A digital quantitative and spatial reference atlas from microscopic 3D images of ovules in *Arabidopsis thaliana*

TECHNISCHE UNIVERSITÄT MÜNCHEN

Dissertation for Doktors der Naturwissenschaften
(Dr. rer. nat.)

Athul Remadevi Vijayan

TECHNISCHE UNIVERSITÄT MÜNCHEN

TUM School of Life Sciences

A digital quantitative and spatial reference atlas from microscopic 3D images of ovules in *Arabidopsis thaliana*

Athul Remadevi Vijayan

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Frank Johannes
Prüfer der Dissertation:	1. Prof. Dr. Kay Heinrich Schneitz 2. Prof. Dr. Alexis Maizel

Die Dissertation wurde am 06.12.2021 bei der Technischen Universität München eingereicht und durch die TUM School of Life Sciences
am 02.05.2022 angenommen

Table of Contents

Lists of figures

Introduction

Materials and Methods

Results

Appendix

Lists of tables

Summary

How an organ attains different shapes and sizes at different developmental stages is an interesting question in biology. This study uses Arabidopsis ovules as a model organ to understand cellular patterns underlying organ morphogenesis in plants. The Arabidopsis ovule arises as a simple finger-like outgrowth from the placental surface, and undergoes complex morphogenesis to form a final curved structure at maturity. This study approaches the shape transformation that occurs during ovule development as a process that requires closer examination in three-dimensional space with detailed qualitative and quantitative cellular descriptions of the organ at different developmental stages using microscopic imaging.

This study resulted in several major methodological improvements that allow precise microscopic imaging followed by instance cell segmentation for quantitative analysis of the 3D cellular architecture of the ovule. Despite the existence of several imaging techniques to follow Arabidopsis ovule development, no easy and straightforward method existed for examining the morphology of internal elements in the ovule at the cellular level. In a first step, I developed an improved method for precise microscopic imaging of the entire organ with fine cellular resolution. In a collaborative effort, the high-resolution images from this study were instrumental to form a convolutional neural network-based image segmentation pipeline “PlantSeg”. The pipeline resulted in a precise instance cell segmentation of the images, it allowed fast processing of image dataset that eventually resulted in a large dataset of Arabidopsis ovules for further analysis. This study also focused on 3D instance nuclei segmentation from challenging nuclei stain images. Applying a novel approach this study came up with significant improvements in available methods for nuclei segmentation using machine learning bases training performed on nuclei stain images. This study has additionally provided a new method for tissue classification using cell layers detection in ovule primordia which can be also applied in other layered organs.

The 3D ovule dataset provided a baseline for further exploration of 3D organ morphogenesis. By being able to analyze the architecture of tissues in ovules with a proper 3D cellular resolution this study could identify previously undefined regions within the organ. Detailed investigation of ovule primordia development resulted in the identification of slanting, the first morphological manifestation of polarity in the early development of ovules and how the slant is translated into the final orientation of the

ovule with the micropyle pointing towards the apex of the pistil. The chalaza of the Arabidopsis ovule could be subdivided into two distinct groups of cells.

This study has obtained a detailed description of different events and features of integument development and how these might be related to the formation of the final curved structure of the organ. A highlight feature is the polarity and differential growth of integument layers along the proximal-distal axis of the organ. Characteristic events during the differential outgrowth are described, most importantly, the outer integument outgrowing the inner integument, the distal enlargement of integument cells, and their differences in internal layers. These events suggest that the outer integument imposes curvature on internal tissues. The detailed 3D analysis identified the exposed cells of the inner layer of the outer integument with differential characteristics and how they are linked to the outer integument interacting with inner integument growth. 3D cell surface analysis of internal tissues reveals the morphological changes induced by contact friction. Quantitative analysis of growth and proliferation rates in integuments reveals different growth pulses during differential outgrowth of integuments and most of the growth pulses correlate with the described events of integument outgrowth. Overall, the study proposed several cellular features of the organ and suggest that the ovule curvature is a complicated multi-step process shaping cells to form these features and that happens in three-dimensional space and time.

Finally, I played a central part in a collaborative effort that led to the establishment of 3DCoordX, a computer pipeline implementing a spatial coordinate system for the annotation and analysis of 3D organ morphogenesis. 3DCoordX can be applied to various organs of different morphological complexity. The pipeline defines spatial coordinates or position of cells in the tissue, proximal-distal, and medial-lateral axis. Taking advantage of 3DCoordX I mapped mitotic cells from the young ovule primordia to the organ coordinates to understand their spatial distribution. I found that most cell divisions were located in the proximal half of the outgrowing primordium indicating the existence of a spatially restricted cell proliferation zone in the young primordium. Moreover, I established polar cell geometries for cells of the integuments.

Overall, this study has contributed to valuable new tools and techniques for studying organ morphogenesis. It then used these tools to extensively understand 3D cellular organization in wild-type Arabidopsis ovule development with fine details. This now stands as a benchmark for understanding various features displayed by the organ at

different developmental stages and how that correlates with the tissue and organ shapes at their respective stages of development.

Zusammenfassung

Wie ein Organ in verschiedenen Entwicklungsstadien unterschiedliche Formen und Größen annimmt, ist eine interessante Frage in der Biologie. In dieser Studie wird die Samenanlage oder Ovule von Arabidopsis als Modellorgan verwendet, um die zellulären Muster zu verstehen, die der Organmorphogenese in Pflanzen zugrunde liegen. Die Ovule von Arabidopsis entsteht als einfacher fingerartiger Auswuchs aus der Plazentaoberfläche und durchläuft eine komplexe Morphogenese, um schlussendlich eine endgültige gebogene Struktur zu bilden. Diese Studie betrachtet die Formveränderung während der Entwicklung der Samenanlage als einen Prozess, der eine genauere Untersuchung im dreidimensionalen Raum mit detaillierten qualitativen und quantitativen zellulären Beschreibungen des Organs in verschiedenen Entwicklungsstadien unter Verwendung mikroskopischer Bildgebung erfordert.

Diese Arbeit führte zu mehreren wichtigen methodischen Verbesserungen, die eine quantitative Analyse der zellulären 3D-Architektur der Ovule ermöglichen. Obwohl es mehrere bildgebende Verfahren gibt, um die Entwicklung der Ovule zu verfolgen, gab es keine einfache und unkomplizierte Methode, um die Morphologie der inneren Elemente der Ovule auf zellulärer Ebene quantitativ zu untersuchen. In einem ersten Schritt habe ich eine verbesserte Methode zur präzisen mikroskopischen Abbildung des gesamten Organs mit feiner zellulärer Auflösung entwickelt. Die Segmentierung von Zellen in diesen mikroskopischen Bildern mit Zellkonturen ist eine weitere Herausforderung im Bereich der quantitativen Organmorphogenese. Die Verbesserung der mikroskopischen Bildgebung von Ovulen führte zu einer weiteren Möglichkeit, die Werkzeuge für die Nachbearbeitung der mikroskopischen Rohbilder zu erweitern, um die Zellsegmentierung zu verbessern. Diese Bilder dienten als wertvoller Trainingsdatensatz für das maschinelle Lernen. Schließlich wurde in einer interdisziplinären Kollaboration eine auf neuronalen Netzen basierende Bildsegmentierungspipeline "PlantSeg" eingerichtet, die die hochauflösenden Bilder aus dieser Studie nutzt. Diese Pipeline führte zu einer präzisen Segmentierung der Zellen und ermöglichte eine schnelle Verarbeitung der Bilddaten, was wiederum dazu beitrug, die Anzahl der Proben für die Analyse der Ovulen zu erhöhen. Ein weiterer Schwerpunkt dieser Studie war die 3D-Segmentierung von Zellkernen aus einfachen Kernfärbebildern. Dies war eine äußerst schwierige Aufgabe, da keine geeigneten

Werkzeuge für die Segmentierung solcher schwachen Kernfärbungen verfügbar waren. Durch die Anwendung eines neuartigen Ansatzes wurden in dieser Studie erhebliche Verbesserungen der verfügbaren Methoden zur Kernsegmentierung mit Hilfe von maschinellem Lernen erzielt, das auf Bildern von Kernfärbungen trainiert wurde. Diese Studie hat außerdem eine neue Methode zur schnellen Erkennung von Zellschichten in Ovulenprimordien geliefert, die auch bei anderen geschichteten Organen angewendet werden kann.

Anschließend leistete ich wichtige Beiträge zu einem Datensatz, der eine Entwicklungsreihe von 3D-zellsegmentierten Ovulen von Arabidopsis enthält. Dieser Datensatz bildete die Grundlage für die weitere Erforschung der 3D-Organmorphogenese. Durch die Möglichkeit, die Architektur der Gewebe in den Samenanlagen mit einer angemessenen zellulären 3D-Auflösung zu analysieren, konnten in dieser Studie bisher nicht definierte Regionen innerhalb des Organs identifiziert werden. Die detaillierte Untersuchung der Entwicklung des Ovulenprimordiums führte zur Identifizierung der Schrägstellung (slanting), der ersten morphologischen Manifestation der Polarität in der frühen Entwicklung der Ovule. Die Chalaza der Ovule konnte in zwei verschiedene Zellgruppen unterteilt werden. Die quantitative Analyse der zellulären Merkmale der digitalen 3D-Ovulen führte zu einem besseren Verständnis des unterschiedlichen Wachstums der Integumentgewebe und der Polarität im zellulären Muster entlang der proximal-distalen Achse des Organs. Diese Studie hat sich auch auf die 3D-Tiefenanalyse der Zellmorphologien an den Oberflächen der inneren Gewebe ausgeweitet und zeigt, wie diese mit der Kontaktreibung zwischen den Integumentgeweben und dem Nuzellus korreliert werden können.

Schließlich war ich maßgeblich an der interdisziplinären Entwicklung von 3DCoordX beteiligt, einer Computerpipeline zur Implementierung eines räumlichen Koordinatensystems für die Annotation und Analyse der 3D-Organmorphogenese. 3DCoordX kann auf verschiedene Organe unterschiedlicher morphologischer Komplexität angewendet werden. Die Pipeline ermöglicht eine umfassende Analyse des Organs in 3D mit einer räumlichen Auflösung von Zellen, die im Wesentlichen die Position der Zellen im Gewebe, die proximal-distale und die medial-laterale Achse definiert. Mit Hilfe von 3DCoordX habe ich mitotische Zellen aus den jungen Ovulenprimordien auf die Organ-Koordinaten abgebildet, um ihre räumliche Verteilung

zu verstehen. Ich fand heraus, dass die meisten Zellteilungen in der proximalen Hälfte des auswachsenden Primordiums stattfanden, was auf die Existenz einer räumlich begrenzten Zellproliferationszone im jungen Primordium hindeutet. Außerdem habe ich polare Zellgeometrien für die Zellen der Integumente ermittelt.

Insgesamt hat diese Studie zu wertvollen neuen Werkzeugen und Techniken für die Untersuchung der Organmorphogenese beigetragen. Anschließend wurden diese Werkzeuge eingesetzt, um die zelluläre 3D-Organisation in der Entwicklung der Wildtyp-Arabidopsis-Eizelle bis ins kleinste Detail zu verstehen. Dies dient nun als Maßstab für das Verständnis der verschiedenen Merkmale, die das Organ in den verschiedenen Entwicklungsstadien aufweist, und wie diese mit den Gewebe- und Organformen in den jeweiligen Entwicklungsstadien korrelieren.

1 Introduction

1.1 Organ Morphogenesis

Cellular growth and division happen in space and time to form an organ. Organ morphogenesis is a complicated process that involves spatiotemporal coordination of cellular and tissue growth and proliferation that results in a multicellular organ of a specific shape and form. How organs of different architecture are formed by a group of initial cells and how they attain their required size and shape at different developmental phases in a reproducible fashion is the basic question in morphogenesis (Gibson et al., 2011; Coen and Rebocho, 2016; Coen et al., 2017; Vijayan et al., 2021)

Unlike animal tissues where cells can move around in space, cells in plant organs are encased by the cell wall and are fixed in their position, preventing any movement (Coen et al., 2004; Coen et al., 2017). Due to the immobility of cells in plant tissues, a combination of key factors including cell cycle control, differential growth and proliferation of tissues and orientation of the division plane, is involved in defining cellular organization in space and time (Meyerowitz, 1997; Huang et al., 2018; Vijayan et al., 2021).

1.2 Arabidopsis as a model organism

Arabidopsis is widely used as a plant model organism; it is an angiosperm from the mustard family (Brassicaceae). It has a rapid life cycle (about 6 weeks from germination to mature seed) where it grows up to 20-25 cm at maturity, It can be easily grown at high density in a glasshouse or culture room. Self-pollination allows easy propagation, thus making it an attractive model organism for experiments. A single plant can produce thousands of seeds. It has several traits that make it a powerful model for understanding the molecular, genetic, and cellular biology of angiosperms. Thus Arabidopsis offers important advantages for basic research in genetics, molecular and developmental biology (Coen and Meyerowitz, 1991; Hülskamp et al., 1994; Jürgens et al., 1991; Masson et al., 2002).

The *Arabidopsis thaliana* life cycle can be divided into two phases: vegetative and reproductive (Kwiatkowska, 2008; Telfer et al., 1997) . During vegetative phase, the shoot apical meristem produces a compact spiral/ rosette of leaves, and the stem

elongates rapidly. In the subsequent period of growth, the SAM becomes a dome-like structure and produces flower buttresses at the periphery. As development proceeds, floral buttresses turn into a floral meristem which later forms the flower. This phase is known as the reproductive phase. These flowers have both female and male reproductive organs (Figure 1.1 A) hence, they can self-pollinate and self-fertilize to form seeds. Through artificial means, cross fertilization is also possible.

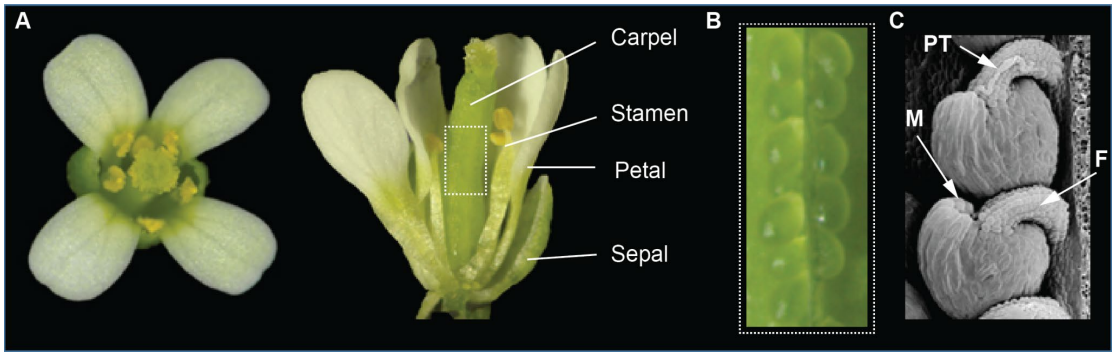

Figure 1.1 Arabidopsis floral organs
Figure representing an Arabidopsis flower containing both male and female reproductive organs (A). Different floral organs are labelled. Box highlights an area of carpel slit open to visualize ovules in (B). Ovules are arranged one below the other and on the two sides of the carpel. A SEM image of an Arabidopsis ovule post fertilization (C). Figure (C) adapted from (Chaudhary et al., 2018). Abbreviations; PT: pollen tube; M: Micropyle; F: Funiculus.

The Arabidopsis flower is composed of four whorls of floral organs (Bowman et al., 1991; Coen and Meyerowitz, 1991). The outermost whorl is composed of sepals and within this whorl are the petals, then the stamens (the male reproductive organs) and finally the innermost whorl is formed by the carpels or female reproductive organs in the center of the flower. Each carpel contains 50-60 ovules which later develop into seeds upon successful fertilization. After seed formation, the growth is terminated. The seed then undergoes dehydration until it gets into the dormant stage; under favorable conditions the seed germinates and the cycle continues.

Arabidopsis thaliana is a suitable system to study organ morphogenesis in plants. Plant cells, being opaque in nature, do not deliver information from deeper tissue layers when subjected to confocal microscopy, resulting in quantitative analysis being limited to organ surfaces or few layers from the surface. Live and time lapse imaging of developing organs has been performed on organ surfaces to study morphogenesis (Barbier de Reuille et al., 2015; Hong et al., 2016; Sapala et al. 2018; Kierzkowski et

al., 2019; Silveira et al. 2021). It gives a good understanding of cellular growth patterns and divisions happening in real time and allows the quantification of growth directions and mapping of daughter cells to parents. This is often performed using a fluorescent reporter for cell outline or with cell wall or membrane stains. Plant organs such as leaf and shoot meristem have been well studied at the organ surfaces in a qualitative and quantitative manner from live time lapse imaging.

1.3 Arabidopsis ovules as model systems for organ morphogenesis

The ovules are the female reproductive organs in plants. In Arabidopsis, they provide an excellent model system to study organogenesis. The Arabidopsis ovule has been established as a model to study several important aspects of tissue morphogenesis. This mainly includes primordium formation, female germ line development and integument development (Schmidt et al., 2015; Chaudhary et al., 2018; Nakajima, 2018; Gasser and Skinner, 2019; Hernandez-Lagana et al., 2020). Ovules show a highly stereotypic development with about 50 ovules per carpel. This allows a large sample size for quantitative or qualitative analysis. Ovule ontogenesis has been well described during the last decades of research (Mansfield et al., 1991; Robinson-Beers et al., 1992; Modrusan et al., 1994; Schneitz et al., 1995; Grossniklaus and Schneitz, 1998; Schneitz et al., 1997; Drews et al., 1998; Schneitz et al., 1998; Gasser et al., 1998; Schiefthaler et al., 1999; Truernit and Haseloff, 2008; Cucinotta et al., 2014). A morphology-based identification of different developmental stages of Arabidopsis ovules and their staging system has been proposed (Schneitz et al., 1995).

Arabidopsis ovule is an organ of complex architecture. It undergoes several morphological changes during growth and development. Distinct processes are involved during its development: identity specification, initiation and outgrowth, pattern formation and morphogenesis. The ovule primordium is a simple and straight structure that grows out from the surface of the placenta. Integuments are lateral determinate organs that get initiated from the epidermis of late-stage ovule primordia. Ovule accommodates newly growing integuments and transforms into a more complex-shaped organ. Integuments outgrow around the organ core in 3D, forming a mass of layered tissues that forms the final curved shape of the ovule. Ovule primordia, at a very early developmental stage can be observed as a radially symmetric structure. At maturity, ovule loosens radial symmetry and becomes close to a bilaterally symmetric organ (Vijayan et al., 2021). This transformation from a simple form to a curved

complex shape in a stereotypical manner makes the ovule a unique plant organ to study morphogenesis in 3D.

1.4 Arabidopsis ovule developmental stages

Morphological and qualitative descriptions of ovule development have been performed on 2D sections revealing the major changes happening during development of this organ (Robinson-Beers et al., 1992; Schneitz et al., 1995; Christensen et al., 1997). Ovule stages are also defined according to morphological events that happen during development (Schneitz et al., 1995)

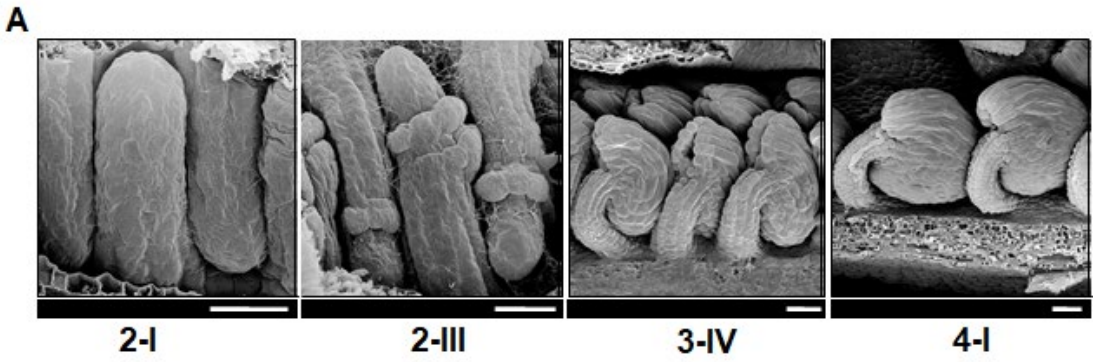

Figure 1.2 SEM images of different stages of Arabidopsis ovule development
Different stages of Arabidopsis ovules (A). From left to right: ovule primordia during outgrowth at about stage 2-I with a typical finger like protruded structure. Stage 2-III of ovule development marked by the initiated outer and inner integument. A mature ovule at about stage 3-IV with integuments encapsulating the nucellus and embryo sac leaving the micropylar clef exposed. A mature ovule during fertilization at stage 4-I. Stages are indicated at bottom. Scale bar 20um. Figure adapted from (Chaudhary et al., 2018)

Ovule primordia are initiated by periclinal divisions in the sub epidermal tissue of the placenta. Early-stage ovules (stage 1-I) become visible as they just bulge out from the placental surface. They can be found as prominent finger-like protrusions orthogonal to the placental surface at stage 1-II. Megasporogenesis is initiated by the appearance of a large megaspore mother cell (MMC) at the L2 distal tip of the primordia; presence of an MMC defines stage 2-I of ovule development. After the emergence of the inner and outer integument (stages 2-II and 2-III), the megaspore mother cell undergoes meiosis (stage 2-IV). With the tetrad formation, MMC divides and generates four haploid gametes (stage 2-V) while the cells of the integuments are growing and extending towards the apex of the nucellus. Out of the four haploid gametes, only one survives and is the functional megaspore.

During stage 3, megagametogenesis takes place. During stage 3-I and 3-II, the integuments are almost enclosing the whole nucellus, leaving a clearly visible gap at the distal tip of the integuments, the micropyle. Stage 3-I is defined by a mono-nuclear embryo sac, which forms out of the functional megaspore. The haploid gamete then undergoes three nuclear divisions followed by cellularization. After the first nuclear division, a two nuclear embryo sac is generated (stage 3-II) and with the second one a four nuclear embryo sac that defines stage 3-IV. In between the first and second nuclear divisions, the vacuole appears that marks stage 3-III. With the third mitosis, the cells are taking special positions determined by their role: three antipodals gather at the chalazal end of the embryo sac, the two synergids together with the haploid egg cell assemble at the micropylar end of the ovule and the diploid central cell is positioned in the middle of the embryo sac (stage 3-V). With the fusion of the nuclei in the central cell and the degeneration of the antipodals, megagametogenesis is completed (stage 3-VI).

During stage 4, the development carries on with double fertilization, embryo formation and endosperm development. Arabidopsis embryo sac is protected by two bilayered integuments, the outer and inner integuments, forming a four-layer tissue of cells around the nucellus containing the embryo sac. Fertilization happens within the ovule, where the pollen tube delivers the sperm cell to the embryo sac. Further, the fusion of the sperm with the egg cell contained in the embryo sac results in the formation of diploid embryo, which grows inside the ovule. Eventually the ovule develops into seeds and the integuments develop into a seed coat.

1.5 Integument Development and ovule curvature

Ovules arise as finger-like protrusions from the placental surface. A mature ovule consists of three functional domains along its proximal-distal axis: the proximal funiculus- a stalk-like structure that connects the ovule to the placenta, the central chalaza- from where the integuments initiate and the distal nucellus- which harbors the embryo sac or female gametophyte (Figure 1.3 A).

Integument outgrowth around the nucellus-containing embryo sac is an interesting morphogenetic process during ovule development. *A. thaliana* carries bitegmic ovules, meaning that each ovule presents two distinct integuments, which originate from the chalaza (Jenik and Irish, 2000; Schneitz et al., 1995; Truernit and Haseloff, 2008). Each integument is composed of two cell layers, an inner or adaxial layer and an outer

or abaxial layer, and both are characterized by laminar planar growth. Integuments are arranged as layers of cells that form a curved hood-like outer structure. Inner layers of integuments can be approximated as curved cylinder-like structures (Vijayan et al., 2021). The inner integument initiates earlier than the outer integument and grows as a radially symmetric structure. The outer integument undergoes asymmetric growth, growing more at the gynbasal or posterior side of the ovule, the side that faces the basal end of the carpel (Schneitz et al., 1995; Grossniklaus and Schneitz, 1998; Schneitz et al., 1997; Balasubramanian and Schneitz, 2000; Villanueva et al., 1999). At maturity, the integuments surround the nucellus and retain a small cleft called micropyle, through which the pollen tube reaches the female gametophyte for fertilization. The integuments develop into seed coat after fertilization. The Arabidopsis ovule is anatropous which means that the micropyle lies close towards funiculus resulting in a curvature at 180 degrees.

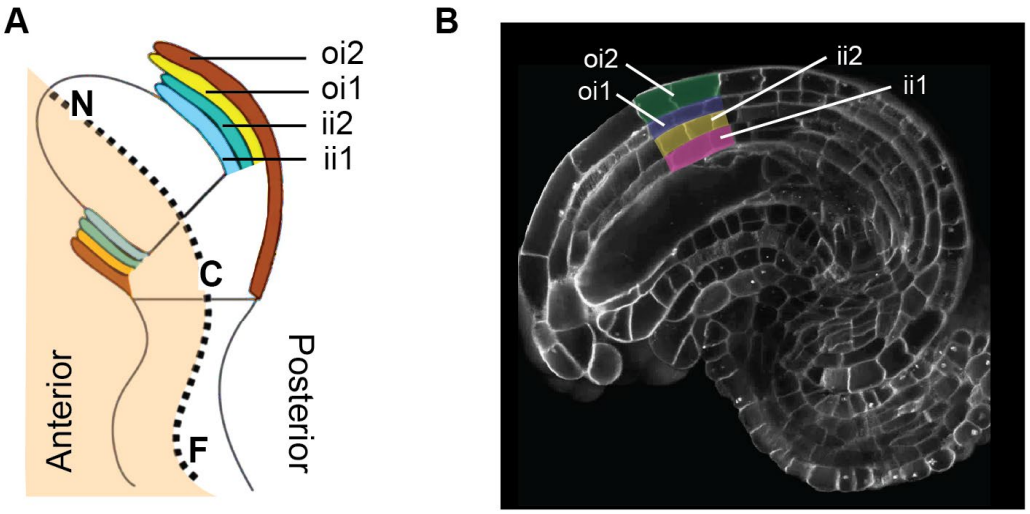

Figure 1.3 Spatial arrangement of integuments in Arabidopsis ovules
A cartoon depicting the mid-sagittal section view of an ovule at about stage 2-V. Integument cell layers are labelled in different colors (A). The central axis of the organ is depicted by a dotted black line; along the axis are the proximal funiculus, central chalaza and distal nucellus. Anterior side of the organ at the sagittal section view is shown with a shaded color. A similar sagittal section view from a 2D confocal microscopic image of an Arabidopsis ovule at a mature stage (B). Cell wall is stained using SR2200 cell wall stain. Integument tissues at the posterior side of the organ are labelled with false color. Abbreviations; oi2: outer layer of outer integument or the abaxial layer of outer integument; oi1: inner layer of outer integument or the adaxial layer of outer integument; ii2: outer layer of inner integument or the abaxial layer of inner integument; ii1: inner layer of inner integument or the adaxial layer of inner integument.

