.5 M EDTA pH 8.0 (1 l):

186.1 g of Na_2-EDTA $2H_2O$ is dissolved in 700 ml ddH_2O, the pH was adjusted to 8.0 with 10 M NaOH, and the final volume was adjusted to 1 l with ddH_2O.

Table 3: Primer combinations used for the PCR screening of T-DNA insertion lines with expected PCR product band size in base pairs (bp).

The top row for each T-DNA line shows the primer pair for the amplification of the genomic wild type copy of the gen. The bottom row shows the primer pair used to amplify the flanking genomic sequence.

Gene	T-DNA lines	Primers	PCR product expected size (bp)
MINI3	Salk_050364	bn68+bn69	1167
		bn78+bn69	803
MINI3	Salk_056336	bn171+bn172	1099
		bn78+ bn172	534
PHE1	RATM11-2982-1_H	bn66+bn67	1045
		bn130+bn67	790
PHE1	GT_5_6830	bn273+bn274	1059
		bn130+bn274	506
PHE1	ET189	bn66+bn67	1045
		bn66+bn277	750
IKU2	Salk_073260	bn60+bn61	1168
		bn78+bn61	900
IKU2	Salk_110598	bn167+bn168	1084
		bn78+ bn168	815
FIE	GK-362D08	bn165+bn166	1178
		bn142+bn166	549
FIE	GK-532F01	bn205+bn206	1033
		bn142+bn206	441

2.5.4 RNA extraction from Arabidopsis siliques.

The REB method modified from (Kay *et al.*, 1987), is used to isolate RNA from tissues rich in polysaccharides**.** Fresh siliques (approximately 20) were harvested in 2 ml microcentrifuge tubes containing stainless steel beads. The samples were immediately frozen and ground in liquid nitrogen using a mixer mill. 1 ml of REB Buffer, 1 ml phenol/chloroform (1:1), and 20 µl β-mercaptoethanol were added to each sample followed by vigorous vortexing and incubation for 5 min at room temperature. After centrifugation for 5 min at 4°C, the aqueous phase was extracted twice, first with phenol/chloroform (1:1), and then with chloroform. The final supernatant was transferred to a fresh sterile tube. A solution of 8 M LiCl (300 µl; DEPC treated) was slowly added to the supernatant to a final concentration of 2 M. After overnight incubation at 4°C, the samples were centrifuged for 15 min at 14000 rpm. The pellets were then rinsed with cold 2 M LiCl followed by resuspension in 70 % ethanol and incubation for 20 min at -20°C. The RNA was precipitated by centrifugation for 10 min at 14000 rpm. The final pellets were dissolved in 50 µl DEPC treated water and stored at -80°C. The RNA concentration was calculated after measuring the absorption at 260 nm (A_{260}) and 320 nm (A_{320}) using the spectrophotometer. An absorbance of 1 unit at 260 nm (A_{260}) corresponds to 40 µg/ml of pure RNA.

RNA concentration (µg/ml) = A_{260} x dilution factor x 40 µg/ml

The purity of the extracted RNA was controlled by calculating the (A_{260}/A_{280}) ratio. A high purity RNA sample has a (A_{260}/A_{280}) ratio between 1.8 and 2. The integrity of the isolated RNA was checked by running 0.5 to 1 µg in 1 % agarose gel with 1 x TE buffer.

10 X TE buffer (1 l):

242 g	Tris-base
57.1 ml	glacial acetic acid
100 ml	0.5 M Na_2-EDTA (pH 8.0)

DEPC- H_2O **(**1 l**)**: 1 ml (0.1%) of diethylpyrocarbonate is added to 1 l H_2O. After vigorous shaking the solution is incubated overnight at 37°C followed by autoclaving to times for 20 min each at 120°C.

2.5.5 First-Strand cDNA synthesis and reverse transcription PCR (RT-PCR)

The first-strand cDNA was synthesised from total RNA using the SuperScript® III kit (Invitrogen). Following the manufacturer's instructions, 5 µg of total plant RNA was mixed with 1 µl oligo dT (50µM), 1 µl of 10 mM dNTP mix, and DEPC-treated water was added to a total volume of 10 µl. Then the mixed solution was incubated at 65°C for 5 min, and placed on ice for at least 1 min. Meanwhile 10 µl cDNA Synthesis Mix was prepared by mixing 2 µl of 10 X RT buffer, 4 µl of 25 mM $MgCl_2$, 2 µl of 0.1 M DTT , 1 µl of RNaseOUT (40 U/µl), and 1 µl SuperScript. III RT (200 U/µl). The cDNA Synthesis Mix was added to the RNA/ primer mixture and mixed gently, the solution was then incubated for 50 min at 50°C, followed by incubation at 85°C for 5 min to terminate the reaction and chilled on ice. The solution was collected by brief centrifugation and treated with 1 µl of RNase H for 20 min at 37°C. The cDNA synthesised was used for PCR amplification.

2.5.6 Preparation and transformation of chemically competent *E. coli* cells

Preparation of competent cells and transformation of *E.coli* was done as described by Inoue and co-authors with small modifications (Inoue *et al.*, 1990). A culture tube with medium (3 ml of LB with 50 µl/ml streptomycin) was inoculated with a single colony of ElectroSHOX competent cells (Bioline) and incubated over night at 37°C with rotation at 180-200 rpm. A 1 l-flask, containing 200 ml SOB medium, was inoculated with 1 ml over night culture and incubated at 37°C with rotation (180-200 rpm) until the OD_{600} was ca. 0.5. Following incubation on ice for 10 min, the cells were poured into 50 ml sterile Falcon tubes and centrifuged at 2500 g at 4°C for 10 min. After discarding the supernatant, cell pellets were resuspended in 80 ml ice-cold TB buffer and incubated on ice for 10 min, then centrifuged again at 2500 g at 4°C for 10 min. Cell pellets were then resuspended with gentle swirling in 20 ml ice-cold TB buffer and DMSO at a final concentration of 7 % and incubated for 10 min on ice. The cell suspension was divided into 100 µl aliquots in 1.5 ml microcentrifuge tubes and immediately frozen in liquid nitrogen and stored at -80°C. In order to perform a transformation, the competent cells were thawed on ice and 0.5-1 µl of the plasmid DNA or 5 µl ligation reaction was added to the tube and incubated on ice for 30 min. The cells were heat shocked at 42°C for 90 sec without shaking and transferred to ice for 2-3 min. 1 ml of LB^+ or SOC medium was added to the tube. After incubation for 1 h at 37°C with vigorous rotation, the cells were plated onto LB plates with the appropriate antibiotic and incubated over night at 37°C.

Blue/white screening

For blue/white screening, IPTG and X-Gal were added to the plates prior to incubation. pGEM-T Easy, pUC18, and other subcloning vectors harbor the *lac* operon which is a prerequisite for blue/white screening.

IPTG: 0.12 g Isopropyl-β-D-thiogalactopyranoside dissolved in 5ml ddH_2O filter sterilized and stored at 4 °C.

X-Gal: 0.1 g 5-bromo-4-chloro-3-indoly-β-D-galactoside dissolved in 2 ml N,N-dimethylformamide (DMF), stored at -20° C.

TB buffer:

10 mM HEPES (or Pipes)
15mM $CaCl_2$
250mM KCl
55mM $MnCl_2$

} Adjust pH to 6.7 with KOH, filter sterilized and stored at 4°C.

SOB	**SOB**	**LB+**
0.5% yeast extract	0.5% yeast extract	5 ml LB
2% tryptone	2% tryptone	50 µl 40% glucose
10 mM NaCl	10 mM NaCl	100 µl 1M $MgCl_2$
2.5 mM KCl	2.5 mM KCl	
10 mM $MgCl_2$	10 mM $MgCl_2$	
10 mM $MgSO_2$	10 mM $MgSO_2$	
	20 mM glucose	

Colony PCR

Colony PCR is designed to quickly screen for plasmid DNA inserts directly from *E. coli* colonies. This technique can be used to determine the size and orientation of the insert using specific primers. Thus, any colony which gives rise to an amplification product of the expected size is likely to contain the correct DNA sequence. To each PCR tube containing 30 µl PCR reaction, a small amount of the colony is added using fine white pipette tips and mixed by pipetting. Afterwards a PCR is performed.

2.5.7 Plasmid DNA isolation from *E. coli*

The isolation of small amounts of plasmid DNA from *E. coli* was done using a modified method described by Riggs and McLachlan (Riggs and McLachlan, 1986). 2 ml of the overnight culture was collected in 2 ml microcentrifuge tubes. The cells were pelleted by centrifuging at 11 000 rpm for 1 min. After discarding the supernatant, the cells were resuspended in 200 BF buffer with lysozyme (1 mg/ml) and incubated for 1-2 min at room temperature. The samples were incubated in the heating block or water bath at 95°C for 1 min followed by immediate cooling on ice for 5 min. After centrifugation for 30 min at 13 000 rpm, the supernatant was transferred into new microcentrifuge tubes containing 480 µl of the IS mix. Then the suspensions were mixed by inverting the tubes 6-8 times and incubated for 1-2 min at room temperature. The plasmid DNA was precipitated by centrifuging for 20 min at 13 000 at room temperature, the pellet was washed twice with 70 % (v/v) ethanol and dried at room temperature. The pellet was resuspended in 80 µl ddH_2O with 1 µl RNase A and incubated for 30 min at 37 °C, then stored at -20°C.

The isolation of highly purified plasmid DNA for sequencing was performed using the NucleoSpin® Plasmid and NucleoSpin® Plasmid Quick Pure kits from MACHEREY-NAGEL following the manufacturer's instructions.

BF buffer:

8 % (w/v)	sucrose
0.5 % (w/v)	Triton X-100
50 mM	EDTA pH 8.0
10 mM	Tris-HCl pH 8.0

Lysozyme: 20 mg/ml lysozyme dissolved in ddH_2O and stored at - 20°C.

IS mix:

400 µl isopropanol

80 µl of 5 M ammonium acetate

2.5.8 Enzymatic modification of cDNA

For overexpression in *Arabidopsis* seeds, cDNAs of corresponding genes were cloned into the binary vector pBinGlyRed1 carrying a seed specific promoter. The cDNAs of candidate genes were obtained from the Arabidopsis biological resource centre (ABRC) and stored in our lab stocks (table4).

Table 4: cDNA clones of candidate genes obtained from stock centres.

Stock N°.	Gene	Vector	cDNA clone	Stock Center
899	*FIE*	pUNI51	U60089	ABRC
900	*PHE1*	pENTR/D-TOPO	PYAT1G65330	ABRC
901	*MINI3*	pDONR 221	DQ446362	ABRC
903	*IKU2*	pFLAG-GATEWAY	X3G19700EFK	ABRC

The *E. coli* clones carrying the cDNAs were grown overnight in LB medium followed by plasmid DNA isolation. 2 µg of each plasmid DNA was dissolved in 5 mM Tris-HCl (pH 8.0) and sent to Eurofins MWG Operon for sequencing using the primers bn58 and bn59 for the vector pFLAG-GATEWAY, and M13 uni(-21) and M13 rev(-29) for other vectors.

2.5.9 PCR amplification of the cDNAs and cloning into pGEM®-T Easy vector

The amplification of the cDNA was done using gene specific primers. The cloning of the PCR products from pGEM-T Easy into the binary vector pBinGlyRed1 was achieved by adding restrictions enzyme sites onto the 5' end of each primer (Table 5). *Pfu* DNA polymerase or other proofreading DNA polymerases were used for amplification.

Table 5: Primers for PCR amplification of cDNAs.

cDNA clone	Primers	Primers 5' modification	PCR product size (bp)
U60089	bn157	MluI	1122
	bn158	XhoI	
PYAT1G65330	bn155	EcoRI, MluI	858
	bn156	XhoI	
DQ446362	bn159	MluI	1470
	bn160	XhoI	

The cloning of the *IKU2* cDNA was achieved by PCR using cDNA and genomic DNA fragments. A 1077 bp cDNA fragment was amplified from the clone X3G19700EFK using the primers bn553 and bn551 with an XhoI restriction site at 5' end of the primer bn551, the genomic fragment of 2265 bp was amplified using the primers bn550 and bn669 with MluI restriction site at 5' end of the primer bn550.

Thermostable DNA polymerases that have a "proofreading" function, such as *Pfu* and *Tli*(a) DNA Polymerases, exhibit 3´-->5´exonuclease activity and produce greater than 95% blunt-end fragments (Lundberg *et al.*, 1991). PCR fragments generated by proofreading enzymes can be ligated into the pGEM®-T vectors if they are first tailed with dATP using *Taq* DNA polymerase. After the PCR amplification, 50 µl of the PCR product is separated on 1 % agarose gel, the expected band is cut out and the PCR product is extracted from the gel using the NucleoSpin® Extract II kit from MACHEREY-NAGEL following the manufacturer's instructions.

A-tailing protocol:

17 µl	PCR product extracted from gel	incubation at 72°C for 30 min
2 µl	10 x Taq PCR buffer	
0.5 µl	dATP (100 uM)	
0.5 µl	Taq poylmerase	

The A-tailing products are stored at – 20°C.

The A-tailed PCR product is ligated to the pGEM- easy vector (Promega) following the manufacturer's instructions. The ligation reaction can be incubated overnight at 4°C to increase the efficiency.

The cDNAs, and DNA (fragment of *IKU2*) cloned into the pGEM-T Easy vector were digested using restriction enzymes specific to sites introduced to the 5' and 3' sites of the PCR amplification products. The pBinGlyRed1 vector was first digested with the restriction enzymes EcoRI and XhoI. Then, the *PHE* cDNA was ligated into these sites thereby generating the construct pBinPHE-Red1. This construct harbours an extra MluI site that is important for the cloning of the other constructs. *The MINI3* and *FIE* cDNAs were ligated into the vector pBinPHE-Red1 previously digested with MluI and XhoI. The cloning of the full length *IKU2* cDNA coding sequence into pBinGlyRed1 vector was done in three steps; first, the cDNA fragment amplified from the clone X3G19700EFK was released from pGEM-T easy using PstI and XhoI. The 726 bp fragment (insert 1) obtained after digestion was purified. The second step consisted of cutting the *IKU2* genomic DNA fragment previously cloned into pGEM-T easy using MluI and PstI, resulting in a 2248 bp fragment (insert 2). In the third step, the two purified fragments were ligated in one reaction into pBinPHE-Red1 previously digested with MluI and XhoI, with a vector/insert 1/insert 2 molar ratio of 1:3:3. The constructs were transformed into competent *E. coli* cells followed by plasmid DNA isolation and digestion with specific restriction enzymes to confirm the

presence of the DNA fragment within the vector. A glycerol stock was made from the correct culture and stored at -80°C.

Glycerol stock:

- Add 800 µl of 70% autoclaved glycerol into 2 ml microcentrifuge tube.
- Add 800 µl from the overnight culture into the tube.
- Gently mix the solution and store at -80°C.

The competent *A. tumefaciens* cells were first thawed on ice (10 µl per transformation). 1 µl of the binary vector carrying the construct of interest was added to the thawed cells followed by gentle mixing and incubation on ice for 3-5 min. The mixture was then transferred to prechilled electroporation cuvettes. The electroporation was done using the following conditions: Voltage, 1250 V; Capacitance, 25 µF; and Pulse length, 5 msec. Following the electroporation, 800 µl of YEP media without antibiotics was added to the cuvettes, and the bacterial suspensions were transferred into 1.5 ml microcentrifuge tubes. The cells were incubated for 2 hours at 28°C with vigorous rotation. 100 µl of the cell suspensions were spread on YEP agar plate containing a final concentration of 50 µg/ml and 60 µg/ml of kanamycin and rifampicin, respectively, and then incubated for 2 to 3 days at 28°C. Single colonies were picked from each transformed plate and incubated over night at 28°C, and then sealed with parafilm and stored at 4°C as stock plates. Glycerol stocks were made from a single colony and stored at -80°C.

2.5.10 The Arabidopsis plants transformation using floral dipping

The transformation of Arabidopsis plants using floral dipping was based on the method described by (Clough and Bent, 1998). *Arabidopsis* plants were grown under long day conditions until they were flowering (4 plants each per 10 cm pot). The first bolts were clipped away to stimulate the proliferation of many secondary bolts. 4-6 days after clipping the plants were ready for transformation. Clipping can be repeated to delay the day for the transformation. A single colony of an *Agrobacterium tumefaciens* cells carrying the gene of interest on a binary vector was used to inoculate an overnight pre-culture that was grown with vigorous shaking at 200 rpm in 5 ml YEP media with antibiotics (60 µg/ml rifampicin and 50 µg/ml kanamicin). The 5 ml pre-culture was added to 200 ml YEP media with antibiotics and incubated overnight with vigorous shaking at 28°C.the *Agrobacterium* cells were harvested by centrifuging the 200 ml overnight culture for 20 min at 5000 x g, the pellet was resuspended in 5% sucrose solution (prepared fresh, no need to autoclave). 400 ml of this

solution was used for two to three 10 cm pots. Before dipping Silwet L-77 was added to a final concentration of 0.05 % (500 µl/l) and mixed well. The plants were dipped into the *Agrobacterium* solution for 2 to 3 seconds, with gentle agitation. The dipped plants were laid on their side and covered overnight with plastic foil. On the next day the plants were transferred to the growth chamber and normally grown to harvest the seeds. The transformed seeds were selected under the binocular fluorescence microscope using green light and a red filter to detect DsRed fluorescence.

2.6 Biochemical and Analytical techniques

The lipid, protein and sugar measurements were performed with 3 or 5 biological replicates. The seeds were obtained from Arabidopsis wild types and transgenic lines cultivated side by side in the same growth chamber.

2.6.1 Fatty acid methyl ester measurements (FAME)

The FAME measurements were done as described by Browse and co-authors (Browse *et al.*, 1986) with some modification. Five seeds of each *Arabidopsis* line including wild type (n= 5 independent replicates) were subjected to direct transmethylation by putting them into an 7 ml glass tube with screw cap with teflon septum. Then 500 µl 1.5 N methanolic HCl and 100 µl of the internal standard solution were added. 100 µl of pentadecanoic acid (15:0; 50 µg/ml in methanol) was used as the internal standard. The samples were incubated for 2 h and 30 min at 90°C. After cooling, 500 µl 0.9 % NaCl and 500 µl hexane were added. Fatty acid methyl esters (FAMEs) were extracted in the hexane phase by vigorous shaking followed by centrifugation at 1000 g for 3 min. 150 µl of each sample was taken from the hexane phase and transferred to a gas chromatography vial. The FAMEs were quantified using the pentadecanoic acid methyl ester standard on the Agilent Technologies 7890A Gas Chromatograph with flame ionization detection (GC-FID) equipped with an autosampler.

2.6.2 Total lipid extraction

The total lipid extraction from was done as described by Dörmann and co-authors (Dormann *et al.*, 1995) with some modifications. 20 mg of *Arabidopsis thaliana* Col-0 seeds (n= 3 independent replicates) were ground with a porcelain mortar in a 7 ml glass tube with screw cap and teflon septum. 2 ml of methanol/chloroform/formic acid (1:1:0.1) and 1 ml of 1 M KCl/0.1 M H_3PO_4 were added to the ground seeds. 100 µl of pentadecanoic acid 15:0 (50 µg/ml in methanol) as internal standard was added to the extraction mixture for

quantification. The homogenised mixture was centrifuged for 2 min at 5000 g, the organic phase was transferred to a new tube, the aqueous phase was re-extracted with 3 ml methanol/chloroform (1:2). The combined organic phases were evaporated under the air stream and re-dissolved in 500 µl chloroform/methanol (2:1). To determine the fatty acid methyl esters (FAMEs), a 50 µl aliquot of the 500 µl lipid solution was transferred into a new screw cup tube. The solvent was evaporated under an air stream and lipids were dissolved in 500 µl 1 N methanolic HCl. The methylation reaction was incubated for 30 min at 80°C. The FAMEs were extracted with 500 µl 0.9 % NaCl and 500 µl hexane. 150 µl of the hexane phase was transferred to a gas chromatography sample vial.

