(Form No. 14)

ABSTRACT OF DOCTORAL THESIS

Name: MOSTAFA AHMED ABOULELA MOHAMED

Title: Analysis of pollen development and pollen wall formation in *Arabidopsis thaliana* (L.) Heynh.

(シロイヌナズナにおける花粉発達と花粉表層形成に関する研究)

Chapter 1

Introduction

Arabidopsis thaliana (L.) Heynh.

Arabidopsis thaliana is a small annual weed belongs to the mustard family Cruciferae (Brassicaceae, Capparales). *A. thaliana* is a cosmopolitan species distributed widely in many areas of the world with preferential existence in central, western and northern Europe. It has also been found in a less extent in Australia, parts of North Africa, Turkey, Korea and Japan [1, 2]. *A. thaliana* resides a wide variety of habitats such as sandy or loamy soils, rocky slopes, river banks, cultivated ground, road sides, waste places, and open areas [3]. Many accessions and ecotypes of *A. thaliana* have been collected from wild populations growing throughout different ecological and geographical regions. From these accessions, Landsberg erecta (Ler), Columbia (Col-0) and Wassilewskija (Ws) are commonly used [1, 4]. Generally Columbia (Col-0) ecotype is considered as the reference genotype based on available data collected from many physiological, biochemical, sequencing and microarray studies [5]. Moreover, many research tools, such as sequence indexed T-DNA insertion lines, were developed in the Col-0 background [6].

For *A. thaliana* having numerous advantages, it was suggested earlier the suitability of *A. thaliana* as a model plant system by Laibach (1943) [7]. *A. thaliana* has been the subject of intense research in recent years and it is the most widely studied plant species with thousands of publications dealt with its biochemistry, molecular genetics, and development and evolution. The focused research on *Arabidopsis* should contribute to better understanding of diverse biological processes not

only in plants but also in animal and other organisms hence many genes are conserved in all eukaryotes.

A. thaliana has been adopted as the organism of choice in many disciplines in plant science for a number of reasons, including its small size that reduce the required space for growth, short generation time (four to six weeks), small genome (only five chromosomes 2n =10), naturally self-fertilizing, the ease of large scale mutagenesis because of its small sized seeds [4, 8, 9]. One of the critical factors for developing *Arabidopsis* as a favoured organism was its amenability to the efficient Agrobacterium-mediated transformation [10]. In addition to availability of the comprehensive research resources such as the Arabidopsis entire genomic sequence [11], the large collection of gene disruptions usually by T-DNA or transposon insertions, and full genome microarrays which are already exist in several fundamental information databases such as TAIR (http://www.arabidopsis.org/) and Arabidopsis Biological Resource Center (ABRC).

Mutagenesis

A well-designed mutagenesis screen at high magnification is a powerful strategy to identify genes affecting many aspects of anther and pollen development. Mutagenesis can be accomplished by exposure to chemical mutagens such as ethyl methanesulfonate (EMS) [12], by treating with ionizing radiation such as fast neutron [13], and by insertional elements such as Agrobacterium-mediated T-DNA disruption [6] and transposon tagging [14]. EMS mutagenesis is the most widely used technique for mutagenesis with several advantages, including the ease of application, the high rate of mutagenicity with low mortality percentages, the random non-biased distribution of mutation in the genome, and the capability of generation novel mutant phenotypes [15, 16]. However, its utility has been limited by the time-consuming mapping required for identification of the mutation responsible-gene. The availability of high-throughput sequencing techniques such as next generation sequencing will help to overcome such limitation and will facilitate the identification of the mutation location. In contrast to EMS mutagenesis, insertional mutagenesis by T-DNA or transposons usually results in low mutation frequencies per plant; this necessitates screening of large numbers of plants to isolate a mutant [17]. Although the generation of large T-DNA insertion collections is time-consuming, the identification of the mutation location is streamlined and can be conducted easily using PCR-based methods [6, 18, 19]. However, through recent large-scale genetic screens using T-DNA insertional mutagenesis, it was found that a high number of T-DNA insertions were not directly linked to the genes causing the phenotypes observed. Instead, single to multiple nucleotide substitutions, deletions, or insertions were responsible for the phenotypes [20, 21].