As mentioned before, unlike animal tissues where cells can move around in space, cells in plant organs are encased by the cell wall and are fixed in their position with connected neighbors, preventing any movement (Coen et al., 2004, 2017). The integument offers a complex tissue to study an exception to this connectivity between cells. Arabidopsis ovules are characterized by two layers of outer integument and two layers of inner integument which are initiated separately and there is no physical connection between these two tissues and between inner integument and nucellus (Robinson-Beers et al., 1992, Schneitz et al., 1995). But anatomically the walls between outer integument and inner integument or similarly inner integument and nucellus are in direct contact because they are positioned one on top of the other, but their cell walls are unconnected allowing them a freedom of growth or movement ideally irrespective of each other. This leads to an interesting phenomenon called contact friction. The contact between the tissues creates friction, a higher order friction would reduce the freedom for two tissues to grow independently, and at least friction allows simple slide over of one tissue on top of the other.

Overall, ovule development is a prime example of complex cellular growth and development that forms asymmetrically initiated layered tissues to grow around in 3D space to build a final curved shape. All these features make the ovule and integuments fascinating organs to study morphogenesis in 3D and to understand how a 3D shape is formed.

1.6 Whole-mount 3D microscopic imaging with cellular resolution

Three-dimensional atlas of plant organs have been developed successfully for organs like shoot and root (Montenegro-Johnson et al., 2015; Hong et al., 2018; Refahi et al., 2021; Willis et al., 2016). Generation of such atlases with cellular resolution is challenging but such atlases provide deep information from 3D microscopic images. A quantitative cellular characterization of individual ovule stages would be a much bigger advancement in understanding the cellular patterns in space and time. It gives us precise information of the number of cells or volume of organs and tissues at individual developmental stages and how they are changing over time. A prerequisite for such a high level of understanding whole-ovule development is to generate datasets of 3D microscopic images with z-stack acquisition with precise cellular resolution that allow us to look at the entire organ at different developmental stages.

Arabidopsis ovule morphogenesis has been traditionally studied from 2D imaging of organ sections. Key advances in microscopy have enabled a different extent of resolution to cellular microscopic images, ranging from widefield to confocal laser scanning microscopy to high resolution multiphoton microscopy. Researchers have shown that a full resolution 3D image is essential for complete representation of the organ structure of interest which is now possible with latest confocal and other advanced microscopes with a z-stack acquisition mode (Moreno et al., 2006; Truernit et al., 2008; Fernandez et al., 2010; Yoshida et al., 2014; Bassel et al., 2014; Tofanelli, Vijayan et al., 2019; Vijayan et al., 2021). Quantitative 3D image analysis is another key component of organ morphogenetic study. This involves digital representation of elements or cells in microscopic images as objects that can be used to form realistic 3D models that allow accurate quantification of shape and size of cells or objects of interest in the image. A higher level of understanding growth and proliferation in 3D requires a whole mount and time lapse imaging of all cells of the organ. This is extremely challenging for ovules as no method has been established to perform such timelapse live imaging of ovules. For organs such as shoot, root and sepals, there are available methods that allow the growth of tissue in a growth medium and then subjected to microscopy at different time points. No such growth media for ovules have been previously published. Overall, an analysis of 3D growth over time provided a better understanding of cellular growth and development (Roeder et al., 2011).

To gain better understanding of how ovules are shaped in 3D, it is essential to visualize the 3D surface and interior morphology, arrangement of constituent cells in 3D and different cellular properties. Accurate imaging of the whole 3D organ is still challenging due to tissue penetration in deeper layers. Previous research has shown that the application of various reagents for clearing the samples allows deeper tissue microscopic imaging by making samples transparent for laser transmission. Modified Pseudo-Schiff propidium iodide stain (mPS-PI) has been classically used with fixed and cleared tissues for precise microscopic z-stacks (Truernit and Haseloff, 2008; Truernit et al., 2008). Clearing is performed using Hoyer's medium containing carcinogenic chloral hydrate and cell outlines are stained using propidium iodide according to the mPS-PI method. 3D architecture of simple organs such as roots can be relatively easily generated by this approach (Bassel et al., 2014; Montenegro-Johnson et al., 2015). This method is time-consuming and also incompatible with fluorescent fusion proteins that limit the study to only understand organ structure.

Tissue clearing has also been performed using other reagents like TOMEI (Hasegawa et al., 2016; Musielak et al., 2016), PEA-CLARITY (Palmer et al., 2015) and ClearSee (Kurihara et al., 2015). These new methods promise similar deep imaging of plant tissues as mPS-PI and are compatible with any fluorescent protein or stains of interest. ClearSee is a well-studied reagent among these, it can be combined with various cell wall fluorescent stains; this would avoid the need for a fluorescent cell outline reporter to understand organ structure in 3D. ClearSee can also be combined with a fluorescent reporter for gene expression, which makes it a better candidate for tissue clearing purposes. It can thus be used to analyze the 3D architecture of entire plant organs with cellular resolution and to investigate gene expression patterns and subcellular protein localization using respective reporters.

This study includes a 3D digital ovule atlas of whole organ development from initiation to maturity and that allows us to identify patterns inside and outside the organ which are never recognizable from 2D images.

1.7 Transforming microscopic images into quantitative cellular descriptors

Whole-mount microscopic imaging of different developmental timepoints of the organ provides a complete overview and qualitative understanding of the shape of the organ. This is limited to the organ surface and different 2D clipping planes or sections of interest. A quantitative analysis of these microscopic images is critical in many biological disciplines (Adams et al., 2004; Slice, 2007; Adams, Rohlf & Slice, 2013; Mitteroecker et al., 2013). For instance, the number of cells present in a developmental stage of an organ and the distribution of cells in different tissues and their cellular characteristics could be extracted from the microscopic images after they are subjected to a 3D image segmentation (Montenegro-Johnson et al. 2019). 3D cell segmentation involves creating digital reconstructions of the shape of a cell from its outline, which can then be used to quantify various features of the cell, such as their volume and shape index.

Segmentation methods are broadly of two types, semantic segmentation, and instance segmentation. Semantic segmentation is the kind where each pixel within an image is associated with a category of the objects present in the image. For example, semantic segmentation of a cell outline image would result in the annotation of two object identities that are the cell boundaries and the background. Ideally, every pixel would have a probability map (PMAP) usually ranging from 0 to 1, indicating the probability

of this pixel belonging to the cell boundary or the background. A value close to 1 would refer to the pixel belonging to a cell boundary. The other segmentation method is called instance segmentation (Hafiz and Bhat, 2020); it is essential for bioimages as it further associates each pixel to an independent object within the image. For example, it groups all the pixels within a cell outline into an object label essentially separating 3D cells from one another and from the image background. This involves locating cell contours and cell interior to identify each cell and separate it from the other (Vicar et al., 2019). Bioimage analysis demands the need for high accuracy cell instance segmentation as it is essential to determine precise 3D cell volumes, shapes, geometry, neighbors and number of cells in regions or tissues (Lei et al., 2020). Like instance segmentation of cells, such segmentation can be also performed on nuclei, given the input nuclei images.

Computational morphodynamics is an emerging field of multidisciplinary research where first whole organ microscopic images are acquired. Then images are segmented using computational software's, and finally a deep biological understanding is comprehended from 3D quantitative image analysis. Further, this information is fed into a model that can be used to test several hypotheses related to growth and development (Roeder et al., 2011).

MorphographX (Barbier de Reuille et al., 2015; Strauss et al., 2019) is an open source software platform that allows 3D visualization of raw microscopic images and also further watershed segmentation of the images using insight segmentation and registration toolkit (ITK) module (Yoo et al., 2002). A few other packages do exist for image segmentation (Fernandez et al., 2010; Stegmaier et al., 2016), but these methods do not take into consideration the latest improvements in the field of computer science for bioimage segmentation.

3D instance segmentation of plant cells from microscopic images is a challenging task with available tools and techniques. This is mainly because of the high-quality demands of the raw microscopic images. Light scatters when imaging deep tissue layers and results in a low signal to noise ratio of the cell outline. Photobleaching and phototoxicity can also be other problematic factors. Additionally, acquisition parameters would also determine the final image quality. Moreover, the whole procedure up to final image segmentation requires enormous human input and manual corrections (Hallou et al., 2021; Tofanelli, Vijayan et al., 2019).

1.8 Application of deep learning tools for bioimage analysis

AI or Artificial Intelligence is an interdisciplinary field of computer science that aims to develop intelligent machines. Machine learning (ML) is an important part of AI; it involves making accurate predictions from very large datasets with the minimum amount of human intervention. Deep learning (DL) is a subfamily of ML which contains models that have demonstrated impressive results with a variety of machine learning tasks, in particular with bioimages (Villoutreix, 2021). DL has revolutionized the bioimage analysis field since the last decade through its ability to reduce manual human inputs, repeat tasks in an automated fashion and reshape the input and final dataset for bioimage analysis (LeCun et al., 2015). A number of recent reviews have addressed the importance of DL and its application in bioimage analysis (Gupta et al., 2019; Wang et al., 2019; Moen et al., 2019; Meijering, 2020; Hoffman et al., 2021; Esteva et al., 2021; Hallou et al., 2021).

Machine learning (ML) allows software packages to perform accurate predictions of their outcome without being explicitly programmed to do so. It defines a broad range of statistical models and algorithms to perform such specific data analysis tasks. Examples of ML tasks include classification of objects, ranking, clustering, regression etc. There are three basic approaches for classical machine learning. They are supervised, unsupervised and reinforcement learning (Murphy, 2012; Villoutreix, 2021). The majority of bioimage analysis is based on supervised and unsupervised learning. In supervised learning, existing human knowledge is used to provide a ground truth reference dataset for every element in the dataset. For example, the ground truth for a cell segmentation would be an extensive hand-corrected segmentation of the existing images. The dataset is then split into two categories, training and testing datasets. Using the training dataset, the ML algorithms are trained to reproduce the provided ground truths or in another way, the training is to produce models that understand the relationship between the input data and the output data. The model is ideally trained for several iterations to have high performance. Once the training is completed, the trained model can be applied to a testing dataset to evaluate the performance. It can be also applied on a new unseen but related dataset to the input image for training the model. The expected result is that the model now predicts elements similar to the ones in the training dataset from the newly provided unseen dataset. Unsupervised learning on the other hand works with unlabeled data as input,

machine learning then identifies patterns in the data without the use of human-provided examples (Hallou et al., 2021).

In bioimage analysis, DL is mainly used to perform three tasks: (1) image restoration, which involves transforming an input image into an enhanced output image; (2) image partitioning, which involves dividing an input image into regions and/or objects of interest; and (3) image quantification, which involves classifying, tracking, or counting objects. An example of image restoration using content-aware image restoration (CARE) (Weigert et al., 2018) on Drosophila wing disc has been demonstrated (Figure 1.4-A).

For cell segmentation from images of cell outlines or boundaries, the boundaries have to be detected as the first step for improvement. Convolutional Neural Networks (CNNs) are currently the most powerful border detectors (Kokkinos, 2015; Xie and Tu, 2015; Shelhamer et al., 2017). The U-Net architecture (Ronneberger et al., 2015) in particular has shown good performance on 2D bioimages, and it has now been expanded to analyze 3D volumetric data (Çiçek et al., 2016).

Overall, the aim of this study is to make use of available machine learning tools or provide new tools to the community for an improved instance segmentation of bioimages. Additionally, the application of machine learning could improve the quality and quantity of realistic templates that can be then fed into computational modeling tasks.

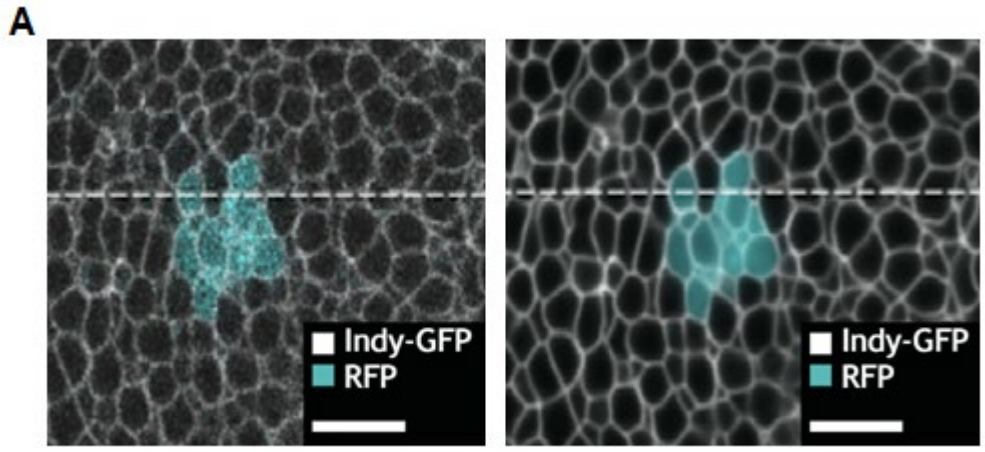

Figure 1.4 Application of content-aware image restoration (CARE) in the developing Drosophila wing disc.
Figure demonstrating the power of machine learning in bioimage analysis (A). From left to right: raw microscopic image of a 2D section of fly wing disc with outlines of cells

marked by Indy-GFP, the right image represents the results of the CARE (Weigert et al., 2018) image restoration after applying their trained model on a new dataset. The results of CARE indicate that the input image is severely improved for further bioimage analysis. Figure adapted from (Sui et al., 2018; Hallou et al., 2021)

1.9 3D organ and tissue coordinates representing cellular position in space

Even though 3D segmentation can be improved by the latest DL tools and techniques, the 3D segmentation of a whole organ of complex architecture is a challenging problem for further exploration. To gain insight into tissue morphogenesis, one must contextualize the cellular data by placing cells in a frame of reference relative to the developmental axes of the tissue or organ (Hejnowicz, 2014; Schmidt et al., 2014; Montenegro-Johnson et al., 2015; Strauss et al. 2022). Before such a frame of reference can be made, more essentially, the tissues in the organ need to be represented by different object labels or identities such that the organ can be understood at tissue scale changes. MorphographX (Barbier de Reuille et al., 2015) provides tools for manually annotating tissue labels to the 3D segmented dataset with possible semi-manual tissue annotations in simple-layered tissue arrangements.

Several computational pipelines such as iRoCS (Schmidt et al., 2014), 3DCellAtlas (Montenegro-Johnson et al., 2015; Schmidt et al., 2014) and 3DCellAtlas Meristem (Montenegro-Johnson et al., 2019) have been developed to provide a tissue-level frame of reference and enable semi-automatic labelling of 3D cellular properties in plant tissue context at cellular resolution. These computational pipelines have been used successfully to label cells and tissues in the main root and hypocotyl, radially symmetric organs with limited curvature, or the SAM, a dome-shaped structure with moderate complexity. However, not all plant organs fall into these simple morphogenetic categories. For instance, curvature limits the usefulness of the analysis strategies implemented in iRoCS and 3DCellAtlas, particularly in indexing the axial position of a cell and determining its absolute distance from a reference. Since ovules exhibit extreme curvature by forming a hood-like outer structure at maturity, the available tools are not applicable for ovules. Thus, new methods are required to contextualize cellular data along the developmental axis of the ovule. This study also includes the development of a new toolkit, 3DCoordX that can be applied on ovules and on other curved organs.

1.10 Objective

Arabidopsis ovules are initiated as simple finger-like protrusions from the placental surface; they undergo complex morphogenesis to form a final hood-like curved structure. This study aims to understand the mechanism and key features of the ovule during the transformation from a simple form to a complex three-dimensional shape. To address this, this study focuses on confocal microscopic imaging and 3D digital reference atlas generation of wild-type Arabidopsis ovules from early to late stage of development before fertilization. An ideal reference atlas would be that with which 3D volumetric and quantitative analysis can be performed with cellular, tissue and spatial resolution.

Generation of 3D digital atlas requires 3D deep imaging of ovules with reliable cellular resolution. This study aims to improve 3D microscopic ovule imaging and to generate a precise 3D digital atlas of ovule development. Different tools and techniques must be developed within the context of 3D digitization of microscopic images that carry cellular, tissue and spatial information. Finally, the digital models of the organ are to be explored in 3D, including quantitative analyses to understand the patterns associated with structural changes at different developmental stages. Overall, this study targets the identification of different phases of 3D tissue growth patterns and understanding of how different tissues or regions within the ovule grow and contribute to its intermediate and final 3D shape forms. The present study is a recapitulation, continuation, and extension of already published findings from Tofanelli, Vijayan et al., 2019, Wolny et al., 2020, Vijayan et al., 2021, Strauss et al 2022 and Vijayan et al., 2021.

1.11 Publications from this study

1. Rachele Tofanelli*, **Athul Vijayan***, Sebastian Scholz & Kay Schneitz 2019. BMC Plant Methods. Protocol for rapid clearing and staining of fixed Arabidopsis ovules for improved imaging by confocal laser scanning microscopy. doi:10.1186/s13007-019-0505-x
2. Adrian Wolny*, Lorenzo Cerrone*, **Athul Vijayan**, Rachele Tofanelli, Amaya Vilches Barro, Marion Louveaux, Christian Wenzl, Sören Strauss, David Wilson-Sánchez, Rena Lymbouridou, Susanne S Steigleder, Constantin Pape, Alberto Bailoni, Salva Duran-Nebreda, George W Bassel, Jan U Lohmann, Miltos Tsiantis, Fred A Hamprecht, Kay Schneitz, Alexis Maizel, Anna Kreshuk. 2020, eLife. Accurate and versatile 3D segmentation of plant tissues at cellular resolution. doi:10.7554/eLife.57613
3. **Athul Vijayan***, Rachele Tofanelli*, Sören Strauss, Lorenzo Cerrone, Adrian Wolny, Joanna Strohmeier, Anna Kreshuk, Fred A Hamprecht, Richard S Smith, Kay Schneitz, 2021, eLife. A digital 3D reference atlas reveals cellular growth patterns shaping the Arabidopsis ovule. doi:10.7554/eLife.63262
4. **Athul Vijayan***, Soeren Strauss*, Rachele Tofanelli, Tejasvinee Atul Mody, Karen Lee, Miltos Tsiantis, Richard S. Smith and Kay Schneitz. 2022 ,Plant Physiology. The annotation and analysis of complex 3D plant organs using 3DCoordX. doi.org/10.1093/plphys/kiac145
5. Soeren Strauss, Adam Runions, Brendan Lane, Dennis Eschweiler, Namrata Bajpai, Nicola Trozzi, Anne-Lise Routier-Kierzkowska, Saiko Yoshida, Sylvia Rodrigues da Silveira, **Athul Vijayan**, Rachele Tofanelli, Mateusz Majda, Emillie Echevin, Constance Le Gloanec, Hana Bertrand-Rakusova, Milad Adibi, Kay Schneitz, George Bassel, Daniel Kierzkowski, Johannes Stegmaier, Miltos Tsiantis, Richard S. Smith. (2022). Using positional information to provide context for biological image analysis with MorphoGraphX 2.0. eLife. doi.org/10.7554/eLife.72601
6. Lorenzo Cerrone, **Athul Vijayan**, Tejasvinee Atul Mody, Kay Schneitz and Fred A Hamprecht 2022, IEEE CVPR/CVF. CellTypeGraph: A New Geometric Computer Vision Benchmark. doi.org/10.48550/arXiv.2205.08166

2 Materials and Methods

2.1 Plant work, plant genetics, and plant transformation

Arabidopsis thaliana (L.) Heynh. var. Columbia (Col-0) was used as wild-type strain. Plants were grown in pots containing soil at controlled environment plant chambers with 75% humidity, 130 µmol white light and at 22°C temperature with 16-hour day and 8-hour night. Seeds were directly germinated on the soil after stratification if they need not have to be selected for resistance. In other cases, seeds were sterilized with bleach solution (3.5% bleach with 0.01% Triton x-100) for two minutes followed by another two minutes of wash with 70% ethanol. Further, 4-5 washes with autoclaved water were done for removing the residual ethanol and bleach from the seeds containing the microcentrifuge tube. Seeds were spread onto an agar plate containing half MS. Care was taken to place seeds a distance apart so that they will not be under stress. Seeds were stratified at 4°C for 72hrs prior to incubation in growth cabinets. Seedlings of about eight to ten days old were screened for antibiotic resistance and were transferred to soil and further grown under a plant growth cabinet. Fluorescent reporters were also passed through epifluorescence screening of the root whenever the reporter was known to express in root tissues. Plants on pots were watered from below thrice in a week and care was taken to give water only when the soil is becoming dry to avoid stress. The light quality, temperature and day-night cycles of the chambers have to be constantly checked for any variations that can affect the plant health.

Wild-type plants were transformed with different constructs using Agrobacterium strain GV3101/pMP90 (Koncz and Schell 1986) and the floral dip method (Clough and Bent 1998). Transgenic T1 plants were selected on respective selection media (20 µg/ml) and transferred to soil for further inspection.

2.2 Recombinant DNA work

For DNA work, standard molecular biology techniques were used. PCR fragments used for cloning were obtained using Q5 high-fidelity DNA polymerase (New England Biolabs, Frankfurt, Germany). All PCR-based constructs were sequenced.

2.3 Live Imaging of ovules for quick preparation

The whole inflorescence containing multiple flowers were cut out from the plant and placed under a dissecting microscope. The flower of the desired stage was selected and separated from the whole shoot and was stuck to a double sticky tape under the dissecting microscope. Floral organs were removed using fine needle syringes, retaining the pistil with stalk. A slit was made at both halves of the outer carpel wall using the syringe and ovules were exposed. The tissue was then transferred to a glass slide containing a drop of half MS liquid media at the middle for sample mounting. Ovules were harvested from the pistil and collected directly on the half MS media. In most cases, the funiculus was cut out from the placenta as we did not aim to image the whole organ attached to the placenta. A coverslip was placed on top of the sample and was ready for immediate use. Live dissected ovules were never stored for further use.

2.4 Cleared sample preparation for confocal imaging

A detailed protocol was recently published (Tofanelli, Vijayan et al., 2019). Fixing and clearing of dissected ovules in ClearSee were done essentially as described (Kurihara et al., 2015). Tissue was fixed in 4% paraformaldehyde in PBS followed by two rounds of washes in PBS before transfer into the ClearSee solution (xylitol (10%, w/v), sodium deoxycholate (15%, w/v), urea (25%, w/v), in H2O). The clearing was done at least overnight or for up to 2–3 days. Staining with SR2200 (Renaissance Chemicals, Selby, UK) was essentially performed as described in Musielak et al., 2015. Cleared tissue was washed in PBS and then put into a PBS solution containing 0.1% SR2200 and a 1/1000 dilution of TO-PRO-3 iodide (Bink et al., 2001; Van Hooijdonk et al., 1994) (Thermo Fisher Scientific) for 20 min. The tissue was washed in PBS for one minute, transferred again to ClearSee for 20 min before mounting in Vectashield antifade agent (Florijn et al., 1995; Vectashield Laboratories, Burlingame, CA, USA). The detailed method follows below. A user protocol for the same is also available at section 2.20.

2.4.1 Fixation and clearing

A flower of the desired floral stage was selected, and other floral organs were removed as done for the live imaging experiment. Slit open the pistil wall and carefully expose the ovules. Care was taken not to damage any of the ovules. The tissue was attached to the double sticky tape, allowing the dissection to be easy. A sharp syringe was used

for making the slit. Don't harvest the ovules, rather just leave them exposed and still attached to the placenta. Try to remove most of the carpel wall. This would make the fixing and clearing more effective. The whole procedure has to be performed in not more than two minutes, else there are chances of the sample getting dried under the dissecting microscope. Quickly transfer the pistils to a microcentrifuge tube containing fixative (4% paraformaldehyde containing 1x PBS, pH 7.4). Don’t let the sample float on the fixative, make use of a rotating device to turn around the sample containing tubes so that the fixation is most effective. Fixation can be done for a minimum of 1.5hr at room temperature. It can be also extended to several days when performed at cold temperature, optimally at 4° celsius. In that case, the fixative solution should be exchanged every fourth day or so. Care was also taken not to overcrowd the 1.5ml tube with more than 5 pistils while fixation and clearing. The fixation is effective when the pistils sink to the bottom of the tube. If fixation was done at 4°C hold the samples at room temperature for at least 30 min before proceeding further. Carefully remove most of the fixative but make sure samples are still submerged. Wash twice the fixed tissues for 1 min in 1 x PBS. The washing step is very important to avoid the formation of precipitates that occur upon the addition of ClearSee to the fixative. Transfer the fixed and washed pistils to a 1.5 ml microcentrifuge tube containing 1 ml ClearSee solution and clear them at room temperature overnight on the rotating device. Overnight is usually sufficient. Slightly better results are obtained upon 2 to 3 days of clearing or up to a week. Change the ClearSee solution after 2 days if samples are stored for extended periods of time. The carpels are stable in ClearSee for several weeks. After clearing proceed to the staining procedure

2.4.2 Staining

Cleared samples were stained with appropriate stains in 1x PBS solution. In the case of ovules, samples were stained with cell wall stain and nuclei stain at the same time. Cell wall stain SR2200 was used at a concentration of 0.1% and nuclei stain TO-PRO®-3 was titrated to a final concentration of 1μM (1:1,000 dilution of 1 mM stock solution). Stains were dissolved in 1x PBS. Stains were more effective when the staining was performed in 1x PBS solution rather than using the ClearSee solution. ClearSee containing samples were taken out of the rotor and checked for the extent of clearing, if the samples don't look transparent, they were further cleared. ClearSee was removed from the microcentrifuge tube using a pipette and the staining solution

containing both the stains were added to the microcentrifuge tube containing samples. Care was taken to remove the maximum amount of ClearSee from the tube before adding the staining solution. At this moment one can also do an additional wash with 1ml of staining solution. Samples were stained at dark and on the rotating holder for the next twenty minutes. The stain solution was removed from the microcentrifuge tube without causing any damage to the samples. The staining solution was twice washed off from the tube with 1x PBS solution. The samples are further cleared for a minimum of 20 minutes before final mounting in the mounting media.