Parameters of the Agilent 7890A Gas Chromatograph:

Column:	Supelco SP-2380
Length:	30 m
Diameter	0.53 mm
Coating thickness	0.20 µm
Carrier gas:	Helium
Flow rate	7 ml/min

Table 6: Temperature program for FAME measurement on the Supelco SP-2380 GC column:

Oven Ramp	Oven temperature	Run time
Initialization	100°C	0 min
25°C/min	160°C	2.4 min
10°C/min	220°C	8.4 min
25°C/min	100°C	13.2 min

Detector parameters:

Temperature	250 °C
Hydrogen (H_2) flow	30 ml/min
Air flow	400 ml/min
Make up (Helium)	19 ml/min
Mode	constant make up of 30 ml/min

2.6.3 Total seed protein extraction and quantification

The total seed proteins were extracted and quantified as described by (Focks and Benning, 1998) with some modifications. In a 1.5ml microcentrifuge tube, 20 seeds (n= 3 independent biological replicates) were ground to a fine powder using the mixer mill and homogenized in 250 ml acetone. Following centrifugation at 16000*g*, the supernatant was discarded, and the vacuum-dried pellet was resuspended in 250 ml of extraction buffer containing 50 mM Tris-HCl pH 8.0, 250 mM NaCl, 1 mM EDTA, and 1 % (w/v) SDS. The resuspended pellet was incubated for 2 h at 25°C, and then centrifuged at 16000*g* for 5 min. 100 µl of the supernatant was used for protein measurements using the Lowry DC protein assay (Bio-Rad) following the manufacturer's instructions.

2.6.4 Soluble carbohydrate analysis

The extraction of soluble sugars was performed as described by (Focks and Benning, 1998). 50 seeds (n= 3 independent biological replicates) were ground in 2 ml microcentrifuge tube using the mixer mill and homogenized in 500 µl 80 % (v/v) ethanol followed by incubation for 90 min at 70°C. Following centrifugation at 16000*g* for 5 min, the supernatant was transferred to a new 2 ml microcentrifuge tube. The pellet was extracted twice with 500 µl of 80% ethanol. The solvent of the combined supernatants (1500 µl) was evaporated at room temperature under a vacuum. The residue representing the soluble carbohydrate fraction was dissolved in 50 µl of water. To quantify soluble sugars, 10 µl of the sugar extract was added to 990 µl of reaction buffer containing 100 mM imidazole pH 6.9, 5 mM $MgCl_2$, 2 mM NADP, 1 mM ATP, and 2 units ml^{-1} of glucose-6-P dehydrogenase. For enzymatic determination of glucose, fructose, and sucrose, 4.5 units of hexokinase, 1 unit of phosphoglucoisomerase, and 2 ml of a saturated fructosidase (invertase) solution were added in succession. The production of NADPH was photometrically followed at a wavelength of

340 nm. The amount of NADPH was calculated based on a molar extinction coefficient of 6.22 (l/mmol.cm).

2.6.5 Determination of seed weight

100 mature Arabidopsis seeds (n= 3 independent biological replicates) were counted and then weighed carefully on a Mettler Toledo XS205 analytical balance. The seeds were harvested from plants grown under the same conditions. Mature seeds of the different mutant lines and their WT controls were dried under vacuum for 48 h before the determination of the mean of seeds weight.

2.6.6 Measurement of seed length, width and area

The measurement of seed length, with and area was performed using the Evaluator software (developed by Dmitry Peschansky, IPK Gatersleben). First, approximately 100 *Arabidopsis* seed were spread well separated on microscope slides and then fixed with sticky transparent tape. The slides and a scale (provided by Dr. Renate Schmidt, IPK Gatersleben) were scanned using a conventional scanner at high resolution (800 dpi) and saved as TIFF files. These TIFF files were exported into bitmap (bmp) format using Adobe Photoshop CS3. The preparation of images and their analysis were performed as described by the software developer.

2.6.7 SDS-PAGE (Polyacrylamide Gel Electrophoresis)

SDS-PAGE is a method used to separate proteins according to their size. SDS (sodium dodecyl sulfate) is a detergent that dissolves hydrophobic molecules and covers them with a negative charge. Following denaturation, the proteins are separated according to their size by running them in a polyacrylamide gel under a constant electric field.

Sample preparation:

1 or 5 *Arabidopsis* seeds were ground using the mixer mill (Retsch) for 5min. 20 µl of the protein sample loading ("Laemmli") buffer (Laemmli, 1970) is added to the sample. The seeds are ground again for 2 min then centrifuged for 1 min at 2000g followed by incubation at 65°C for 20 min. The sample can be used directly or stored at -20°C. 11.5 µl and 20 µl of each sample are loaded in the gel when extraction was done from five or from single seeds, respectively.

Laemmli buffer (1X)

200 mM Tris-HCl pH 6.8	6 ml 1M Tris-HCl pH 6.8
8 % SDS	2.4 g
0.4 % bromophenol blue	0.12 g
40 % glycerol	13.8 ml (87% stock)

ddH_2O was added to a final volume of 120 ml.

DTT (Dithiothreitol) is freshly added to each sample prior to loading into the gel (0.0154g DTT for each 1 ml Laemmli buffer).

17% polyacrylamid gel preparation (protocol for 4 minigels)

Separating Gel:

Acrylamide stock (30 % acrylamide, 0.8 % bisacrylamide;	28.3 ml
Protogel from National Diagnostics)	
1.5 M Tris-HCl, pH 8.8	12.5 ml
10 % SDS	0.5 ml
H_2O	8.2 ml
10 % (100 mg/ml) amonium persulfate (APS)	0.5 ml
TEMED (1,2-bis-dimethylamino-ethane)	20 µl

The polymerization started shortly after adding TEMED. After gentle mixing, the gels are poured into the chamber up to a level of 4 cm to the top (ca. 0.5cm from the comb). To provide a smooth surface on the top of the gels, 1 ml of isopropanol is added to the top of each gel. The gels are completely polymerized after ca. 30 to 40 min. The surface of the gels is rinsed with water before adding the stacking gel.

Stacking Gel:

Acrylamide stock (30 % acrylamide, 0.8 % bisacrylamide; Protogel from National Diagnostics)	3.4 ml
1 M Tris-HCl, pH 8.8	2.5 ml
10 % SDS	2 ml
H_2O	13.6 ml
10 % (100 mg/ml) amonium persulfate (APS)	0.2 ml
TEMED (1,2-bis-dimethylamino-ethane)	20 µl

Following gentle mixing after adding TEMED, the chambers are filled with stacking gel and combs are inserted carefully avoiding air bubbles bellow the combs' teeth. The gels are fully polymerized after ca. 60 min. The gels can be stored at 4 °C no longer than two weeks after wrapping them with wet paper or nylon.

Electrophoresis:

The gels are run with 1 X tank buffer, with constant voltage of 60 mA until the blue dye reaches the end of the gel.

5 x Tank buffer (1000 ml)

15 g Tris base	125 mM
72 g glycine	960 mM

2.7 Light microscopy

Green Arabidopsis seeds at late stage of development were used to check the mature embryo size and morphology under the light microscopy. Mature green siliques were opened using fine forceps under the binocular microscope and 10 to 15 seeds were gently picked and placed in 1 ml of the fixative solution. The samples were incubated in the vacuum for 1 hour at room temperature. After discarding the fixative solution, fresh fixative solution was added to the samples and kept in the vacuum overnight at 4°C. The next day, the samples were washed twice for 30 min with 1X phosphate buffered saline buffer (PBS) pH 8.0. The samples were then subjected to series of dehydration as indicated below:

30 % ethanol	15 min
50 % ethanol	30 min
70 % ethanol	30 min
80 % ethanol	1 hour
90 % ethanol	1 hour
100 % ethanol	1 hour twice

Following dehydration, the samples were subjected to pre-infiltration by incubation for 2 hours at room temperature in 100 % ethanol/ Technovit 7100 solution (1:1). The pre-infiltration solution was discarded, and then 1 ml of the solution A (infiltration solution) was added to the seeds followed by vigorous shaking for 1 min and incubation up to 24 hours with shaking in the solution A. Afterward, the solution A was discarded and 1 ml of embedding solution (solution B) was added to the samples. The mixture was poured into the embedding forms (Histoform S) and kept overnight at room temperature for polymerization. In order to section the samples, Histoforms S containing samples were attached to sample holders (Histobloc) using the Technovit 3040. Using a microtome, samples were sectioned (3-4 µm thickness) and stretched at room temperature in distilled water on microscope slides. The sections were afterward stained with 0.05% toluidine blue followed by 4-5 washes with water, and then dried for 10 min at 42°C and. The embryo structure and size were observed using the light microscope.

10X 0.1 M PBS buffer (1 l)

$Na2HPO_4$	14.4 g
KH_2PO_4	2.4 g
NaCl	80 g
KCl	2 g

After dissolving, adjust the pH to 8.0 using HCl and filled the final volume to 1 l with ddH_2O, the buffer is autoclaved or filter sterilized and stored at room temperature.

Fixative solution:

4 % paraformaldehyde, 0.2 % glutaraldehyde in 0.1 M PBS buffer pH 8.0. (It takes several hours to dissolve.)

Infiltration solution (solution A):

1 g of Hardener I is added to 100 ml Technovit 7100, the solution is mixed every well for 1 min and can be stored for 4 weeks at 4°C.

Embedding solution (solution B):

Solution A/Hardener II (15:1). The mixture is mixed very well for 1 min using 15 ml falcon tube.

2.8 Statistics

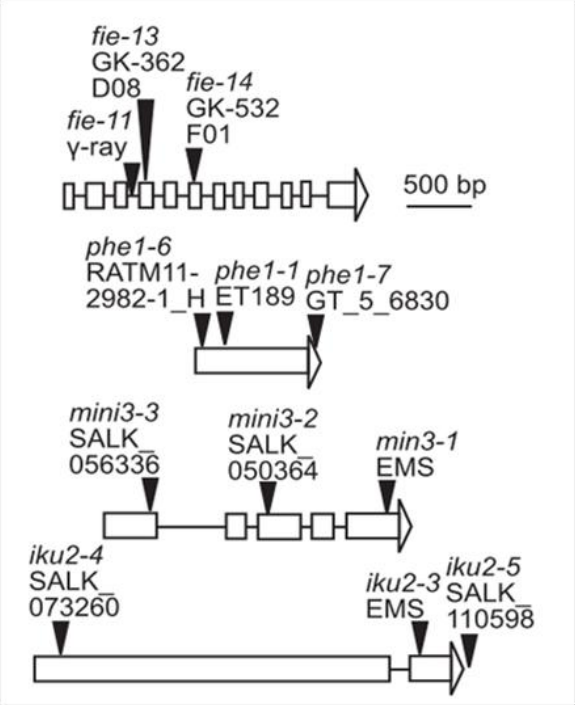
The data were statistically evaluated using the TTEST function of Microsoft Office Excel 2007. The settings tails = 2 and type = 3 (Two-sample unequal variance/heteroscedastic) were employed for the calculation of the p value.

3 Results

3.1 Characterization of Arabidopsis seed size candidate genes

3.1.1 Isolation of homozygous mutant lines

To investigate the role of Arabidopsis seed development genes in seed size determination and storage product deposition, four candidate genes were chosen for analysis. *FIE* (*fertilization independent endosperm*, At3g20740), *PHE1* (*pheres1*, At1g65330), *MINI3* (*miniseed3*, At1g55600), and *IKU*2 (*haiku2*, At3g19700). These candidate genes were previously isolated as key seed development genes (Ohad *et al.*, 1996; Garcia *et al.*, 2003; Köhler *et al.*, 2003a; Luo *et al.*, 2005). Figure 13 shows the gene maps and localization of mutations of each candidate gene used in this study. Homozygous *fie* seeds are embryo lethal because the *fie* mutant allele cannot be transduced through the female gamete. Thus, *fie* mutants can only segregate as heterozygous plants (*Fie*/*fie*) (Ohad *et al.*, 1996). Homozygous lines of *phe1*, *mini3-2*, *mini3-3*, *iku2-4* and *iku2-5* were isolated via PCR genotyping of heterozygous plants grown from seeds obtained from stock centers. Homozygous lines of *mini3-3*and *iku2-3* (EMS mutants) were obtained directly from Dr. W.J. Peacock, CSIRO, Australia. Expression of the candidate genes in homozygous mutant plants was analyzed via semi-quantitative rt-PCR of RNA of siliques obtained 5 days after flowering (Figure 14).



FIE, At3g20740

PHE1, At1g65330

MINI3, At1g55600

IKU2, At3g19700

Figure 13. Schematic representation of the genomic structure of the mutations of the candidate genes used in this study. Exons and introns are depicted as boxes and lines, respectively.

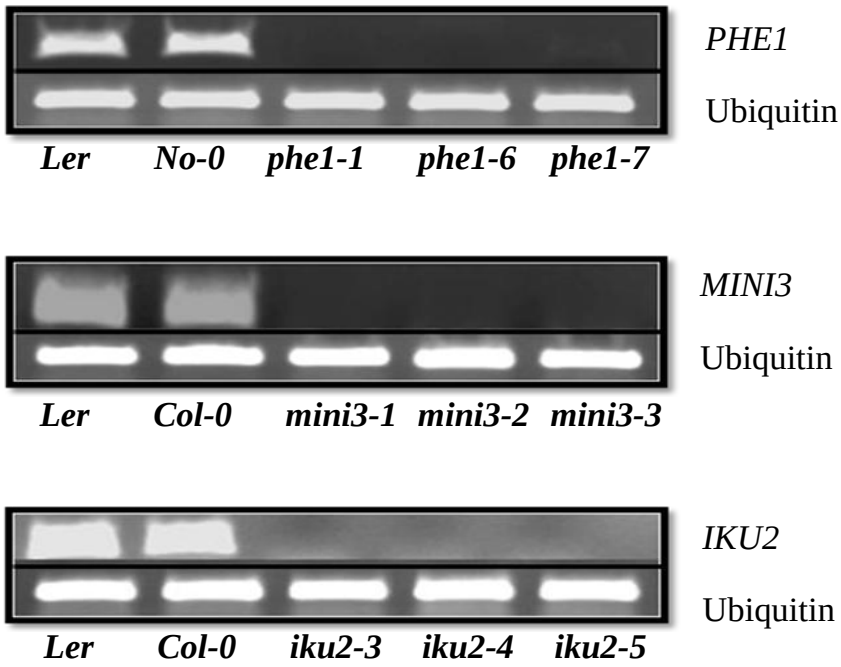


Figure 14. Expression of *PHE1, MINI3,* and *IKU2* in developing siliques of the different mutant lines. rt-PCR was performed from RNA isolated from siliques obtained 5 days after flowering. PCR amplifications of cDNA fragments of *PHE1, MINI3,* and *IKU2* were achieved using gene specific primers. PCR reactions with primers specific to ubiquitin were done as controls. The amplified DNA was separated in agarose gels and stained with ethidium bromide.

PHE1 expression was abolished in *phe1-1* and *phe1-6* lines compared to wild type; however there is a low expression of *PHE1* in the mutant line *phe1-7*. *MINI3* and *IKU2* were completely repressed in all lines compared to the corresponding controls.

To study the role of these candidate genes in seed size and seed storage reserve deposition, homozygous mutant lines, and the wild type plants were grown under the same conditions. The mature dry seeds of the plants were subjected to different analytical biochemical measurements.

3.1.2 Seed sizes, weight, and morphology

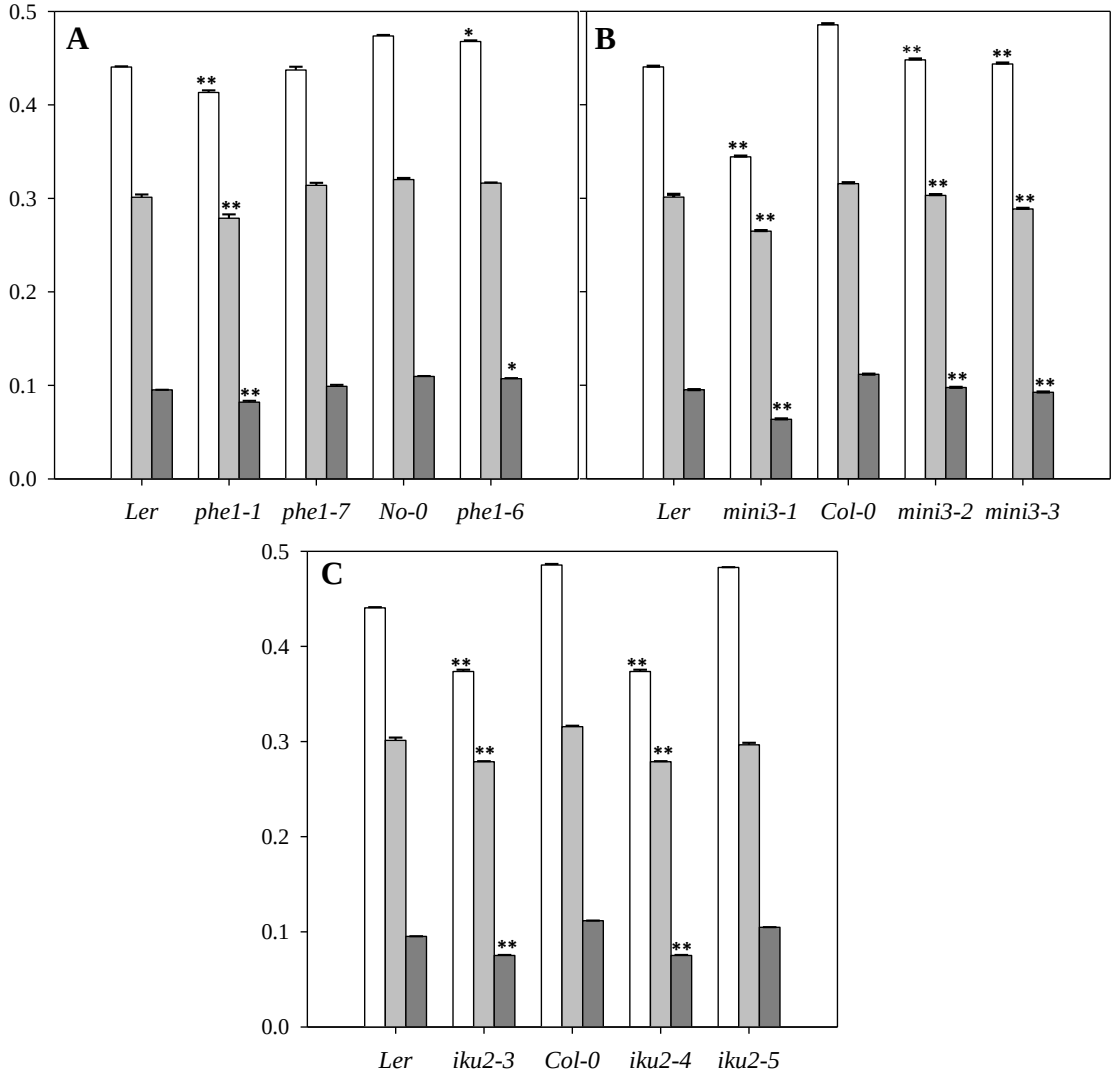


Figure 15. Seed size measurements of the mutant lines (A: *phe1*, B: *mini3*, and C: *iku2*) using the Evaluator software. The results for the mutant lines are depicted at the right of their corresponding WT controls. (□)Seed length (mm), (▭) width (mm), (■) area (mm^2). (Mean and SD, n ≈100). Asterisks indicate statistically significant differences compared with the corresponding WT control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The seed sizes of *phe1*-1 and *phe1-6* lines were slightly reduced compared to the wild type controls *Ler* and *Col-0* respectively, whereas *phe1-7* revealed a similar seed size compared to the corresponding WT (Figure 15). Homozygous mutant lines of *mini3-1*, *mini3-2*, and *mini3-3* showed a reduction in seed size compared to their wild type controls. *mini3-1* showed the strongest reduction in seed size, i.e. approximately 20%, 12%, and 33% reduction in length, width, and area respectively. The mutant lines *iku2-3* and *iku2-4* had a reduced seed size of approximately 15%, 7%, and 21% in length, width, and area respectively. The mutant line *iku2-5* displayed a similar seed size compared to the wild type control *Col-0*.