EMS induces chemical modifications and alkylation of guanine (G) nucleotides resulting in base mispairing; G will pair with thymine (T) instead of cytosine (C) and during DNA repair, an

amino acid change will occur (G-C pair will be replaced with A (adenine)-T pair) [12]. Occasionally, EMS induces replacement of G-C to C-G, G-C to T-A or A-T to G-C [22]. In EMS Mutagenesis of Arabidopsis, the M1 generation is allowed to self-fertilize and the mutants are screened and isolated among M2 generation. For the recessive mutants, a ratio of 7:1 (divided as 4:0 or 3:1) in an M2 population is expected, depending on the sector (+/+ or m/+) from which the seeds were collected (Fig. 1-1).



Fig. 1-1 Mutagenesis and segregation of recessive mutations. After treatment with a mutagen (e.g., EMS), seeds are collected and germinated to obtain M1 mutagenized plants. The M1 generation is allowed to self-fertilize and the resulting M2 generation are screened for mutants. In case of the recessive mutations, a ratio of 7:1 is expected in Arabidopsis. [Cited from Page D, Grossniklaus U. (2002) *Nat Rev Genet.* 3(2):124-36)]

Anther development and dehiscence

In *A. thaliana*, as in most flowering plant species, anther development initiates with the emergence of the stamen primordia in the meristemic floral third-whorl. In the stamen primordia, multiple cell-specification and differentiation events result in the formation of mature anther cell types (the epidermis, endothecium, middle layer, and tapetum) and generate the characteristic morphology of the anther and the filament [23]. In *A. thaliana*, the anther has a four-lobed structure inside which the pollen grains develop (Fig. 1-2). During development, the filament elongates, the anther enlarges and expands, the tapetum and middle layer degenerate, and the anther starts a dehiscence program [23].

Anther dehiscence (Fig. 1-2) is a multiple-step process that concludes with the release of pollen grains. At the beginning, expansion of the endothecium and accumulation of lignified materials in the walls of the endothecial cells occur [24]. Next, septum and stomium cells go through a cell degeneration program. The septum degenerates first, creating a bilocular anther. Then, a longitudinal weakness in the epidermis breaks at the stomium region of the anther wall. Finally, retraction of the anther wall leads to the full opening of the stomium and pollen release.



Fig. 1-2 Schematic representation for Anther structure and dehiscence. A An illustration of an anther showing its structure before and after dehiscence. **B** An illustration of transverse sections of anthers showing the key steps in dehiscence process. A, anther; F, filament; Ov, ovary; P, petal; PG, pollen grains; S, sepal; Sg, stigma; Sm, septum; St, stomium. [Adopted from Sanders PM *et al.* (2000) *The Plant Cell* 12 (7):1041-1061]

Pollen development in Arabidopsis

The development of the male gametophyte (pollen) (Fig. 1-3) is initiated in the anther when the callose-encased microsporocyte (diploid) goes through two meiotic divisions and cytokinesis to form haploid microspores arranged in a tetrad. The microspores are released from tetrads by the act of callase, an enzyme contributed by the tapetum. The microspores enlarged in size and produced a single large vacuole (vacuolated microspores). This large vacuole pushes the microspore nucleus to migrate to the periphery of the cell. Subsequently, each microspore undergoes the first mitosis (pollen mitosis I) to generate bicellular pollen with a generative cell (small) surrounded by a large vegetative cell. Then, the small generative cell undergoes a second mitosis (pollen mitosis II) to complete the last stage of pollen development and generates two sperm cells [25]. Upon pollination, the two sperm cells migrate to the ovule in the female gametophyte to form the zygote.



Fig. 1-3 Pollen development in *A. thaliana* **showing major division events.** The meiocyte divides meiotically to produce four identical microspores. Microspores undergo two mitotic divisions to produce mature pollen having two sperm cells and one vegetative cell. [Cited from Twell D, et al. (1998) *Trends in Plant Science* **3** (8):305-310]

Pollen wall stratification and pollen wall development in Arabidopsis

The *Arabidopsis* pollen wall has an architecturally-complex structure (Fig. 1-4) consists of two layers; the inner pectocellulosic-based intine and the outer sporopollenin-based exine. The exine wall covers the entire pollen surface except for apertures (the places specified for tube germination) where it is absent or greatly reduced. The exine wall is divided into two layers; inner nexine and outer sexine, which is further, subdivided into two structures the bacula and the tectum [26]. Both structures are responsible for the characteristic and taxon-specific architecture of the exine which is reticulate in *A. thaliana*. The nexine is composed of two layers, an outer nexine I, which represents the base for the bacula, and an inner nexine II [27, 28]. Additionally, in dry stigma species including *A. thaliana*, a lipid-based pollen coat is formed as a third wall component covering the exine layer.