2.4.3 Mounting

Mounting was performed in a different way than regular sample preparation for other tissues or 2D imaging. Ovule imaging in 3D demands to preserve the sample intact, for this reason, two coverslips were glued on top of a microscopic glass slide using nail polish. A small distance of around 1.5cm was left between the two coverslips where the sample can be mounted. Care was taken to use a 170µm thick coverslip so that the working distance of the microscope objective is handled. In this case, the microscope working distance is 250µm, which would enable imaging up to 250µm deep from the inner surface of the sample coverslip. Ovules are generally 60µm deep structure, but because of their orientation in microscopic slides, one might have to acquire a z-stack of a maximum of around 120µm deep. Sample mounting was taken care that all this would be possible, and samples won't be squeezed.

Place a drop of mounting media (VECTASHIELD® antifade mounting medium) in between the two coverslips. Gently pick up the pistil from the ClearSee solution and place it on the mounting media. Care was taken to not transfer any ClearSee solution to the mounting media. Dissect out the ovules from the pistil. Care was taken when the experiment requires the entire organ attached to the placenta. In that case, the carpel wall was first removed from the pistil and the whole pistil was torn apart leaving behind two halves of the pistil with ovules attached to the placenta. Further, the ovules were cut out, retaining their attachment to the placenta. Care was taken to not make any bubbles on the mounting media. A coverslip was placed gently on top of the two supporting coverslips and was stored in a glass slide book. The sample is ready for immediate use, it can be also stored at 4^{o} Celsius for afterwards use. Nuclei stain might not be intact after a few days of storage, so it's not recommended to store for weeks if imaging the nuclei is also of interest. Sample dissection can be also done in a simple

manner when the whole organ attached to the placenta is not required. In that case, the ovules were just collected using a fine syringe and instead of retaining their attachment to the placenta.

2.5 Microscopy and image acquisition

Confocal laser scanning microscopy (CLSM) was performed on upright Leica TCS SP8 X WLL2 HyVolution 2 (Leica Microsystems) equipped with two GaAsP (HyD) detectors and two photomultiplier (PMT) detectors. The system was equipped with an Argon laser which can be excited at five wavelengths. It also included a 405-diode laser and white laser. The white laser can be excited at a broad spectrum up to 670nm. Overall, the system provides four detectors and almost all excitation wavelength possibilities for regular fluorescent dyes or stains.

A sample slide was mounted onto the CLSM and looked at first using a 10x objective. Ovules position inside the slide were marked using the XY grid in the microscope touch screen. This allows an easy switch between samples at a higher magnification. Sample quality was visually evaluated by looking at the epifluorescence channel of DAPI, which eventually fluoresces the SR2200 cell wall stain in the stained samples. Care was taken to evaluate the best oriented and stained samples and position them in the XY coordinates for further looking at them on higher magnification. Samples were again looked at 63x glycerol objective (HC PL APO CS2 63x/1.30 GLYC, CORR CS2) and aligned to the center. Ovule images were acquired at the same magnification using a HyD in most cases to increase the image quality. Signal intensity was first checked at a lower excitation power of 0.1% laser power and was gradually increased to a point where the signal is not oversaturated. Usually, the cell wall stain requires a maximum of about 1.5% laser power and a minimum of 0.1 when the samples are well stained. Cell wall stain was excited using the 405-diode laser and the emission was detected at 420-480nm wavelength with a minimum detector gain of 10 and maximum of 200. Nuclei stain was excited using the white laser 641 nm and the emission was detected at 655-670 nm wavelength. Laser power for white laser sometimes has to be increased up to 10% to image the faint nuclei stain. The white laser is overall weak compared to argon and diode lasers.

3D z scan was done using the bidirectional scan to perform a fast acquisition and a scan speed of 400 Hz was used to compensate with the image quality, the pinhole was set to 0.8 Airy units to have a sharp image, line average between 2 and 4 to further

smooth the signal, and the digital zoom between 1 and 2 depending on the size of the sample. Z stacks were acquired with positioning the beginning and endpoints of the sample slice using the microscopic z joystick and the scan interval was sliced according to the system optimized format of 0.24µm in most cases. For fine images, the Z stack was never more than 0.35µm. Z scan was performed in XYZ mode where the XY format was chosen with the system optimized manner following the Nyquist criterion. A high-quality z stack used for 3D cell segmentation without any machine learning tools was of final voxel size 0.063µm x 0.063µm x 0.24µm in the XYZ axis. 3D images were saved as 12 or 16bit instead of 8-bit images to increase the voxel information that would further improve the post-processing and segmentation of the image.

With the advantage of machine learning for processing the raw images, the z stacks could be of less fine quality and the 3D cell segmentation can be achieved similar to the original method which demands high image quality. Z stacks were acquired with XYZ voxels 0.125µm x 0.125µm x 0.24µm where we took advantage of the PlantSeg-trained model to generate equal standard cell segmentation as with images with fine voxels. This was possible because the PlantSeg model ‘generic_confocal_3d_unet’ was trained on down sampled original images and ground truths. The model now requires raw images whose voxels are scaled to the trained dataset so that it generates the best cell boundary predictions. Overall, raw images captured with 2x down sampled voxels were helpful in that they simplified the rescaling step in PlantSeg and allowed us to generate raw images in less time without compromising segmentation quality.

A sequential scan was performed on the reporter line whenever needed to avoid any bleed through from different channels. Z stack was also optimized with z compensation either manually or programmed where the detector gain or the excitation power was adjusted to compensate for the scan in the z-direction. The z scan image can be looked at at the z begin and z end, the intensity at z end is always lower because of the laser penetration to deeper layers through the sample. Z compensation was applied to complement this lowering of signal intensity so that the cell wall channel is equally bright even at the z end position. This is also critical for further segmentation of the dataset.

Image acquisition parameters for the pCUC3::CFP line were the following: SR2200; 405 diode laser 0.10%, HyD 420–480 nm, detector gain 10. CFP; 514 nm Argon laser 2%, HyD 525-550 nm, detector gain 100. TO-PRO-3; 642 nm White Laser 2%, HyD 660–720 nm, detector gain 100. Image acquisition parameters for pUBQ::H2B:tdTomato were the following: SR2200; 405 diode laser 0.10%, HyD 420–480 nm, detector gain 10. tdTomato; 561 nm Argon laser 2%, HyD 575-625 nm, detector gain 100. TO-PRO-3; 642 nm White Laser 2%, HyD 660–720 nm, detector gain 100. In each case, the sequential scan was performed to avoid any cross talk.

2.6 3D cell segmentation

3D cell segmentation was performed using different methods. ITK segmentation was performed using MorphographX. PlantSeg is the platform where raw images are processed via machine learning to improve the image information and are further segmented using the GASP method. PlantSeg-MGX hybrid method is another method bridging the traditional ITK segmentation lifted by the cell boundary predictions from the machine learning training.

2.6.1 ITK cell segmentation using MGX

ITK segmentation can be performed on a raw cell wall or boundary image. It uses an auto seeded ITK watershed segmentation of cells in 3D giving labels to individual cells and the image background. The method is adapted from (Barbier de Reuille et al., 2015). The major issue in cell segmentation is to have accurate labels of a cell without under or over segmenting it. Watershed is often prone to leaking in 3D because of faint walls that merge adjacent cells on segmentation resulting in under segmentation. One way to prevent watershed leaking is to blur the raw images that prevent watershed leaks. Finally, segmentation was done on the processed raw image where it was brightened (*Stack/Filter/brighten darken*) by a factor of two and gaussian blurred (*Stack/Filter/gaussian blur stack*) by a factor of 0.2 x 0.2 x 0.2 (xyx) for three to five times depending on raw image quality. The extent of blurring could be visualized in a clip plane mode. The blurred stack was segmented using the process *Stack/ITK/Segmentation/ITK watershed Auto Seeded.*

2.6.2 Machine learning based boundary prediction and cell segmentation using PlantSeg pipeline

PlantSeg (Wolny et al., 2020) is a machine learning-based cell segmentation pipeline. PlantSeg uses raw cell boundary image and passes it through the pipeline to give out 3D segmented images. There are several intermediate steps in the PlantSeg pipeline. First, the raw image is optimized for the whole pipeline, for which the raw image has to be of a scale of right voxels such that the neural network model can understand it. There are several neural network trained models available in PlantSeg of different sizes. The optimal model is selected on the CNN prediction module of the PlantSeg graphical user interface (gui). “Generic_confocal_3d_unet” is the best performing neural network model which was trained on raw images of voxel size 0.15µm x 0.15µm x 0.235µm in xyz directions.

For a raw image of equivalent size, rescaling the raw images to this size is optimal for best performance in cell segmentation and boundary probability map. This is done in the data preprocessing module of the PlantSeg gui. The module CNN prediction contains the parameters for generating the cell boundary probability map (b-pmap) or boundary predictions which are the output of a U-Net-based convolutional neural network (CNN). The b-pmap looks like an improved raw cell wall image which contains more information than just the raw image. Every pixel in the b-pmap image has a resultant value between 0 and 1 which indicates if this pixel belongs to a cell wall or background. This is the result from the neural network prediction of cell boundaries. Overall, the b-pmap looks just like a sharp cell wall image with a high signal to noise ratio. Generally, a patch size of 96 x 256 x 256 was used for generating b-pmap. This would make patches or individual 3D blocks or tiles within the 3D dataset, eventually making small subsets of the entire 3D image and generating predictions of the small areas, further all the patches or tiles are merged to form the final image in 3D. Stride parameter was set to Accurate, It involves extensive mirror padding on the patches. At the end generating a precise boundary prediction.

3D cell segmentation is done directly on the b-pmap z stack. PlantSeg has several possibilities to perform cell segmentation. GASP (Bailoni et al. 2019) method is the best of them when there is only b-pmap as input image. GASP segmentation was run on the PlantSeg default parameter with watershed in 3D in prediction threshold 0.5. The segmentation and prediction outputs are saved as TIFF file format from PlantSeg using the post processing module in the PlantSeg gui. PlantSeg always uses H5 or

HDF5 file format to run boundary prediction or segmentation. These files are part of the intermediate steps of the pipeline and could be also retrieved back from the pipeline folder.

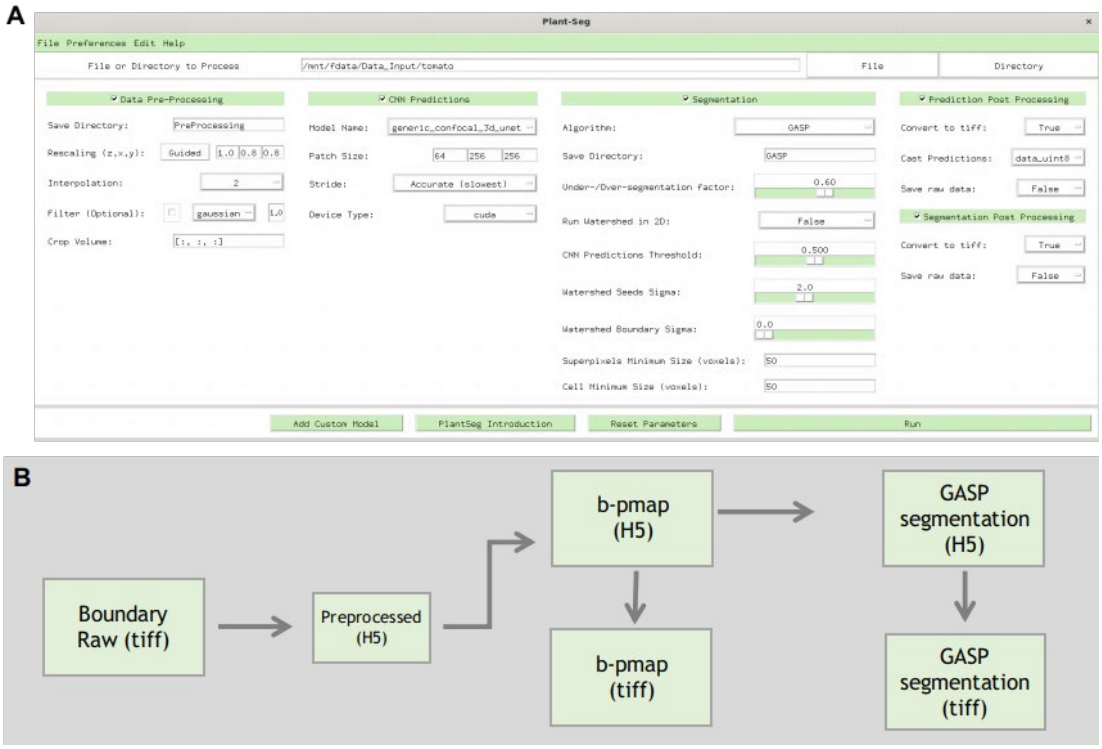

Figure 2.1 PlantSeg GUI and scheme of the pipeline.
Graphical user interface of PlantSeg (A). Different modules include Data Pre-processing, CNN Prediction, Segmentation, Prediction Post-Processing and Segmentation Post-Processing. Users can click on different modules to activate them for the pipeline. Scheme of the pipeline for cell boundary segmentation from inputting the raw boundary image (B). The pipeline starts with the raw boundary image of the cell wall, which is scaled to form the preprocessed H5 file which is further used for the pipeline. The b-PMAP and segmentation is upscaled according to the factor of scaling done in the data preprocessing menu.

This pipeline performs for near perfect 3D cell segmentation as only a small number of errors such as over segmented cells had to be corrected by visual inspection of the segmented stack in MGX. In critical cases, this included cross-checking the TO-PRO-3 channel which included the stained nuclei. Image acquisition of mature ovules takes about 15 min for both channels (SR2200/cell contours; TO-PRO-3/nuclei), running the PlantSeg pipeline requires about 25 min on our computer hardware (1x Nvidia Quadro P5000 GPU), and manual correction of segmentation errors takes less than 5 min with the improved method lifted by machine learning. A schematic representation of the PlantSeg pipeline and the gui interface can be found in figure 2-1.

2.6.3 PlantSeg-MGX hybrid segmentation

PlantSeg-MGX hybrid is as the name suggests a method bridging the PlantSeg b-pmap and classical ITK segmentation with MGX from raw images. The hybrid method just back blends the b-pmap of PlantSeg with the raw boundary image to create a better raw image which can be further segmented with ITK watershed method in MGX. The b-pmap itself can be also used for ITK segmentation, but the b-pmap sometimes doesn't make faint walls which are present as a weak signal in the raw boundary image. Overall combining the b-pmap with the raw image makes sharp boundaries wherever the b-pmap is strong while retaining the weak signals of faint walls from the raw image. The z-stacks of raw cell wall images and PlantSeg b-pmaps were combined in MGX using the process *Stack/Multistack/Combine Stack/Max* to generate a merged image stack, which was further blurred twice with a radius of 0.3 × 0.3 × 0.3 using the process *Stack/Filter/Gaussian Blur Stack*. The processed image was further segmented in MGX by auto seeded ITK watershed with the default threshold of 1500 using *Stack/ITK/Segmentation/ITK Watershed Auto Seeded*. PlantSeg-MGX hybrid segmentation was time consuming, but the results were comparable to the graph-partitioning-based 3D segmentation of PlantSeg.

2.7 3D nuclei segmentation

3D segmentation of nuclei can be performed using different methods. Cellpose is now a wide accepted machine learning based method for nuclei segmentation. PlantSeg also provides a possibility for generating a nuclei probability map. Here are a few methods used in the study for comparison.

2.7.1 Cellpose nuclei segmentation

Nuclei segmentation using CellPose (Stringer et al., 2021) was performed using the default parameters. Tiff files of raw nuclei images were loaded in CellPose and 3D nuclei segmentation was performed with the CellPose nuclei model. Segmentation was saved as a tiff file from CellPose and was loaded in MGX for 3D visualization and further correction or processing.

2.7.2 StarDist nuclei segmentation

StarDist nuclei segmentation was performed using the ovule trained StarDist model. Parameters were set as default. Raw images were 2x down sampled to perform the

StarDist segmentation. StarDist segmentation results were saved as H5 file format. H5 files were opened in ImageJ Fiji and using HDF5 viewer and were exported as tiff files from Fiji with right voxel information in the image metadata. Tiff files were then loaded in MGX for 3D visualization and further correction or processing.

2.7.3 PlantSeg-Cellpose hybrid method of Nuclei segmentation

PlantSeg-Cellpose hybrid method uses nuclei probability map (n-pmap) or nuclei predictions as an input for CellPose nuclei segmentation. Nuclei probability maps or semantic segmentation are generated in PlantSeg using the same tools as in the PlantSeg cell segmentation pipeline but uses a nuclei model instead of the cell boundary model on the CNN prediction module. Nuclei model was a result of convolutional neural network training of nuclei stain images and their respective nuclei segmentation images from ovule datasets. Nuclei trained model (confocal_unet_bce_dice_nuclei_stain_ds1x) understands raw images of voxel size 0.1µm x 0.1µm x 0.35µm. Nuclei stain or reporter images were the input images of PlantSeg. n-pmap were generated with the PlantSeg pipeline module CNN predictions. The results of PlantSeg nuclei prediction was saved as a tiff file and was further imported to CellPose for final segmentation. n-pmap also contains the probability map of nuclei ranging from 0 to 1 with 0 being the background and 1 being the nuclei. Nuclei prediction greatly improves the signal to noise ratio on the raw images and allows to 3D segment even any faint nuclei image stack. Based on simple thresholding, nuclei can be also segmented within the PlantSeg after generating the n-pmap, this segmentation could adversely affect the size and shape of nuclei. Whereas the CellPose segmentation of n-pmap is more accurate in terms of size and shape of nuclei.

2.8 Method for improved cell segmentation based on nuclei seeds

Improving the cell segmentation based on the presence of a nucleus inside the cell was done in mainly three methods for comparison of the best. The first method was just to proofread an existing segmentation. This is a script on PlantSeg tools (https://github.com/hci-unihd/plant-seg-tools). The script requires the boundary prediction file and segmented image along with nuclei probability map or nuclei segmentation. The script would export a modified segmentation file where the correction is done only on the cells where there were more than one nucleus inside the segmented cell. Essentially, the script maintains the original label of the stack while

correcting the labels which had errors. The script can be found under the useful scripts of PlantSeg tools "*fix_over_under_from_nuclei.py*".

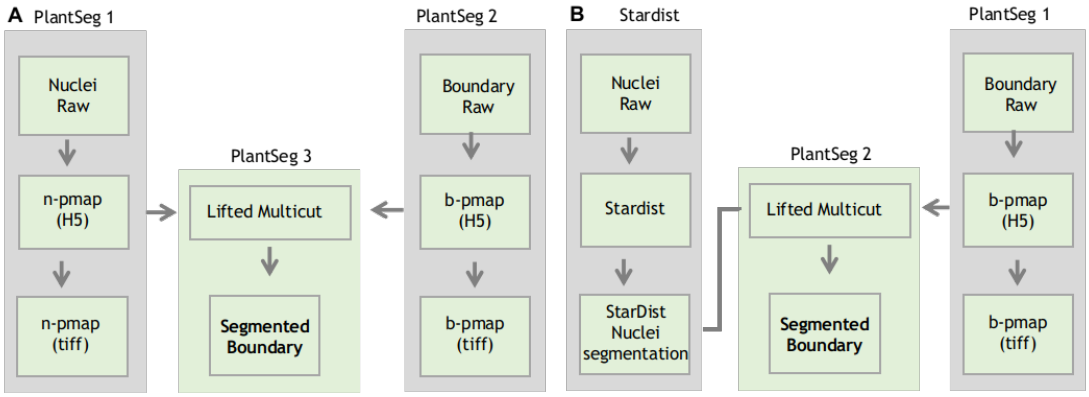

Figure 2.2 Scheme of Lifted multicut segmentation of cell boundary images. Scheme of lifted multicut where the n-pmaps are used for nuclei seeds (A). The method requires running PlantSeg three times. The first PlantSeg is for generating the n-pmap, second one for b-pmap and the third one for running lifted multicut with the results of n-pmap and b-pmap. Scheme of lifted multicut where the nuclei segmentation is used as seeds for cell segmentation (B). This pipeline can be also divided into three parts, the first part involved generating StarDist nuclei segmentation images. The second part involves cell b-pmap generation with PlantSeg. The final step involves running lifted multicut from the results of nuclei segmentation and b-pmap.

The second method is the lifted multicut method (Pape et al., 2019; Horňáková et al., 2017). Lifted multicut can be done using the n-pmap or nuclei segmentation as seeds for the cells. Lifted multicut can be done using PlantSeg with three steps. The first step is to generate cell boundary predictions, the second step is to generate n-pmap and the third step is to run the lifted multicut with the results of the b-pmap and n-pmap. pmap is generated as explained in section 2.6.2 and 2.7.3. They are saved as H5 files instead of tiff files. The third round of PlantSeg runs with both the n-pmap and b-pmap as input images. Lifted multicut can be performed only in the command line PlantSeg

where one has to enter the parameters in the YAML file. These parameters include the path to n-pmap, path to b-pmap, parameters for segmentation and postprocessing module. An example YAML file can be found on PlantSeg GitHub (https://github.com/athulrv/3D-Ovule-Atlas.gitl). A scheme of lifted multicut can be found in figure 2.2-A.

The third method is like the second one but uses the nuclei segmentation instead of n-pmap. This still requires the first step of lifted multicut, where the b-pmap is generated in PlantSeg. Precise nuclei segmentation can be exported from the StarDist model.

Further, the path to the nuclei segmentation is added to the YAML file and the third step of lifted multicut is performed like the one in method two. This method is more robust because it uses the nuclei segmentation instead of n-pmap. A scheme of lifted multicut can be found from the figure 2.2-B.

2.9 Generation of 3D and surface cell meshes

3D meshes were generated on the proofread segmented stack. Error in segmentation can't be corrected in the mesh. 3D cell meshes were generated in MGX using the segmented image stacks using the process "Mesh/Creation/Marching Cube 3D" with a cube size of 1. A cube size of 0.8 was sometimes used for fine corners. All the cell annotation was performed on the 3D cell meshes in MGX. 3D cell mesh can be saved as an MGXM mesh file or can be also exported as .obj, .stl etc. The 3D mesh file can be directly used for volumetric analysis of cells or organs or tissues once there are specific labels to the tissues or regions.

Organ surface mesh or surface mesh refers to a mesh made on the outer surface of the organ, which is used for surface estimation and further measurements for various feature extraction as part of different workflows. Surface mesh can be also segmented, and the cell area can be quantified from such a surface mesh. The surface mesh is mainly used as a bridge between 3D cell mesh and organ surface. The 3D cell mesh doesn't contain positional information. MGX can be used with two meshes at the same time. 3D cell mesh in mesh 1 and surface mesh in mesh 2. This allows further definition for cells of the cell mesh which is intrinsically absent in them. Other features like distance to the surface for inner cells and cell axis are also relying on organ surface mesh. Surface mesh is mainly required for the method of semi-automatic cell type labelling of the mature ovule. The organ surface mesh is generated from the segmented stack. The segmented stack was first gaussian blurred using the process *Stack/Filter/Gaussian Blur Stack* with a radius of 0.3 in xyz. The smooth stack was used to generate organ surface mesh using *Mesh/Creation/Marching Cube Surface* with a cube size of 1 and threshold 1. The generated surface mesh was then smoothed several times using the process *Mesh/Structure/Smooth mesh* with 10 passes. 3D cell mesh for further data extraction requires several parameters to be quantified. MGX process *Mesh/heatmap/analysis/cell analysis 3D* was performed on the 3D cell mesh, this adds more information to the raw mesh. It includes cell volume, area, neighbors cell label, number of cell neighbors, cell centroid and other attributes. This is also a

requirement for other processes that do several annotations or quantification like tissue labelling or distance coordinates.

2.10 Classification of tissue types

MGX allows to cluster or group cells while retaining their identity or cell label. Grouped cells come under the “parent label” in MGX. Parents or parent labels were originally part of a two time point mesh dataset where the daughter cells get the ID of the parent cell. The same tools are used here in a different manner where there is just one time point and cells that belong to the same tissue are grouped under the parent ID.

For young ovules, the cell layers L1, L2 and L3 are annotated by a semi-automatic method where the 3D mesh of ovule primordia and the respective segmented stack is used for performing the classification. Classification is done using the process *Mesh/cell atlas 3D/ovule/detect cell layer stack* with the parameter for the number of layers as 3 in the case of primordia. The detect cell layer process requires the segmented stack in working stack and 3D cell mesh in working mesh. The outcome of the process is L1, L2 and L3 labelled ovule primordia. To further subcluster different organs in the image, another process called label ovules was run (*Mesh/cell atlas 3D/ovule/label ovule*). This requires a user input which marks the distal end of the organ. As a last step, these two labels are merged using the process *Mesh/Lineage tracking/Unique parents from attributes.* 3D Mesh tools in MGX were used for proofreading the labelled mesh before exploring any of the valuable attributes. Overall, the semi-automatic method used the stack labels on top of the 3D cell mesh and first quantifies the outside wall area ratio. This is used as a criterion for cell classification. The outer cells of ovule primordia have a significantly higher outside wall area ratio because of the outer exposed wall. This feature is used to first cluster them as L1 cells. Further a distance parameter is quantified by the process, and it looks for the distance of all cells to the selected L1 cells. This would allow the subclassification of L2 and L3 layers in the primordial dataset based on the distance threshold. This saves a lot of time in annotating the cell types which is otherwise done manually.