Mature seeds of the different mutant lines and their WT controls were dried under vacuum for 48 h before the determination of the seed weight (mean of 100 seeds using three independent biological replicates) (Figure 16).

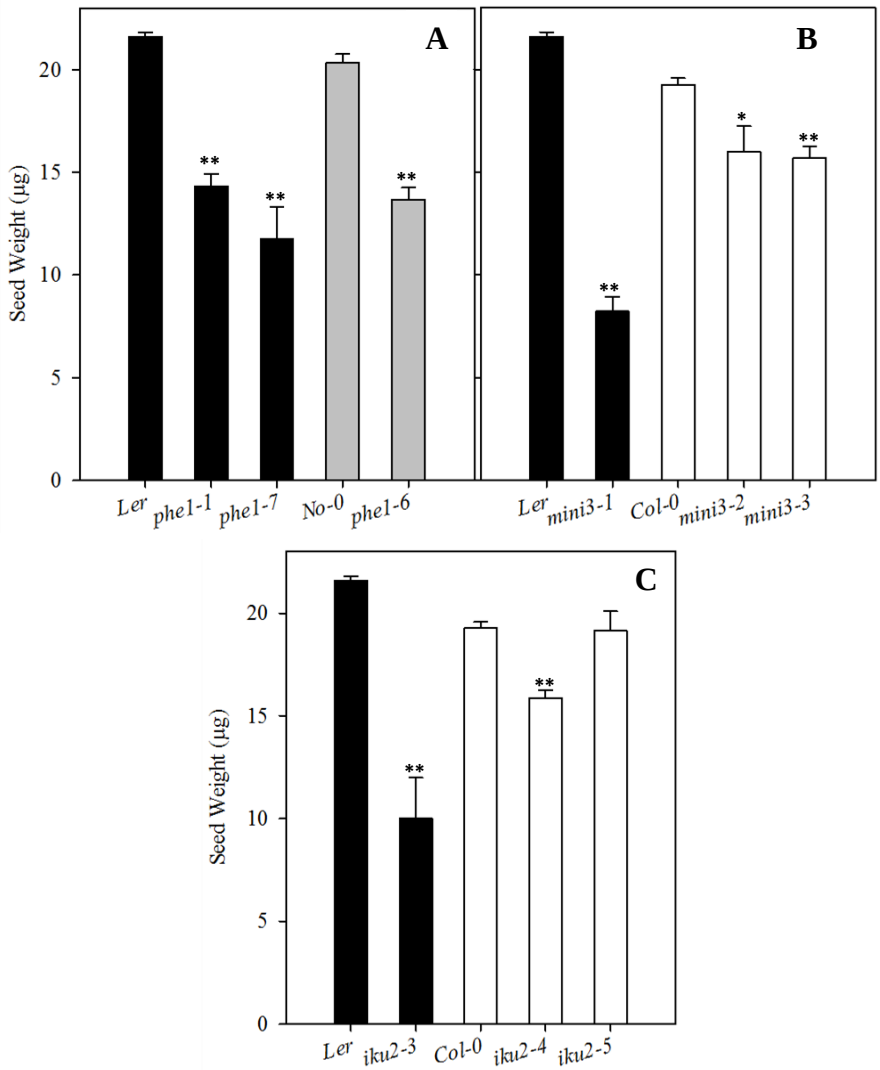


Figure 16. Seed weight of *phe1* (A), *mini3* (B), and *iku2* (C) lines. The mutant lines are depicted at the right of the corresponding WT controls. The bars show the mean and SD of three measurements of 100 seeds each. Asterisks indicate statistically significant differences compared with the corresponding WT control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The *phe1*, *mini3*, *iku2-3*, and *iku2-4* lines revealed a reduction of approximately 60% to 18% in seed weight compared to the corresponding WT controls. However, *iku2-5* showed no changes in seed weight in comparison with WT *Col-0*. The mutant lines *mini3*-1 and *iku2-3* showed the strongest reduction in seed weight, approximately 60% and 54% respectively.

Figure 17. Mature seeds of the different mutant lines and the WT controls. (Bar = 0.5 mm).

Figure 17 shows that plants heterozygous for *fie* produce 3 types of seeds, (i) aberrant seeds (on the left side of *fie* panels in Figure 17), (ii) wild type like seeds (right), and (iii) small seeds with shrivelled surface (middle) compared to the corresponding WT controls. *phe1*, *mini3*, *iku2-3* and *iku2-4* lines produce seeds of reduced size and shrivelled surface in comparison with the corresponding controls. *mini3-1* and *iku2-3* lines produce seeds with spherical shape in contrast with the oblong wild type seeds of *Ler*. Homozygous plants of *iku2-5* produce 2 types of seeds, (i) shrivelled seeds with reduced size, and (ii) wild type like seeds. Because the insertion of *iku2-5* is located 3' of the stop codon, the mutation most likely has a low influence of gene activity and this mutant was excluded from the analysis of sugars, proteins, and oil.

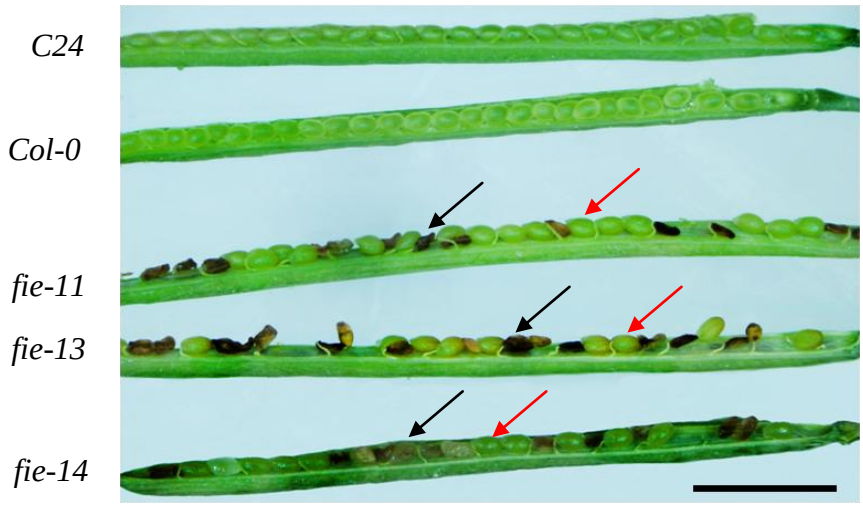


Figure 18. Developing siliques (12 DAF) of different heterozygous *fie* lines and WT controls. The aberrant seeds and normal developing seeds are indicated by black and red arrows, respectively. (Bar = 2 mm).

As previously reported (Ohad *et al.*, 1996), heterozygous *fie* plants produce 50 % aberrant seeds which carry the *fie* mutation in the female gametophyte. Developing siliques of *fie* lines revealed the presence of these aberrant seeds (black arrows), in addition to wild type like seeds (red arrows), whereas siliques of WT controls *Col-0* and *C24* showed only normal developing seeds (Figure 18).

3.1.3 Analysis of seed storage reserves

Developing seeds of Arabidopsis accumulate starch at the early stage of development. The starch is later degraded and it is absent in the mature dry seeds (Focks and Benning, 1998; Baud *et al.*, 2002). To analyse the starch content in the mature seeds of *phe1*, *mini3*, and *iku2* lines, the starch accumulating mutant *sex1* (Caspar *et al.*, 1991) was used as positive control (Figure 19). The presence of starch in the mature seeds is indicated by the change in colour from brown to black after staining with potassium iodide/iodine solution.

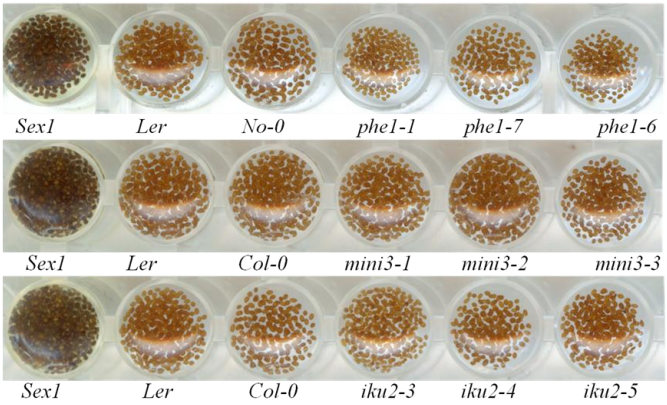


Figure 19. Analysis of starch content in the mature seeds of the different mutant lines for *phe1*, *mini3*, *iku2*, and *sex1* (control) after staining with potassium iodide/iodine solution.

In contrast to the mature seeds of *sex1* which accumulate a high amount of starch indicated by the black colour after staining with potassium iodide/iodine solution, no starch was detected in the seeds of the mutant lines *phe1*, *mini3*, and *iku2* at the mature stage. Thus, starch degradation was not impaired by the different mutations of *phe1*, *mini3*, and *iku2*.

Mature seeds of Arabidopsis wild type contain sugars, mainly in the form of sucrose, of approximately 0.5 µg/seed (Focks and Benning, 1998; Baud *et al.*, 2002). The sucrose content in *phe1*, *mini3*, *iku2* lines and the corresponding WT controls was measured enzymatically using a photometer based assay, using 3 biological replicates of 50 seeds (Figure 20).

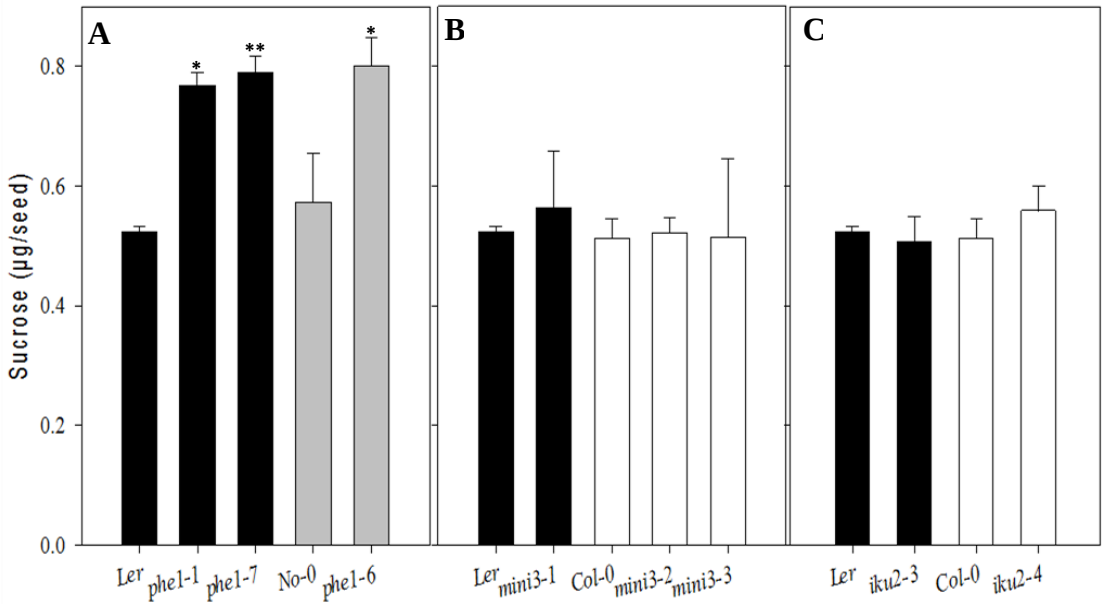


Figure 20. Sucrose content in the seeds of the mutant lines of *phe1* (A), *mini3* (B), *iku2* (C). n= 3. The mutant lines are depicted at the right of the corresponding WT controls. The bars show the mean and SD of three measurements of 50 seeds each. Asterisks indicate statistically significant differences compared with the corresponding WT control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The sucrose content in the seeds of the *phe1* lines was increased by approximately 20% compared to the WT controls, whereas *mini3* and *iku2* lines showed no significant changes in sucrose content in comparison with the corresponding controls.

Mature seeds of Arabidopsis wild type accumulate approximately 9 µg protein per seed (Focks and Benning, 1998; Baud *et al.*, 2002). Seed storage protein content in the different mutant lines and the corresponding controls was measured photometrically (Figure 21). The main seed storage proteins are 12S globulins (α and β subunits) and 2S albumins (large and small subunits). To investigate the effect of the different mutations on the composition of seed storage proteins, soluble proteins extracted from mature seeds of the mutant lines were subjected to SDS-PAGE and Coomassie blue staining (Figure 22).

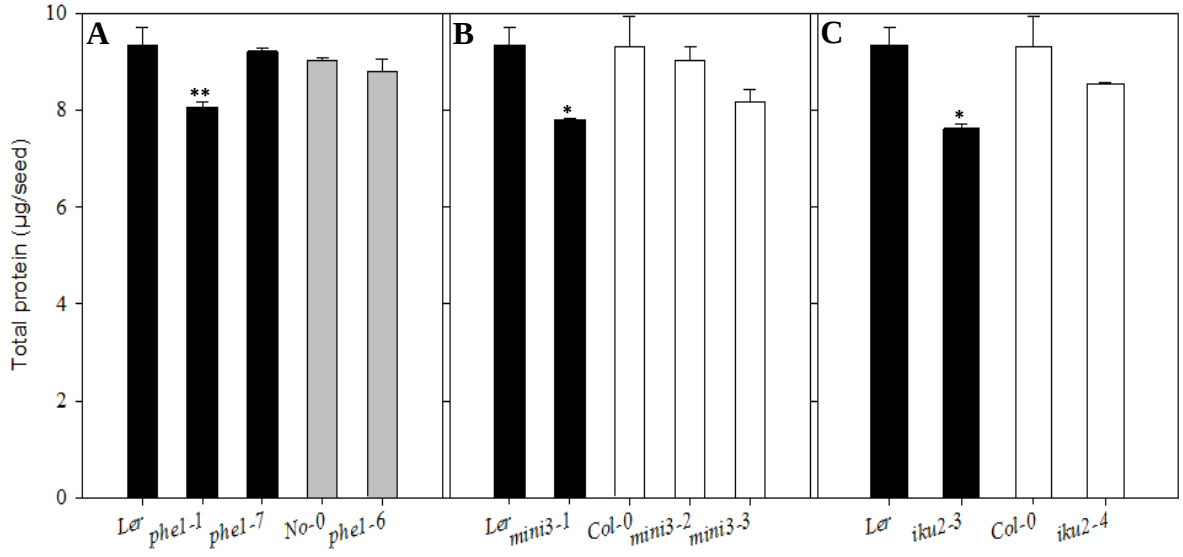


Figure 21. Total protein content in *phe1* (A), *mini3* (B), *iku2* (C) lines, and the corresponding controls. The bars show the mean and SD of three measurements of 20 seeds each. Asterisks indicate statistically significant differences compared with the corresponding WT control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The total protein content was slightly reduced in *mini3*, *iku2*, and *phe1-1* lines. Protein content in *mini3-1*, *mini3-3*, *iku2-3*, and *phe1-1* was reduced by approximately 1-2 µg compared to the control *Ler*, whereas the reduction in protein content in *mini3-4* and *iku2-4* was less than 1 µg. No change in protein content was detected in *phe1*-6 and *phe1-7* lines, compared to the WT controls.

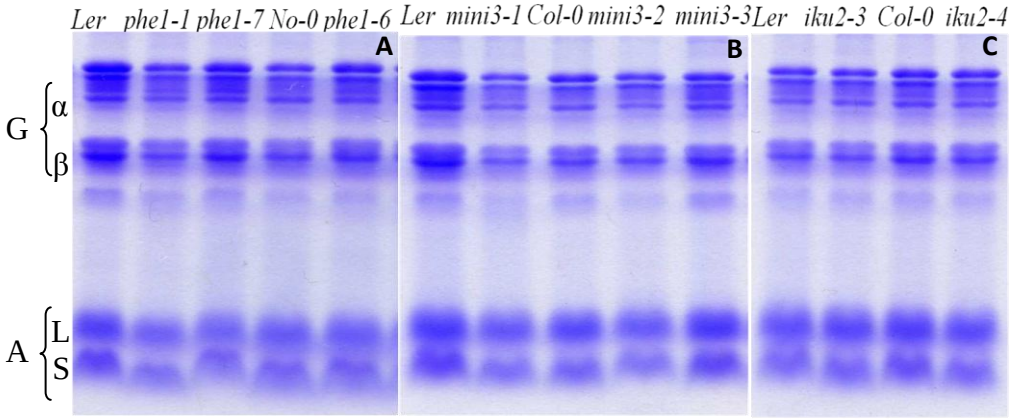


Figure 22. SDS-PAGE of proteins extracted from mature seeds of *phe1* (A), *mini3* (B), *iku2*(C) lines, and the corresponding controls. Each lane contains the total protein of 5 seeds. The gel was stained with Coomassie blue.

SDS-PAGE analysis of protein extract from Arabidopsis wild type seeds showed the presence of α and β subunits of 12S globulin in addition to large (L) and small (S) subunits of 2S albumin. The seeds of *phe1*, *mini3* and *iku2* lines showed no changes in storage protein composition in comparison with the corresponding WT controls.

The oil content in Arabidopsis seeds is variable between ecotypes, and it is highly affected by growth conditions (O'Neill *et al.*, 2003; Li *et al.*, 2006). The seed oil content and fatty acid composition of total lipids in *fie*, *phe1*, *mini3*, and *iku2* lines were determined by the analysis of fatty acid methyl esters (FAME) extracted from 5 seeds using gas chromatography. Five biological replicates were measured for each line (Figures 23 and 24). Because of the small size of the aberrant *fie* seeds, FAME analysis for *fie* was performed using 25 aberrant seeds.

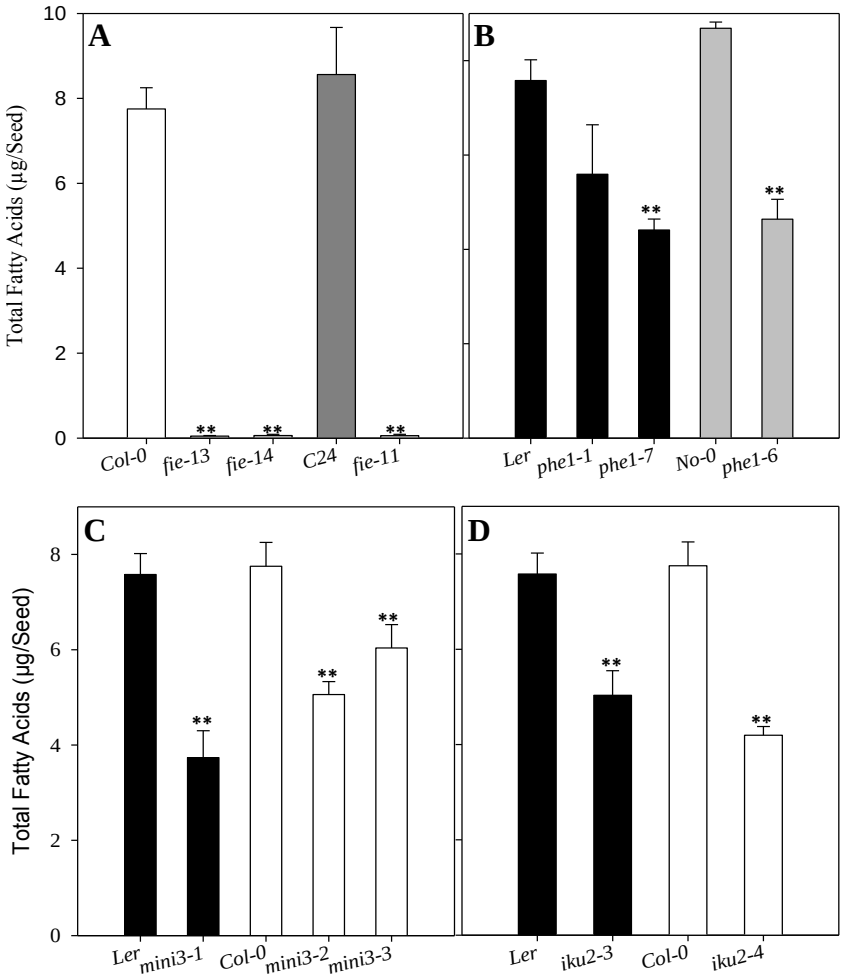


Figure 23. Total fatty acid content in *fie* (A), *phe1* (B), *mini3* (C), and *iku2* (D) lines. The bars show the mean and SD of five measurements of 5 seeds each. Asterisks indicate statistically significant differences compared with the corresponding WT control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The total fatty acid content of homozygous *fie* lines is extremely low, approximately 0.05 µg per seed, in comparison with ca. 8 µg in *Col-0* and *C24* seeds. In the seeds of the other mutants, the content of total fatty acids was reduced to various degrees, between approximately 15 and 50% in *phe1*, *mini3*, and *iku2* lines. The most drastic reduction in oil content was observed for seeds of *mini3-1* (ca. 50%), while *phe1-1* and *mini3-3* displayed the lowest reduction in storage oil (ca. 15%).