Fig. 1-4 Pollen wall structure in *A. thaliana*. Mature *Arabidopsis* pollen grains have a typical wall consisting of inner intine (surrounding the plasma membrane) and outer exine. Exine comprises two layers, sexine (subdivided into tectum and bacula) and nexine (subdivided into nexine I and nexine II). At the last stage, pollen coat materials fill the spaces in between the bacula. [Cited from Suzuki T, *et al.* (2008) *Plant and Cell Physiology* 49(10): p. 1465-1477]

Pollen wall development requires contributions from both the sporophytic and gametophytic tissues. For example, the exine and pollen coat constitutes are derived from the surrounding sporophytic tapetum cells while the intine is manufactured by microspore itself (gametophytic origin) [26]. Pollen wall development (Fig. 1-5) can be summarized as follows. At the tetrad stage, the four haploid microspores are surrounded by the callose wall. Next, primexine layer (matrix of polysaccharides) is deposited between the microspore plasma membrane and the callose wall. Then, the microspore plasma membrane develops an undulated structure. The undulated membrane represents anchoring sites of sporopollenin forming what known probacula and protecta [28-30]. After the microspore surface forming the bacula and the tectum. By the end of this stage, a pectocellulosic intine is developed around the microspores [31]. The bacula, tectum, and intine continue in development in the bicellular stage. At the tricellular stage, pollen coat materials (also known as tryphine or pollen kitt) fill the space in between bacula giving the pollen wall its characteristic structure.



Fig. 1-5 Current model of Arabidopsis pollen wall development. Sporopollenin synthesis occurs in the tapetum and starts to accumulate around the callose wall surrounding the microspores at the tetrad stage. Primexine deposition, probacula, and protectum polymerization occur subsequently on the microspores. At the uninucleate stage, true bacula and tectum are formed and the intine layer is initiated. At the bicellular stage, bacula and tectum become longer and intine layer continues in growing. At the tricellular stage, pollen coat (tryphine) is deposited as a final pollen wall component. [Adopted from Zhang D, *et al.* (2016) *Subcell Biochem.* 86:315-37]

Coat protein complex II (COPII) assembly and vesicle formation

The COPII vesicle formation includes the sequential recruitment of five proteins, Sar1, SEC23/24, and SEC13/31 [32, 33] (Fig. 1-6). The COPII vesicle formation initiated by recruitment of Sar1 to the ER membrane via the activity of its guanine nucleotide exchange factor SEC12 [34, 35]. Activated Sar1 further recruits the SEC23/24 complex (the "inner" coat) by the direct interaction with SEC23 and forms a "prebudding complex" [36]. The prebudding complex captures the cargo protein and initiates vesicle curvature. Then, SEC13/31 heterotetramer (the "outer" coat) is finally recruited onto the prebudding complex, by the interaction between SEC31 and SEC23, to complete the vesicle formation process by promoting further membrane curvature and fission [37, 38].



Fig. 1-6 Schematic representation of COPII vesicle formation. The process of vesicle formation is accomplished by the sequential recruitment of five proteins, Sar1, SEC23/24, and SEC13/31. Lipids and proteins assembled in the detached COPII vesicles are shuttled from the ER to the Golgi. [Adopted from D'Arcangelo JG, *et al.* (2013) *Biochimica et Biophysica Acta - Molecular Cell Research* 1833 (11):2464-2472]

Gateway cloning technology

In recent years, the Gateway cloning system [39] has proved to be extremely useful for cloning of foreign genes in high-throughput investigations and for constructing large cDNA libraries. Gateway cloning technology is based on a specific site recombination technology in which the integration and excision of λ phage DNA into and from *E.coli* genome occurs through two reversible clonase reactions named BP reaction and LR reaction [39, 40]. The *att*P sites of the λ phage are recombined with the *att*B site of the *E.coli* in a BP reaction to obtain the integrated λ phage genome flanked by *att*L and *att*R. This reaction is catalyzed by the BP Clonase enzyme mix. In the LR reaction, the phage DNA is excised from the bacterial chromosome by recombination between *att*L and *att*R sites. The LR reaction is catalyzed by the LR clonase enzyme mix (Fig. 1-7).