Mature ovule cell type annotation is more complicated as it must include different layered integuments and other internal and external tissues. The pattern is obvious by eye, but there exists no precise tool for such a classification of cell types in this complicated organ. The 3D mesh of the ovules was annotated with labels for cell type (abaxial outer integument, adaxial outer integument, abaxial inner integument, adaxial

inner integument, nucellus, chalaza, subepidermal outer integument, funiculus, embryo sac). With the existing tools, we can just achieve about 60% true labels. The rest is manually proofread. For mature ovules cell type annotation, we used the MGX process *Mesh/Cell Atlas 3D/Ovule/Detect Cell Layer* (a modified 3DCellAtlas Meristem tool [Montenegro-Johnson et al., 2019]) with the 3D cell meshes in the mesh 1 workspace and the tissue surface mesh in the mesh 2 workspace using a cone angle parameter of 1.2. We manually corrected mis-annotations and labeled the rest of the cells by using the mesh tools in MGX. We used the "Select Connected Area" tool to select individual cells of different layers in 3D and proofread the cell type labeling with *Mesh/Cell Types/Classification/Set Cell Type*. Each cell layer of the integuments was consecutively shaved off and proofreading was performed using 3D surface view. We further used the processes *Mesh/Cell Types/Classification* to save all labels, load labels and select cell types as required. The saved cell types csv file was reloaded onto the original 3D cell mesh and final proofreading was performed in the section view.

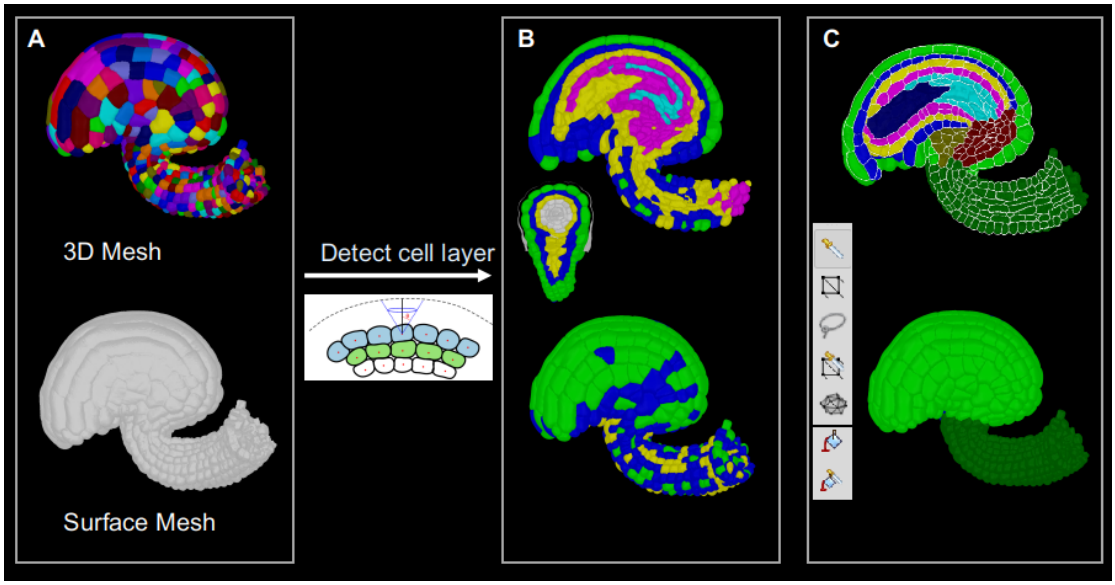

Figure 2.3 Method for classification of cell types in mature ovules.
Input dataset for the workflow include the 3D cell mesh and organ surface mesh (A). Detect cell layer stack function uses a cone angle parameter for identifying individual cell layers (B). Images represent the 2D section view on top and the 3D view at bottom. A transverse section shows how effective the tool is working on just three layer detection. The results from the automatic detect cell layer are further taken for user interactive proofreading of individual tissue types with the help of mesh tools (C). The mesh tools used for interactive tissue annotation are also included.

2.11 Annotation and data extraction from 3D cell meshes using MorphographX

Different annotations and quantifications were done on 3D cell mesh using MGX. The quantifications are always exported as a csv attribute file from *mesh/attributes/save to csv*. Annotations include tissue labels or labelling several regions within the organ or tissue. All such labels were exported as an attribute and saved to the 3D cell mesh too. Overall, the 3D cell mesh contains several attributes of a label or cell. All the attributes were exported at the same time and a long file format excel file was made from different cell meshes with information containing the ovule ID or image ID, ovule stage, volume of cells, tissue identity label, anterior-posterior label, distance coordinates and many more. Cells less than 30μm^3 in volume was excluded from any analysis as it would be part of segmentation errors or artefacts in cell mesh creation or void spaces between cells.

Several quantifiable measures were also exported to attributes. The “Neighbor Voxel Stack” command was used to differentiate cells with the exposed outer surface. Lobyness measures including sphericity, convexity, compactness, and solidity were quantified using the process *mesh/3D cell atlas/lobyness 3D.* For OI2 cell geometry measurement, depth of cells was first quantified by creating cell depth axis from the organ surface mesh and segmented stack using the process *Mesh/Cell axis 3D/Custom/Create Surface Directions* (Custom X: None; Custom Y: Surface Normal; Custom Z: None). Cell Length was quantified by creating a length axis, essentially the proximal-distal axis of the cell using a bezier curve. A bezier was first made from the selected central file of OI2 cells with *Misc/bezier/Bezier from Cell File*. Cell length directions were generated using the process *Mesh/Cell axis 3D/Custom/Create Bezier grid direction* (Custom X: Bezier X; None; None). Lastly, cell width or the medial-lateral length of a cell was quantified from the cell width directions. Cell width directions were generated using the process *Mesh/Cell axis 3D/Custom/Create Orthogonal Direction* (Primary Axis: X; Secondary Axis: Y; Change Secondary Axis: Yes). To add the measurements to the other extractable attributes the “Cell Length Custom” process was used.

2.12 3D organ coordinate system

Organ coordinates in 3D were performed in different methods to identify the best method for a curved organ like an ovule. Organ coordinates require a coordinate origin. For straight organs, a bezier curve was placed at the centre of the organ and that acted

as an origin. Organ coordinates were quantified using the process *mesh/heatmap/measure3D/location/bezier coord*. Organ coordinates where a group of cells act as the origin were quantified using the process *mesh/heatmap/measure3D/location/cell distance.* The cell distance process can quantify distances (in µm) of all cells to the origin cells through the shortest path, this can be done with the parameter Euclidean. The same process can be run to quantify the shortest path between cells to the origin cell where it quantifies how many numbers of cells is the gives cell apart from the nearest origin cell. It also has a parameter to restrict or not the measurement within the parents or tissue labels provided in the mesh while running this process.

Distance coordinates for ovules were quantified using a new process. For the primordia, the coordinate origin was placed at the distal tip of the organ. The coordinate process requires only a cell mesh and bezier ring. Bezier ring was made by the process *Misc/bezier/new bezier ring from selected cells* (radius xy 3x3) after selecting a distalmost cell of primordia. This places the ring approximately at the distal end of the primordia. The position of the ring can be also adjusted in an interactive manner using the control key interaction set to “cut surface” in MGX. Once the ring is placed at the distal end of the primordia, the distance coordinates can be extracted using the process *Mesh/heatmap/measure 3D/location/cell distance bezier ring*. This process contains several parameters, firstly the type of coordinates, distance coordinates or cell coordinates. Secondly, if the measurement has to be restricted within the parent labels or tissues. When the distance is restricted within the parents, the “direct distance limit” parameter can be used to assign the cells where the distance seeds are first assigned. The direct distance limit first finds the cells within this limit to the origin, then it finds how far are they. Once the distances are assigned to these direct cells, they act as distance seeds for the remaining cells within their tissue. The distance coordinate of another cell far from the direct distance limit is quantified by finding its distance to the nearest neighbour cell which was within the direct distance limit and within the same tissue type. Essentially the cells within the direct distance limit act as distance seeds to the remaining cells within their tissue. “Selected as direct” allows to include any manually selected cell to the direct distance limit which might not be within the distance limit. The direct distance limit for ovule primordia was set to a minimum value of 7µm. The “Consider ring orientation” parameter allows to include a positive and negative direction to the cell coordinates across the 2D plane of the ring. This is irrelevant for the ovule primordia where the bezier ring is positioned at the distal end.

For primordia coordinates, the 3D mesh where the cell distance bezier ring was quantified has the tissue labels of anterior and posterior L1, L2 and L3 cells. This allows to restrict the distance measurement within the anterior or posterior halves of the organ and not just within the layers.

For mature ovules, tissue labels were subdivided to anterior and posterior integument layers. OI2 was grouped into also lateral cells and a central group of cells for tissue restriction. *Mesh/heatmap/measure 3D/location/cell distance bezier ring* was again used for the mature ovules, where the ring is placed at the proximal end of integument initiation at the surface of OI1. Placing the ring was already a two-step process. First, a 3D mesh was loaded, the OI2 layer was removed from the mesh to visually select the proximal OI1 cells where the ring must be placed. The ring was made using the same process as in the primordia *Misc/bezier/new bezier ring from selected cells* (but with a higher radius of 15 x 15 xy). The bezier ring in ovule primordia was of a small radius, such that it can act as a point origin, here the ring acts as a wide origin. Once the ring was made, it was saved using *File/Save-as.* The original 3D mesh with appropriate tissue labels for the coordinates was loaded to mesh 1 for running the “cell distance bezier ring” process. The direct distance limit for the mature ovule was set to 15 µm. Consider ring orientation was set to yes. This generated the coordinates below the ring (proximal region of ovule including funiculus and proximal OI2) with negative values and coordinates above the ring with positive values allowing.

2.13 Primordia length and slanting

Length along the surface of the anterior and posterior sides of primordia was quantified for individual ovule primordia. A file of surface cells at the mid-section of two halves of the ovule was extracted and a bezier grid of size 3 x 3 (xy) was placed on the distal tip of the primordia surface. MGX process *Mesh/Heatmap/Measure 3D/Distance to Bezier* was used to quantify the length. The measured values are the distance from the Bezier grid to individual cell centroids through the file of connected cells. Primordia height was quantified by averaging the two values. Slanting was quantified by obtaining the difference of the maximal values at the anterior and posterior sides, respectively (MM-FIG 2). After the advancement of the ovule coordinate system, the primordia slanting can be also directly inferred from the coordinate values at the proximal end of anterior and posterior L1 cells.

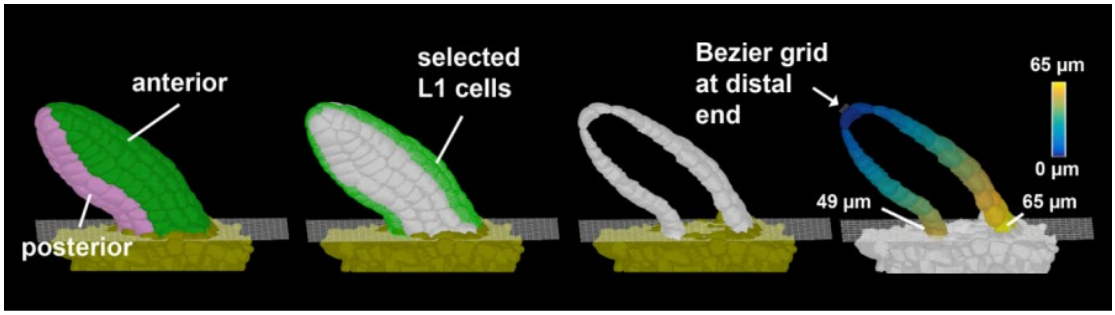

Figure 2.4 Primordium length and slant measurement.
3D surface view of primordia displaying the method for length measurement. 3D vew displays the anterior and posterior half of ovule primordia. L1 cells at the mid file are selected and extracted for slant measurement. The heatmap on the extracted cells depicts the quantified distance value between individual measured cells to the distal tip of primordia. Figure from Vijayan et al., 2021.

2.14 Statistical analysis and visualization

The distribution of cell attributes along the distance coordinates was visualized in RStudio with the ggplot function, these plots were mainly used to detect interesting differences which were then analyzed further with Tukey's post-hoc test through anova and Tukey HSD. For each analyzed group the mean attribute value of all cells belonging to that group was calculated for each ovule, since 5 ovules were analyzed per stage the population size was therefore always 5. Results with a p-value < 0.05 were accepted as significant. GraphPad PRISM was used to visualize the results with graphs.

2.15 Growth and proliferation rate calculation

Growth and proliferation rates for integuments was calculated as a normalized value with respect to the overall growth of only the integument tissues. For proliferation rates, the increase in total number of integument cells from one stage to the next was calculated as a ratio. For example, the increase in total number of integument cells from 100 to 140 is 1.4 times increase. Similarly, the increase in the number of cells in individual tissues was calculated. For example, if the oi2 cells increase from 40 to 60, which makes it 1.5 times increase. Finally, a ratio between these two values was plotted as a heatmap. In this case, it would be the (1.5/1.4= 1.07). The resulting value indicates if the tissue is growing at a higher rate compared to the increase in the overall growth of integuments. Similarly, the growth rates were also calculated where the overall increase in all integument tissue volume is compared over the increase in the different integument tissues. Essentially, the final values on the heatmap above one

represent if the tissue is growing at a higher rate than the overall growth of integuments. A value equal to one indicates that the tissue grows at the same rate to that of overall growth of integuments. A value below one indicates that it is slow growing with respect to overall growth of integuments. The principle is the same for proliferation and growth rate calculations.

2.16 Datasets

The dataset encompassing the segmented wild-type 3D digital ovules was described earlier (Vijayan et al. 2021). The z-stack of the Utricularia gibba trap was obtained from the fixed and modified pseudo-Schiff-stained (Truernit et al. 2008) specimen presented in Fig. 3C in (Lee et al. 2019). Dataset including 3D cemm meshes and images used in this study are available in a biostudies public repository. Additionally, the analysis excel files are also available in the below biostudies repository.

Vijayan A, Tofanelli R, Strauss S, Cerrone L, Wolny A, Strohmeier J, Kreshuk A, Hamprecht FA, Smith RS, Schneitz K (2020) BioStudies ID S-BSST475. WT Arabidopsis ovule atlas: A 3D digital cell atlas of wild-type Arabidopsis ovule development with cellular and tissue resolution.
https://www.ebi.ac.uk/biostudies/studies/S-BSST475

Vijayan A, Tofanelli R, Strauss S, Cerrone L, Wolny A, Strohmeier J, Kreshuk A, Hamprecht FA, Smith RS, Schneitz K (2020) BioStudies ID S-BSST513. 3D qualitative attribute dataset of Arabidopsis ovule atlas.
https://www.ebi.ac.uk/biostudies/studies/S-BSST513

2.17. Artificial templates of cells

Artificial templates of radial cell layers were created in MorphoDynamX (MDX, https://morphographx.org/morphodynamx/). MDX plugin “cell make” was used to generate radial cell layers of straight and bent tissue. Tissue annotations to radial cell layers were performed manually.

2.18 Media and solutions

Half Murashige-Skoog medium (for plant tissue culture):

Dissolve the following in distilled water and autoclave for solid media (w/v). Liquid media can be made without agar.

- 0.22g MS medium powder
- 1% sucrose
- 0.9% Agar (plant cell culture tested)

10x PBS:

Dissolve the following in 800ml distilled H2O.

- 80g NaCl (1.37 M)
- 2.0g KCl (27 mM)
- 14.4g Na2HPO4 (100 mM)
- 2.4g KH2PO4 (18 mM)
- Adjust pH to 7.4
- Adjust volume to 1L with additional distilled H2O
- Sterilize by autoclaving
- Prepare 1x PBS solution by diluting 10x PBS stock solution 1:10 in distilled H2O and adjust pH to 7.4

Paraformaldehyde fixative

1. Prepare 4% paraformaldehyde in 1x PBS solution to a final volume of 100 ml.
2. Heat and stir the solution to approximately 60°C. Take care that the solution does not boil! Don't go over 70°C!
3. The powder will not immediately dissolve. Slowly raise the pH by adding NaOH or KOH dropwise from a pipette until the solution clears
4. Once the paraformaldehyde is dissolved, recheck the pH, and adjust it with small amounts of HCl to approximately 6.9 pH
5. Cooldown the solution before use

Note: The fixative can be aliquoted and kept at +4ºC or - 20Cº. Always use fresh PFA to prepare the fixative. The fixative can be stored for up to one week (4°C) or two weeks (-20°C)!

ClearSee solution

Dissolve the following in distilled H2O (w/v)

- Xylitol [final 10% (w/v)]
- Sodium Deoxycholate [final 15% (w/v)]
- Urea [final 25% (w/v)]
- Water to the final volume

Note: Sodium Deoxycholate is volatile. One has to weigh this in a chemical hood. Ideal when urea and xylitol are first dissolved in water. Care has to be taken to mix the chemicals in a minimum amount of water as the chemicals themselves weighs 50 % (w/v). Stir the solution until all the chemicals are dissolved and the solution looks transparent.

2.19 User guide for annotation and analysis of complex 3D plant organs using 3DCoordX

This user guide will follow a step-by-step procedure to analyze and annotate organ coordinates for complex 3D plant organs using 3DCoordX toolbox in MorphographX (MGX, www.morphographx.org). 3DCoordX is an add-on for the MGX platform (Barbier de Reuille et al., 2015; Strauss et al., 2021). For general MGX questions such as installation or usage please refer to the user guide of the software (available online or from the MGX user interface via Help-User Guide).

2.19.1 Ovule sample datasets

3DCoordX can be used for creating organ-centric coordinate systems of any 3D organ. This guide uses two data sets of young and mature ovules of *A. thaliana*, respectively. Each data set consists of different 3D stack tif files (SR2200 cell wall signal (Tofanelli, Vijayan et al., 2019), PlantSeg CNN predictions of the cell wall and the 3D instance segmentation using GASP method in PlantSeg (Wolny et al., 2020)), mesh files (for the analysis in MGX), csv files (with stored cellular data) and a mgxv file to load the necessary files for this user guide.

The young ovule dataset (“dataset_598”) consists of four multiple stage ovule primordia. Mature ovule dataset (“dataset_401”) consists of one stage 3-IV mature ovule. To load a sample data set, either drag and drop the corresponding file into MGX or open the mgxv file. The attributes .csv files containing heatmap values have to be loaded using the MGX process (`Mesh/Heat Map/Heat Map Load`). Attributes .csv file containing cell type information can be loaded using process in MGX (`Mesh/Cell Types/Load Cell Types`).

2.19.2 Quantification of general cellular attributes and data export

MGX allows the quantification of 3D cellular properties, such as geometry or location. For a complete list of measures and methods we refer to the MGX user guide or (Strauss et al., 2021). Basic attributes of 3D cells such as cell volume, area, neighbours, outside wall area ratio etc can be automatically quantified using the ñMesh/CoordX/A Cell Analysis 3Dò process. More advanced geometric features such as cell length along a custom direction can be quantified by first assigning a cell axis that represents the measurement and then using the process ñMesh/Cell Axis 3D/Shape Analysis/Compute Shape Analysis 3Dò. For cell length, a Bezier curve was placed along the central axis of the cell and a measure of length along the central axis was quantified as a

next step using the cell axis 3D process. (Cell Axis 3D/ Custom/Create Bezier Line direction). Any computed cellular property can then be exported, together with the organ coordinates created using this guide, using the process Mesh/Attributes/Save to CSV.

2.19.3 The 3D CoordX AddOn

Within MGX, the 3D CoordX toolbox is a collection of processes that can be found in the “Process” tab and “Mesh” subtab in the “CoordX" folder. The goal of the 3D CoordX pipeline is the creation of organ-centric coordinate systems for 3D segmented organs of different complexities. In the following section, this process is described using the example of young and mature ovules.

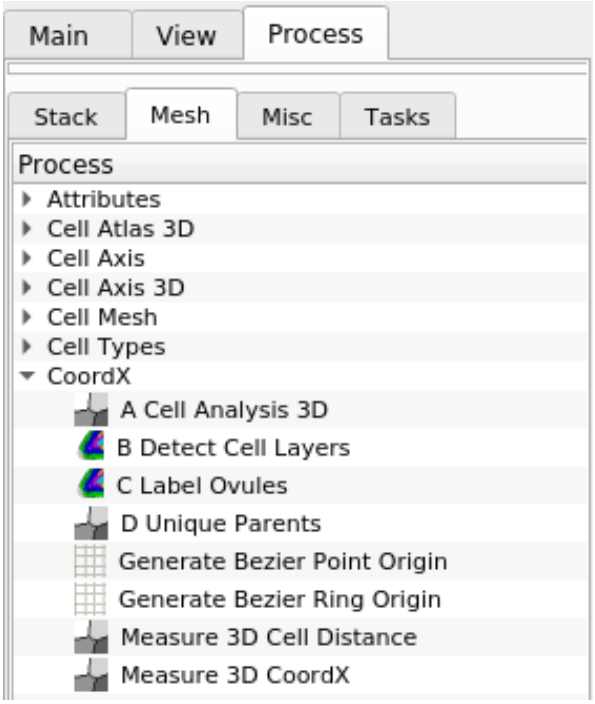

Figure 2.5 MorphographX process sidebar displaying the 3DCoordX add-on under the mesh section. 3DCoordX associated new processes are listed. Listing A-D are the processes for layer detection in organs. Processes for bezier ring origin creation and to measure the 3DCoordX organ coordinates are listed below. Other processes mentioned in this user guide are already existing ones and can be found in their specific locations as mentioned in the MGX user guide.

2.19.3.1 Workflow for young ovules data set using 3DCoordX

Young ovules are connected and are finger-shaped organs extruded from the placenta. The sample data set consists of 4 ovules which will be labeled by cell type and organ. For each of the ovules, cell layer labeling and a proximal-distal coordinate system can be created.

1. Download and open the “Young_ovules” folder. Load the file “Analysis_young_ovules.mgxv” into MGX.
2. 3D cell layer detection
    a. Layer detection is working best when the segmented image stack is also provided. When opening the mgxv file as described in step 1, it is already loaded in Stack1/Main. Alternatively, it can also be loaded using

the Stack1->Open menu. It is also possible to perform the layer detection with just the 3D cell mesh as an input in mesh1.

b. Run the process "`Mesh/CoordX/A Cell Analysis 3D`" to obtain general cellular attributes and the connectivity map.

c. Run the process: "`Mesh/CoordX/B Detect Cell Layers`". Set the parameter "Use Stack" depending on whether you are using the stack or not. The other parameters typically can be left on the default values. Results of cell layer detection can be visualized on the cell mesh by enabling the parent label view of the cell mesh. (Mesh/labels, Parents/on, hide the stack from display). The number of layers to be detected can be also altered in the parameter field.

d. Note that MGX saves cell types such as the layers in the "parent label" which acts as a secondary cell label. Only one set of parent labels can be active for visualization, but additional ones can be saved in attribute maps.

3. 3D Organ detection

a. Select a cell at the distal tip of each organ primordia. In a case with multiple ovules attached to the same fragment, one can select the distal-most cell in all the ovules together.

b. Run the process: "`Mesh/CoordX/C Label Organs`" with a parameter defining the number of cells along the proximal-distal axis of the ovule. For this sample data, 8 is a good value.

c. Results of ovule ID detection can be visualized on the cell mesh by enabling the parent label view of the cell mesh.(Mesh/labels, Parents/on, hide the stack from display).

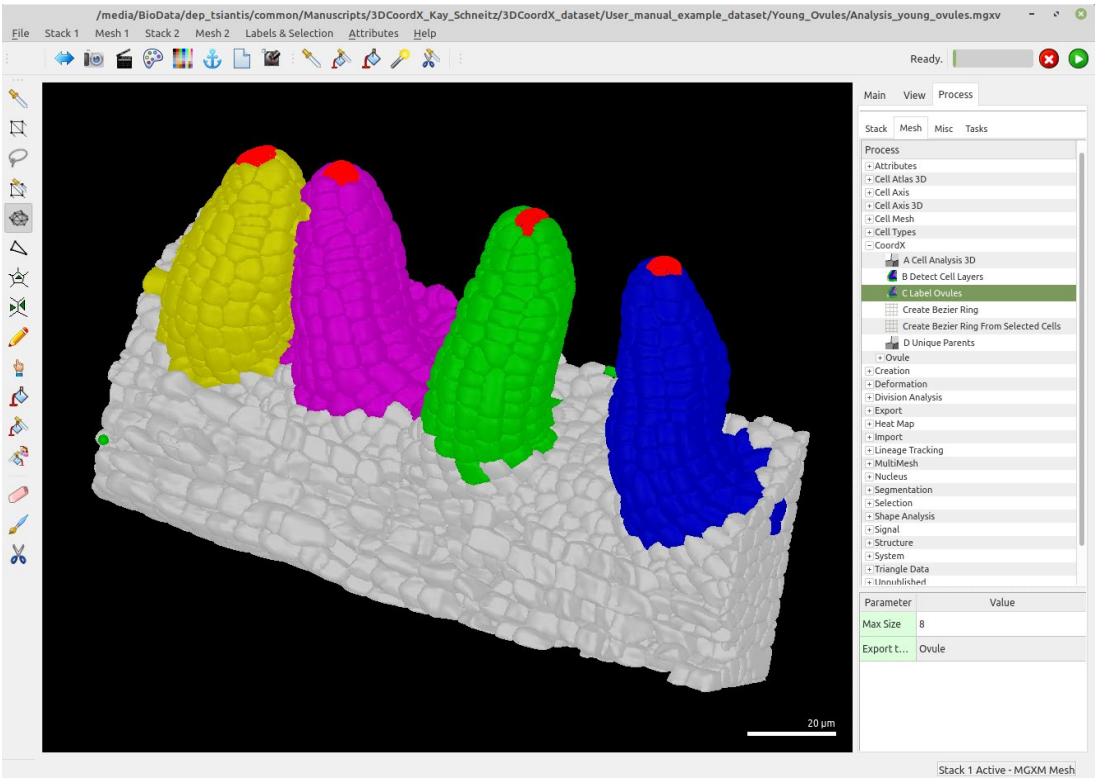

Figure 2.6: **MorphographX GUI loaded with the young ovule data set.** Different colors indicate the results of organ detection and separation into different IDs as in step 3 of 3.1.