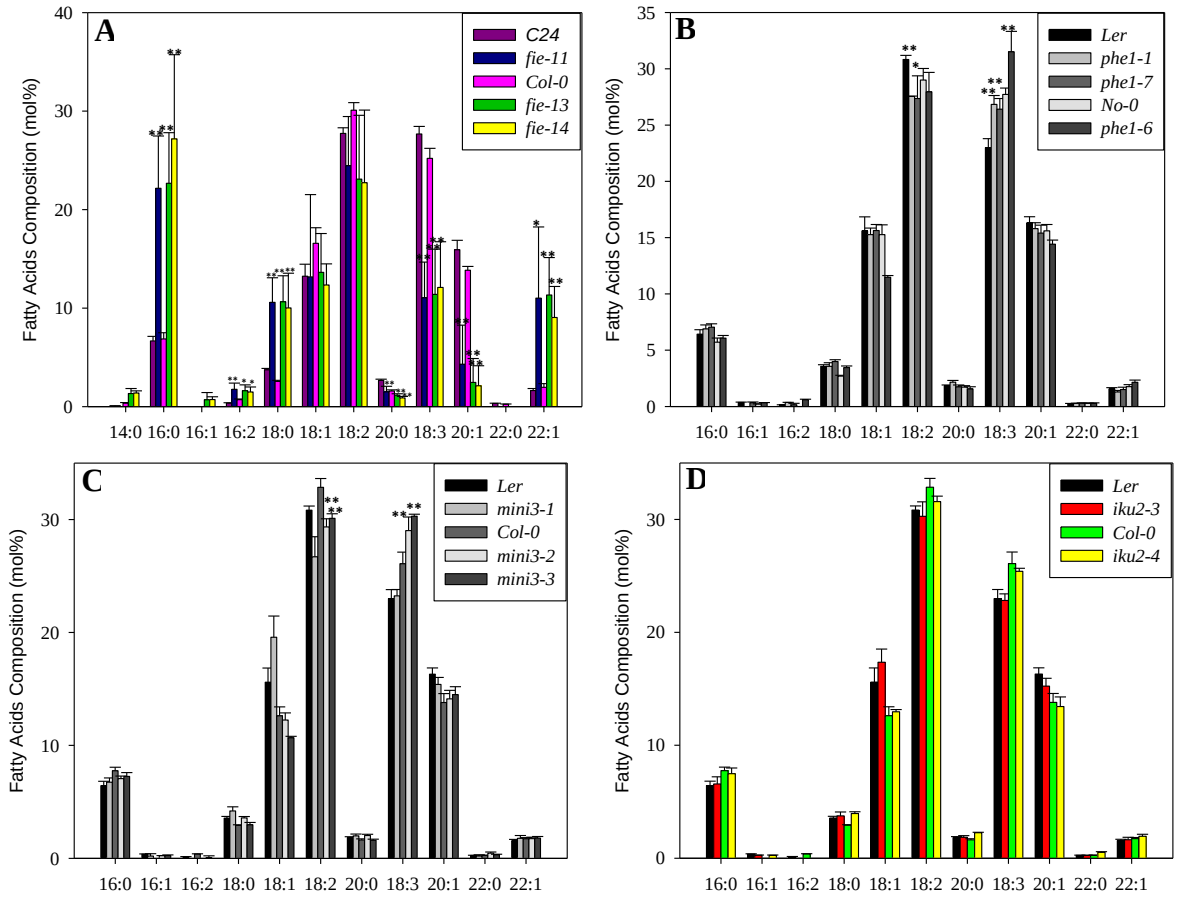


Figure 24. Total fatty acid composition of seeds of *fie* (A), *phe1* (B), *mini3* (C), and *iku2* (D) lines and the corresponding controls. The bars show the mean and SD of five measurements of 5 seeds each. Asterisks indicate statistically significant differences compared with the corresponding WT control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

Seeds of *fie* lines revealed a drastic increase in the content of the saturated fatty acids palmitic acid (16:0) and stearic acid (18:0), and a concomitant decrease in linoleic acid (18:2), α-linolenic acid (18:3), and arachidic acid (20:0) compared to the controls *C24* and *Col-0*. In addition, these lines revealed a strong increase in erucic acid (22:1) accompanied with a decrease in eicosenoic acid (20:1). By contrast, fatty acid composition in seeds of *phe1*, *mini3*, and *iku2* lines was similar to the corresponding controls. There was a tendency that the content of 18:2 was slightly reduced accompanied by an increase of 18:3 in *phe1* seeds, and also in *mini3-2* and *mini3-3*.

3.2 Generation of transgenic Arabidopsis seeds overexpressing seed size candidate genes

To investigate the effect of overexpression of *FIE*, *PHE1*, *MINI3* and *IKU2* on seed size and storage reserve deposition, the four candidate genes were overexpressed in Arabidopsis plants using the glycinin promoter from soybean which is embryo specific. The

overexpression constructs used in this study harbour a DsRed marker gene. Thus, transgenic seeds are easily selected for their red fluorescence under green light using a fluorescence microscope. For each construct, at least 15 independent transformed seeds were isolated. Three independent T1 lines were selected for each construct, and the expression of the transgenes *FIE*, *PHE1*, *MINI3* and *IKU2*; in comparison with an empty vector control, was measured via semiquantitative rt-PCR of RNA isolated from siliques harvested approximately 13 DAF (Figure 25).

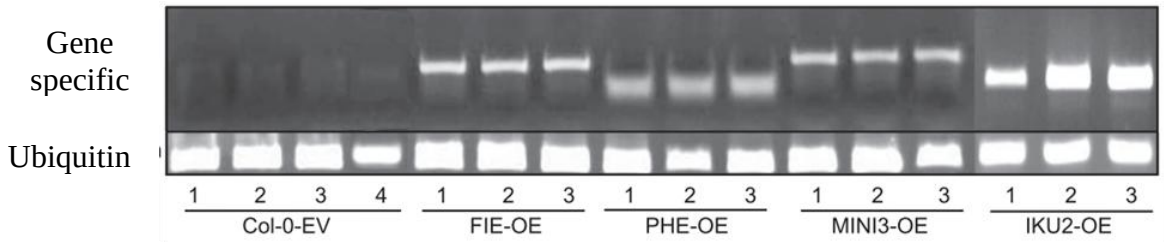


Figure 25. Expression of *FIE*, *PHE1*, *MINI3*, and *IKU2* in the overexpression lines compared to the empty vector controls. The gene expression was investigated in 3 independent lines for each construct. The empty vector control shows the expression of (1*) FIE*, (2) *PHE1*, (3) *MINI3*, and (4) *IKU2*. Semi-quantitative rt-PCR was performed with RNA isolated from siliques obtained 13 days after flowering. PCR amplifications of cDNA fragments of *FIE1*, *MINI3*, and *IKU2* were achieved using gene specific primers. PCR reactions with primers specific to ubiquitin were used as controls. The amplified DNA was separated in agarose gels and stained with ethidium bromide.

The expression level of *FIE*, *PHE1*, *MINI3*, and *IKU2* in the siliques of the three independent lines for each construct was strongly elevated. However, the expression level of the four genes was below detection limit in the siliques of the empty vector control.

3.3. Characterization of the seeds of transgenic Arabidopsis overexpression lines

Transgenic T2 seeds with red fluorescence harvested from three independent T1 lines of each construct were dried and subjected to analytical biochemical measurements as previously done for the mutant lines of the candidate genes. The T2 seeds showed a 3:1 segregation of fluorescent to non-fluorescent seeds (wild type) indicating a heterozygous integration of the T-DNA (Figure 26).

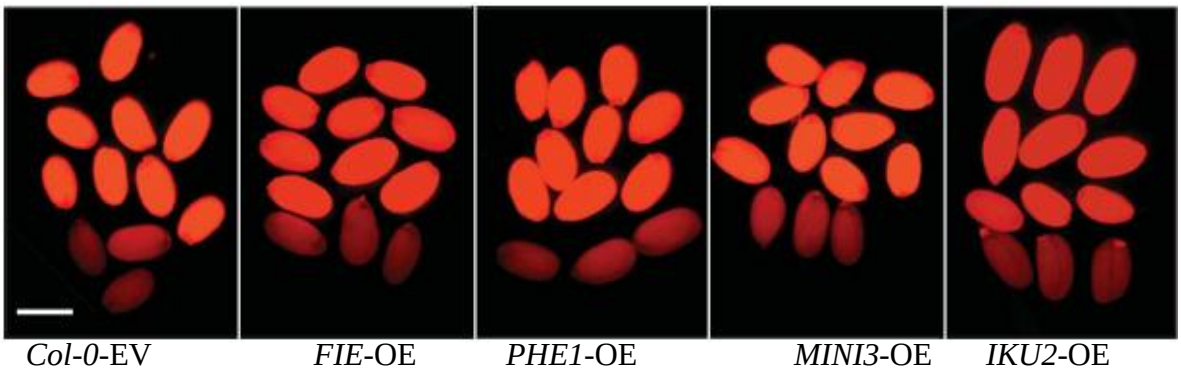


Figure 26. Red fluorescence of T2 seeds from one representative line for each overexpression construct. Transgenic seed are fluorescent under green light (top), while wild type segregant seeds showed only weak fluorescence (bottom). Bar = 5 µm.

The transgenic T2 seeds of the overexpression lines *FIE*-OE, *PHE1*-OE, and *MINI3*-OE, revealed a similar seed size compared to their non-transgenic segregant seeds and the empty vector control *Col-0*-EV. However, *IKU2* overexpression lines displayed two types of transgenic seeds, large seeds (50%), and seeds with similar size to *Col-0*-EV (25%). Interestingly, non-transgenic *IKU2*-OE segregant seeds also showed an increase in size.

3.3.1 Determination of seed size and seed weight

The seed sizes of the transgenic lines were determined using the Evaluator software. Approximately 100 transgenic (fluorescent) mature seeds of three independent T2 lines were used for each construct (Figure 27).

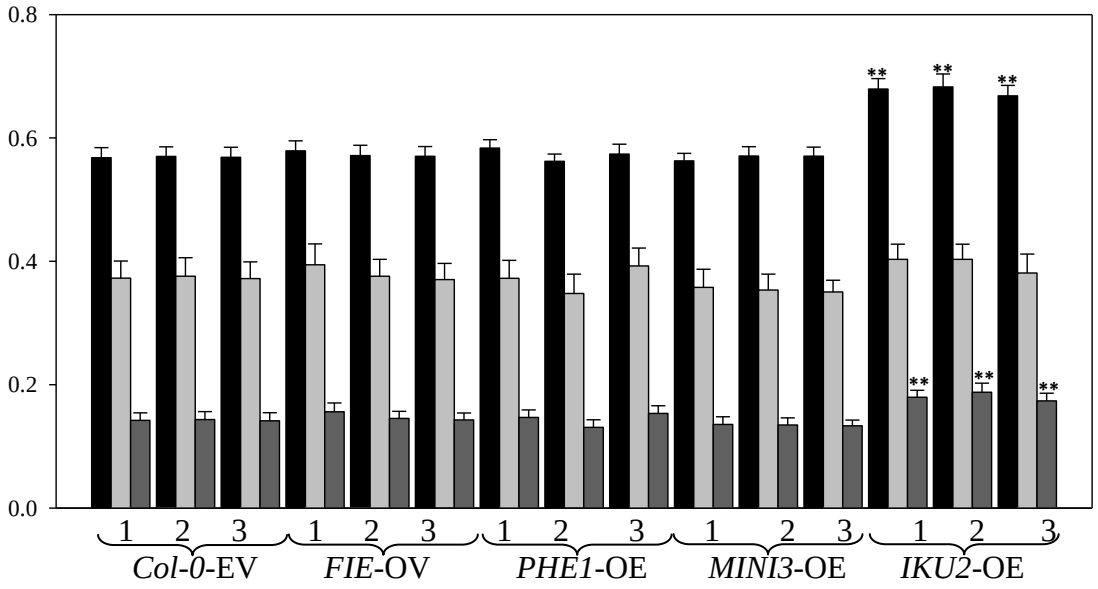


Figure 27. Seed sizes of transgenic T2 seeds of three independent overexpression lines of *FIE, PHE1, MINI3,* and *IKU2* compared to the empty vector control Col-0-EV.
〔□〕Seed length (mm), 〔■〕seed width (mm), 〔▒〕seed area (mm^2). (Mean, SD, n = 100). Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The overexpression lines *FIE*-OE, *PHE1*-OE, and *MINI3*-OE revealed a similar seed size in comparison with the empty vector *Col-0*-OE. By contrast, *IKU2* overexpression lines revealed an increase in seed sizes. *IKU2*-OE seeds displayed an increase of approximately 20% and 26% in seed length and area, respectively. Moreover, these lines showed a slight increase in seed width, ca. 5% compared to *Col-0*-EV, but this was not significant.

The measurement of seed weight of the transgenic seeds was done by determining the mean weight of 100 seeds using three independent biological replicates for each line (Figure 28).

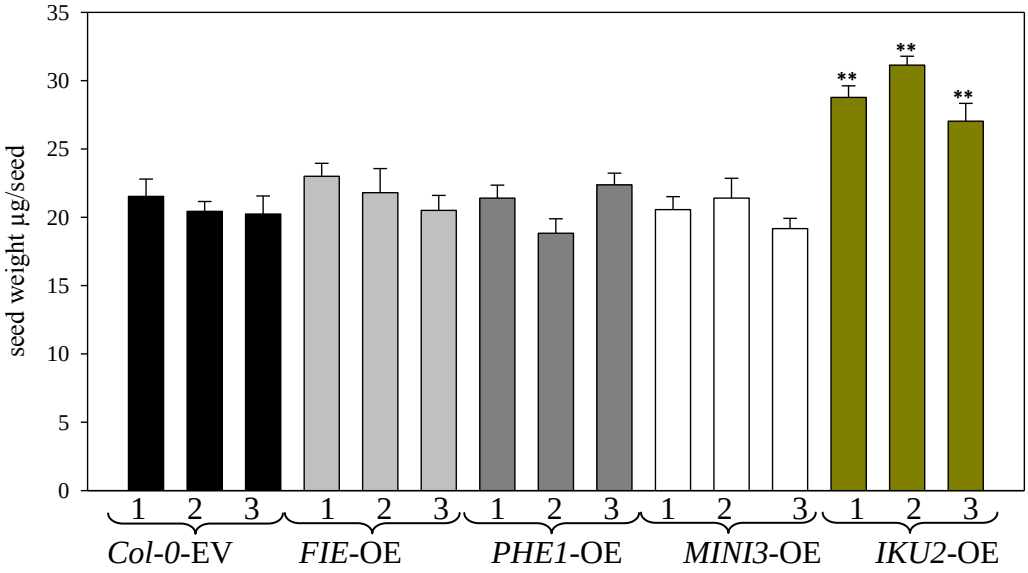


Figure 28. Seed weight of transgenic T2 seeds of the overexpression lines *FIE*, *PHE1*, *MINI3*, and *IKU2* compared to the empty vector control *Col-0*-EV. Seeds of three independent T1 lines were used for each construct. The bars show the mean and SD of three measurements for 100 seeds each. Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The seeds of the transgenic lines *FIE-OE*, *PHE1*-OE, and *MINI3*-OE showed no changes in weight compared to the empty vector control *Col-0*-EV. However the seed weight of the transgenic *IKU2*-OE lines was markedly increased by approximately 30%.

3.3.2 Measurements of seed storage products

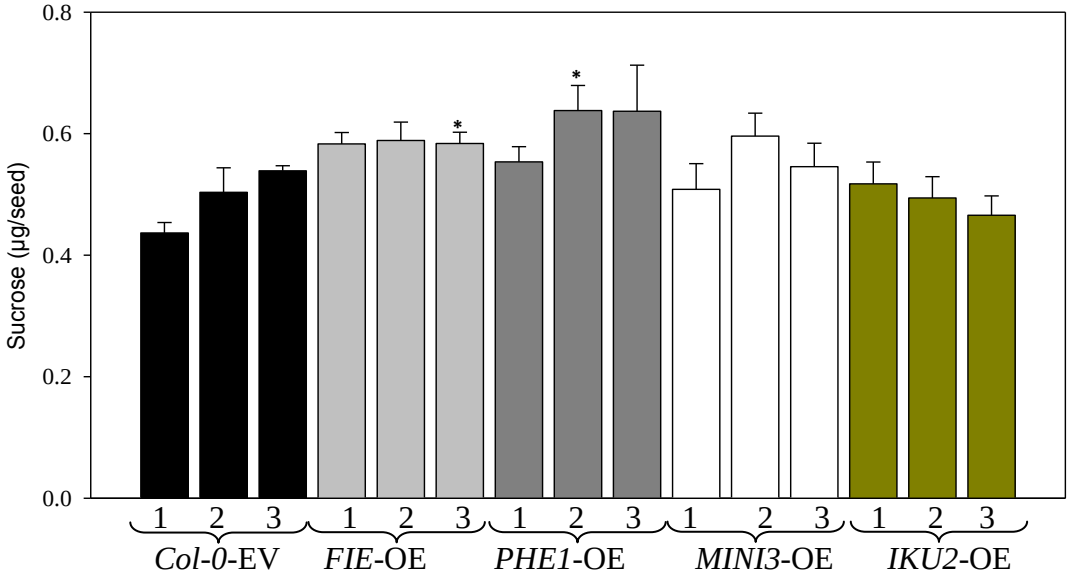


Figure 29. Sucrose content in three independent overexpression lines *FIE*, *PHE1*, *MINI3*, and *IKU2* in comparison with the empty vector control *Col-0*-EV. The data were obtained by measuring three replicas of 5 seeds each. Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The amount of sucrose in the seeds of the overexpresison lines was similar to *Col-0*-EV (Figure 29). The sucrose content was slightly increased in the seeds of the overexpression lines *FIE* and *PHE*. Seeds of *MINI3*-OE and *IKU2*-OE lines revealed similar sucrose content compared to the control *Col-0*-EV.