Fig. 1-7 Outline of Gateway cloning showing the site specific recombination and the BP and LR reactions. [Cited from (Nakagawa T, et al. (2009) Plant biotechnol. 26: 275-284]

The adaptable and streamlined Gateway cloning methodology represents a significant advance over the classical restriction approach and overcomes many of its limitations especially the speed and ease with which recombinant constructs can be generated and the convenience of transfer of these constructs between vectors regardless of their sequence. Increasing the number of *att* recombination signals to six different high specified ones [41] has enabled the simultaneous subcloning of multiple DNA fragments in a single LR reaction and has made the recombination of multiple expression elements such as promoters, ORFs, terminators, and reporters much easier [42]. The MultiRound Gateway technology [43, 44] and the Gateway recycling cloning system [45] have been developed as alternative applications of multiple *att* sites. These systems enable the step-by-step repetitive cloning of an expression cassette into a vector to make a multi-gene binary construct using multiple rounds of LR reactions. Although these are outstanding methods to clone an unlimited number of expression cassettes into a binary vector, they are limited by the laborious traditional cloning steps required to prepare a promoter:ORF construct on a prerequisite donor vector.

Stomatal development

The development of stomata is an ideal model for examining intra- and intercellular signaling networks, cell polarity, and cell-type differentiation [46]. In *A. thaliana*, the development of stomata goes through a specialized cell lineage (Fig. 1-8), which consists of the following five cell types; meristemoid mother cells (MMCs), meristemoids, stomatal lineage ground cells (SLGCs), guard mother cells (GMCs) and guard cells (GCs) [47-49]. All stomata are developed through at least one asymmetric and one symmetric division. A protodermal cell turns to MMCs which, in turn, goes an asymmetric division to produce a small triangular meristemoid cell and a larger cell called SLGC. Meristemoid divisions are called amplifying divisions and can occur up to three times or four times [50, 51]. The meristemoid cells lose their stem cell activity and develop into GMCs (characterized by their oval shape). A GMC divides once symmetrically to yield two GCs.



Fig. 1-8 The stomatal lineage development in *Arabidopsis*. A protodermal cell differentiates to a meristemoid mother cell (MMC) or a further pavement cell. MMCs divide asymmetrically to form a meristemoidal cell and a stomatal-lineage ground cell (SLGC). The meristemoidal cell undergo a limited number of asymmetric amplifying divisions and eventually a guard mother cell (GMC) is formed. The GMC divides once symmetrically to two identical guard cells (GCs). [Cited from Pillitteri L and Torii K (2012) *Annual review of plant biology* 63, 591-614]

Screening for mutations affecting anther/pollen development and pollen wall formation in *Arabidopsis thaliana*

AtSEC23A and AtSEC23D, two Arabidopsis COPII components, are essential for pollen wall development and exine patterning

Development of an R4 dual-site gateway cloning system for simultaneous cloning of two desired sets of promoters and open reading frames in a binary vector for

plant research

A dual-site gateway cloning system for simultaneous cloning of two genes for plant

transformation

Proposed conclusions

The author proposes the conclusions of this thesis as follows:

(i) Several mutants with a range of defects in anther/pollen development and exine wall formation were recovered, some of them showing novel phenotypes.

The large scale screening ($\sim 10,000$) using scanning electron microscopy at a high-magnification level was successful to isolate multiple mutant plants defective in anther/pollen development and exine wall formation. This screening strategy was effective to isolate mutants with mild defects in the exine surface structure that would be neglected with using other screening strategies e.g., visual screen by the naked eye, dissecting microscope-based, or confocal microscope-based screens.

(ii) Twenty-three mutations were successfully mapped to specific regions of particular chromosomes of the Arabidopsis genome.