4. Combination of layer and organ detection to create unique layer labels
    a. In the previous steps we created labelings for cell layers (step 2) and ovule IDs (step 3). Both labels were saved in attribute maps with names that were provided as parameters (by default:"Layers" and "Ovule"). In this step both labeling maps are merged to create a new labeling with unique ID for cell layer/ ovule labels.
    b. Use the process: "`Mesh/CoordX/D Unique Parents`". Before running the process make sure the names of the "Parent Attr" parameters are matching the names of the previously generated label maps. In our case set them to "Layers" and "Ovule", then run the process. The result of the unique parent labels can then be displayed again by enabling the parent view of the mesh.
    c. At this stage it is a good time to perform proofreading of the unique label results to optimise the dataset for further analysis.
    d. Any cell type or ID label can be exported to attribute maps or saved separately as a .csv file using the process "Mesh/Cell Types/Save Cell Types".

5. Separation of the anterior and posterior region
    a. Anterior and posterior separation can be done by manually subgrouping the existing labels of the primordia from the previous step. For the sample data this has been done already and saved to a csv-file containing the cell type labels. Nevertheless, the next step explains how to do this.
    b. To assign cell type labels to individual cells or groups of cells, they need to be selected using the mesh tools “Add label to selection” or “Select connected component”. After selecting cells their parent/cell type label can be set using the process “`Mesh/Cell Types/Set Cell Type`”. Different tissues are annotated one after the other.
6. Creation of CoordX PD organ coordinates for a young ovule. (Note that the following steps create the coordinates for a single ovule. If you are interested in creating the coordinates for multiple organs in the same mesh, as it is the case for the sample data, then the steps need to be executed for each organ.)
    a. Run the process: ”`Mesh/CoordX/Create Point Origin`” which creates a small point-like Bezier ring. The Bezier needs to be placed at the distal tip of the ovule primordia of interest. The Bezier is located in the “Cutting Surface” and can be moved by two different methods: 1) Go to the “View” tab and activate “CutSurf” in the “Control-Key-Interaction” menu. Now the Bezier can be moved and rotated when holding the “Ctrl”-key together with pressing the left or right mouse button and moving it. 2) The Bezier points (in yellow) can be selected using the “Select Points in Mesh” mesh tool. Once selected they turn red and can be moved with the “Alt” and right mouse key held. Please see the MGX user guide for more details about the general controls.

b. To create the PD-coordinates of the organ use the process: Mesh/CoordX/ Measure 3DCoordX. This process contains several parameters. Firstly. the type of coordinates must be defined, such as distance coordinates (parameter wall weight = Euclidian) or cell coordinate parameter (wall weight = 1). Secondly, decide if the measurement has to be restricted within the parent labels or tissues. When the distance is restricted within the parents, the “direct distance limit” parameter can be used to assign the cells where the distance seeds are first assigned. The direct distance limit first finds the cells within this limit to the origin, then it determines their distances to the Bezier ring. Once the distances are assigned to these direct cells, they act as distance seeds for the remaining cells within their tissue. The distance coordinate of another cell far from the direct distance limit is quantified by finding its distance to the nearest neighbour cell which was within the direct distance limit and within the same tissue type. Essentially the cells within the direct distance limit act as distance seeds to the remaining cells within their tissue. “Selected as direct” allows to include any additional manually selected cell to the direct distance limit which might not be within the direct distance limit. The direct distance limit for ovule primordia was set to a minimum value of 7µm. The “Consider ring orientation” parameter allows to include a positive and negative direction to the cell coordinates across the 2D plane of the ring. This is irrelevant for the ovule primordia where the Bezier ring is positioned at the distal end. For primordia coordinates, the 3D mesh where the cell distance Bezier ring was quantified has the tissue labels of anterior and posterior L1, L2 and L3 cells. This allows to restrict the distance measurement to the anterior or posterior halves of the organ, respectively, and not just within the layers.

c. The distance coordinates heatmap can then be exported as an attribute using the process: “`Mesh/Heat Map/Operators/Export Heatmap to Attribute Map`”.

d. The same Bezier ring is then moved to the distal tip of the other ovule primordia of interest to annotate the distance coordinates of this ovule.

7. Further data export (e.g. volume or signal quantification)

    a. Run process: "`Mesh/Attributes/Export to CSV`". A collection of all the attributes is then exported as a .csv file. (See also the section "Quantification of general cellular attributes and data export").

2.19.3.2 Workflow for mature ovule data set using 3DCoordX

1. Download and open the "Mature_ovule" folder and load the file "Analysis_mature_ovule.mgxv" into MGX.
2. Cell type assignment
    a. For mature ovules, tissue labels were manually performed by mesh selection tools. Integument tissues were labelled as adaxial or abaxial layers (oi1, oi2, ii1, ii2) (Vijayan et al., 2021). For the sample data set a cell type labeling is already provided, which can be used for the following steps. Alternatively, cell types can be manually assigned by selecting cells and using the process: "`Mesh/Cell Type/Set Cell Type`" or using the cell type classification tools (see the MGX user guide and Strauss et al. 2021).
3. Creation of CoordX medial-lateral (ML) organ coordinates
    a. A thin clipped plane of about 5 µm was enabled at the mid-sagittal section of the ovule to restrict the mesh vertice selection for the ML coordinate annotation.

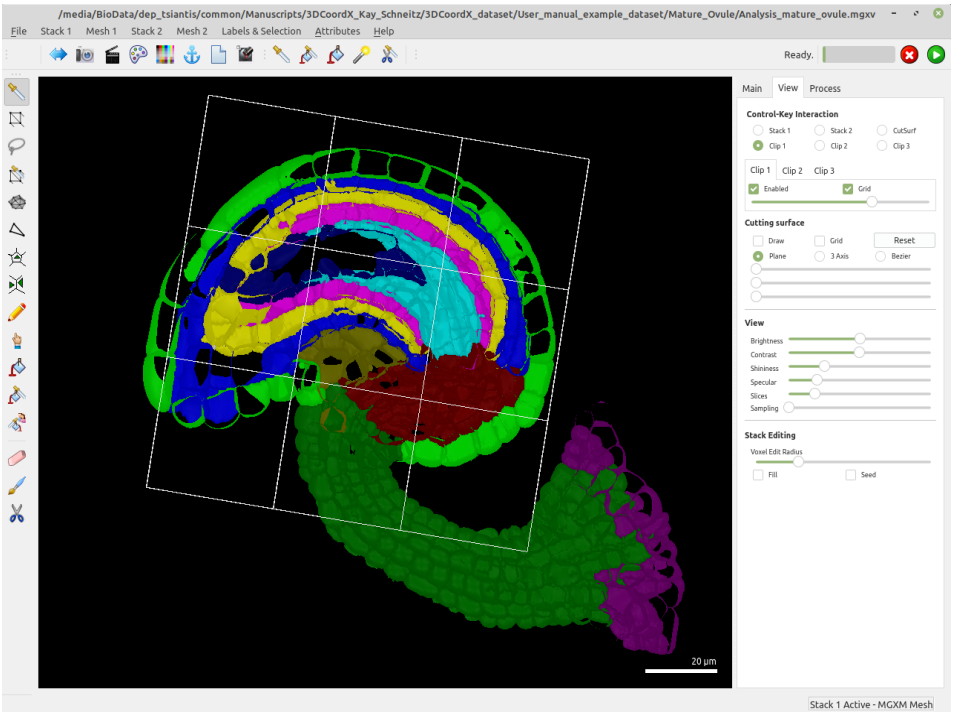

Figure 2.7: A sagittal section of mature ovule dataset displaying the cell types with mesh view. A clip plane is enabled to restrict the origin selection for ML organ coordinates to the central group of cells at the midline cells.

b. Mesh vertex selection: The lasso selection tool for mesh vertices is used to select the vertices of the cells of the central file along the posterior midline of the ovule. Selection is done only on the midline cells of all four layers of integuments.

c. Extend connectivity: The midline selection was then extended to entire cells using the process: "`Mesh/Selection/Extend by connectivity`".

d. Run the process: "`Mesh/CoordX/Measure 3D Cell Distance`" to create the ML coordinates: Selected cells are used as origins of the ML coordinate system and the coordinate values of the number of cells are measured from the cell distance process. The parameter "`Restrict connectivity to the same cell type`" must be enabled and the wall weight parameter is set to 1 as the distances in terms of the number of cells are quantified here.

4. Creation of CoordX proximal-distal (PD) organ coordinates

a. Run process: "`Mesh/CoordX/Generate Bezier Ring`" with a radius of 15µm (parameter "Radius X" and "Radius Y"). The Bezier ring was semi-automatically placed at the proximal end of integument initiation by selecting the appropriate cells and setting the parameter "Place Ring at Selection" to "Yes".
b. Placing the ring was already a two-step process. First, a 3D mesh was loaded and the OI2 layer was removed from the mesh to visually select the group of ring-like proximal OI1 cells where the Bezier ring has to be placed. Unlike in the young ovules, here the Bezier ring is larger and acts as a wide origin and is placed inside the organ. Once the Bezier ring is placed, it is saved with the mgxv file using File/Save-as. The original 3D mesh with appropriate tissue labels for the coordinates was reloaded to mesh 1 to obtain the organ coordinates in the next step.
c. Run the process: "`Mesh/CoordX/Measure 3DCoordX`". The direct distance limit for the mature ovule was set to 15 µm. "`Consider Ring Orientation`" was set to yes. Other parameters were set the same to that with ovule primordia. This generated the coordinates below the ring (proximal region of ovule including funiculus and proximal OI2) with negative values and coordinates above the ring with positive values allowing.

5. Further data export (e.g. volume or signal quantification)
   a. Run process: "`Attributes/Export To CSV`". A collection of all the attributes is then exported as a .csv file. (See also the section "Quantification of general cellular attributes and data export").

2.20 Detailed protocol to sample preparation and microscopy of Arabidopsis ovules (Adapted from Tofanelli, Vijayan et al., 2019)

2.20.1 SOLUTIONS AND MATERIALS

- Fixative solution
- Clearsee solution
- Washing solution (1x PBS)

- Coverslips 22x22 mm, 0.17 mm thickness (No. 1.5H, Paul Marienfeld GmbH & Ko. KG, Lauda-Königshofen, Germany, CAT No: 0107052)
- Insulin Syringes U-40 (1 ml/40 I.U.) (B. Braun Melsungen AG, Melsungen, Germany, CAT No: Inject® 40 Duo)
- Microscope Glass Slides 76x26x1 mm (Paul Marienfeld GmbH & Ko. KG, Lauda-Königshofen, Germany, CAT No: 1000000)
- Petri dish large 94/16 mm (Zefa, CAT No: 10029880)
- Petri dish small 35/10 mm (Opti-Lab GmbH, Munich, Germany, Petri dishes, CAT No: 6055567)
- Double sticky tape (Kaut-Bullinger, München, Germany, CAT No: 055446)

2.20.2 STAINING SOLUTION: SR2200 and TO-PRO®-3

Prepare a combined staining solution containing 0.1% SR2200 and 1µM TO-PRO®-3 (1:1,000 dilution of 1 mM stock solution) in fresh 1x PBS.

NOTE: Avoid exposure to light. For best results, prepare a fresh solution. SR2200 stock solution should be allotted to avoid repetitive handling of the original stock. The solution showed a tendency to crystallize. Staining could be also done for individual stains rather than combined.

2.20.3 PROTOCOL FOR FIXATION AND CLEARING

1. Check the inflorescence and select the flowers at the proper stage.
2. Harvest the flower and place it under the stereomicroscope on a double-sided tape on the bottom of an inverted petri dish. Excise the carpel by using tweezers. With the help of a needle slightly open the carpels to expose the ovules.
3. Quickly transfer the pistils to a small petri dish with a double-sided tape fixed at the bottom and containing fixative. If necessary, pistils can be attached to the double-sided tape to prevent them from floating. The fixation is effective when the pistils sink to the bottom of the petri dish.
4. Fix for at least 1 to 2 hours at room temperature with gentle agitation or overnight at 4°C.

5. Samples should be transferred to 1.5 ml microcentrifuge tubes containing fixative for overnight or long-term storage (up to one month).
6. If fixation was done at 4°C put the samples at room temperature for at least 30 min before proceeding
7. Carefully remove most of the fixative but make sure samples are still submerged. Don't let samples dry out!
8. Wash twice the fixed tissues for 1 min in 1 x PBS. The washing step is very important to avoid the formation of precipitates that occur upon addition of ClearSee to the fixative!
9. Transfer the fixed and washed carpels to a 1.5 ml microcentrifuge tube containing 1 ml ClearSee solution and clear them at room temperature overnight with gentle agitation. Overnight is usually sufficient. Slightly better results are obtained upon 2 to 3 days of clearing. Change the ClearSee solution after 2 days if samples are stored for extended periods of time. The carpels are stable in ClearSee for several weeks. After clearing, proceed to the staining procedure.

2.20.4 PROTOCOL FOR STAINING

1. Wash the cleared tissue for 1 minute in 1x PBS solution containing 0.1% SR2200 stain.
2. Transfer the pistils to another small petri dish filled with 1 ml of combined staining solution (0.1% SR2200 and TO-PRO®-3 (final dilution of 1:1,000)).
3. Stain at room temperature with gentle agitation for 20 minutes.
4. Wash the stained tissues for 1 minute in 1 x PBS.
5. Transfer the pistils into ClearSee solution for 20 minutes with gentle agitation for final clearing.
6. Proceed immediately with mounting.

NOTE: Use tweezers to transfer the pistil from one solution to another. Avoid mixing the solutions.

2.20.5 PROTOCOL FOR SAMPLE MOUNTING

1. Stick two coverslips with transparent nail polish on a microscopy slide leaving around 1.5 cm between them.
2. Place a drop of VECTASHIELD® antifade mounting medium on the slide.
3. Gently pick up a pistil from the ClearSee solution and place it on the slide with a minimum of ClearSee solution transferred.
4. Dissect the ovule from the pistil with the help of syringe needles.
5. Gently place a coverslip onto the sample and store the slide in a slide box at 4 °C.
6. Let the samples settle for 1 to 2 hours before imaging.

NOTE: For best results, imaging should be done 1 to 2 hours after mounting and on the same day. However, slides with samples mounted in VECTASHIELD® can be stored for up to a week with minor reduction in image quality.

2.20.6 GENERAL NOTES

- It is important to carry out every step under the stereomicroscope with minimal light exposure to avoid drying and shrinkage of ovules.
- Tissue in fixative can be stored for several weeks.
- Prior to staining samples can be stored in ClearSee solution for several days to optimize clearing.
- Do not store stained tissue in ClearSee solution for more than 2 hours prior to mounting in VECTASHIELD®) to avoid degradation of image quality.
- The protocol can be combined with fluorescent reporters. However, in such a case it is recommended to fix the samples for only 1 or 2 hours and perform clearing only overnight. Otherwise, the fluorescent signal is weaker or lost completely.
- The support coverslips prevent the mounted ovules from squeezing. The coverslip thickness may not exceed the working distance of the objective used for imaging.

3 Results

3.1 Method development for generating 3D Arabidopsis ovule atlas with cellular and tissue resolution

Arabidopsis ovules have long been established as a model system for the study of organogenesis in plants. Different imaging methods allow to follow the development of Arabidopsis ovules, but there exists no precise method for straightforward investigation of 3D cellular morphology of the whole organ including the internal tissue architecture with cellular resolution. This chapter focuses on method development for generating reliable 3D digital images of Arabidopsis ovules. Improvements are made in different directions. With advancement in method for sample preparation and microscopy for whole mount organ imaging, machine-learning-based image improvement and instance cell segmentation (Tofanelli, Vijayan et al., 2019, Wolny et al., 2020, Vijayan et al., 2021), this study was able to generate precise 3D models of Arabidopsis ovules which can be used for further detailed investigation. Additionally, nuclei segmentation was also optimized for a faint nuclei channel that has application in further proofreading cell segmentation.

3.1.1 3D whole organ microscopic imaging

Investigation of 3D organ architecture with cellular resolution requires precise imaging of cells in 3D. This section is based on the publication from this study (Tofanelli, Vijayan et al., 2019). I initially focused on mature ovules with a live imaging approach. A plasma membrane-localized FP (pSUB::SUB:EGFP) STRUBBELIG (SUB) which encodes a receptor kinase involved in tissue morphogenesis was used for visualization of the cell membrane (Chevalier et al. 2005; Vaddepalli et al. 2011). The outer surface of the organ can be clearly visualized with the 3D surface view of reporter images, but this approach lacks cellular resolution as one goes deep inside the organ (Figure 3.1-A). The approach following live imaging was to image the same sample at multiple angles to attain full 3D reconstruction (results not shown). The overall method seems more complicated as it might need to involve stitching from multiple angle images and the results are also not promising full 3D resolution with precise cell boundaries that can be used for further processing.

The second approach was to devise a method that allows imaging the whole organ from just one direction and that avoids any image stitching. Tissue clearing is widely used on different samples to make them transparent such that laser penetration at deeper layers is possible. TDE based clearing is one such approach for organ imaging as has been shown in (Aoyagi et al. 2015). TDE clearing first involves sample fixing in fixative (4% Paraformaldehyde containing 1x PBS solution) followed by clearing in TDE and further staining and mounting. However, TDE clearing was not satisfying for Arabidopsis ovules (Figure 3.1-B). TDE clearing improved the surface visualization of ovules compared to the live images, but the internal cellular architecture can't be detailed with TDE cleared samples which are stained with any fluorescent dyes. For the study, a widely used bright fluorescent stain SR2200 (Musielak et al. 2015) was used. The results from TDE cleared images still indicated that fixing the samples helped in improving the whole procedure than a live tissue.

Modified Pseudo-Schiff propidium iodide (mPS-PI) method is a classical method for cleared tissue imaging that involves clearing with Hoyer's medium (which includes carcinogenic chloral hydrate as a clearing agent) (Berleth and Jürgens 1993; Grossniklaus et al. 1998). The method is promising and was successful in imaging the whole Arabidopsis ovule at a mature stage (Figure 3.1-C). The method involves fixation with paraformaldehyde and further washing and clearing with multiple solutions that takes almost a week-long procedure. mPS-PI method severely improved the image quality for morphological studies, but it's incompatible with fluorescent stains and fluorescent reporters. This restricts the usefulness of this method for any gene expression analysis or reporter activity analysis in 3D.

ClearSee is another method for clearing plant tissues for confocal imaging. ClearSee is essentially a clearing solution with three components. Detailed information in the materials and methods (Section 2.4, 2.18). ClearSee has been proven to be compatible with any fluorescent reporter-based line. ClearSee clearing of Arabidopsis ovules was excellent compared to TDE or live images.

This study found that the original ClearSee method described in (Kurihara et al., 2015) is not optimal for precise whole organ imaging of Arabidopsis ovules. It could be because the protocol uses ClearSee solution for dissolving fluorescent cell wall

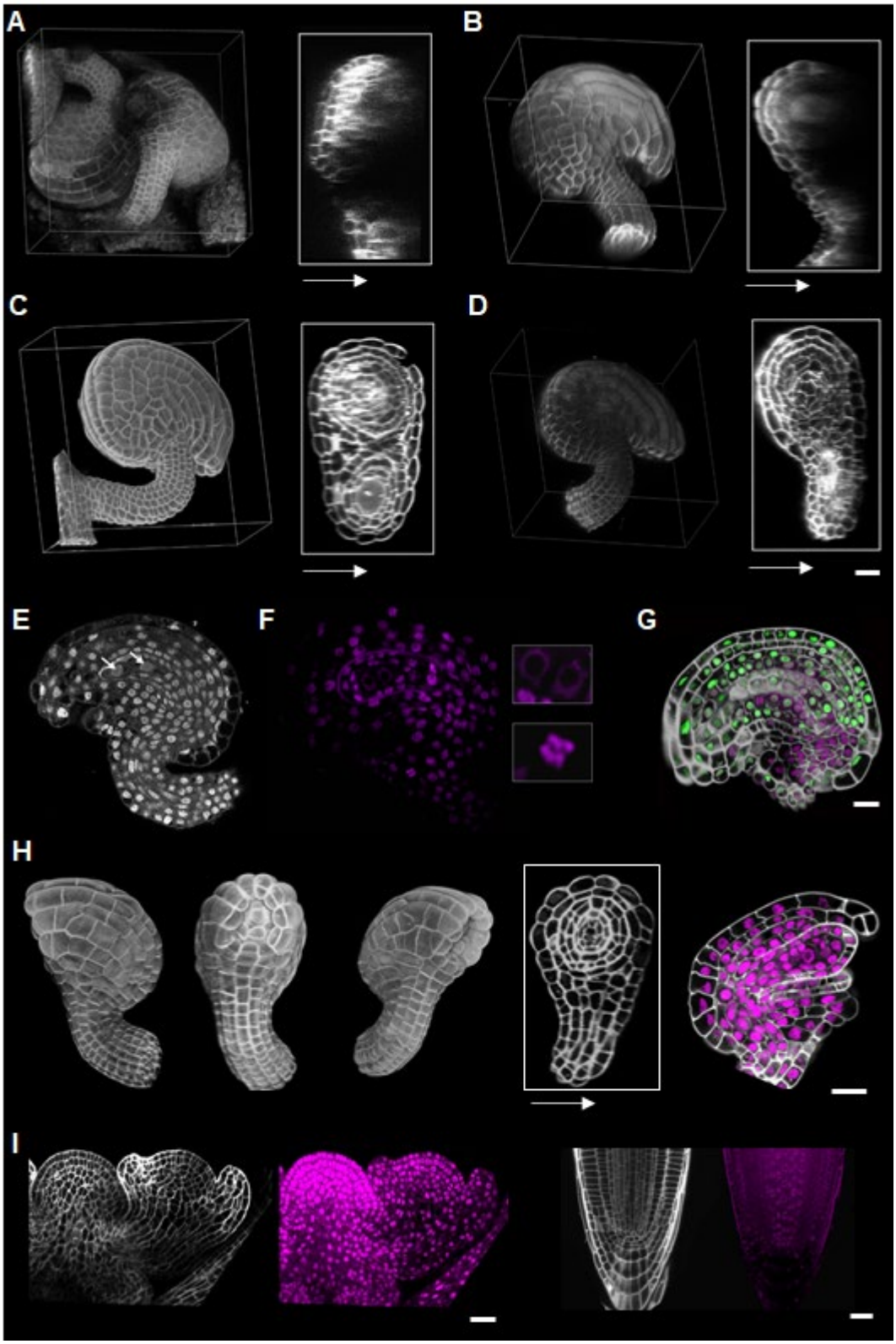

Figure 3.1 Improving 3D whole-mount confocal imaging of Arabidopsis ovules
3D surface view and 2D orthogonal section view of XZ plane from multiple imaging approaches (A-D). Live ovule image (A), Fixed and TDE cleared ovule (B), Classical m-PS PI method (C), ClearSee original method (D). A 3D bounding box is projected

on the left images displaying the 3D surface view. Right image represents an xz section view within a box, white solid arrow indicates the z direction and with declined signal intensity in all these methods except the mPS-PI method. 2D midsagittal section view of a DAPI stained mature Arabidopsis ovule (E). Arrows indicate the embryo sac nuclei. 2D midsagittal section view of a TO-PRO3 stained sample (F). Highlighted boxes indicate the embryo sac nuclei and mitotic nuclei. Different views of ovule stained with the improved method from this study (G, H, I). Results demonstrating the application of fluorescent reporter imaging along with cell wall and nuclei stain in the improved method (G). Cell wall stain in white, nuclei stain in magenta, fluorescent reporter activity in green. (H) From left to right, 3D rendered image at multiple viewing angles from the z stack obtained from one-directional imaging using the advanced imaging method, a 2D section of an orthogonal XZ plane displaying less variable signal intensity along the z direction indicated by solid whote arrow, a 2D section displaying the overlay of cell wall stain with the nuclei stain. The method has wide applicability, it can be applied on Arabidopsis inflorescence meristem and Arabidopsis roots as an example (I). Scalebar 20 µm. Some images of the figure are modified from Vijayan et al., 2021 and Tofanelli, Vijayan et al., 2019.

stains. The issue is with consistent stain penetration to deeper tissues. Visual analysis of ovules imaged with the original ClearSee method displays a variable cell wall stain quality and highly reduced signal intensity along the z direction (Figure 3.1 D). This is expected due to z scanning in one direction.

Essentially, the laser penetration through the sample is a little improved in the original method such that at least a weak signal can be detected at the bottom z direction of the sample. But such weak signals don't help when it comes to whole organ image segmentation. For precise 3D reconstruction and further segmentation of the organ, the signal intensity throughout the sample has to be optimal.

Original ClearSee methods describe samples to be mounted in ClearSee for final imaging. This study found that there is severe photobleaching when samples are mounted in ClearSee. Only one ovule could be images from a microscopic slide containing about forty ovules from the same carpel for this reason. Overall the method is much more reliable than other described imaging methods, but it needs further optimization for imaging ovules with fine quality requirements.

3.1.1.1 Improving 3D confocal imaging of the whole organ for cell segmentation

The below section describes the improvement in microscopic imaging of Arabidopsis ovules based on this study and the publication from the same study (Tofanelli, Vijayan et al., 2019). This study focuses on the ClearSee method and in improving the same for precise microscopic imaging. The issue with cell wall stain penetration is of priority. This study proposes 1x PBS as the staining solution instead of ClearSee for staining procedure. Results indicate significant improvement in image quality (Figure 3.1 H).

The signal to noise ratio is highly improved and the faintly stained walls observed with staining in ClearSee solution was improved when the staining was performed in 1xPBS containing stains.