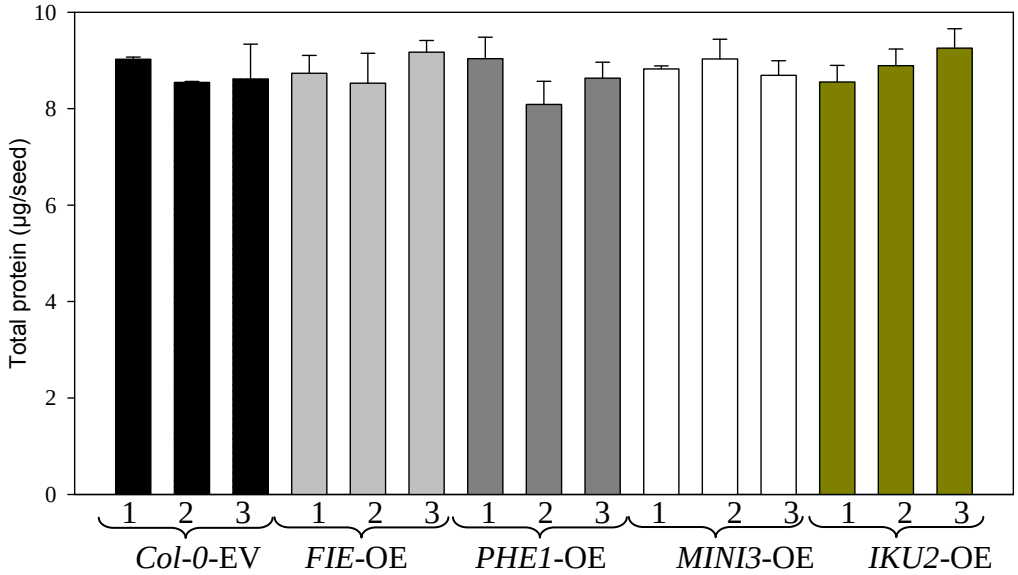


Figure 30. Protein content in the overexpression lines *FIE*, *PHE1*, *MINI3*, and *IKU2* compared to the empty vector control *Col-0*-EV. Three lines of each construct were used for measurements; the bars show the mean and SD of three measurements of 20 seeds each. Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

No significant change was observed in the protein content in the seeds of the overexpression lines *FIE*, *PHE1*, *MINI3*, and *IKU2* compared to the empty vector control (Figure 30). The amount of total seed protein in all lines was ca. 9µg/seed. The analysis of storage proteins of the transgenic lines using SDS-PAGE gel revealed no changes in protein composition in comparison with the empty vector control *Col-0*-EV (Figure 31).

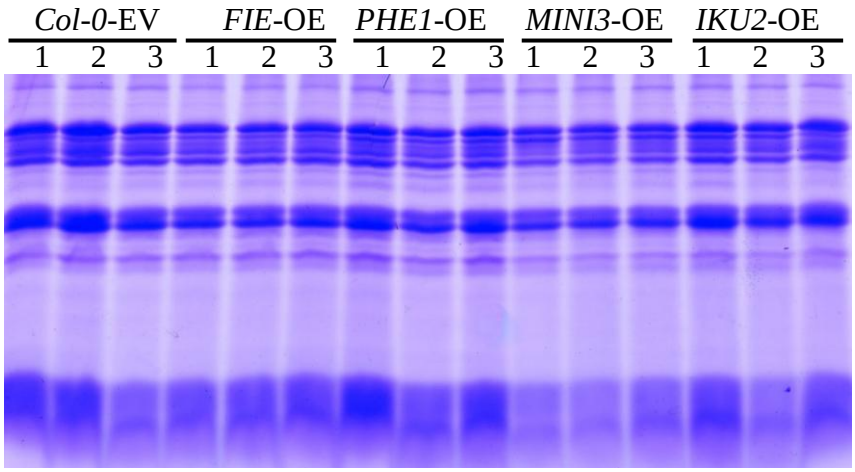


Figure 31. SDS-PAGE of storage proteins extracted from 5 mature seeds of the empty vector control *Col-0*-EV and the overexpression lines *FIE*, *PHE1*, *MINI3* and *IKU2*. Total protein isolated from 5 seeds each was loaded per lane. Three independent lines of each construct were used. The gel was stained with Coomassie blue.

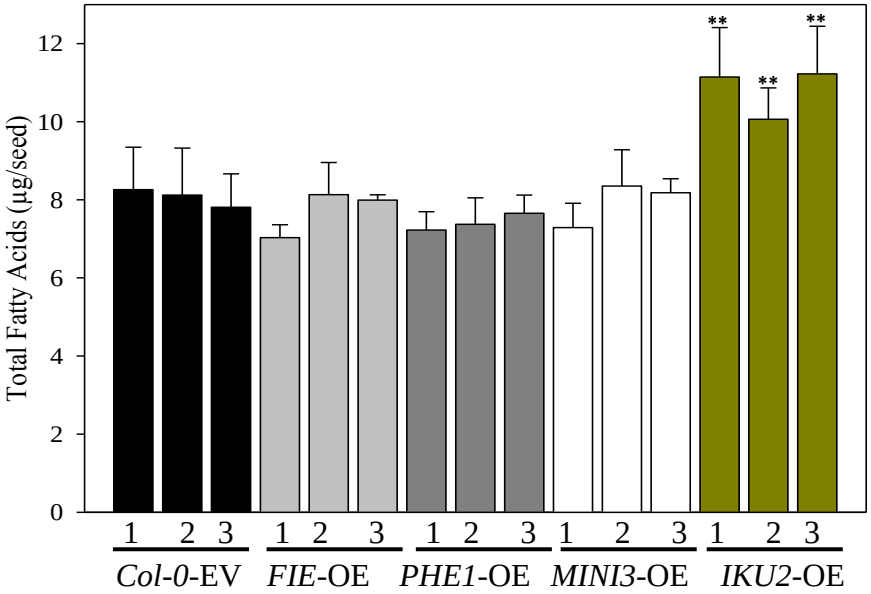


Figure 32. Total fatty acid content in the seeds of the overexpression lines *FIE*, *PHE1*, *MINI3*, and *IKU2* compared to the empty vector control *Col-0*-EV. Three lines of each overexpression experiment were used for measurements. The bars show the mean and SD of five measurements of 5 seeds each. Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

The total FA content in the overexpression lines *FIE, PHE1, MINI3* was similar compared to *Col-0*-EV, with ca. 8 µg/seed. However, the FA content was increased by approximately 35% in *IKU2* overexpression lines (Figure 32).

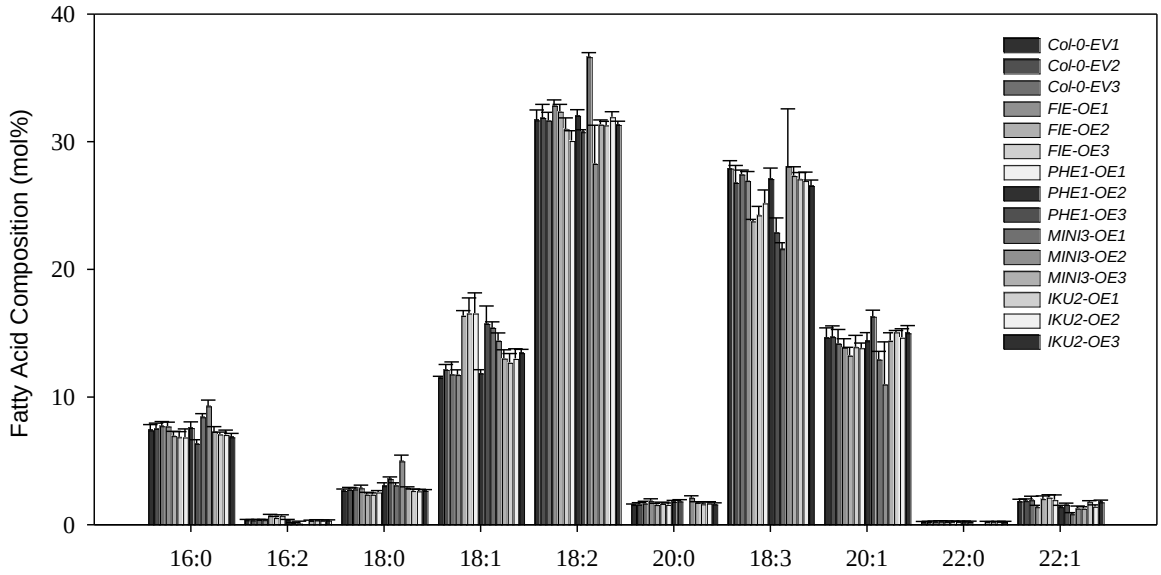


Figure 33. Fatty acid composition of the seeds of the overexpression lines *FIE, PHE1, MINI3, IKU2,* and the empty vector control *Col-0*-EV. Three lines of each overexpression experiment were used for the measurements. The bars show the mean and SD of five measurements of 5 seeds each.

The analysis of the total fatty acid composition of the four overexpression lines revealed no difference compared to the empty vector control (Figure 33).

3.4 Characterization of *IKU2*-OE plants

As shown in Figure 26, the T1 lines of *IKU2*-OE produce two types of T2 transgenic (fluorescent) seeds, (i) large seeds (in a ratio of about 2/3) and (ii) small seeds (in a ratio of about 1/3) with a size comparable to the control *Col-0*-EV. The segregation ratio of the transgenic seeds of 2:1 of large seeds to normal seeds suggests that the large seeds could represent *IKU2-OE* heterozygous seeds, while the normal-size seeds represent the transgenic homozygous seeds. Further investigation of T2 lines of *IKU2*-OE showed that the siliques of heterozygous *IKU2*-OE plants are thicker compared to homozygous *IKU2*-OE plants, and the controls *Col-0*-EV or *Col*-0 (Figure 34).

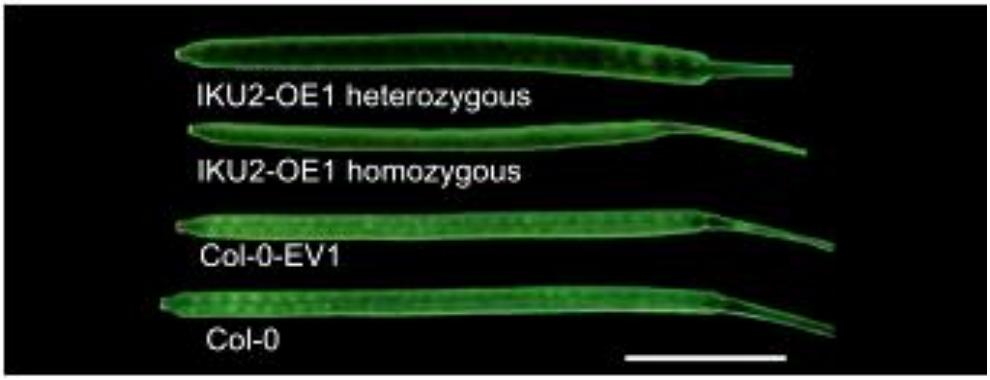
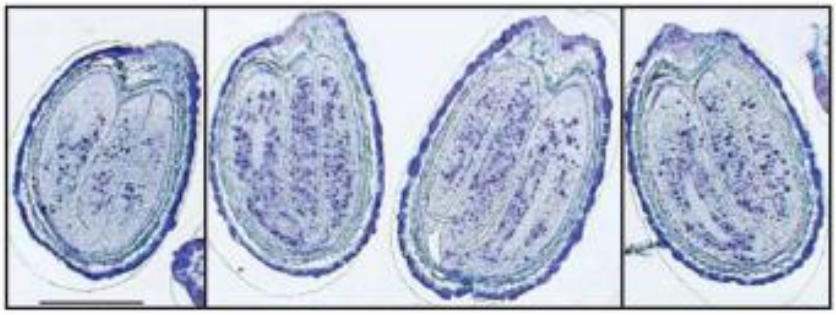


Figure 34. Developing siliques of heterozygous and homozygous *IKU2*-OE1 plants, compared to the controls *Col-0*-EV and *Col-0*. One representative silique of each line is shown (ca. 13 DAF). Bar = 0.5 mm

The developing seeds of the late cotyledon stage (ca. 13 DAF) from the siliques of homozygous and heterozygous *IKU2*-OE1 plants and the control *Col-0*-EV were microscopically analyzed after thin-sectioning and toluidine blue staining (Figure 35).



Col-0-EV *IKU2*-OE1 (heterozygous) *IKU2*-OE1 (homozygous)

Figure 35. Developing seeds of homozygous and heterozygous *IKU2*-OE1 plants and the empty vector control *Col-0*-EV (ca 13 DAF). The seeds in the siliques were analyzed by light microscopy after thin-sectioning and toluidine blue staining by light microscopy. Representative seeds are shown. Bar = 0.2 mm.

The developing seeds of homozygous *IKU2*-OE1 plants showed a size comparable to *Col-0*-EV, whereas heterozygous *IKU2*-OE1 plants produce seeds with different embryo sizes, large embryos (right site in the *IKU2*-OE1 heterozygous panel), and small embryos (left site in the *IKU2*-OE1 heterozygous panel) with a size similar to *Col-0*-EV and homozygous *IKU2*-OE1.

Homozygous *IKU2*-OE1 and *Col-0*-EV plants have seeds with similar sizes, whereas heterozygous *IKU2*-OE1 plants produce large and small transgenic seeds with a ratio of 2:1. Moreover, the non-transgenic segregant seeds of the heterozygous plants are increased in size (Figure 36A). To confirm these results, seeds size measurements of mature dry seeds from

homozygous and heterozygous *IKU2*-OE1 plants were done and results compared to *Col-0*-EV (Figure 36B).

Seed size measurements revealed that mature homozygous *IKU2*-OE1 seeds displayed no changes in length, width and area compared to *Col-0*-EV, whereas heterozygous transgenic and non-transgenic seeds of *IKU2*-OE1 revealed an increase of approximately 18%, 6%, and 28% in length, width, and area, respectively.

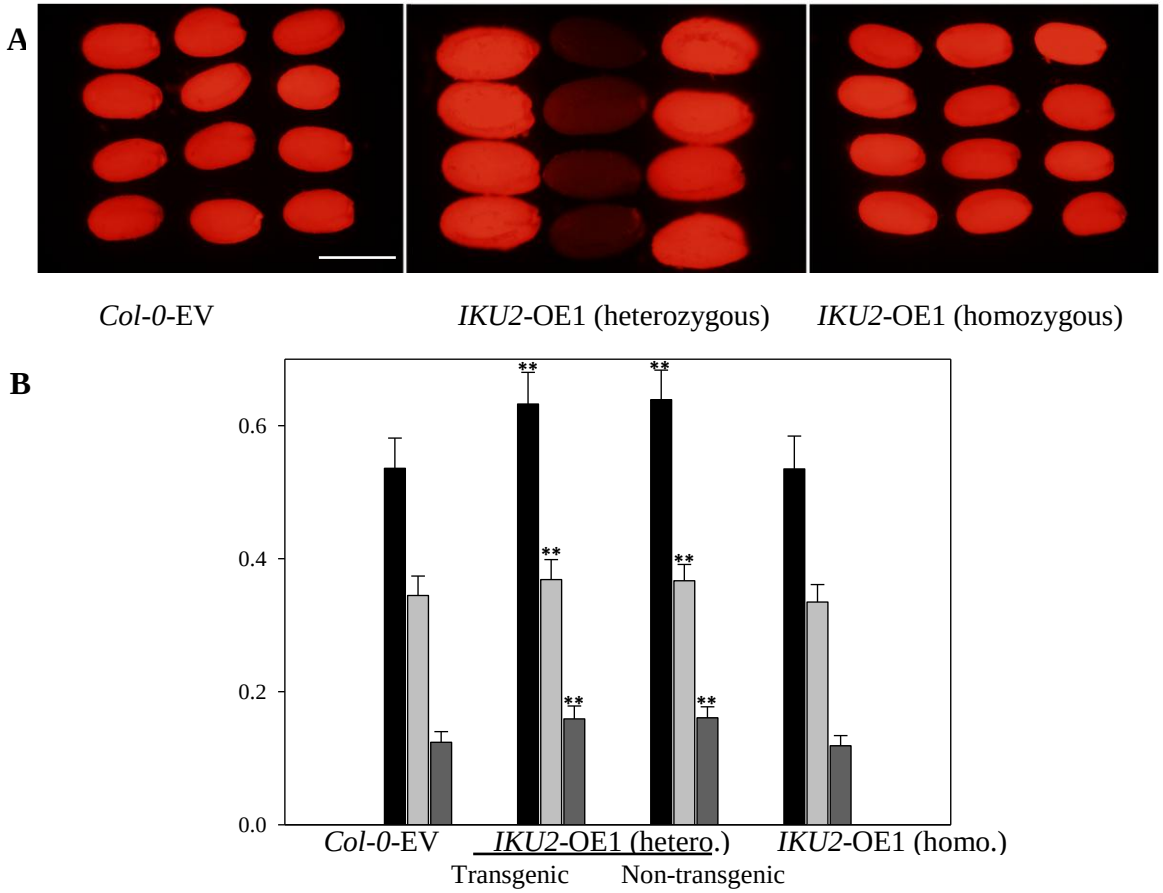


Figure 36. Analysis of seed size of homozygous and heterozygous *IKU2*-OE1.
A: Fluorescence derived from the DsRed marker protein in mature seeds observed under the fluorescence microscope. Seeds in the *IKU2*-OE1 panel are organized in three groups, i.e. normal-size transgenic (left), non-transgenic segregants (center), and large transgenic (right).Bar= 0.5 mm**. B:** seed size measurement using the Evaluator software. 〔■〕Length (mm), 〔▭〕width (mm), 〔▬〕area (mm^2). (Mean, SD, n=100). Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

To investigate whether the increased seed size of heterozygous *IKU2*-OE1 plants, which is suppressed in homozygous plants, can be de-repressed in the F1 generation, a reciprocal cross of a homozygous *IKU2*-OE1 plant with WT *Col-0* was performed.

Visual inspection of F1 seeds from the reciprocal crosses of a homozygous *IKU2*-OE1 plant with WT *Col-0* showed an increase in size of the transgenic seeds compared to seeds of wild type control *Col-0,* regardless of the direction of the cross (Figure 37A). This result was

confirmed by seed size measurements (Figure 37B). Heterozygous F1 seeds of *IKU2*-OE derived from the reciprocal crosses *Col-0* x *IKU2* and *IKU2* x *Col-0* were increased in length, width and area compared to *Col*-0-EV. Thus, the increase in seed size was only observed in heterozygous *IKU2*-OE seeds, but not in homozygous *IKU2*-OE seeds.

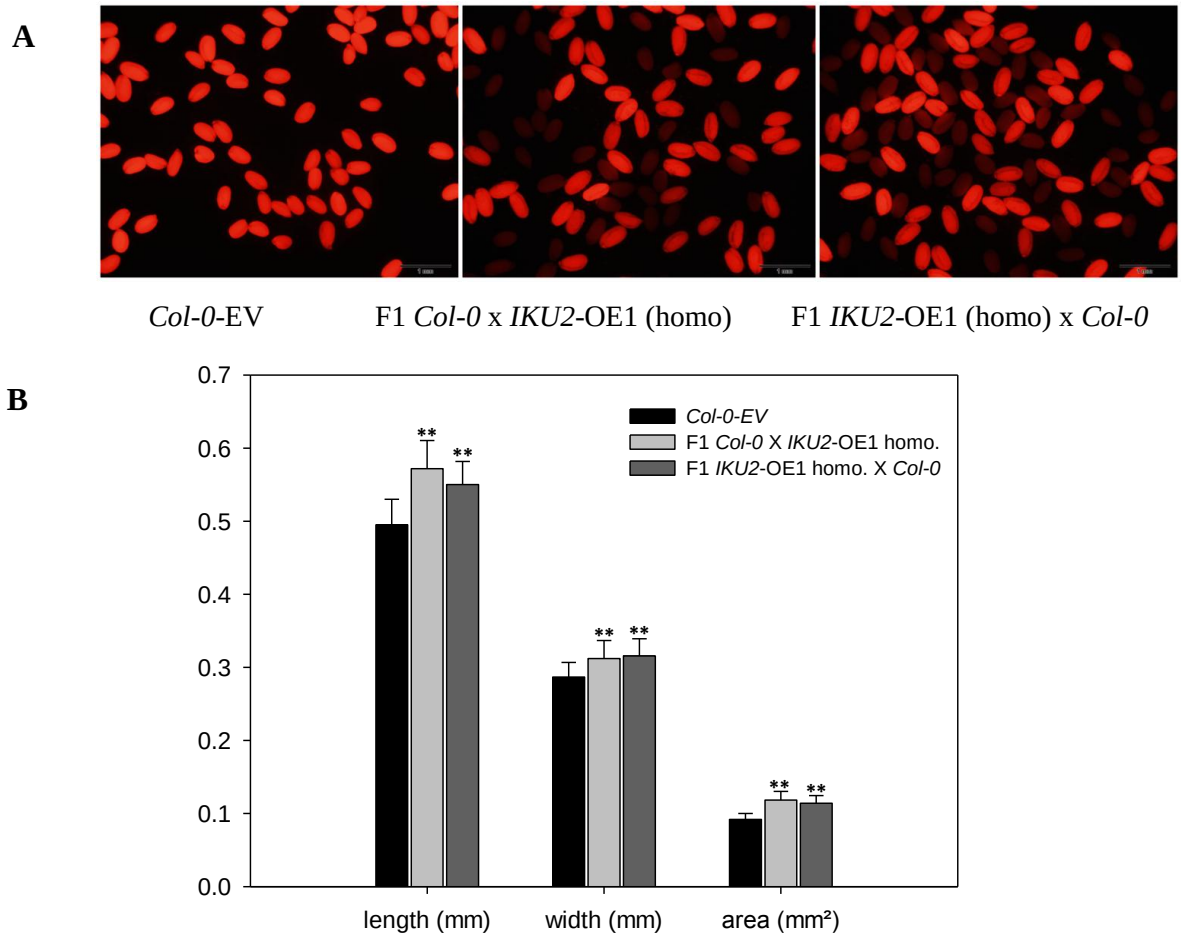


Figure 37. Analysis of F1 generation seeds from reciprocal cross of homozygous *IKU2*-OE1 with WT *Col-0.* A*:* Red fluorescence of the seeds under the fluorescent microscope, Bar= 1 mm. **B:** Seed sizes measurement using the Evaluator software. 〔■〕Length (mm), 〔□〕width (mm), 〔▒〕area (mm^2). (Mean, SD, n=100). Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

To investigate the molecular origin for the suppression of the increased seed size in homozygous *IKU2*-OE lines, *IKU2* expression was measured via semi-quantitative rt-PCR of RNA isolated from 13 days old siliques of *Col-0*-EV plants, heterozygous *IKU2*-OE1 plants and homozygous *IKU2*-OE1 plants (Figure 38).