The identification of the mutation location was performed by bulked-segregant analysis using a high-throughput sequencing technique (next generation sequencing) rather than using the traditional mapping methods. Only in one case, the identification was identified by capillary sequencing in the *quartet-like* mutant.

(iii) The isolated mutants provide an additional resource for plant researchers.

Characterization and analyzing of the genes responsible for these mutants will help to dissect the processes of anther/pollen development and exine wall formation.

(iv) *AtSEC23A* and *AtSEC23D* are required for proper pollen wall formation, exine patterning, and tapetum development.

Single *atsec23a and atsec23d* mutant plants exhibited a compromised pollen wall formation with less sporopollenin deposition in the outer exine layer. Double *atsec23ad* mutant plants showed defective pollen walls, and altered development of tapetal cells with morphological abnormalities of the ER, Golgi, elaioplasts, and tapetosomes.

(v) *AtSEC23A* and *AtSEC23D* may organize pollen wall development and exine patterning by regulating ER export of lipid and proteins necessary for pollen wall formation in tapetal cells.

AtSEC23A and *AtSEC23D* were highly expressed in the tapetal cells. *AtSEC23A* and *AtSEC23D* are functional COPII proteins exhibiting the characteristic COPII localization at ER exit sites. To

investigate the expression and co-localization patterns of *AtSEC23A* and *AtSEC23D* necessitate developing new systems to deliver the two expression cassettes simultaneously.

(vi) The developed vector systems are valuable experimental tools with multiple applications in plant research.

The author successfully developed two new vector systems named R4DS Gateway cloning system and DS Gateway cloning system. These two systems have multiple flexibility features, e.g., multiple types of tags and 4 kinds of plant selection markers. Both systems are effective multipurpose tools in plant research, enabling the simultaneous co-expression of two genes, the co-localization analysis of two proteins, and facilitating the investigation of protein-protein interactions *in vivo* by BiFC or co-immunoprecipitation.

(vii) The developed vector systems are suitable for exploring the functions and molecular mechanisms underlying various biological processes including pollen development and exine wall formation.

The R4DS cloning system was successfully used in exploring the expression and co-localizations patterns of two Arabidopsis genes required for pollen wall development, *AtSEC23A* and *AtSEC23D*, as well as in the co-localization of each of *AtSEC23A* and *AtSEC23D* with other organelle markers (e.g., endoplasmic reticulum, Golgi, and plastid markers).

Summary

The research described in this thesis aimed to improve our understandings of the important processes of pollen development and exine wall formation and to isolate and characterize new genes involved in such complicated processes. Vast numbers of genes work cooperatively to regulate various biological processes including pollen development and exine wall formation. In order to understand the functions and molecular mechanisms underlying these processes, analyses of transgenic plants that concomitantly express two protein-coding genes are often required.

In Chapter 1, the author presents a general introduction to the work. Chapter 2 describes a high-magnification level screening aimed to isolate new genes affecting anther and pollen development in Arabidopsis thaliana. Unlike most of the previous screens which were mainly based on isolating male sterile and semi-sterile plants neglecting a large number of mutants with severe to mild defects in pollen structure, however, still fertile; the screen described here using scanning electron microscopy could isolate such mutants. Several mutants with a range of defects in anther/pollen development and exine wall formation were recovered; some of them exhibit novel surface-structure phenotypes. A total of one hundred and one mutant plants were isolated (some of them were found to be allelic) and classified according to their phenotypic characters into three classes including multiple subclasses and types. The identification of the mutation location was performed mainly by the bulked-segregant analysis using next generation sequencing rather than using the traditional mapping methods. Twenty-three mutations were successfully mapped to specific regions of particular chromosomes of the Arabidopsis genome; thirteen of them were linked to specific genes. This work illustrates that high magnification level screens (i.e., using scanning electron microscopy) are highly beneficial and may even be required to isolate mutants with novel phenotypes, (i.e., most mutants with severe defects that can be found by naked-eye visual screen or dissecting microscope-level screen most likely will be alleles of previously identified genes). Characterization of the genes responsible for mutants isolated in this screen will expand our knowledge of the molecular mechanism underlying several vital processes in pollen development, including pollen wall formation, sporopollenin synthesis, polymerization and transport, and exine patterning. This screen provides an additional resource for plant researchers to analyze functions of genes involved in anther/pollen development and exine wall formation processes and will help unrevealing previously unknown players in such processes.