Arabidopsis mature ovules from stage 3-I to 3-VI are staged according to the number of nuclei at the embryo sac. This makes it complicated that a nuclei stain must be included in the staining method always. Also, a nuclei stain can be used to confirm the presence of an enclosed cell in 3D when the cell wall stain is faint to distinguish between cells. DAPI nuclei stain was performed on ClearSee cleared samples, the stain works well, and the embryo sac nuclei can be also identified with DAPI (Figure 3.1-E). Nevertheless, the widely used bright cell wall stain SR2200 can't be combined with DAPI stain as they come under the same excitation and emission spectra. Another fluorescent nuclear stain TO-PRO-3 iodide (TO-PRO-3) was also a good alternative for ovule tissues. It can be used as a counterstain with the cell wall stain SR2200 as its excitation is at 641 nm. TO-PRO-3 can be also used to identify mitotic cells that show a different pattern compared to regular nuclei (Figure 3.1-F). Nevertheless, TO-PRO-3 stain was sensitive to ClearSee mounting and that resulted in the rapid disappearance of the stain after mounting and in a failure to image the complete 3D structure of ovules simultaneously co-stained with SR2200 and TO-PRO-3.

To further improve the method for precise 3D imaging of ovules where the signal from both the channels are preserved, the mounting media was altered. Vectashield is an antifade mounting media that is widely used in fluorescent imaging. Finally, ovules were fixed in paraformaldehyde fixative, cleared in ClearSee, stained with SR2200 and TO-PRO-3 in 1xPBS solution, recleared in ClearSee and mounted in vectashield. Detailed protocol for the improved whole imaging procedure can be found under materials and methods section 2.4 and 2.20. Vectashield mounting also helped in avoiding any mismatch in refractive index as the sample mounting and microscope objective immersion are both glycerol-based. Vectashield mounting has additionally improved the resolution in the z-direction, overall image quality and retained both the cell wall stain and nuclei stain intact after mounting. The antifade property of the mounting media now allowed to image multiple ovules from the same slide. Overall, the resulting images were sharp and they passed the quality requirements for further 3D morphological analysis with cellular resolution.

Prepared samples can be also stored in vectashield for a few days where the nuclei stain would be retained, samples could be stored for up to eight months or more when only cell wall stain is required. As a final refinement, the image quality was set to a maximum possible threshold with only a little oversampling. This included imaging the samples using a HyD detector and acquiring 12 bit or 16bit images instead of 8-bit images and with a line average of 2-4. The voxel size of images used in the dataset was according to the Nyquist criterion. This was about 0.063 x 0.063 x 0.24 μm^3 XYZ directions. Overall, all these factors have significantly contributed to the full refinement of images in 3D and its usability for further processing.

Moreover, we could confirm that this method has wide applicability in other plant tissues and all stages of Arabidopsis ovules could be imaged precisely with this improved ClearSee method. Arabidopsis shoot apical meristem is widely studied at the epidermis or up to two to three layers deep. With this advanced method, this organ can be now studied in much more detail. The same applies to Arabidopsis root, using this method, the root can be imaged through and whole 3D morphological analysis could be made rapidly (Figure 3.1 I). The method is also proved to be applicable with ovules from other organisms like Tomato, Lotus, Cardamine etc (Results not shown).

3.1.2 3D whole organ instance cell segmentation

This section focuses on methods developed for transforming microscopic images into quantitative cellular descriptors. In other terms, this section aims to optimize the 3D instance cell segmentation from raw microscopic images obtained from section 3.1.1. This section is adapted based on the publication from this study (Wolny et al., 2020, Vijayan et al., 2021).

A 3D dataset of microscopic images is a great descriptor for the overall morphology and shape of the organ, but a 3D volumetric and quantitative analysis can't be performed on raw images. 3D cell segmentation is critical for quantitative understanding of the 3D volumetric information in the images. Instance cell segmentation is achieved when the pixels enclosed by the 3D cell boundaries or cell wall are assigned by an object identity label, called a segmentation label (Figure 3.2-A). Segmentation involves locating the cell contours and cell interiors such that each cell within the image may be identified as an independent entity (Vicar et al. 2019). This is usually performed in the whole 3D image stack in an automatic manner where the results would be another image stack with segmented cells (Figure 3.2-B). This segmented image is used for creating 3D cell mesh around each labelled object or cells (Figure 3.2-A). 3D mesh would reproduce the cell boundaries and the whole organ with connected cells through mesh vertices. 3D cell mesh is more advanced than a segmented image stack in that it can handle more pieces of information and annotations which can be easily exported for quantitative analysis. Accurate cell segmentation is essential for quantifying significant biological and morphological information from 3D microscopic images including cell volume, shape, geometry, tissue features, growth rates etc.

3D cell segmentation of plant tissues is classically performed using ITK watershed segmentation algorithm integrated into MorphographX (Barbier de Reuille et al., 2015; Strauss et al., 2019). Raw cell boundary image itself is not used for segmentation, but a few images processing is performed to enhance the cell boundary connectivity at the pixel level (Figure 3.2-B). The image processing includes several rounds of Gaussian blurring and brightening the pixels such that any watershed leaks are prevented in 3D (Detained method can be found under materials and methods 2.6.1). The processed image is auto seeded, and watershed segmented with a threshold for under or over-

segmentation. The resulting segmentation requires laborious hand correction of errors in segmentation.

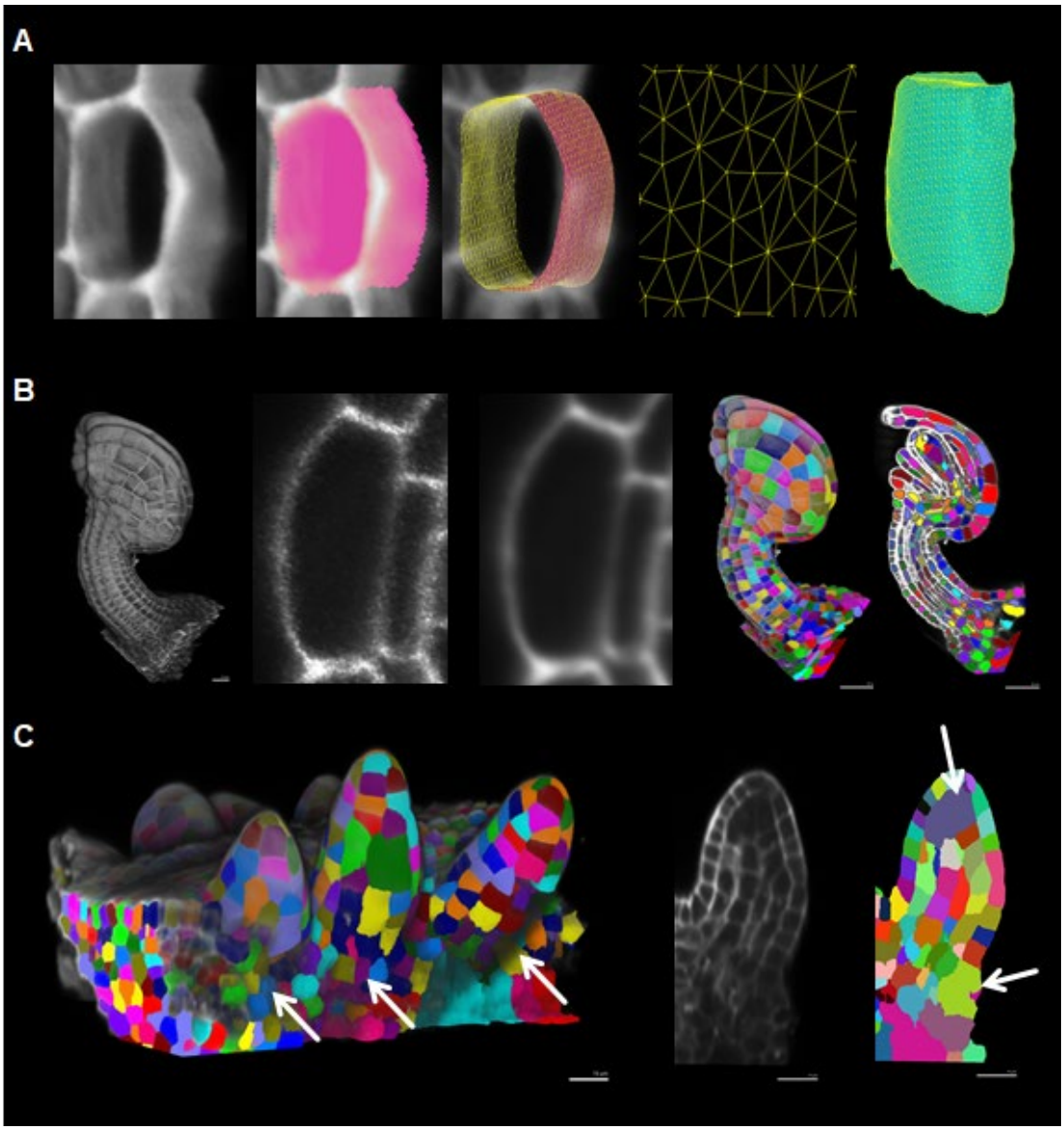

Figure 3.2 Classical ITK 3D cell segmentation of Arabidopsis ovule images

Segmentation and mesh creation from a raw cell boundary image (A). From left to right, raw cell wall image clipped from a 3D image stack zoomed into a single cell, the same cell with instance segmentation, the same cell with mesh view highlighting the vertices around the segmented label, zoomed view of few mesh vertices, 3D view of a single segmented cell mesh with an overlay of fine mesh vertices in yellow color. MGX ITK 3D segmentation of a processed cell boundary image (B). From left to right, raw 3d outer view of the image, raw 2D section view, processed 2D section view, segmented 3D outer view and segmented 2D section view. Image highlighting typical errors in segmentation (C). Under Segmentation error in an early-stage ovule where few cells are missing from the organ surface. Arrows point to regions or cells with these segmentation errors from MGX ITK segmentation method.

Despite the improvement in microscopic image acquisition, the possibility to precisely segment the images is largely dependent on final image quality. This is largely due to an unavoidable faint signal of the cell wall at some regions of the image. A typical error includes a missing cell, or several cells merged or a faint wall causing a boundary between two cells less intact. All these result in incorrect segmentation of cells which can't be used for further quantifications. Over segmentation error in 3D can be corrected, but under segmentation can't be corrected. Overall, segmentation of the whole organ with good precision is a challenge without any other modern machine learning tools.

3.1.2.1 PlantSeg deep learning based pipeline for precise 3D instance cell segmentation

The workflow of digital atlas generation is partially achieved at this stage with a high level of manual proofreading on the ITK cell segmented images of high-resolution microscopy. To further optimize and uplift the segmentation pipeline, a machine-learning-based approach is used. This study has extensively contributed datasets to set up PlantSeg, a deep learning-based pipeline for volumetric instance segmentation of dense plant tissues at single-cell resolution (Wolny et al., 2020). PlantSeg CNN (Convolutional neural network) pipeline has three modules: CNN training, CNN prediction, and 3D cell segmentation (Figure 3.3 A). The CNN training module is the beginning of the whole pipeline, this has to be ideally performed once, which generates a trained model. The model is then applied in the CNN prediction module to run on any new dataset. PlantSeg CNN prediction module processes raw microscopic images with a powerful boundary detector that is based on Convolutional Neural Networks (CNNs) (Long et al., 2015; Kokkinos, 2015; Xie and Tu, 2015). The results from the CNN prediction module are taken for 3D instance cell segmentation using the 3D segmentation module.

Once a trained model is generated, the PlantSeg pipeline is a two-step process, it takes raw images as input and produces a semantic segmentation of cell boundaries which can be instance segmented within the same pipeline (Figure 3.3 A). Semantic segmentation is visually represented as predictions or probability map of the input image (Figure 3.3 B). Here, it's called b-pmap (boundary probability map) or boundary predictions and is generated from the PlantSeg pipeline. Visually the predictions look like an improved cell boundary image of the provided input, but it contains information

about pixels in the image if it belongs to a boundary or background, hence called semantic segmentation. Instance 3D cell segmentation is when the pixels inside the enclosed semantic segmentation are further grouped into objects labels. This is performed using the GASP instance cell segmentation method (Bailoni et al., 2019). This allows performing any volumetric 3D quantification with cellular resolution.

3.1.2.2 Establishing PlantSeg pipeline and CNN network trained models

This study has improved microscopic imaging to a greater extent that now allows to produce a near-perfect cell segmentation from microscopic images after extensive manual correction. A highly proofread gold standard dataset of Arabidopsis ovules generated from this study was used as a major input for PlantSeg CNN model training using the CNN training module. First establishing the CNN training requires high-end computer knowledge and was not part of this study, this was performed by group of computer scientists in a collaborative manner (Wolny et al., 2020). This study is rather the data bank for establishing the whole pipeline. The performance of such a pipeline is greatly dependent on the trained model. An ideal model is when it can understand every pixel from the provided raw image in an accurate manner. For the same reason, a large number of input images are required for training a good-performing model. The training dataset included raw cell boundary images and their respective ground truths or instance cell segmentation. 48 volumetric confocal images stacks of Arabidopsis ovules were provided from this study. Additionally, 27 images of developing lateral roots from light-sheet microscopy was provided for training from another study. This combination has generated a versatile trained CNN model “generic_confocal_3d_unet” with good performance. Additionally, different models are generated with a combination of these datasets that allow users to perform PlantSeg predictions on input images of different voxel resolutions. The pipeline is available for public download under the PlantSeg GitHub page (Wolny et al., 2020). Detailed workflow for this PlantSeg machine-learning-based boundary prediction and segmentation can be found under materials and methods section 2.6.2. Currently, the PlantSeg model “generic_confocal_3d_unet” is of high performance for ovule image datasets. PlantSeg boundary predictions heavily uplifted the quality of final instance cell segmentation and adapted the whole procedure from raw image to cell segmentation in an automated manner (Figure 3.3 B, C). Overall, the PlantSeg pipeline is a huge success for instance cell segmentation from cell outline images. This study has valuably contributed to the different phases of establishing the PlantSeg

pipeline mainly by performing quality control and by providing datasets for training. PlantSeg now serves as a benchmark for a deep learning-based plant cell segmentation pipeline using the latest advancements in computer science

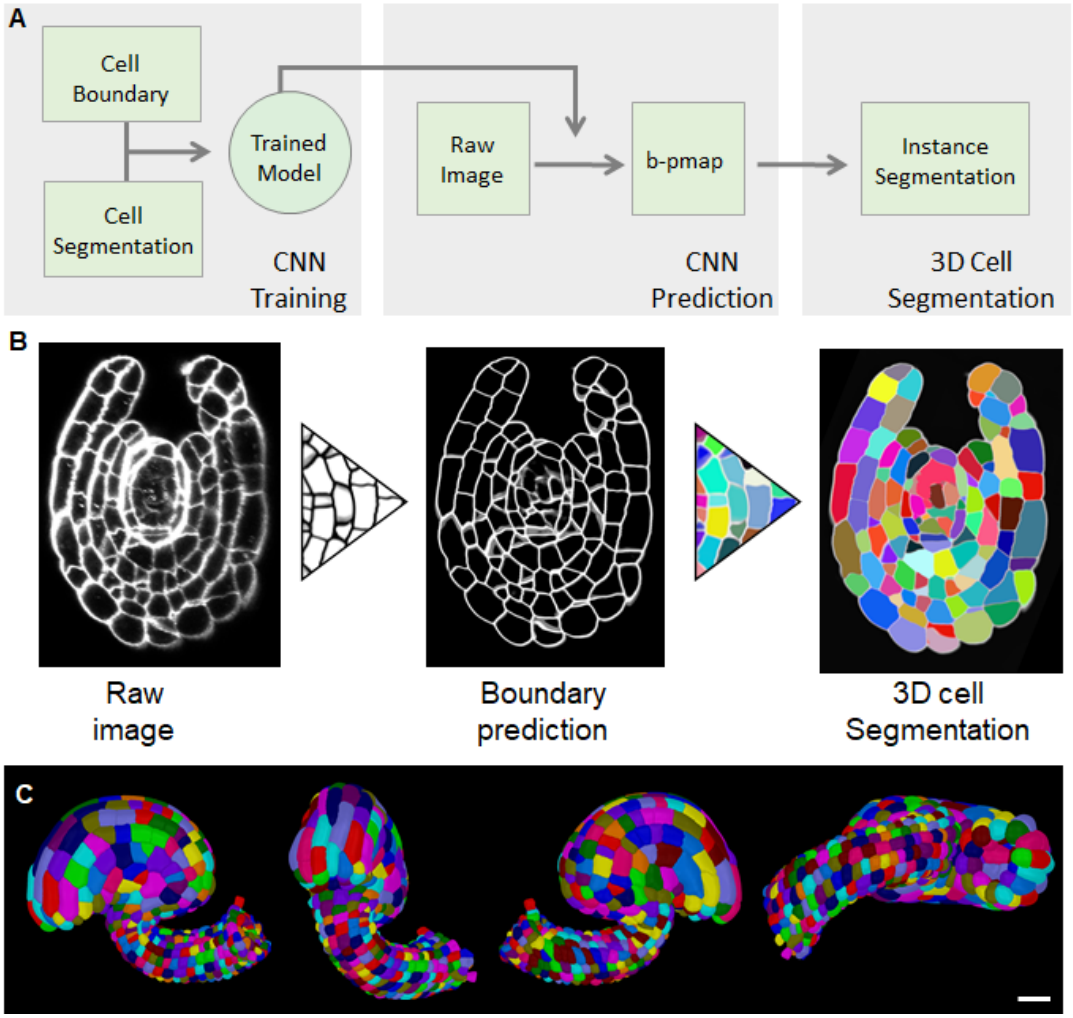

Figure 3.3 PlantSeg - Machine-learning-based image segmentation pipeline
Schematic outline of the PlantSeg cell segmentation workflow (A). The pipeline includes three modules, CNN training, CNN prediction and segmentation. Training is performed to create a model. The trained model is then used to generate b-pmap using the CNN prediction module. The resulting b-pmap is then further processed to obtain the final instance segmentation in the segmentation module. 2D orthogonal XZ section of raw image and the corresponding b-pmap and segmentation using the PlantSeg pipeline (B). A mature Arabidopsis ovule at multiple viewing angles after 3D GASP segmented using the PlantSeg pipeline.

Even after extensive optimizations and uplifting the raw image with PlantSeg machine learning-based segmentation, the procedure still resulted in a few mistakes in segmentation. In young ovules, the MMC and its direct lateral neighbors at stage 2 is the main source of error. In mature ovules, the embryo sac cell boundaries to the nucellus tissue are weak and that results in few cell errors in this region. The errors

are mainly due to the extremely poor staining of cell boundaries at these regions. It could be also due to the cell walls being particularly thin or of a different biochemical composition allowing less binding to the cell wall stain which is not detected by the PlantSeg model. Minor segmentation errors were partially corrected and were taken for further analysis.

3.1.2.3 PlantSeg-MGX hybrid method as an alternative instance cell segmentation method

PlantSeg b-pmap predicted accurate boundaries from the provided raw image. This study took that as an advantage to formulate another method for instance cell segmentation. This is a PlantSeg-MGX hybrid method for segmentation. The hybrid method uses the b-pmap from PlantSeg and blends it to the raw boundary image before ITK segmenting in MGX. The PlantSeg GASP segmentation requires a precise b-pmap. For some images, the generated b-pmap might not be accurate because of the anisotropy in raw image quality and acquisition parameters. A less accurate b-pmap doesn't perform well with GASP segmentation, whereas the b-pmap, in this case, can be merged back to the raw image and can be further segmented using the MGX-ITK segmentation. This was also performed similarly to the classical method of MGX ITK watershed segmentation where the input image is gaussian blurred to prevent any watershed leaks. The segmentation results were similar to the GASP segmentation, only that this workflow requires manual loading of the dataset from PlantSeg to MGX and processing's have to be performed in a user-selected manner. A detailed protocol of the PlantSeg MGX hybrid segmentation can be found under materials and methods section 2.6.3. The detailed method is also available to public after the publication from this study (Vijayan et al., 2021)

3.1.2.4 PlantSeg allows instance cell segmentation from low resolution microscopic images

The extensive improvement in the machine-learning-based cell segmentation pipeline now provides some additional benefits such that the quality requirements for raw images are less demanding. Several models were created by neural network training on datasets of different input sizes, this was possible because the training datasets could be down sampled to generate a low-resolution raw image and its segmentation. The best trained model "generic_confocal_3d_unet" was trained on 2.4 times low

resolution image dataset than the original image requirements for segmentation without any machine-learning-based inputs. Now, with the advancement of PlantSeg, the model requires raw images equivalent to its down sampled training dataset of voxel size 0.150µm x 0.150µm x 0.235µm in xyz axis. Whereas the original image requirement without PlantSeg was of voxel size 0.063µm x 0.063µm x 0.234µm (xyz).

To the end, new images were acquired with a compromised voxel size of 0.12µm x 0.12µm x 0.234µm in the xyz axis with a minimum line average of 3. This helped in fast imaging of a less fine raw input image processed via PlantSeg. The resulting segmentation was as good as with the original imaging requirements using the PlantSeg pipeline. Detailed protocol for the improved imaging could be found under materials and methods section 2.5. Overall, using PlantSeg pipeline, ovule imaging is now quick, and this has significantly helped in increasing the number of image datasets for the study.

3.1.3 3D Instance nuclei segmentation and its application

3.1.3.1 3D nuclei instance segmentation of the nuclei stain and fluorescent tdTomato signal

The 3D ovule image dataset includes a nuclei stain channel (To-PRO-3). To-PRO-3 detects double-stranded nucleic acids and hence can be a useful tool for nuclear DNA quantification and nuclei volume extraction from the dataset. This stain is faint and is sensitive to photobleaching. The signal intensity is extremely weak such that the raw image fails to precise nuclei segment with any available tools like Cellpose (Stringer et al. 2020) (Figure 3.4-A). Major issues include erroneously fused nuclei, this is mainly because of these nuclei being close in 3D and the faint signal on their background connecting them fused. Nuclei stain is absent in the nucleolus resulting in an uneven nuclei surface and its segmentation looking like a hole extruded from the surface of nuclei (Figure 3.4-B). Overall, the challenges are in fixing the size, shape and the exact number of nuclei. This study also focuses on improving the faint nuclei stain segmentation and its further application.

Segmenting this nuclei stain is extremely challenging for any available tools and available trained neural network models. Whereas a nuclei reporter pUBQ:H2B:tdTomato produces strong fluorescence that can be segmented using most of the available tools. Cellpose was used here for nuclei segmenting the tdTomato signals from the H2B reporter images. The raw H2B reporter images were processed before segmenting them in Cellpose (Figure 3.4-C). This enhanced the segmentation quality by size, shape, and number of individual nuclei in the 3D image (Figure 3.4-E). Overall, the H2B nuclei reporter segmentation could be used to quantify nuclei volume and other nuclear morphological features of interest, but for any comparison, the reporter line must be generated in separate genotypes such that the nuclei are segmentable. But it can't be used for direct nuclear DNA content estimation as the reporter represents H2B signals. All these add several complications when it comes to mutant genotypes, additionally, the nuclei reporter is prone to create any plant phenotypes depending on where it gets integrated into the genome (Troadec et al. 2019). To avoid all these, the nuclei stain segmentation can be improved with any new tools such that they are also segmentable.

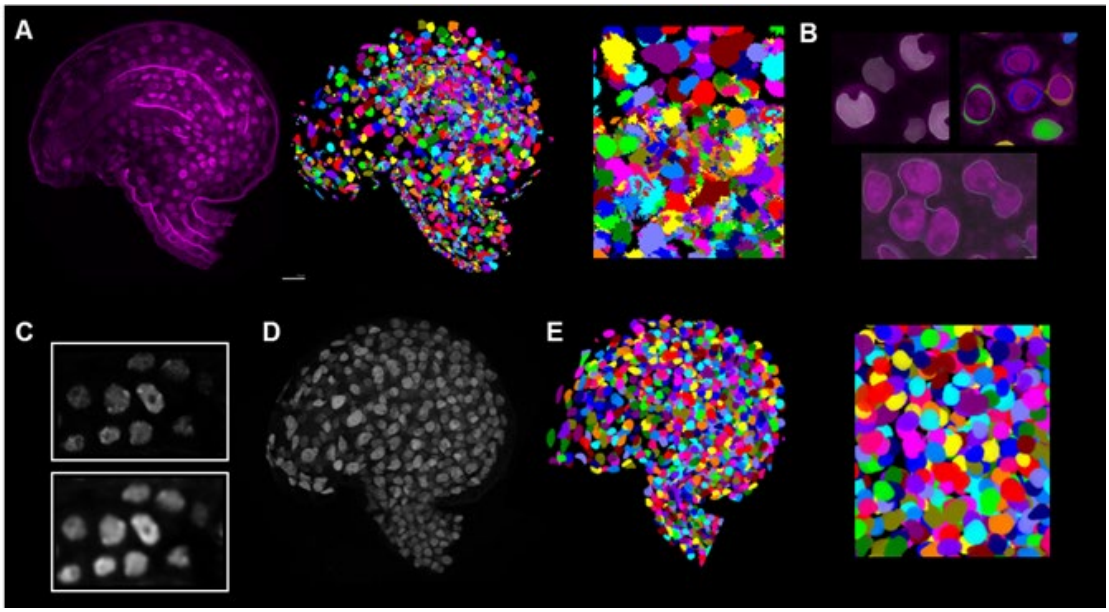

Figure 3.4 Nuclei stain and fluorescent tdTomato nuclei segmentation
2D section view of To-PRO-3 nuclei image, raw nuclei segmentation in 3D and a zoomed view displaying the erroneous segmentation (A). Typical segmentation error in the nuclei stains segmentation resulting in improper size, shape and number of nuclei (B). Fluorescent tdTomato raw image processing (C). Top row: displaying the raw image, bottom row: processed image (brightened and gaussian blurred with a small radius of 0.2 in xyz). 3D view of the processed tdTomato signal (D). 3D outer view and zoomed view of Cellpose segmentation of the bright tdTomato nuclei fluorescence (E).