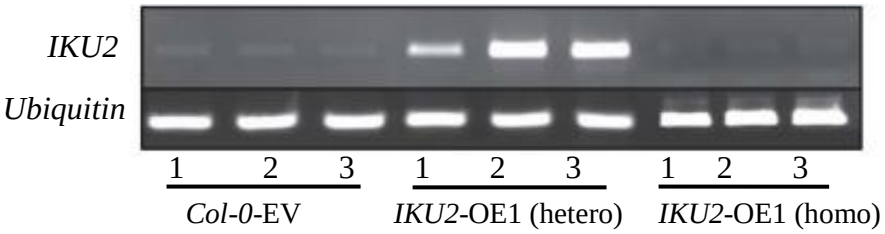


Figure38. *IKU2* expression in developing siliques (13 DAF) of *Col-0*-EV, heterozygous *IKU2*-OE and homozygous *IKU2*-OE. Three independent lines were analysed for each construct. Semi-quantitative rt-PCR was performed with RNA isolated from siliques obtained 13 days after flowering. PCR amplifications of cDNA fragments of *IKU2* were achieved using gene specific primers. PCR reactions with primers specific to ubiquitin were used as controls. The amplified DNA was separated in agarose gels and stained with ethidium bromide.

IKU2 expression is low in the developing seeds of homozygous *IKU2*-OE1 lines, similar to the empty vector control *Col-0*-EV. However, *IKU2* expression is strongly increased in the siliques of heterozygous *IKU2*-OE1 lines.

To analyse the consequences of the increased seed size and oil content of the heterozygous *IKU*-OE lines on seed yield, the silique length, number of seeds per silique, and the total amount of seeds per plant were determined for three independent heterozygous *IKU2*-OE lines and compared to the empty vector control (Figure 39).

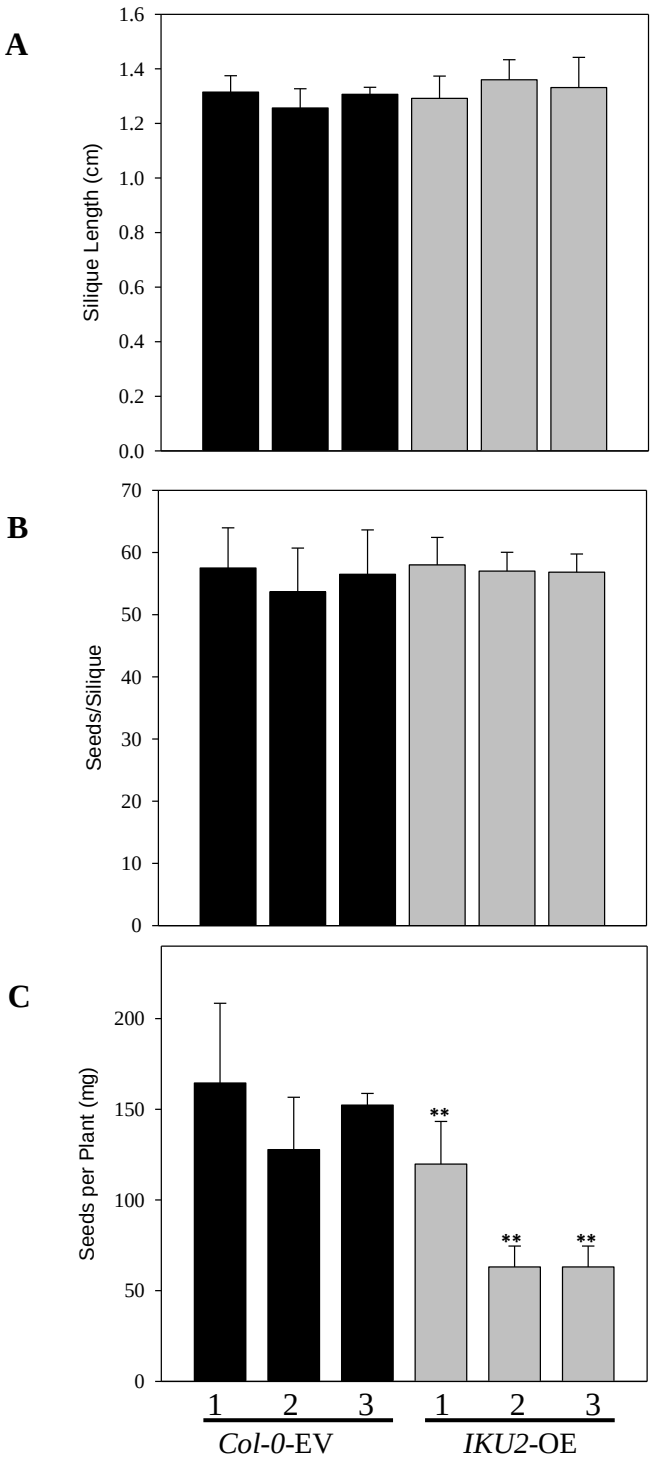


Figure 39. **Silique length (A), number of seeds per silique (B), and seed yield per plant (C) of three independent heterozygous *IKU2*-OE lines and the control *Col-0*-EV.** The bars show the mean and SD after measuring 10 replicas. Asterisks indicate statistically significant differences compared with the control (Student's t test: * $P < 0.05$; ** $P < 0.01$).

Heterozygous *IKU2*-OE lines showed no alteration in the silique length and number of seeds per silique compared to the control *Col-0*-EV. However, the seed yield per plant was strongly reduced by approximately 25-50% in *IKU2*-OE lines compared to *Col-0*-EV.

4 Discussion

Plant seeds are an important component of the world's diet. For example, cereal grains represent ca. 90 % of all cultivated seeds and contribute up to half of the global nutritional energy intake. Thus, seed biology is one of the most important research areas in plant biology (Bewley, 1997). Seeds are chimeras developed from fertilised ovules. In flowering plants, the double fertilisation of the embryo sac by sperm nuclei initiates the development of zygotic tissues consisting of the diploid embryo and the triploid endosperm. These tissues are protected by the maternally derived seed coat, which comprises several cell layers (Baud and Lepiniec, 2010). The seeds contain the genetic information of the new plant generation. Moreover, they accumulate storage reserves at the maturation stage; mainly in the embryos, to support the growing seedling until it gains it function as an autotrophic organism via photosynthesis. The main storage compounds in Arabidopsis seeds are oil in the form of TAGs and proteins (Bewley, 1997; Baud *et al.*, 2002; Baud and Lepiniec, 2010). Several studies have shown that alterations in embryo development have significant effects on the final amount of seed storage reserves. On the other hand, seed storage reserve deposition and metabolism are under the control of hormonal and developmental signals from the embryo (Brocard-Gifford *et al.*, 2003; Hill *et al.*, 2003).

4.1 Characterization of mutant alleles for *fie*, *phe1*, *mini3* and *iku2*

To investigate the effects of alterations in the expression of Arabidopsis seed development genes on seed size and storage reserve accumulation, mutants of the four candidate genes *FIE*, *PHE1*, *MINI3* and *IKU2* were submitted to analytical biochemical analyses. The mutant alleles *fie-11*, *phe1-1*, *mini3-1*, *mini3-2*, and *iku2-3* were previously isolated (Guitton *et al.*, 2004; Köhler *et al.*, 2005; Luo *et al.*, 2005). In this study, new mutations in the four candidate genes were identified: *fie-13*, *fie-14*, *phe1-6*, *phe1-7*, *mini3-3*, and *iku2-4*. As previously reported (Ohad *et al.*, 1996), heterozygous *fie* plants produce 50 % aberrant seeds which carry the *fie* mutation in the female gametophyte. In addition, these plants produce wild type like seeds, and small seeds with shrivelled surface (Figures 17 and 18). Köhler and co-authors (Köhler *et al.*, 2005) have previously isolated the *phe1-1* mutant, reporting that no defects in vegetative or reproductive development were observed. However, in their study, no data on *phe1-1* seed size (length, width) or weight were presented. Figure 15A shows that homozygous *phe1* mutants (*phe1-1*, *phe1-6*, and *phe1-7*) produce seeds with length, width, and area similar to the WT controls *Ler* and *No-0*. However, seeds produced by homozygous *phe1* mutants have a reduced weight of ca. 32-45% as compared to WT (Figure

16A). This apparent contradiction can be explained by the fact that *phe1* seeds have a shrivelled surface (Figure 17), indicating that storage reserve synthesis and deposition during development and consequently seed weight is reduced. Therefore, in contrast to Köhler and co-authors (Köhler *et al.*, 2005), the results of this study show that the development of *phe1* mutant seeds is affected. This apparent difference might be explained by differences in growth conditions.

Plants homozygous for *mini3-3* and *iku2-4* produce seeds of reduced size (Figure 15A, B) and weight (Figure 16A, B). These findings are similar to the results previously reported for *mini3-1*, *mini3-2*, and *iku2-3* (Luo *et al.*, 2005). The reduction in seed size of *mini3* and *iku2* mutants is caused by the early arrest of endosperm growth and therefore early cellularization and restraint of cell proliferation of the mutant embryos (Luo *et al.*, 2005).

Arabidopsis embryos accumulate high amounts of starch at early stages of development (3-14 DAF) corresponding to high photosynthesis activity. This transient starch is later degraded during embryo maturation, and only traces are found in the mature seeds (da Silva *et al.*, 1997; Focks and Benning, 1998; Baud *et al.*, 2002). Several Arabidopsis mutants deficient in starch mobilization (*starch excess*) were previously isolated. These mutants accumulate high amounts of starch in the photosynthetic organs after extended incubation in the darkness (Caspar *et al.*, 1991; Zeeman *et al.*, 1998; Zeeman *et al.*, 1998). The starch mutant *sex1* (Caspar *et al.*, 1991) accumulates high amounts of starch in the mature seeds, and in leaves. Thus, the staining of mature *sex1* seeds with potassium iodide/iodine solution led to a change in colour from brown to black. However, no change in colour of the mature seeds of the mutant alleles of *phe1*, *mini3*, and *iku2* was observed after staining with potassium iodide/iodine solution (Figure 19), indicating that the starch was totally degraded in the seeds. This finding reveals that the expression of genes encoding the enzymes of starch degradation was not affected by the different mutations in *phe1*, *mini3*, and *iku2*. It is not clear whether the maximal amount of transient starch produced during seed development was impaired by the mutations. To answer this question, it would be necessary to compare the amount of starch synthesized in the mutant seeds and the corresponding WT controls throughout seed development.

Sucrose is synthesized in photosynthetic organs, and it is the main form of sugars exported to different sink tissues. In the developing seeds of Arabidopsis, sucrose is used to fuel the growth of the endosperm and the embryo, and it is the primary substrate for the biosynthesis of seed storage reserves such as TAGs and starch (Fallahi *et al.*, 2008). Sugar measurements (Figure 20) showed that the amount of sucrose in the seeds of *mini3* and *iku2*

was similar to the corresponding WT controls *Ler* and *Col-0*, with ca. 0.5 µg/seed. The total sucrose content in the seeds of the WT controls is consistent with previous reports (Focks and Benning, 1998; Baud *et al.*, 2002). The seeds of *phe1* lines showed an increase in total sucrose of ca. 20% compared to the WT controls. Interestingly, Arabidopsis *wri1* mutant seeds deficient in the *Wrinkled1* (*WRI1*) gene, a key regulator of seed oil biosynthesis, are characterized by a reduction of seed oil content by ca. 80%, and a concomitant increase in total sucrose by two fold compared to WT (Focks and Benning, 1998). Thus, the increase in total sucrose in the *phe1* lines may be explained by the fact that mutations in seed development genes affect the conversion of carbohydrates into precursors of FA and TAG biosynthesis resulting in an accumulation of sucrose.

During the early maturation phase (6-10 DAF); storage lipid and protein synthesis is initiated in Arabidopsis embryos at a low rate. Between 11 to 16 DAF, the synthesis of TAGs and storage proteins markedly increases resulting in an increase in seed dry weight. At the late maturation phase (17-20 DAF), storage reserve synthesis is arrested and the embryo becomes metabolically quiescent. In the mature dry seeds, the storage oil and proteins constitute ca. 40% of total seed dry weight each (Focks and Benning, 1998; Baud *et al.*, 2002; Fait *et al.*, 2006; Baud *et al.*, 2008). To investigate the effects of the reduction in weight in the seeds of *phe1*, *mini3* and *iku2* on storage compound accumulation, total protein content, protein composition, and total fatty acid content were analyzed. The protein contents in the *phe1*, *mini3*, and *iku2* lines were slightly reduced compared to the corresponding WT controls (Figure 21), while SDS-PAGE analysis showed no changes in protein composition (Figure 22). Total FA measurements (Figure 23) showed a reduction in seed oil content by ca. 15-50 % in *phe1*, *mini3* and *iku2*. These results are in line with data reported for *wri1* which showed a reduction of ca. 80% in FA content and no change in storage protein content and composition (Focks and Benning, 1998). Hence, the alteration in expression of seed development genes does not influence carbon deposition in the form of storage proteins and the sorting of these proteins. However the carbon flow from carbohydrates to FA biosynthesis is affected. To confirm the role of *PHE1*, *MINI3*, and *IKU2* in the regulation of carbon flow to FA biosynthesis in the developing seeds of Arabidopsis, it is necessary to determine the rate of FA synthesis through development in the mutant lines and the WT controls. Moreover, genes encoding seed-specific glycolytic enzymes have to be characterized and their expression compared in WT and in *phe1*, *mini3* and *iku2* lines.

Despite the variations in seed storage oil content between different *Arabidopsis thaliana* ecotypes, the set of FA found in the seed TAG is conserved in most ecotypes

regardless of the geographic locations and growth conditions. On the other hand, there is quite some variation in the composition of seed fatty acids, in particular of 18.2, 18:3 and 20:1 (Millar and Kunst, 1999; O'Neill *et al.*, 2003). Total fatty acid composition in seeds of *phe1, mini3,* and *iku2* lines showed a slight increase in linolenic acid (18:3) and a decrease in linoleic acid (18:2) content compared to the WT controls. A similar change in fatty acid composition has previously been observed in two Arabidopsis mutants, *fat1/fat2* and *wri1* which accumulate reduced amounts of storage oil. The increase in the amount of the end product of FA desaturation, 18:3, and the concomitant decrease in the amount of 18:2, can be explained by the reduction in the substrate pool size of FA available for the enzymes of oil synthesis, including desaturases, in the seed oil mutants resulting in an enhanced conversion into the desaturated products (Focks and Benning, 1998; Moreno-Perez *et al.*, 2012).

Female gametophytes carrying the *fie* mutation are embryo lethal, and the embryo development is arrested at the heart stage (Ohad *et al.*, 1996; Ohad *et al.*, 1999; Luo *et al.*, 2000). The oil content in the aberrant seeds of *fie* lines was extremely low with ca. 0.05 µg per seed, in comparison with ca. 8 µg in *Col-0* and *C24* seeds (Figure 23). The total fatty acid composition of *fie* lines showed a strong increase in the saturated fatty acids 16:0 and 18:0, and a concomitant decrease in 18:2 and 18:3 compared to WT (Figure 24). This result is in agreement with previous data obtained by Baud and co-authors (Baud *et al.*, 2002) showing that developing Arabidopsis WT seeds at the heart stage contain high amounts of 16:0 and 18:0 and low amounts of 18:3 and 20:1 in comparison with mature seeds. However, the aberrant seeds of *fie* lines revealed a strong increase in the amount of 22:1 and a decrease in 20:1 compared to WT controls, which is similar to what has previously been reported in the fully developed mature seeds of *wri1* (Focks and Benning, 1998).

4.2 Characterization of transgenic Arabidopsis seeds overexpressing seed size candidate genes

To investigate the effects of the overexpression of *FIE, PHE1, MINI3*, and *IKU2* on seed size and storage reserve accumulation, cDNAs encoding the candidate genes were overexpressed in *Arabidopsis thaliana Col-0* using the promoter of the storage protein glycinin from soybean (*Glycine max*) (Nielsen *et al.*, 1989). The expression of seed storage proteins is tissue specific, i.e. mainly in the embryo and, in some cases, in the endosperm of developing seeds. These proteins are never expressed in vegetative tissues. In soybean, glycinin represents ca. 40% of total seed protein and is restricted to the embryo and

cotyledons, indicating that the glycinin promoter is embryo specific (Nielsen *et al.*, 1989; Utsumi, 1992; Thomas, 1993). The glycinin promoter is active during the mid to late phase of embryogenesis corresponding to the seed filling stage of soybean. However, it is repressed during the early phase of embryogenesis and prior to seed dormancy (Nielsen *et al.*, 1989; Golombek *et al.*, 2001; Rosche *et al.*, 2005; Fait *et al.*, 2006).

The Arabidopsis seed development genes *PHE1*, *MINI3* and *IKU2* are expressed during the early stage in the developing wild type seeds (1-4 DAP). *PHE1* expression is restricted to the embryo and chalazal region of the endosperm (Köhler *et al.*, 2003a), *MINI3* is expressed in embryo and endosperm, whereas *IKU2* expression is restricted to the endosperm (Luo *et al.*, 2005). In contrast, *FIE* is expressed in reproductive and vegetative tissues before and after fertilization (Ohad *et al.*, 1999). Thus, the expression of the transgenes which are under control of the glycinin promoter, and of the endogenous genes, is spatially and temporally distinct. This difference is confirmed by rt-PCR from RNA isolated from developing siliques (14 DAF) of the overexpression lines and controls. The expression level of the transgenes *FIE, PHE1, MINI3,* and *IKU2* recorded by semi-quantitative rt-PCR of RNA from whole siliques of the overexpression lines is strongly upregulated compared to the empty vector controls (Figure 25).