COPII proteins are involved in lipid and protein transport events from the endoplasmic reticulum to the Golgi in the early secretory pathway. **Chapter 3** provides phenotypic characterization of two components of the Arabidopsis coat protein complex II (COPII) proteins, *AtSEC23A* and *AtSEC23D*, and shows how these genes are involved in regulating plant growth and development. The study revealed that both *AtSEC23A* and *AtSEC23D* are essential for normal pollen

and tapetum development. Single *atsec23a and atsec23d* mutant plants, although normally fertile, showed a compromised pollen wall formation with less sporopollenin deposition in the outer exine layer. Double *atsec23ad* mutant plants were semi-sterile and exhibited a wide range of developmental defects in pollen and tapetal cells. Pollen grains of *atsec23ad* plants had defective exine and intine layers and their cytoplasm have shown signs of degeneration or leakage. Moreover, the development of tapetal cells in *atsec23ad* plants was altered, with structural abnormalities in inner organelles (i.e., ER, Golgi, elaioplasts, and tapetosomes). Both AtSEC23A and AtSEC23D exhibited the characteristic localization pattern of COPII proteins, indicating that they are functional COPII components. The results indicated that *AtSEC23A* and *AtSEC23D* may only share partial redundant functions and may have differing substrate preferences. Thus, this work suggests that *AtSEC23A* and *AtSEC23D* organize pollen wall development and exine patterning by regulating ER export of lipid and proteins necessary for pollen wall formation in tapetal cells. This work provides a direct link between the tapetum early secretory pathway and the processes of sporopollenin deposition and exine patterning. In addition, the results presented in this Chapter shed light on the functional heterogeneity of SEC23 homologs in plants.

To better understand the functions and molecular mechanisms underlying various biological processes including pollen development and exine wall formation, developing new systems to deliver two genes simultaneously are often required. For example, exploring the expression patterns and co-localizations of the two SEC23 proteins described in Chapter 3 required utilizing vector systems that enable simultaneous cloning of two expression cassettes. For these reasons, two vector systems allowing simple cloning of two expression cassettes simultaneously in plant cells were developed and presented in Chapter 4 and Chapter 5. Chapter 4 describes a vector system that enables the simultaneous cloning of two desired sets of promoters and open reading frames, named the R4 dual-site gateway cloning system. Chapter 5 presents a vector system allows the simultaneous cloning of two genes under control of the moderate nopaline synthase promoter, named the dual-site gateway cloning system. These vectors have been used successfully not only in exploring the expression patterns and co-localizations of the two SEC23 proteins but also in the co-localization of each of AtSEC23A and AtSEC23D with other organelle markers (e.g., endoplasmic reticulum, Golgi, and plastid) as described in Chapter 3. These vectors also facilitate the protein-protein interaction analysis by bimolecular fluorescence complementation (BiFC) assay and co-immunoprecipitation. Investigating protein-protein interactions is required in many cases to dissect the functions and molecular mechanisms underlying multiple biological processes including pollen development and exine wall formation. In addition, the two developed vector series are versatile cloning tools with multiple flexibility features, including multiple types of tags and 4 kinds of plant selection markers. Both cloning systems will be invaluable experimental tools with multiple

applications in plant research, e.g., analysis of co-localization patterns of transgenes and investigation of protein-protein interactions by BiFC or co-immunoprecipitation.

In the **final Chapter**, the author provides proposed conclusion remarks and discusses the findings presented in this study.

In sum, the large scale screening (~ 10,000) using scanning electron microscopy at a high-magnification level was successful to isolate mutants with mild defects in the exine surface structure that would be neglected with using other screening strategies. Multiple mutants with novel phenotypes were recovered. Two genes, AtSEC23A and AtSEC23D, were characterized in detail; both were found to be required for proper pollen wall development and exine patterning and they may function by regulating ER export of lipid and proteins necessary for pollen wall formation in the tapetal cells. Two novel vector systems were developed for analysis of the genes revealed in this study. These systems are beneficial tools in exploring the functions and molecular mechanisms underlying various biological processes including pollen development and exine wall formation.

"In addition, some of the figures, etc. have been omitted."