3.1.3.2 Machine-learning-based nuclei instance segmentation for precise nuclei identification in CLSM dataset

Like the approach used for improving cell segmentation, machine-learning-based new methods or for further optimization of available methods requires 3D image datasets for training an ideal model which can be used for the method. The aim here is to generate a new variant of a convolutional neural network trained model (like the PlantSeg cell boundary prediction model) to be able to predict the nuclei from the faint nuclei stain images. The key requirement to the nuclei 3D U-Net training is the dataset of nuclei segmentation representing the faint nuclei stain images. Since the nuclei stain image segmentation was not precise (Figure 3.4 A), this can't be used to generate a ground truth segmentation for training. The possibility to segment pUBQ:H2B:tdTomato (Figure 3.4 C, D, E) fluorescence is taken as an advantage here.

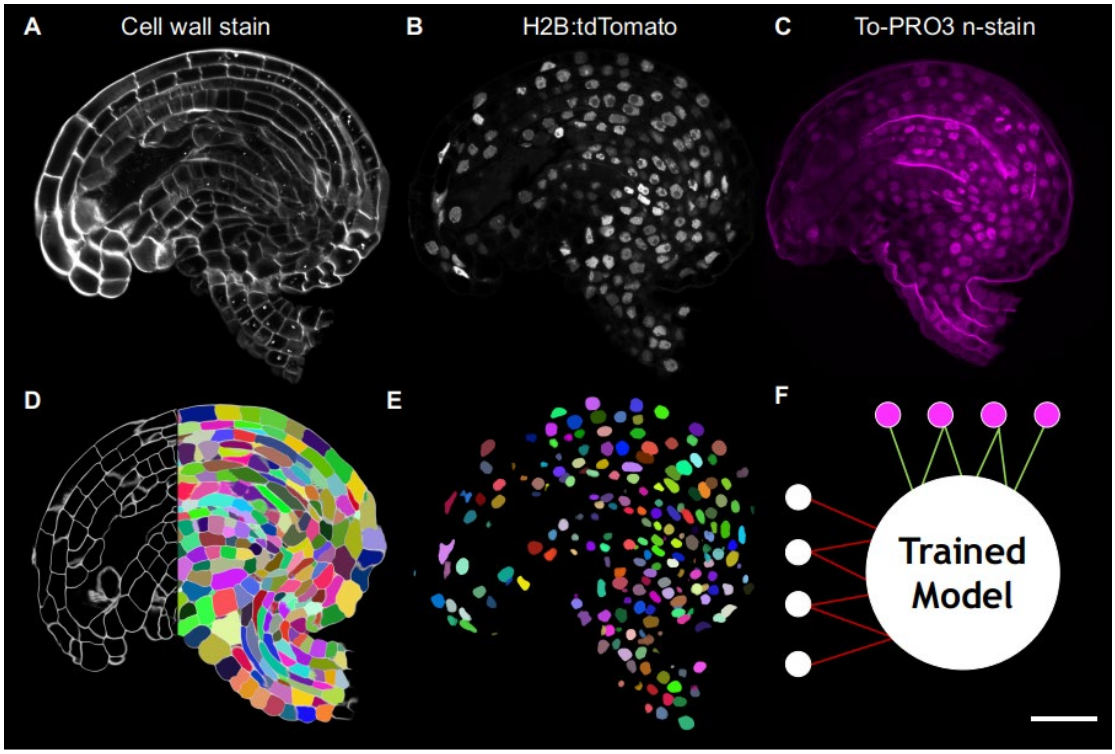

Figure 3.5 Dataset for 3D nuclei training
The training dataset includes SR2200 cell wall stain channel (A), H2B:tdTomato channel (B) and To-Pro3 nuclei stain channel (C). Cell boundary image was processed through PlantSeg (D). 2D image of the section of raw image was fractionated with the view from PlantSeg boundary prediction, the other half representing the PlantSeg GASP segmentation. Nuclei segmentation of the tdTomato channel using Cellpose (E). Scheme for the training the CNN model where the segmentation from tdTomato is used as a proxy for the nuclei stain image segmentation (F).

This study proposes a novel training dataset composed of three different image channels: SR2200 cell wall stain, H2B:tdTomato nuclei reporter and To-PRO-3 nuclei stain. SR2200 cell wall stain was processed with the PlantSeg pipeline to generate a cell boundary prediction and segmentation. tdTomato nuclei reporter signal was segmented using Cellpose as described in section 3.1.3.1 and 2.7.1 (Figure 3.5 A-E). Nuclei segmentation from the tdTomato channel could pseudo-sketch the ideal nuclei segmentation of the nuclei stain channel. Hence, the segmentation from tdTomato fluorescence was used as a proxy for the original nuclei stain segmentation. The training dataset includes four 3D image stacks encompassing 8833 nuclei. Only the raw images of To-Pro-3 nuclei stain and the segmentation from the tdTomato channel was required for the 3D U-Net model training. The available cell segmentation was used as a piece of additional information contained within the dataset that can be used for further application after having a precise nuclei segmentation. The training dataset

was set up from an otherwise not segmentable dataset of faint nuclei images making the model an exceptional one than the models available with other tools. Training was performed using the PlantSeg module “Training on New Data”. More details of the method used for training can be found in Wolny et al., 2020. A successful trained nuclei model can be now found under the general list of PlantSeg models, and it’s named “confocal_unet_bce_dice_nuclei_stain_ds1”. Again, this study has not performed the training, it has generated the reporter line and provided image stacks for training and performed quality control for the whole workflow.

The trained nuclei model was then applied on raw nuclei stain images unseen by the trained network using the same CNN prediction module of PlantSeg (Figure 3.6 A, B). The resulting nuclei probability map (n-pmap), otherwise called foreground-background probability image, has really improved the signal to noise identification (Figure 3.6-B). For instance segmentation of the n-pmap, simple hysteresis thresholding can be also performed in PlantSeg, but the results were not precise segmentation. Finally, the n-pmap images were instance segmented in 3D using Cellpose (Figure 3.6-C,E), where every nucleus is provided with an object label. The method is hence a PlantSeg-Cellpose hybrid method for nuclei segmentation as it uses PlantSeg nuclei model “confocal_unet_bce_dice_nuclei_stain_ds1” to generate n-pmap which is further exported to Cellpose for instance nuclei segmentation. The faint nuclei were greatly recovered with the newly trained model. The nucleoli issue is also solved as the segmented nuclei surface is devoid of any artefacts like an extruded hole as in the raw nuclei image segmentation. Overall, the trained model performs excellently on the raw images of nuclei stains.

For comparison of the best performing model, another CNN training was performed using StarDist (Schmidt et al. 2018; Weigert et al. 2020). The StarDist 3D model was trained on the same training dataset as with the PlantSeg nuclei U-Net model as described before. StarDist uses a neural network to predict the shape representing star-convex polyhedra for 3D nuclei. It identifies distance to the object boundaries along several radial directions (Star-Convex distances) and additionally an object probability map to determine which pixel is part of the nucleus.

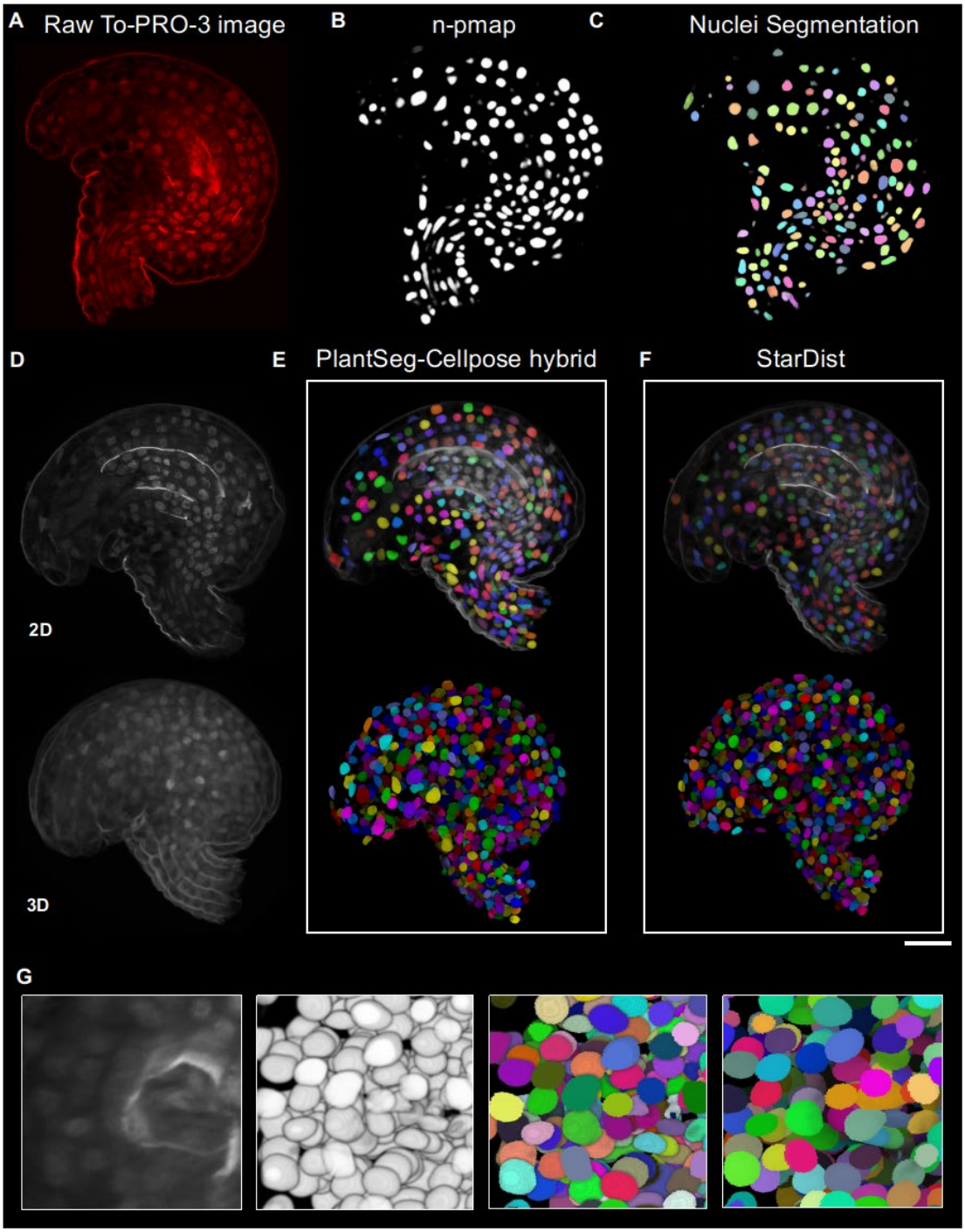

Figure 3.6 Machine-learning-based precise nuclei segmentation for fain nuclei stain images

2D section image from a raw To-Pro-3 3D nuclei stains z stack (A). Nuclei probability map (n-pmap) generated using the PlantSeg nuclei model (B). Nuclei segmentation of the n-pmap from PlantSeg using Cellpose (C, E). Comparison of raw image to instance segmentation results from the proposed two methods (D, E, F). Top row is a 2D clip plane section image from the bottom 3D image. PlantSeg-Cellpose hybrid method uses PlantSeg for n-pmap generation, further, it uses Cellpose for instance segmenting them (E). StarDist segmentation results are using the StarDist trained model generated using the dataset from this study (F). Magnified view of a clip plane

in 3D along the orthogonal axis of the images (G). From left to right: raw nuclei stain image, n-pmap image, PlantSeg-Cellpose hybrid segmentation and StarDist segmentation. Scale Bar 20µm

Final instance segmentation is performed by non-maximum suppression of these values. The StarDist trained model using the dataset from this study performed equally well compared to the results from the PlantSeg-Cellpose hybrid method where it uses the PlantSeg trained model. Both these methods now allow precise nuclei segmentation from the faint nuclei stain images (Figure 3.6 G). Detailed protocol for these two improved methods of nuclei segmentation can be found under materials and methods section 2.7. Overall, this study has significantly contributed to an exceptionally good training dataset which could be a baseline for next generation improvement in very faint signals from nuclei stains and its application on different images.

3.1.3.3 Uplifting the cell segmentation based on nuclei domain knowledge

Despite the heavy improvement in the cell boundary segmentation pipeline with PlantSeg CNN boundary predictions, the final image segmentation still includes some error in specific regions of the images where the cell wall stain is poor. An example is the faint walls around the MMC in young ovules. From the raw cell wall stain images, it's almost impossible to identify the presence of these faint walls (Figure 3.7 A). The processed raw images (brightened) along with the nuclei stain clearly displays the faint wall and the presence of multiple nuclei in this region (Figure 3.7 B, C). PlantSeg b-pmap fails to detect boundaries on these poorly stained regions of the images (Figure 3.7 D) and the PlantSeg GASP cell segmentation results in a false merge error (Figure 3.7 E). The presence of multiple nuclei in this cell confirms that this is a segmentation error and is largely due to PlantSeg not being able to detect these poorly stained walls.

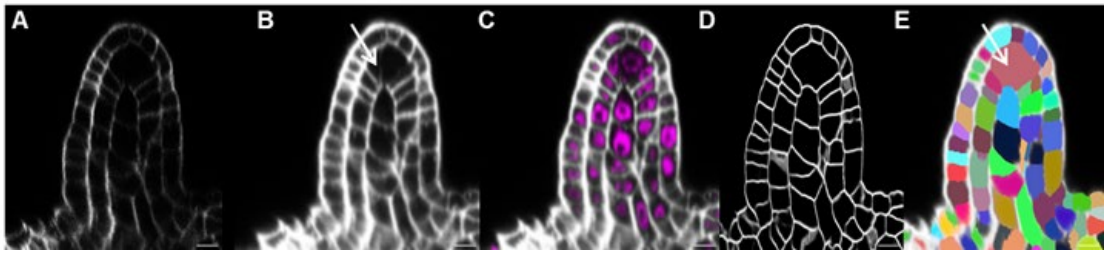

Figure 3.7 PlantSeg boundary detection fails on poorly stained walls
2D section image of raw SR2200 cell wall stain image of a young ovule from a 3D image stack (A). Processed cell wall stain image (B). Processed cell wall stain image overlaid with the nuclei stain image (C). PlantSeg boundary prediction from raw cell

wall stain image and its segmentation using GASP method (D, E). Arrows indicate the faint wall and the segmentation error. Scale Bar 5µm

This study proposed a new workflow to make use of the available nuclei segmentation to auto proofread cell instance segmentation. This is available as a script for proofreading the cell segmentation based on the presence of nuclei. Its currently available on public repository in GitHub under PlantSeg-tools (https://github.com/hci-unihd/plant-seg-tools). This script takes the cell b-pmap, cell segmentation and nuclei segmentation as input images (Figure 3.8 A-C). It automatically finds the erroneous cell segmentation by first quantifying the number of nuclei within a cell. When it finds more than one nucleus, a bounding box is approximated in 3D around this cell. Further corrections are only made within the bounding box. Corrections are made by resegmenting the erroneous cell using watershed segmentation with nuclei as seeds. This method now proofreads the segmentation error in most cases and leaves other cells without segmentation errors untouched. This method doesn't perform when the segmentation error is a missing cell instead of an under segmented cell. This study didn't involve in generating this script for correction of these segmentation errors, it has only performed quality control in the workflow and adapted it after identifying the usability in the dataset. Final instance segmentation is performed by non-maximum suppression of these values. The StarDist trained model using the dataset from this study performed equally well compared to the results from the PlantSeg-Cellpose hybrid method where it uses the PlantSeg trained model. Both these methods now allow precise nuclei segmentation from the faint nuclei stain images (Figure 3.6 G). Detailed protocol for these two improved methods of nuclei segmentation can be found under materials and methods section 2.7. Overall, this study has significantly contributed to an exceptionally good training dataset which could be a baseline for next generation improvement in very faint signals from nuclei stains and its application on different images. Final instance segmentation is performed by non-maximum suppression of these values. The StarDist trained model using the dataset from this study performed equally well compared to the results from the PlantSeg-Cellpose hybrid method where it uses the PlantSeg trained model. Both these methods now allow precise nuclei segmentation from the faint nuclei stain images (Figure 3.6 G). Detailed protocol for these two improved methods of nuclei segmentation can be found under materials and methods section 2.7. Overall, this study has significantly contributed to an exceptionally good training dataset which could be a baseline for next

generation improvement in very faint signals from nuclei stains and its application on different images.

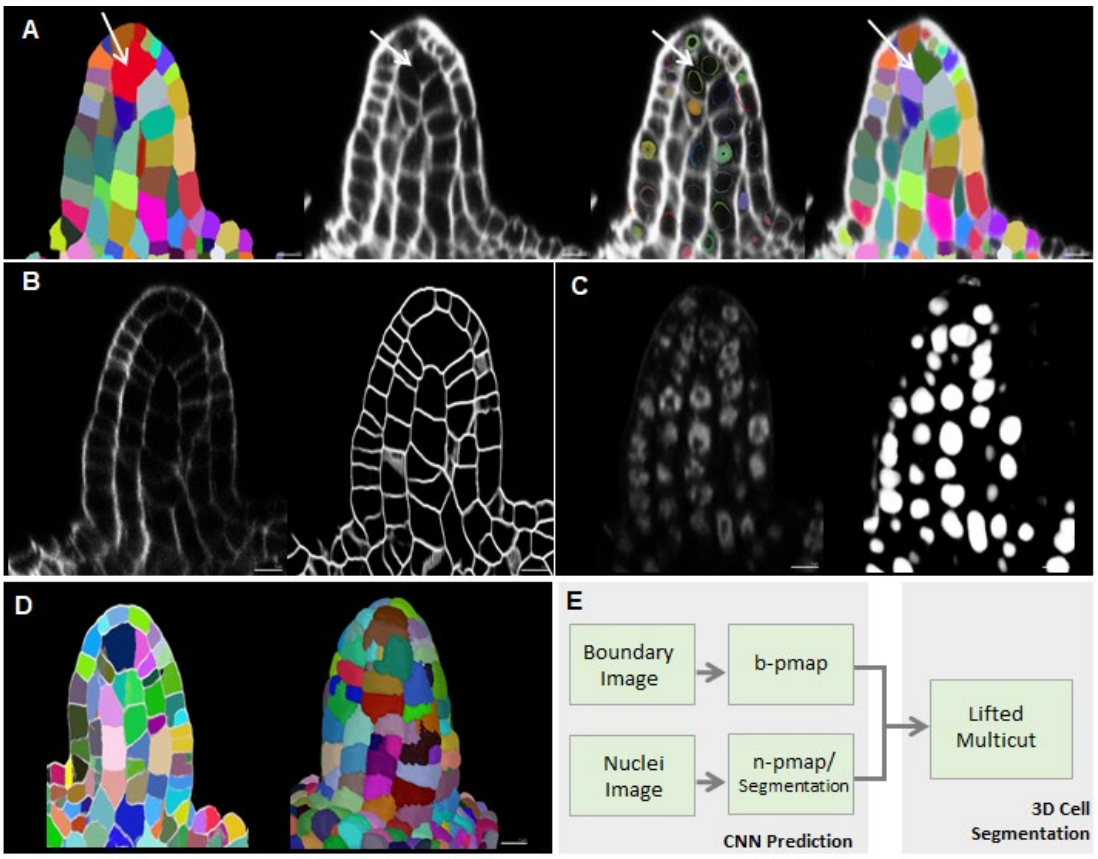

Figure 3.8 Uplifting cell segmentation based on nuclei domain knowledge
PlantSeg tools script proofread cell segmentation (A). From left to right: GASP cell segmentation from PlantSeg with error in the MMC region, brightened cell boundary image, cell boundary image outlined with segmented nuclei mesh and script corrected nuclei segmentation. Arrow indicates the region of interest. Input images for Lifted multicut cell segmentation (B, C). Left image is the raw image, the right image is the result of the PlantSeg boundary or nuclei prediction. Raw cell boundary image is used to generate cell boundary predictions (b-pmap) (B). Raw nuclei image is used to generate nuclei prediction (n-pmap) (C). Lifted multicut is performed in PlantSeg, the results of LMC (D). Left image is the 3D surface view, the right image is the sagittal section view. Figure demonstrating the pipeline of LMC cell segmentation that involves generating boundary and nuclei prediction at first step which is then used for cell segmented (E).

A second method was also proposed to perform the cell segmentation from scratch where the segmentation algorithm uses the information of nuclei segmentation or nuclei probability map to perform the cell segmentation in the first place. The method takes raw cell boundary images and nuclei images as input and processes them through the corresponding boundary and nuclei CNN model to generate the pmap of nuclei and boundary (Figure 3.8-E). Further, the segmentation is performed using the

Lifted multicut approach (LMC) (Horňáková et al., 2017; Pape et al. 2019) where the nodes in the graph corresponding to different nuclei segmentation are used as domain knowledge to introduce additional repulsive edges for cell boundary segmentation. Essentially, the nuclei pmap or nuclei segmentation serves as a basis to force each cell segmentation to only contain one nucleus. Lifted multicut can be also performed with nuclei segmentation from other sources like StarDist nuclei segmentation. This study did not implement anything new on lifted multicut pipeline, only that it applied the lifted multicut on this new available dataset for comparison of results. Results suggest that some improvements are necessary for lifted multicut segmentation using nuclei and cell predictions as the segmentation is not optimal (Figure 3.8 D). Overall, the improvement in nuclei detection and segmentation with the proposed new nuclei segmentation methods for the nuclei stain now allows a possible application in improving the cell segmentation using the nuclei domain knowledge but would require further optimization when it comes to LMC. The PlantSeg script would still perform well on this dataset. The whole procedure for nuclei segmentation or lifted multicut has wide applicability in different image datasets from other organs too (results not shown). This study has not scored the results of the segmentation from these two methods, only qualitative study is performed.

3.1.4 Whole organ tissue annotation

3D cell meshes are quantitatively informative with all geometric attributes and regular features that can be studied, but to further look inside the organ, cells must be clustered or annotated to different labels representing their tissue type. A surface visualization of internal tissues is essential to understand any pattern of cell volumes as can be seen at the surface of the organ. Tissue annotation can be performed manually, where one must select all cells from one tissue and annotate it with a common label representing the tissue. This is labor intensive as one has to manually select all cells from the mesh view. An automatic method was well described for other organs like root and shoot meristems (Montenegro-Johnson et al., 2015; Montenegro-Johnson et al., 2019; Schmidt et al., 2014). Tissue classification method in shoot meristem (3D cell atlas meristem tools) is simple as it contains a proper layered arrangement of cells into L1, L2 and L3.

3.1.4.1. A semi-automatic method for ovule primordia cell layer detection

The ovule is a composite of three clonally distinct radial layers (Jenik and Irish, 2000; Schneitz et al., 1995). Ovules follow a general principle of plant organ architecture to organize itself into L1 (epidermis), L2 (subepidermal layer beneath L1) and L3 layers (innermost layer) (Satina et al., 1940). The available automatic method can be applied on the ovule primordia, but the challenge is that it requires an organ surface mesh for performing the cell type classification. Organ surface mesh is a 2.5D curved mesh at the surface of the organ, it wouldn't have a cell segmentation. It's used for surface estimation of all cells in the organ and thereby cluster them into different tissues. It first quantifies the shortest path from individual cell centroids to the nearest vertices on the organ surface mesh, then it approximates a cone angle from the cell centroid to the surface mesh and quantifies how many cell centroids are inside the proposed cone of a cell. When a cell has no other cell centroid in its cone, it's a cell at the outer surface or L1 (Figure 3.9-A, B). An approximated cone for a cell in L2 would contain just one centroid of the cell above it. When it contains two cell centroids, it belongs to L3. This method works well when a perfect organ surface mesh is available, but occasionally, the surface of the organ could have cells that are in contact with the outgrowing organ primordia. The resulting surface meshes fail to outline the surface of the cells that are in contact (Figure 3.9-B). It's labor intensive to recreate a surface mesh that outlines the organ surface in such cases. The same issue exists for ovule primordia as most

primordia are attached to one another and a perfect organ surface mesh for layer detection requires it to be all around the 3D surface of the organ.

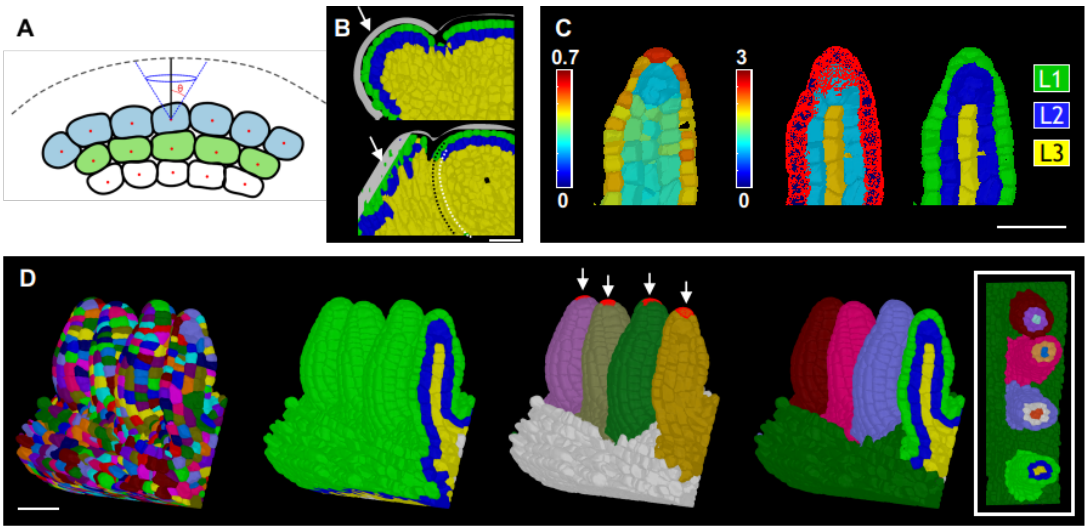

Figure 3.9 Semi-automatic method for L1, l2 and L3 layer detection
Illustration of surface distance estimation and cone angle of a layered structure using 3D cell atlas meristem tools(A) (Illustration adapted from Montenegro-Johnson et al., 2019). Red dots indicate cell centroid, dotted lines indicate organ surface mesh, angle indicate the cone angle. Results of 3D cell atlas meristem layer detection in a shoot apical meristem dataset (B). Arrows indicate the surface mesh. Green, blue and yellow colors in cells represent the L1, L2 and L3 annotations provided to the 3d cells. Top image represents an ideal case where the layer detection works well. Bottom image represents a case where the surface cells are in contact with the outgrowing organ primordia, which results in improper surface mesh creation and layer detection. Dotted black and white lines represent the outline of the continuation of L1 and L2 cells respectively which are erroneously detected as L3 cells using the 3D cell atlas meristem tools. The new generalized tool used in the study doesn't involve a surface mesh creation (C, D). Figure C from left to right: heatmap of outside wall area ratio, heatmap of cell distances from the L1 cells and the results of L1,L2 and l3 layer detection using the new proposed method. Heatmap of outside wall area ratio ranges from 0 to 0.7, the values represent the ratio of unshared wall area to shared wall area of individual cells. Heat values of cell distance represent how many cells away is an individual cell from the L1 cells. L1 cells are selected (marked red) to detect the cell distances. The usability of the new proposed method (D). From left to right: unlabelled 3D cell mesh as an input, layer detected mesh using the new method, organ detection by selecting the distal tip of the organ, unique labels of individual ovules. Arrows represent the user selected distalmost cell that is used for organ separation. Box represents a transverse section view representing the radial layers and organ identity labels. Scale Bar 20µm.