FIE is a member of the PcG protein complex which controls gene expression in the female gametophyte via imprinting during seed development in Arabidopsis (Wang *et al.*, 2006; Pien and Grossniklaus, 2007; Hennig and Derkacheva, 2009). The gene expression of *PHE1* is directly regulated via imprinting by the *FIS*/*MEA* PcG protein complex (Köhler *et al.*, 2003a; Villar *et al.*, 2009). Therefore, it is unlikely that the overexpression of *PHE1* will have a significant effect on seed development, in agreement with the results of this study. Furthermore, the embryo-specific overexpression of *FIE* in transgenic plants did not affect seed size or storage compound allocation. Ohad and co-authors (Ohad *et al.*, 1999) previously reported that *FIE* is expressed not only in reproductive organs, but also in vegetative tissues. This finding was confirmed by the result that although transgenic plants with reduced *FIE* expression obtained by co-suppression showed morphological aberrations, they developed to maturity, indicating that the co-suppression took place after the germination (Ohad *et al.*, 1999). Moreover, *FIE*-silenced plants did not display the characteristics of *fie* mutants, i.e. the development of endosperm and seed-like structures in the absence of fertilization and the embryo abortion after fertilization (Katz *et al.*, 2004). Interestingly, Vinkenoog and co-authors (Vinkenoog *et al.*, 2000) reported that that pollination of maternal *fie* mutants with hypomethylated pollen bearing the wild type *FIE* allele is able to rescue the endosperm

development in these mutants. In our study, the embryo specific overexpression of *FIE* has no effects on seed development. This finding can be explained because *FIE* is repressed by DNA methylation, and therefore *FIE* overexpression might result in hypermethylation and might not have an impact on seed development. We tried to complement the female *fie* mutant gametophytic lethality by introduction of the *FIE* overexpression construct into heterozygous *fie* plants. However, in agreement with the result that *FIE* is down-regulated by DNA methylation, the *fie* mutant phenotype could not be complemented (data not shown).

4.3 Characterization of *IKU2*-OE plants

SHORT HYPOCOTYL UNDER BLUE 1 (SHB1)(Zhou *et al.*, 2009), *IKU2* (Garcia *et al.*, 2003), and *MINI3* (Luo *et al.*, 2005) are major regulators of Arabidopsis seed development. The gain-of-function overexpression mutation *shb1-D* induces an increase in seed size, whereas the seed size in the *shb1* mutant, a loss-of-function allele, is reduced. The increase in seed size of *shb1-D* was associated with increased expression of *MINI3* and *IKU2*. Interestingly, the embryo specific overexpression of *MINI3* did not have an effect on seed size (Figure 27). *MINI3* encodes a WRKY family protein, a putative transcription factor regulating endosperm growth. Expression of *MINI3*::GUS is up-regulated in the endosperm of homozygous *mini3-1* plants carrying the construct compared to heterozygous or homozygous *MINI3* plants carrying the same construct, indicating that *MINI3* is auto-repressive (Luo *et al.*, 2005). The negative auto-regulatory effect of *MINI3* might explain the results obtained in this study. In contrast, the embryo specific overexpression of *IKU2* resulted in increase in seed size, seed weight, and oil content (by ca. 35%) in heterozygous *IKU2*-OE lines (Figures 27, 28, 32). The amounts of storage protein and sucrose in seeds of *IKU2*-OE lines were similar to *Col-0*- EV, in agreement with the fact that oil represents the main sink for storage compounds in Arabidopsis seeds.

Heterozygous T1 lines of *IKU2*-OE plants produce two types of transgenic T2 seeds, (i) large seeds (in a ratio of about 2/3) and (ii) small seeds (in a ratio of about 1/3) with a size comparable to the control *Col-0*-EV. The segregation ratio of 2:1 of transgenic large seeds to transgenic normal-size seeds suggests that the large seeds represent the heterozygous *IKU2*-OE seeds, whereas the normal-size seeds represent the transgenic homozygous seeds. Interestingly, the non-transgenic *IKU2*-OE segregant seeds also showed an increase in size (Figure 36B), indicative for a maternal effect of *IKU2* overexpression in heterozygous plants on the non-transgenic progeny. For the time being, the molecular basis for this effect is

unknown. However, the side activity of the glycinin promoter might result in a minor increase in *IKU2* expression in maternal tissue, and this might have a positive effect on development of the seeds. To confirm this hypothesis, the expression levels of *IKU2* should be measured in the silique and seed coat of heterozygous *IKU2*-OE lines and compared with the empty vector control. The increase in seed size of heterozygous *IKU2*-OE lines was shown to be caused by an increase in embryo size (Figure 35). Previous studies have shown that there is a positive correlation between the increase in embryo size and increase in embryonic cells, and size and/or number of oil bodies (Murphy and Cummins, 1989; Siloto *et al.*, 2006; Liu *et al.*, 2010). However, it is unknown whether the overexpression of *IKU2* results in an increase in the sizes of embryonic cells and number or the sizes of the oil bodies in the transgenic lines.

A strong co-suppression was observed in plants homozygous for the *IKU2*-OE construct, which was manifested by a reduction in *IKU2* expression to levels similar to the *Col-0*-EV, whereas it was strongly increased in heterozygous *IKU2*-OE seeds. This co-suppression was released by the reciprocal crosses of homozygous *IKU2*-OE plant with WT *Col-0,* which showed an increase in size of the transgenic F1 seeds compared to seeds of wild type control, regardless of the direction of the cross (Figure 37). In accordance with this finding, the increase in seed size, weight and oil content was only observed in seeds of a heterozygous plant, but not in homozygous seeds. Hence, the increase in seed size, seed weight and oil content that would be observed for a collection of heterozygous *IKU2*-OE seeds free of homozygous seeds would be expected to be even higher than the results shown in Figures 27, 28, and 32, because the measurements of size, weight and oil content in *IKU*-OE transgenic seeds were performed with seeds from heterozygous transgenic *IKU2*-OE lines which contain a mixture of heterozygous and homozygous IKU2-OE seeds in a ratio of about 2:1.

IKU2-OE lines showed a strong reduction in seed yield per plant compared to *Col-0*-EV (Figure 39C). Thus, the increase in carbon deposition per seed is associated with reduction in seed yield. On the other hand, the siliques size and number of seeds per siliques were similar to the control (Figure 39A, B). Therefore, the number of siliques per plant must be reduced. The fact that the total weight of seeds per plant is reduced indicates that in *IKU2*-OE plants, less carbon resources are employed to produce seeds as compared to *Col-0*-EV. Provided that there is no negative effect on photosynthetic capacity in *IKU2*-OE plants, the same amount of carbon should be available for seed filling. Therefore, the reduction in seed yield cannot be caused by a general limitation in carbon supply. However, it is possible that

the expression of the *IKU2* protein under the control of the strong glycinin promoter has negative effects on overall plant development, such as flower initiation, which would explain the reduced number of siliques per plant. Consisting with this theory, it has been recently reported that the overexpression of *LEAFY COTYLEDON1* (*LEC1*) and *LEC1-LIKE* (*L1L*) in Arabidopsis and rapeseed under control of the strong seed specific napin promoter also led to an increase in seed size; however the transgenic seedling showed severe developmental abnormalities (Tan *et al.*, 2011). The authors reported that the overexpression of *LEC1* and *L1L* under the control of modified napin promoters, significantly increases the seed size of the transgenic rape seeds, without any detrimental effects on plant development or on major agronomic traits.

It would be desirable to obtain transgenic *IKU2* overexpression plants that show an increase in seed size and weight and in the amount of oil per seed without having the detrimental effect on total seed yield per plant. This strategy could be useful to improve the yield of different crop plants including rape or soybean. In order to achieve this goal, *IKU2* should be overexpressed in the seeds using different promoters. In analogy with results obtained by Tan and co-authors (Tan *et al.*, 2011) on overexpression of *LEC1* and *LIL*, it is possible that the overexpression of *IKU2* using truncated or mutated versions of the glycinin promoter can result in the generation of transgenic *IKU2*-OE plants which produce seeds with increased size, but without reduction of the total seed yield per plant. This strategy would imply that the expression level obtained with modified versions of the glycinin promoter in non-seed tissues would be weaker, but still resulting in increased seed size without reduction in seed yield per plant. Another possibility would be to overexpress *IKU2* using different promoters with specificity for the endosperm. Thus, *IKU2* would be overexpressed in the tissue where the endogenous *IKU2* gene is expressed.

5 Summary

Seeds are major components of the human diet and seed oil represents a source of renewable and environmentally friendly replacements for fossil-based raw materials. During maturation seeds accumulate storage reserves to support the growth of the seedling before it gains its photosynthetic activity. The main storage reserves are oil in the form of triacylglycerols, proteins, and carbohydrates in the form of starch. Increasing the seed size and therefore storage reserve content is of major economical and agricultural importance. In this study we investigated the effects of alteration in seed development gene expression on storage allocation in seeds of *Arabidopsis thaliana*. New alleles of the seed development genes *FIE1*, *PHE1, MINI3,* and *IKU2* were isolated and the homozygous mutant lines of the candidate genes were characterized. While the amount of storage proteins and sucrose was similar to the WT controls, the oil content was strongly reduced in all mutants to varying extents. This finding indicates that the carbon flow from carbohydrates to fatty acids biosynthesis is affected. To investigate the effects of the overexpression of *FIE, PHE1, MINI3,* and *IKU2* on seed size and storage reserve accumulation, cDNAs encoding the candidate genes were overexpressed in *Arabidopsis thaliana* ecotype *Col-0* using the embryo-specific glycinin promoter from soybean. Thus, the transgenes under control of the glycinin promoter, and the endogenous genes are differentially expressed, spatially and temporally. While the overexpression of *FIE, PHE1,* and *MINI3* did not have any effects on seed size and storage reserve accumulation, *IKU2*-OE lines displayed a strong increase in seed size and oil content, but no change in protein and sucrose content. Moreover, non-transgenic segregant seeds of a heterozygous *IKU2*-OE line also showed an increase in seed size and oil content indicating a maternal effect of *IKU2*-OE overexpression in the seed coat or silique of the heterozygous plants. The increase in seed size and oil content was only observed in seeds of heterozygous *IKU2*-OE lines but not in homozygous seeds. RT-PCR analysis showed that *IKU2* is highly expressed in seeds of heterozygous *IKU2*-OE lines, but is co-suppressed in homozygous *IKU2*-OE lines, which show similar low expression level as the control *Col-0*-EV. The increase in seed size of heterozygous *IKU2*-OE seeds was associated with a reduction in total seed yield per plant compared to the empty vector controls, while the siliques size and number of seeds per silique were not affected. It has been previously shown that the seed specific overexpression of seed development genes using strong promoters induces an increase in seed size associated with developmental abnormalities in the transgenic plants. However, the overexpression of the same genes under

control of truncated or mutated promoters induces an increase in seed size without detrimental effects on plant development. Therefore, it might be possible to overcome the negative correlation between seed size and seed yield per plant by overexpression *IKU2* using a modified glycinin promoter. The future goals of this study are the overexpression of *IKU2* in oilseed crops such as soybean and rapeseed with the aim to increase the seed size and oil content in the transgenic lines.

.

6 References

Alvarez, H., M. and A. Steinbuchel (2002). "Triacylglycerols in prokaryotic microorganisms." Appl Microbiol Biotechnol **60** (4): 367-376.

Baud, S., J.-P. Boutin, M. Miquel, L. Lepiniec and C. and Rochat (2002). "An integrated overview of seed development in Arabidopsis thaliana ecotype Ws." Plant Physiol. Biochem. **41**: 151–160.

Baud, S., B. Dubreucq, M. Miquel, C. Rochat and L. Lepiniec (2008). "Storage reserve accumulation in Arabidopsis: metabolic and developmental control of seed filling." Arabidopsis Book **6**: e0113.

Baud, S. and L. Lepiniec (2008). "Compared analysis of the regulatory systems controlling lipogenesis in hepatocytes of mice and in maturing oilseeds of Arabidopsis." C R Biol **331** (10): 737-745.

Baud, S. and L. Lepiniec (2010). "Physiological and developmental regulation of seed oil production." Prog Lipid Res **49** (3): 235-249.

Baud, S., S. Wuilleme, R. Lemoine, J. Kronenberger, M. Caboche, L. Lepiniec and C. Rochat (2005). "The AtSUC5 sucrose transporter specifically expressed in the endosperm is involved in early seed development in Arabidopsis." Plant J **43** (6): 824-836.

Beisson, F., Y. Li, G. Bonaventure, M. Pollard and J. B. Ohlrogge (2007). "The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis." Plant Cell **19** (1): 351-368.

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

Breiling, A., L. Sessa and V. Orlando (2007). "Biology of polycomb and trithorax group proteins." Int Rev Cytol **258**: 83-136.

Brocard-Gifford, I. M., T. J. Lynch and R. R. Finkelstein (2003). "Regulatory networks in seeds integrating developmental, abscisic acid, sugar, and light signaling." Plant Physiol **131** (1): 78-92.

Browse, J., P. J. McCourt and C. R. Somerville (1986). "Fatty acid composition of leaf lipids determined after combined digestion and fatty acid methyl ester formation from fresh tissue." Anal Biochem **152** (1): 141-145.

Buleon, A., P. Colonna, V. Planchot and S. Ball (1998). "Starch granules: structure and biosynthesis." Int J Biol Macromol **23** (2): 85-112.

Caspar, T., T. P. Lin, G. Kakefuda, L. Benbow, J. Preiss and C. Somerville (1991). "Mutants of Arabidopsis with altered regulation of starch degradation." Plant Physiol **95** (4): 1181-1188.

Cernac, A. and C. Benning (2004). "WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis." Plant J **40** (4): 575-585.

Chapman, K. D. and J. B. Ohlrogge (2012). "Compartmentation of triacylglycerol accumulation in plants." J Biol Chem **287** (4): 2288-2294.

Clough, S. J. and A. F. Bent (1998). "Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*." Plant J **16** (6): 735-743.

Cove, D. J. and C. D. Knight (1993). "The Moss Physcomitrella patens, a Model System with Potential for the Study of Plant Reproduction." Plant Cell **5** (10): 1483-1488.

da Silva, P. M. F. R., P. J. Eastmond, L. M. Hill2, A. M. Smith and S. Rawsthorne (1997). "Starch metabolism in developing embryos of oilseed rape." Planta **203**: 480-487.

Dehesh, K., H. Tai, P. Edwards, J. Byrne and J. G. Jaworski (2001). "Overexpression of 3-ketoacyl-acyl-carrier protein synthase IIIs in plants reduces the rate of lipid synthesis." Plant Physiol **125** (2): 1103-1114.

Dormann, P., S. Hoffmann-Benning, I. Balbo and C. Benning (1995). "Isolation and characterization of an Arabidopsis mutant deficient in the thylakoid lipid digalactosyl diacylglycerol." Plant Cell **7** (11): 1801-1810.

Dumas, C. and H. L. Mogensen (1993). "Gametes and Fertilization: Maize as a Model System for Experimental Embryogenesis in Flowering Plants." Plant Cell **5** (10): 1337-1348.

Eastmond, P. J., A. J. van Dijken, M. Spielman, A. Kerr, A. F. Tissier, H. G. Dickinson, J. D. Jones, S. C. Smeekens and I. A. Graham (2002). "Trehalose-6-phosphate synthase 1, which catalyses the first step in trehalose synthesis, is essential for Arabidopsis embryo maturation." Plant J **29** (2): 225-235.

Fait, A., R. Angelovici, H. Less, I. Ohad, E. Urbanczyk-Wochniak, A. R. Fernie and G. Galili (2006). "Arabidopsis seed development and germination is associated with temporally distinct metabolic switches." Plant Physiol **142** (3): 839-854.

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.

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.

Garcia, D., J. N. Fitz Gerald and F. Berger (2005). "Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in Arabidopsis." Plant Cell **17** (1): 52-60.

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.

Golombek, S., H. Rolletschek, U. Wobus and H. Weber (2001). "Control of storage protein accumulation during legume seed development." Journal of Plant Physiology **158** (4): 457-464.

Gomez, L. D., S. Baud, A. Gilday, Y. Li and I. A. Graham (2006). "Delayed embryo development in the ARABIDOPSIS TREHALOSE-6-PHOSPHATE SYNTHASE 1 mutant is associated with altered cell wall structure, decreased cell division and starch accumulation." Plant J **46** (1): 69-84.

Grimaud, C., N. Negre and G. Cavalli (2006). "From genetics to epigenetics: the tale of Polycomb group and trithorax group genes." Chromosome Res **14** (4): 363-375.

Guitton, A. E., D. R. Page, P. Chambrier, C. Lionnet, J. E. Faure, U. Grossniklaus and F. Berger (2004). "Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in Arabidopsis thaliana." Development **131** (12): 2971-2981.

Haranishimura, I., T. Shimada, N. Hiraiwa and M. Nishimura (1995). "VACUOLAR PROCESSING ENZYME RESPONSIBLE FOR MATURATION OF SEED PROTEINS." Journal of Plant Physiology **145** (5-6): 632-640.

Heath, J. H., R. Weldon, C. Monnot and D. W. Meinke (1986). "Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of Arabidopsis thaliana." Planta **169** (3): 304-312.

Hennig, L. and M. Derkacheva (2009). "Diversity of Polycomb group complexes in plants: same rules, different players?" Trends Genet **25** (9): 414-423.

Herman, E. M. and B. A. Larkins (1999). "Protein storage bodies and vacuoles." Plant Cell **11** (4): 601-614.

Hill, L. M., E. R. Morley-Smith and S. Rawsthorne (2003). "Metabolism of sugars in the endosperm of developing seeds of oilseed rape." Plant Physiol **131** (1): 228-236.

Hills, M. J. (2004). "Control of storage-product synthesis in seeds." Curr Opin Plant Biol **7** (3): 302-308.

Huang, A. H. (1996). "Oleosins and oil bodies in seeds and other organs." Plant Physiol **110** (4): 1055-1061.

Hughes, D. W. and G. A. Galau (1989). "Temporally modular gene expression during cotyledon development." Genes Dev **3** (3): 358-369.

Inoue, H., H. Nojima and H. Okayama (1990). "High efficiency transformation of Escherichia coli with plasmids." Gene **96** (1): 23-28.

Jako, C., A. Kumar, Y. Wei, J. Zou, D. L. Barton, E. M. Giblin, P. S. Covello and D. C. Taylor (2001). "Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight." Plant Physiol **126** (2): 861-874.

Katavic, V., D. W. Reed, D. C. Taylor, E. M. Giblin, D. L. Barton, J. Zou, S. L. Mackenzie, P. S. Covello and L. Kunst (1995). "Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in Arabidopsis thaliana affecting diacylglycerol acyltransferase activity." Plant Physiol **108** (1): 399-409.

Katz, A., M. Oliva, A. Mosquna, O. Hakim and N. Ohad (2004). "FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development." Plant J **37** (5): 707-719.

Kay, R., A. Chan, M. Daly and J. McPherson (1987). "Duplication of CaMV 35S Promoter Sequences Creates a Strong Enhancer for Plant Genes." Science **236** (4806): 1299-1302.

Köhler, C., L. Hennig, C. Spillane, S. Pien, W. Gruissem and U. Grossniklaus (2003a). "The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1." Genes Dev **17** (12): 1540-1553.

Köhler, C., D. R. Page, V. Gagliardini and U. Grossniklaus (2005). "The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting." Nat Genet **37** (1): 28-30.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227** (5259): 680-685.

Leroy, O., L. Hennig, H. Breuninger, T. Laux and C. Kohler (2007). "Polycomb group proteins function in the female gametophyte to determine seed development in plants." Development **134** (20): 3639-3648.

Li, Y., F. Beisson, M. Pollard and J. Ohlrogge (2006). "Oil content of Arabidopsis seeds: the influence of seed anatomy, light and plant-to-plant variation." Phytochemistry **67** (9): 904-915.

Lin, Y., J. E. Cluette-Brown and H. M. Goodman (2004). "The peroxisome deficient Arabidopsis mutant sse1 exhibits impaired fatty acid synthesis." Plant Physiol **135** (2): 814-827.

Lin, Y., L. Sun, L. V. Nguyen, R. A. Rachubinski and H. M. Goodman (1999). "The Pex16p homolog SSE1 and storage organelle formation in Arabidopsis seeds." Science **284** (5412): 328-330.

Liu, J., W. Hua, G. Zhan, F. Wei, X. Wang, G. Liu and H. Wang (2010). "Increasing seed mass and oil content in transgenic Arabidopsis by the overexpression of wri1-like gene from Brassica napus." Plant Physiol Biochem **48** (1): 9-15.

Lorberth, R., G. Ritte, L. Willmitzer and J. Kossmann (1998). "Inhibition of a starch-granule-bound protein leads to modified starch and repression of cold sweetening." Nat Biotechnol **16** (5): 473-477.

Lu, C. and M. J. Hills (2002). "Arabidopsis mutants deficient in diacylglycerol acyltransferase display increased sensitivity to abscisic acid, sugars, and osmotic stress during germination and seedling development." Plant Physiol **129** (3): 1352-1358.

Lund, A. H. and M. van Lohuizen (2004). "Polycomb complexes and silencing mechanisms." Curr Opin Cell Biol **16** (3): 239-246.

Lundberg, K. S., D. D. Shoemaker, M. W. Adams, J. M. Short, J. A. Sorge and E. J. Mathur (1991). "High-fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus." Gene **108** (1): 1-6.

Luo, M., P. Bilodeau, E. S. Dennis, W. J. Peacock and A. Chaudhury (2000). "Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds." Proc Natl Acad Sci U S A **97** (19): 10637-10642.

Luo, M., E. S. Dennis, F. Berger, W. J. Peacock and A. Chaudhury (2005). "MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis." Proc Natl Acad Sci U S A **102** (48): 17531-17536.

Luo, M., M. Z. Luo, D. Buzas, J. Finnegan, C. Helliwell, E. S. Dennis, W. J. Peacock and A. Chaudhury (2008). "UBIQUITIN-SPECIFIC PROTEASE 26 is required for seed development and the repression of PHERES1 in Arabidopsis." Genetics **180** (1): 229-236.

Maisonneuve, S., J. J. Bessoule, R. Lessire, M. Delseny and T. J. Roscoe (2010). "Expression of rapeseed microsomal lysophosphatidic acid acyltransferase isozymes enhances seed oil content in Arabidopsis." Plant Physiol **152** (2): 670-684.

Michael A. Innis, D. H. G., John J. Sninsky, Thomas J. White (1990). PCR Protocols: A Guide to Methods and Applications. San Diego, CA, Academic Press.

Millar, A. A. and L. Kunst (1999). "The natural genetic variation of the fatty-acyl composition of seed oils in different ecotypes of Arabidopsis thaliana." Phytochemistry **52** (6): 1029-1033.

Molina, I., G. Bonaventure, J. Ohlrogge and M. Pollard (2006). "The lipid polyester composition of Arabidopsis thaliana and Brassica napus seeds." Phytochemistry **67** (23): 2597-2610.

Moreno-Perez, A. J., M. Venegas-Caleron, F. E. Vaistij, J. J. Salas, T. R. Larson, R. Garces, I. A. Graham and E. Martinez-Force (2012). "Reduced expression of FatA thioesterases in Arabidopsis affects the oil content and fatty acid composition of the seeds." Planta **235** (3): 629-639.

Muller, K., S. Tintelnot and G. Leubner-Metzger (2006). "Endosperm-limited Brassicaceae seed germination: abscisic acid inhibits embryo-induced endosperm weakening of Lepidium sativum (cress) and endosperm rupture of cress and Arabidopsis thaliana." Plant Cell Physiol **47** (7): 864-877.

Murphy, D. J. and I. Cummins (1989). "BIOSYNTHESIS OF SEED STORAGE PRODUCTS DURING EMBRYOGENESIS IN RAPESEED, BRASSICA-NAPUS." Journal of Plant Physiology **135** (1): 63-69.

Nielsen, N. C., C. D. Dickinson, T. J. Cho, V. H. Thanh, B. J. Scallon, R. L. Fischer, T. L. Sims, G. N. Drews and R. B. Goldberg (1989). "Characterization of the glycinin gene family in soybean." Plant Cell **1** (3): 313-328.

North, H., S. Baud, I. Debeaujon, C. Dubos, B. Dubreucq, P. Grappin, M. Jullien, L. Lepiniec, A. Marion-Poll, M. Miquel, L. Rajjou, J. M. Routaboul and M. Caboche (2010). "Arabidopsis seed secrets unravelled after a decade of genetic and omics-driven research." Plant J **61** (6): 971-981.

O'Neill, C. M., S. Gill, D. Hobbs, C. Morgan and I. Bancroft (2003). "Natural variation for seed oil composition in Arabidopsis thaliana." Phytochemistry **64** (6): 1077-1090.

Ohad, N., L. Margossian, Y.-C. Hsu, C. Williams, P. Repetti and R. L. Fischer (1996). "A mutation that allows endosperm development without fertilization." Proc Natl Acad Sci U S A **93**: 5319–5324.

Ohad, N., R. Yadegari, L. Margossian, M. Hannon, D. Michaeli, J. J. Harada, R. B. Goldberg and R. L. Fischer (1999). "Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization." Plant Cell **11** (3): 407-416.

Ohlrogge, J. and J. Browse (1995). "Lipid biosynthesis." Plant Cell **7** (7): 957-970.

Penfield, S., E. L. Rylott, A. D. Gilday, S. Graham, T. R. Larson and I. A. Graham (2004). "Reserve mobilization in the Arabidopsis endosperm fuels hypocotyl elongation in the dark, is independent of abscisic acid, and requires PHOSPHOENOLPYRUVATE CARBOXYKINASE1." Plant Cell **16** (10): 2705-2718.

Pien, S. and U. Grossniklaus (2007). "Polycomb group and trithorax group proteins in Arabidopsis." Biochim Biophys Acta **1769** (5-6): 375-382.

Riggs, K. G. and A. McLachlan (1986). "A simplified screening procedure for large nummber of minipreparation." Biotechnique **4**: 310-313.

Rosche, E. G., D. Blackmore, C. E. Offler and J. W. Patrick (2005). "Increased capacity for sucrose uptake leads to earlier onset of protein accumulation in developing pea seeds." Functional Plant Biology **32** (11): 997-1007.

Ross, J. H. E., J. Sanchez, F. Millan and D. J. Murphy (1993). "DIFFERENTIAL PRESENCE OF OLEOSINS IN OLEOGENIC SEED AND MESOCARP TISSUES IN OLIVE (OLEA-EUROPAEA) AND AVOCADO (PERSEA-AMERICANA)." Plant Science **93** (1-2): 203-210.

Ruuska, S. A., T. Girke, C. Benning and J. B. Ohlrogge (2002). "Contrapuntal networks of gene expression during Arabidopsis seed filling." Plant Cell **14** (6): 1191-1206.

Rychlik, W. and R. E. Rhoads (1989). "A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA." Nucleic Acids Res **17** (21): 8543-8551.

Shen, B., W. B. Allen, P. Zheng, C. Li, K. Glassman, J. Ranch, D. Nubel and M. C. Tarczynski (2010). "Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize." Plant Physiol **153** (3): 980-987.

Shewry, P. R., J. A. Napier and A. S. Tatham (1995). "Seed storage proteins: structures and biosynthesis." Plant Cell **7** (7): 945-956.

Shimada, T., K. Fuji, K. Tamura, M. Kondo, M. Nishimura and I. Hara-Nishimura (2003). "Vacuolar sorting receptor for seed storage proteins in Arabidopsis thaliana." Proc Natl Acad Sci U S A **100** (26): 16095-16100.

Shimada, T., Y. Koumoto, L. Li, M. Yamazaki, M. Kondo, M. Nishimura and I. Hara-Nishimura (2006). "AtVPS29, a putative component of a retromer complex, is required for the efficient sorting of seed storage proteins." Plant Cell Physiol **47** (9): 1187-1194.

Shimada, T., K. Yamada, M. Kataoka, S. Nakaune, Y. Koumoto, M. Kuroyanagi, S. Tabata, T. Kato, K. Shinozaki, M. Seki, M. Kobayashi, M. Kondo, M. Nishimura and I. Hara-Nishimura (2003). "Vacuolar processing enzymes are essential for proper processing of seed storage proteins in Arabidopsis thaliana." J Biol Chem **278** (34): 32292-32299.

Siloto, R. M., K. Findlay, A. Lopez-Villalobos, E. C. Yeung, C. L. Nykiforuk and M. M. Moloney (2006). "The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis." Plant Cell **18** (8): 1961-1974.

Simon, J. A. and J. W. Tamkun (2002). "Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes." Curr Opin Genet Dev **12** (2): 210-218.

Smith, A. M., K. Denyer and C. R. Martin (1995). "What Controls the Amount and Structure of Starch in Storage Organs?" Plant Physiol **107** (3): 673-677.

Tan, H., X. Yang, F. Zhang, X. Zheng, C. Qu, J. Mu, F. Fu, J. Li, R. Guan, H. Zhang, G. Wang and J. Zuo (2011). "Enhanced seed oil production in canola by conditional expression of Brassica napus LEAFY COTYLEDON1 and LEC1-LIKE in developing seeds." Plant Physiol **156** (3): 1577-1588.

Thomas, T. L. (1993). "Gene expression during plant embryogenesis and germination: an overview." Plant Cell **5**(10): 1401-1410.

Tzen, J. T. C., Y. Z. Cao, P. Laurent, C. Ratnayake and A. H. C. Huang (1993). "LIPIDS, PROTEINS, AND STRUCTURE OF SEED OIL BODIES FROM DIVERSE SPECIES." Plant Physiology **101** (1): 267-276.

Utsumi, S. (1992). "Plant food protein engineering." Adv Food Nutr Res **36**: 89-208.

Vigeolas, H., T. Mohlmann, N. Martini, H. E. Neuhaus and P. Geigenberger (2004). "Embryo-specific reduction of ADP-Glc pyrophosphorylase leads to an inhibition of starch synthesis and a delay in oil accumulation in developing seeds of oilseed rape." Plant Physiol **136** (1): 2676-2686.

Villar, C. B., A. Erilova, G. Makarevich, R. Trosch and C. Kohler (2009). "Control of PHERES1 imprinting in Arabidopsis by direct tandem repeats." Mol Plant **2** (4): 654-660.

Vinkenoog, R., M. Spielman, S. Adams, R. L. Fischer, H. G. Dickinson and R. J. Scott (2000). "Hypomethylation promotes autonomous endosperm development and rescues postfertilization lethality in fie mutants." Plant Cell **12** (11): 2271-2282.

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.

Wang, D., M. D. Tyson, S. S. Jackson and R. Yadegari (2006). "Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis." Proc Natl Acad Sci U S A **103** (35): 13244-13249.

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

Yu, T. S., H. Kofler, R. E. Hausler, D. Hille, U. I. Flugge, S. C. Zeeman, A. M. Smith, J. Kossmann, J. Lloyd, G. Ritte, M. Steup, W. L. Lue, J. Chen and A. Weber (2001). "The Arabidopsis sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter." Plant Cell **13** (8): 1907-1918.

Zeeman, S. C. and T. Ap Rees (1999). "Changes in carbohydrate metabolism and assimilate export in starch-excess mutants of Arabidopsis" Plant, Cell & Environment **22** (11): 1445–1453.

Zeeman, S. C., F. Northrop, A. M. Smith and T. Rees (1998). "A starch-accumulating mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme." Plant J **15** (3): 357-365.

Zeeman, S. C., T. Umemoto, W. L. Lue, P. Au-Yeung, C. Martin, A. M. Smith and J. Chen (1998). "A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulates both starch and phytoglycogen." Plant Cell **10** (10): 1699-1712.

Zhou, Y., X. Zhang, X. Kang, X. Zhao and M. Ni (2009). "SHORT HYPOCOTYL UNDER BLUE1 associates with MINISEED3 and HAIKU2 promoters in vivo to regulate Arabidopsis seed development." Plant Cell **21** (1): 106-117.

Zou, J., Y. Wei, C. Jako, A. Kumar, G. Selvaraj and D. C. Taylor (1999). "The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene." Plant J **19** (6): 645-653.

7 Appendix

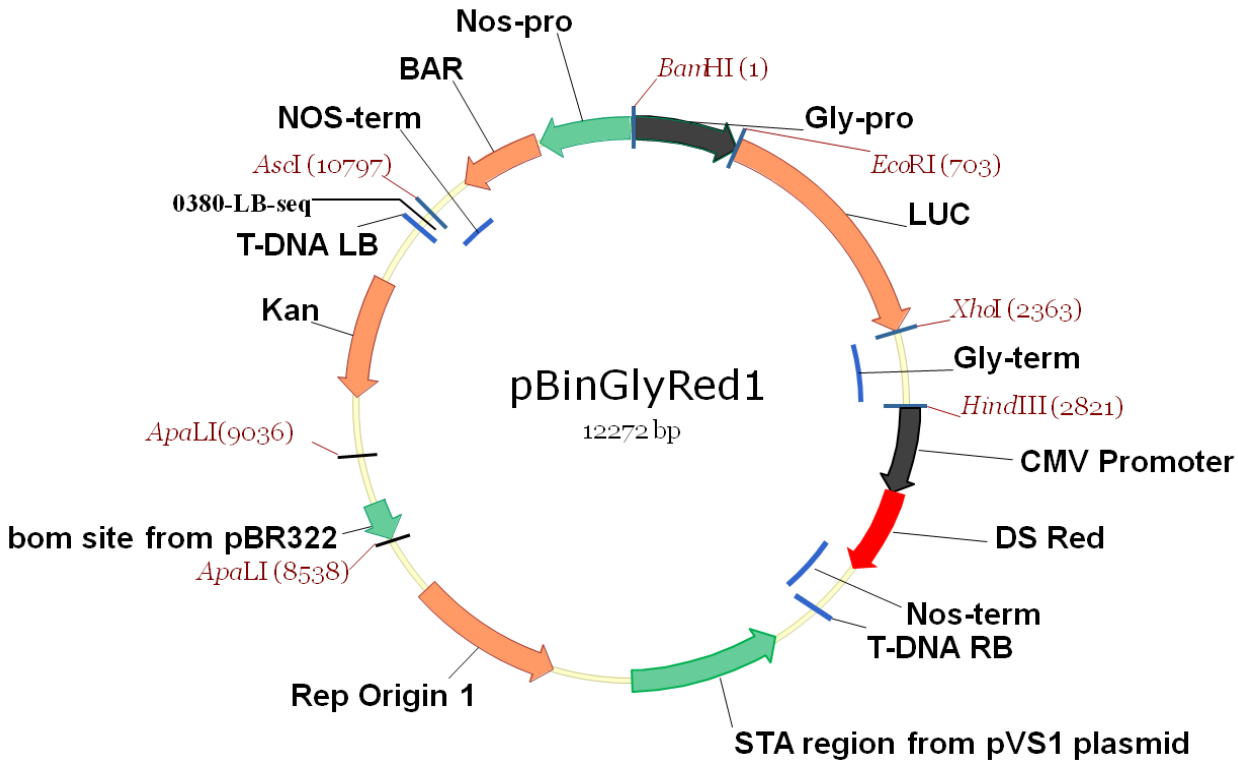
7.1 Synthetic oligonucleotides

Primer	Sequence (5'-3')	Gene, Plant line1
Bn78	ATTTTGCCGATTTCGGAAC	SALK T-DNA LBb1.3 primer
Bn130	TCCGTTCCGTTTTCGTTTTTAC	RIKEN Ds5-2a primer GUS-side
Bn277	CCGTCCCGCAAGTTAAATATG	CSHL Ds3-4 transposon primer
Bn142	CCCATTTGGACGTGAATGTAGACAC	GABI-Kat LB T-DNA primer
Bn165	TCTTTCAGTTGTGGTTCCCAC	5' genomic primer *fie-13*
Bn166	ATGTTTCACTGAGGCCATTTG	3' genomic primer *fie-13*
Bn205	CCCATCTGCAAAAGACAAGAC	5' genomic primer *fie-14*
Bn206	TTCTCTGTGTATTTGGGTGGG	3' genomic primer *fie-14*
Bn66	AGCATGTGCTCACATTTTGTG	5' genomic primer *phe1-6*
Bn67	GATGAGACACGTCCACCTCTC	3' genomic primer *phe1-6*
Bn273	GAAGGCGTTGAAGAAGTGATG	5' genomic primer *phe1-7*
Bn274	TTTGAATCCAACAAAAATATGGC	3' genomic primer *phe1-7*
Bn68	TGTGATTGAACCAATATTTGGG	5' genomic primer *mini3-2*
Bn69	TCCAGCGATAACCATCGTTAG	3' genomic primer *mini3-2*
Bn171	AGGCAGATCAACATTGTACGG	5' genomic primer *mini3-3*
Bn172	ATTTGCTGCTAGCGTTTTCTG	3' genomic primer *mini3-3*
Bn60	TCTTTAAGAACCGCAGCTCTG	5' genomic primer *iku2-4*
Bn61	GTTGTTTCGCCTACAATGACC	3' genomic primer *iku2-4*

Primer	Sequence (5'-3')	Function
Bn157	ACGCGTATGTCGAAGATAACCTTAGGGAAC	5' *FIE* primer with *MluI*
Bn158	CTCGAGCTACTTGGTAATCACGTCCCAGCG	3' *FIE* primer with *XhoI*
Bn155	GAATTCACGCGTATGAGGGGGAAGATGAAGTTATC	5' *PHE1* primer with *EcoRI, MluI*
Bn156	CTCGAGTCAGAGATCATTGATGATGTTAGG	3' *PHE1* primer with *XhoI*
Bn159	ACGCGTATGAGTGATTTTGATGAAAACTTC	5' *MINI3* primer with *MluI*
Bn160	CTCGAGCTACATGTCGACACCAAACTT	3' *MINI3* primer with *XhoI*
Bn550	ACGCGTATGCTCCGGCTACTATTTATC	5' *IKU2* genomic primer with *MluI*
Bn669	CGTTATGCTGCAGAACAACTTC	3' *IKU2* genomic primer
Bn553	AAGATAAGACGAGATAAGTTGAATAAAAC	5' *IKU2* cDNA primer
Bn551	CTCGAGTTATACAACTTTAGTAATCTCATCATT	3' *IKU2* cDNA primer with *XhoI*

Primer	Sequence (5'-3')	Function, PCR product size
Bn876	AACAATTGGAGGATGGTCGT	5' ubiquitin primer (573 bp)
Bn877	CAAGTTTCGCAGAACTGCAC	3' ubiquitin primer
Bn971	TTTGCTGGAGCTGGAGGTCATC	5' FIE primer (575 bp)
Bn972	GTCCTCGCAGCAAGCAAGAATC	3' FIE primer
Bn969	TTCGGTGTTGGACCGGACCAAG	5' PHE1 primer (576 bp)
Bn970	TGGTGACGGTAGCGAGACAATC	3' PHE1 primer
Bn1369	GCTCTCACTGTCCTCAATGGAG	5' MINI3 primer(238 bp)
Bn1370	TTCCGGTGAAGACAATGCAGCG	3' MINI3 primer
Bn1371	CGTGTGAGACAAGCGTTAGC	5' IKU2 primer(496 bp)
Bn1372	GAGGAGACTTGTCCGTGCAT	3' IKU2 primer

7.2 Vector Map



Acknowledgements

Firstly I would to sincerely express my special thanks to my supervisor Prof. Dr. Peter Dörmann for giving me the chance to make this PhD on his lab and also for his endless encouragement, support, continuous assistance and guidance during my research work.
I thank Prof. Dr. Dorothea Bartels for accepting to be the co-referee of my PhD thesis.
I am grateful to the German Federal Department of Education and Research (BMBF) for the funding of the GABI OIL project.
My special thanks are giving to the following: Dr. Anna Zbierzak for her valuable help and precious comments and discussions during this project. Dr. Anastassia Boudichevskaia for evaluator software and seed size measurements. Dr. Georg Hölzl for cloning. Katharina vom Dorp for TAGs analysis.
I would like to thank Meike Siebers for help with the preparation of the samples for microscopical analysis and for producing the seed photos.
Appreciation must be extended to all members of the group, Dr. Felix Lippold, Dr. Simone Zäuner, Marlies Becker, Helga Peisker, Brigitte Dresen-Scholz, Veronika Behnen, Barbara Kalisch, Vera Wewer, Carolin Hohn, Ernst Peter Jonen, Andreas Ahrens, Tanja Fuchs, thanks for good company, valuable discussion, collaboration, and for good work atmosphere.
I am thankful for Mrs Ellen Schulz for the help and assistance with administrative issues.
Lastly and most importantly I would like to express my deepest gratitude to my mother Habiba, my father Abderahman, my sisters Hasna, Khadija, Fatiha, my brothers Rachid and Hassen, Nadine and her parents Petra and Ronald. Thank you all for your love, moral support and encouragement.