In this study, a new generalized method is proposed that clusters cells into L1, l2 and L3, but without a need for a surface mesh. The method is very simple, as a first step, L1 cells are clustered by a cell feature that is distinct for them. An obvious feature is that the cells at the outer surface of the organ don’t have a neighboring cell at the outer surface, in other words, it would have a high outside cell wall area unshared with any

neighboring cell (outside wall area ratio) (Figure 3.9-C). This is used to identify and cluster L1 cells. Once the L1 cells are clustered, L2 and L3 cells are found by their relative distances to the L1 cells. A network of cell centroids is established and the shortest number of centroids a cell must cross to reach to the nearest L1 cell is estimated with the cell distance measurement. Essentially, the results of cell distance for a cell of interest would be how many cells it is away from L1. This can be directly used to cluster these cells to L2 and L3 as L2 is one cell away from L1 and L3 is more than 1 cell away from L1.

Finally, all these features mentioned above were made into a user-friendly tool in the MGX “Detect Cell Layer Stack” that performs layer detection in just one click (Figure 3.9-D). This method uses the segmented stack for quantifying the cells with voxels on the surface or in other words, unshared outside wall area in the first place. As a second step for ovule primordia tissue labelling, individual ovules must be separately marked such that attributes of an organ can be compared to another. For organ separation, another semi-automatic method was used. The user must select a cell at the distal end of the organ and from different ovules at the same time to run another process (“Label Ovules”) that clusters all cells from the organ into one class. Further another process was used for unique label creation of individual ovules that contains the L1, L2, L3 information and the ovule identity information. The proposed method does not completely solve the issue when the surface cells are in full contact, but when there is partial contact that still leaves behind a significant outside unshared cell wall area, the process works well. A detailed protocol for the method can be found under materials and methods section 2.10. Overall, the proposed new method for layer detection is easy and fast as it does not require surface mesh generation and the results are near perfect layer detection with minimum user input. The method also has wide applicability in other organs.

3.1.4.2. Mature ovule cell type identification

3D cell mesh was partially annotated for tissue labels using cell atlas meristem tools. Automatic identification resulted in labelling around 60% of the population of cells. Further, the corrections were made using interactive mesh tools in MGX. Overall, the whole procedure was performed in a combined effort of a semi-automatic and manual method. Cell types were identified by mainly looking at the positional layer depth of the cell of interest, additional factors like 3D cell morphology and cellular connectivity with

neighbors were also used as hints to identify and cluster them. Detailed protocol for tissue labelling can be found under materials and methods section 2.10.

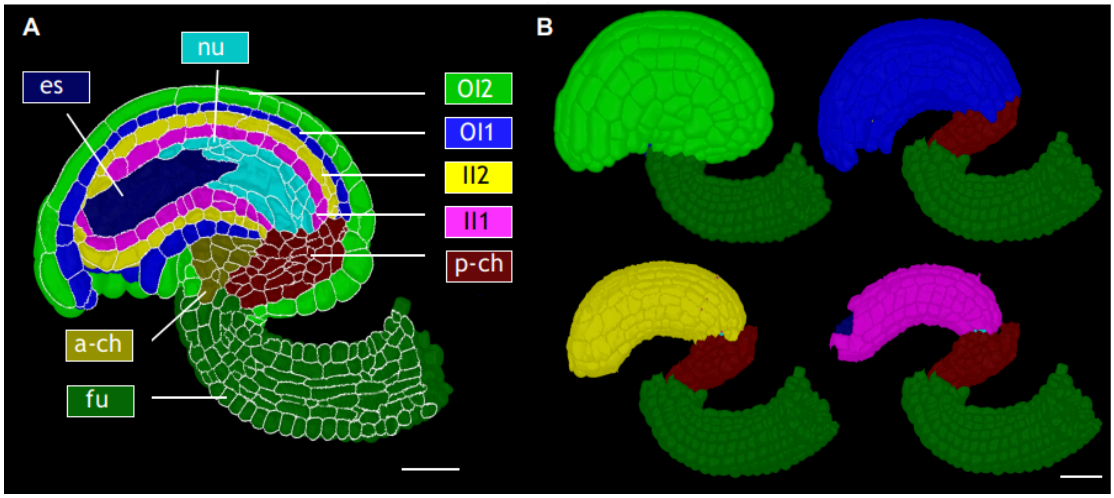

Figure 3.10 3D Cellular organization of different tissues in a mature Arabidopsis ovule
Mid sagittal section view of a mature ovule at stage 3-IV (A). 3D surface view of tissues extracted from the cell mesh using MGX mesh tools (B). Top left: outer surface of the organ, basically the outer surface of OI2 and funiculus are the visible tissues from outside. Top right: surface view after removing the OI2 layer of cells and the anterior chalaza, visualizing the surface of OI1, p-chalaza and funiculus. Bottom left: surface view after removing the OI2, OI1 layer of cells and the anterior chalaza, visualizing the surface of II2, p-chalaza and funiculus. Bottom right: surface view after removing the OI2, OI1, II2 layer of cells and the anterior chalaza, visualizing the surface of II1, p-chalaza, funiculus and partially extruded embryo sac. Different colors indicate different tissues labels annotated to the organ. Abbreviations es: embryo sac, nu:nucellus, OI2: outer layer of outer integument, OI1: inner layer of outer integument, II2:outer layer of inner integument, II1:inner layer of inner integument, fu:funiculus, a-ch:anterior chalaza, p-ch:posterior chalaza. Scale bar 20µm.

A fully automated method for mature ovule cell labeling is extremely challenging, mainly because of the complex organization of tissues in 3D. The core of the organ is composed of embryo sac (es) and nucellus (nu) (Figure 3.10-A). The embryo sac or female gametophyte contains the egg cell and it later develops into the embryo upon successful fertilization. There are two sheet-like lateral determinate organs of epidermal origin surrounding the nucellus and embryo sac, the outer and inner integument (OI and II). Each of the two integuments forms a bilayered structure, essentially forming four layers of cells surrounding the nucellus and embryo sac. Two layers of outer integument (OI1, OI2) and two layers of inner integuments (II1, II2). Eventually, the inner integument forms a third layer (ii1'). Chalaza is the central region of the ovule from where the integuments are initiated at an early stage. Chalaza can

be subdivided into anterior (a-ch) and posterior chalaza (p-ch). Funiculus (fu) is the stalk-like structure that connects the ovule to the placenta (Figure 3.10-A).

Tissue labeling now allows extensive morphological examination of the organ in 3D. MGX mesh tools allow to remove or separate tissues from the 3D cell mesh view, this opens more freedom to understand the shape of individual tissues at their surface view (Figure 3.10-B). This was never possible with any traditional or high-end microscopic approach like electron microscopy, but now with the advanced digital model of microscopic images. The outer layer of outer integument (OI2) is a hood-like structure surrounding the inner tissues. OI2 has proximal regions unshared with other integument layers. OI1, II2 and II1 form a curved cylinder-like shape from the outside surface. There is a visible bent in the central axis of inner tissues suggesting that the inner tissue follows the shape observed on the outer surface but does not extend as long as the outer surface of OI2.

3.2 Stage-specific 3D digital ovules atlas with cellular and tissue resolution

After all the refinement with the imaging method and machine-learning-based image improvement and segmentation, the wild-type 3D Arabidopsis ovule dataset was generated that is composed of images of ovules from different developmental stages which are finally processed to 3D segmented cell meshes with cellular resolution. This study mainly focused on methods development for generating these precise models from raw microscopic images on section 3.1. Section 3.2 focuses on further exploration of the dataset and quantitative analysis with tissue resolution.

A full length understanding of the whole series of development can be explored from this dataset and in 3D. This 3D digital model serves as a next generation basic 3D atlas of the organ. Moreover, several models from the same developmental stage were generated to increase the sample number for general quantitative and qualitative analysis (Figure 3.11 A). Number of 3D digital ovules scored: 10 (stages 2-III, 2-IV, 2-V, 3-I, 3-II, 3-IV, 3-VI), 11 (3-III, 3-V), 13 (stage 2-II), 23 (stage 1-I), 49 (stage 2-I), 66 (stage 1-II).

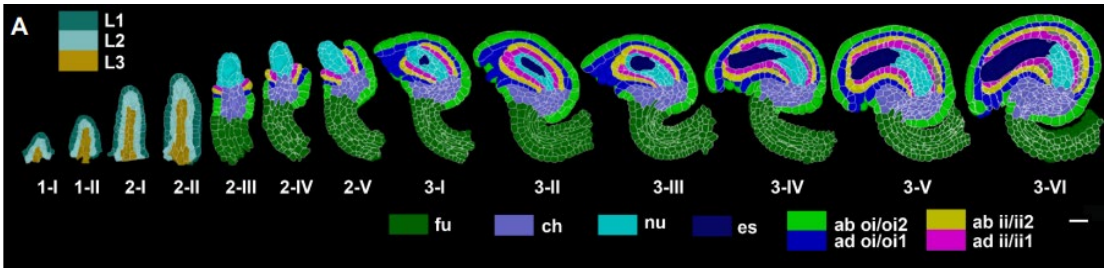

Figure 3.11 Stage specific 3D digital models of wild type Arabidopsis ovules
Mid-sagittal section of ovules from stages 1-I to 3-VI showing the cell type organization in wild-type ovules. Stages 1-I to 2-II includes radial L1, L2, L3 labeling. From stage 2-III, individual cell type labels are assigned according to the specific tissue. Abbreviations: fu:funiculus, ch: chalaza, nu:nucellus, es: embryo sac, OI2: outer layer of outer integument, OI1: inner layer of outer integument, II2:outer layer of inner integument, II1:inner layer of inner integument. Scale bar 20µm. Figure modified from Vijayan et al ., 2021 eLife.

3.2.1 3D wild-type Arabidopsis ovule dataset

The atlas covers all stages from early primordium outgrowth to the mature pre-fertilization ovule. The high-quality dataset includes 158 wild-type 3D digital ovules across stage 1-I to 3-VI. The additional dataset contains another 85 early-stage 3D

digital ovules with potentially a few segmentation errors. Overall, the improvement in nuclei segmentation and its based uplifting of cell segmentation was a later development such that the wild-type Arabidopsis ovule dataset used in this study still contains a few segmentation errors but that of an acceptable number of errors. The high-quality dataset was extremely proofread manually for segmentation errors from stages 1 to 2-II and 3-I to 3-IV. For stages 2-III to 2-V, ovules containing no more than five under-segmented were included for analysis. From stage 3-I to 3-IV, ovules devoid of segmentation error in the sporophytic tissue were used for the study.

3.2.2 3D morphological overview of Arabidopsis ovule development

Improved ovule imaging method now allows to capture every stage of Arabidopsis ovule development with cellular resolution. Ovule development in Arabidopsis thaliana is well described at the gross-morphological and cellular levels (Robinson-Beers et al. 1992; Schneitz et al. 1995; Hernandez-Lagana et al. 2021; Vijayan et al. 2021) (Figure 3.12). The earliest ovule stage 1 is when they appear to be initiated as finger-like protrusions from the placental surface. (Staging according to (Schneitz et al. 1995; Vijayan et al. 2021). As soon as they grow into a straight structure, three pattern elements can be recognized along the proximal-distal (PD) axis of the primordia: proximal funiculus, central chalaza and distal nucellus. The funiculus is a stalk-like structure, harbors the vascular strand and that connects the ovule to the placenta. Central chalaza is where the integuments are initiated from its flanks during stage 2.

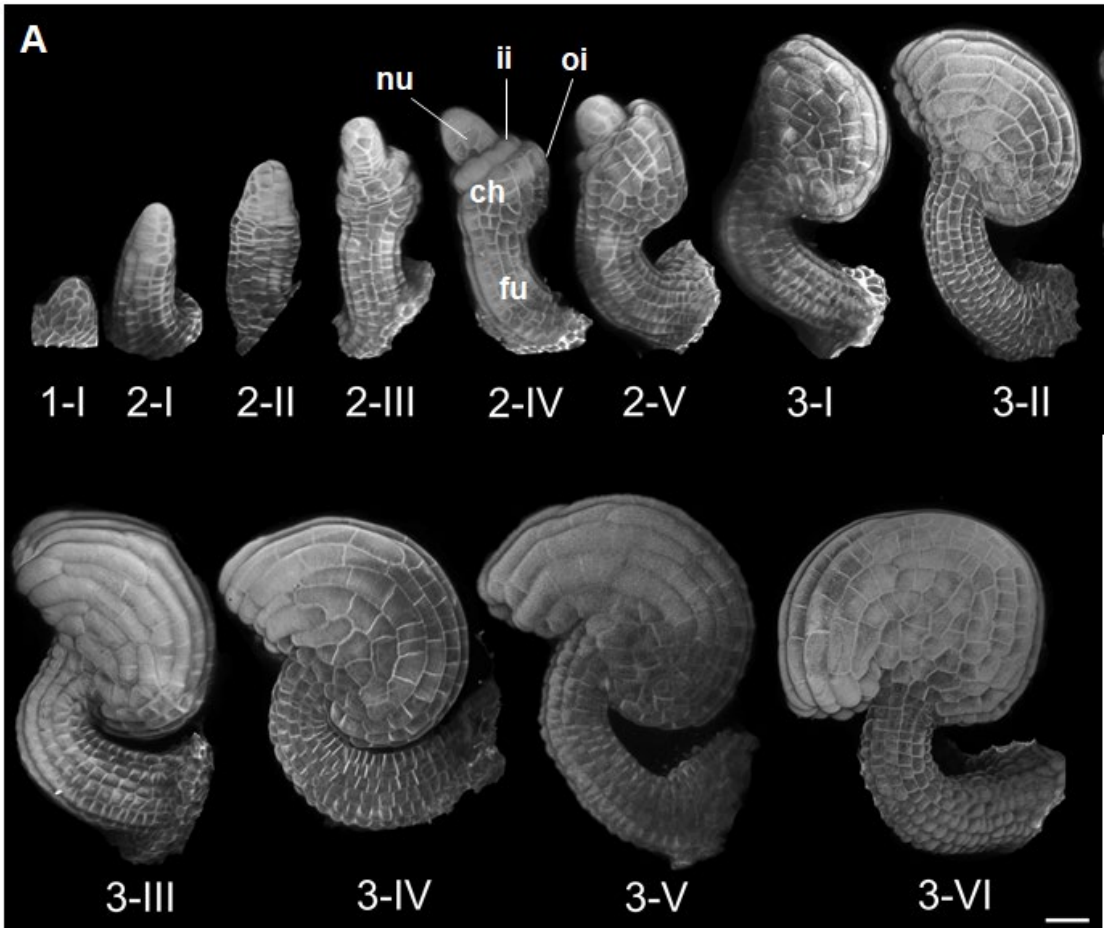

Figure 3.12 3D morphological overview of Arabidopsis ovule development
3D rendered image of the outer surface of Arabidopsis ovules from microscopic z stacks imaged with the improved ClearSee method (A). The figure represents ovules at different developmental progression starting from stage 1-I to stage 3-VI. Abbreviations fu:funiculus, nu:nucellus, ii:inner integument, oi:outer integument. Scale bar 20µm. Figure modified from Vijayan et al., 2021.

The outer integument represents a bilayered structure while the inner integument eventually consists of three cell layers. Nucellus generates a large megaspore mother cell (MMC) in the L2 that will undergo meiosis during stage 2. During stage 3 one of the meiotic products eventually develops into the haploid embryo sac or female gametophyte carrying the egg cell proper. The chalaza is characterized by two epidermally-derived integuments, determinate laminar tissues that initiate from its flanks during stage 2. The two integuments grow around the nucellus at later stages of development, and they leave open a small cleft, the micropyle, through which the pollen tube can reach the embryo sac. Finally, the curved shape of the mature Arabidopsis ovule represents a prominent aspect of its morphology. It is caused in part by the integuments bending around the nucellus during stage 3 until their tips

eventually locate next to the funiculus (anatropy). In addition, differential growth causes the funiculus to bend as well. Overall, the mature ovule exhibits a characteristic doubly curved structure (Figure 3.12).

3.2.3 Polarity and overall assessment of ovule development

Arabidopsis ovule developmental stages were defined using the morphological features from Schneitz et al., 1995. 3D ovules were looked at at the organ surface with a heatmap for cell volumes. Overall, the organ undergoes a high amount of growth as the early ovule primordia are particularly tiny compared to a mature ovule at stage 3-VI before fertilization. The fundamental question here is if the growth is mainly by cell divisions or enlargements or combined action of both.

The heatmap of cell volumes (Figure 3.13-A) illustrates that cell enlargement is happening as the organ grows from the young to mature stages. Young stage ovules are composed of cells of comparatively smaller size than older stages. Attributes from the 3D cell mesh now allow performing a general quantitative analysis on the total count of cells in the organ and cell volumes of individual cells in the organ which can be further used to quantify the total size or volume of the organ in 3D. The early ovule primordia consist of nearly a few cells of around 40 numbers. At stage 3-VI, the numbers gradually rise to an average of about 1900 cells (1897 ± 179.9 (mean ± SD)) (Table 1). The volume of the organ increases about 100 times from 0.5 x 10^4 μm^3 (0.5 x 10^4 ± 0.09) at stage 1-I to about 50 × 10^4 µm3 (49.4 × 10^4 ± 7.2 × 10^4). This suggests that the observed growth is imparted because of the combined effort of cell divisions and enlargement of individual cells. Attributes of cells also included mitotic cell information by manual annotations from the nuclei stain channel. The number of mitotic cells was not constant throughout the development, but the overall percentage of mitotic cells was on average always less than 1.8 %.

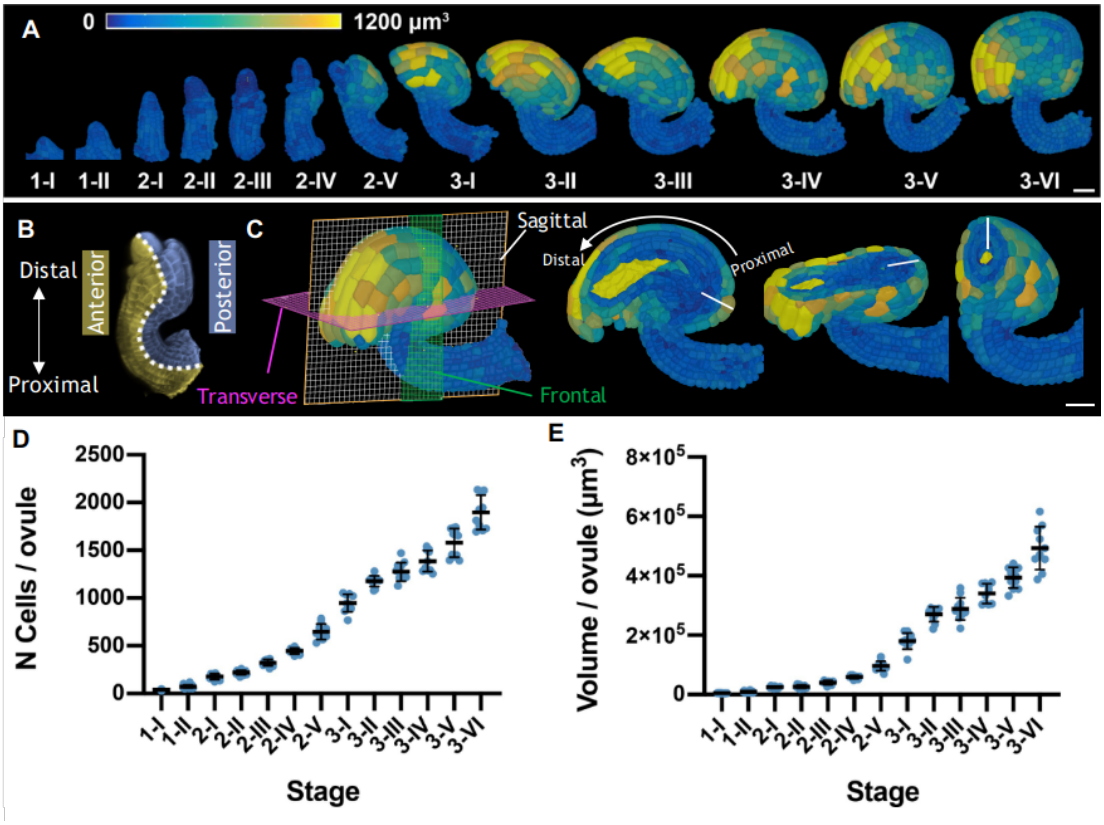

Figure 3.13 3D overall assessment of wild-type Arabidopsis ovule development
Surface view of the organ at different developmental stages displaying heatmaps of cell volumes from the 3D segmented cell mesh. Heat values ranging from 0 to 1200 µm3 (A). Ovule central axis: proximal, distal, anterior, posterior (B). Heatmap of cell volumes of a mature ovule at different views (C). From left to right: 3D outer view illustrating the different planes of the organ, mid-sagittal section view, 3D clip plane section on a transverse plane and 3D clip plane section on a frontal plane. White lines indicate the 3D radial axis from the organ center (D, E) Plots depicting the total number of cells and total volume of individual ovules from early to late stages of development, respectively. Number of 3D digital ovules scored: 10 (stages 2-III, 2-IV, 2-V, 3-I, 3-II, 3-IV, 3-VI), 11 (3-III, 3-V), 13 (stage 2-II), 23 (stage 1-I), 49 (stage 2-I), 66 (stage 1-II). Mean ± SD is shown. Scale bar: 20 µm. Figure modified from Vijayan et al., 2021.

Stage*	N Cells	Volume ($x10^4$ μm^3)	N Mitotic cells	% Mitotic cells
1-I	39.6 ± 5.3	0.5 ± 0.09	1.0 ± 0.0	0.7 ± 1.2
1-II	74.0 ± 17.1	1.0 ± 0.2	1.3 ± 0.5	0.7 ± 0.9
2-I	176.9 ± 31.5	2.5 ± 0.4	3.1 ± 2.1	1.8 ± 1.2
2-II	220.6 ± 24.9	2.7 ± 0.6	2.7 ± 1.6	1.1 ± 0.7
2-III	324.1 ± 32.9	4.1 ± 0.7	3.6 ± 1.7	1.0 ± 0.7
2-IV	447.1 ± 30.7	5.9 ± 0.6	4.1 ± 1.7	0.9 ± 0.4

2-V	648.7 ± 81.5	9.7 ± 1.6	7.3 ± 3.0	1.1 ± 0.5
3-I	948.1 ± 92.5	18.1 ± 2.7	6.4 ± 3.0	0.7 ± 0.3
3-II	1178.0 ± 58.0	27.0 ± 2.5	10.4 ± 4.4	0.9 ± 0.4
3-III	1276.0 ± 97.7	28.9 ± 3.8	10.7 ± 2.8	0.9 ± 0.2
3-IV	1387 ± 111.9	34.0 ± 3.3	5.36 ± 1.8	0.4 ± 0.1
3-V	1580.0 ± 150.7	39.4 ± 3.5	7.9 ± 5.3	0.5± 0.3
3-VI	1897.0 ± 179.9	49.4 ± 7.2	11.1 ± 2.7	0.6 ± 0.2

Table 1. Cell numbers and total volumes of ovules at different stages
[a]Number of 3D digital ovules scored: 10 (stages 2-II- 3-II, 3-IV, 3-VI), 11 (stages 2-I, 3-III, 3-V), 13 (stage 2-II), 14 (stage 1-I), 28 (stage 1-II).
Values represent mean ± SD. Table modified from Vijayan et al., 2021.

Studying ovules in 3D now allows to properly define the different planes and axes of the organ in 3D. Ovule has a primary proximal-distal axis and an anterior and posterior axis (Figure 3.13-B, C). The organ can be also looked at in different 3D planes: sagittal plane, transverse plain and the frontal plane (Figure 3.13-C). Interestingly, cell volumes at the mature organ surface follow a spatial pattern, cells at the proximal region are composed of small cells whereas the distal region is composed of enlarged cells (Figure 3.13-C). This suggests that there might be proximal distal polarity in the cell volume pattern on the organ surface at maturity. 3D clip plane allows further understanding of the cell volumes of internal tissues at the section (Figure 3.13-C). Heatmap from the section views illustrates that the cell volumes of internal tissues are not as enlarged as the outer layer. Essentially, the polarity is observed on the outer layer of outer integument cells which are located at the organ surface. The funiculus doesn't seem to be part of this polarity in cell volumes. Although the section analysis gives a good hint of what is occurring inside the organ, the full picture of the pattern is not clear. To further understand the pattern in fine details, the 3D cell mesh has to be annotated for tissue labels. A tissue label would allow further extraction of cell volumes from individual tissues with cellular resolution.

3.2.4 Continuous growth of early ovule development

Ovules arise as finger-like protrusions from the placental surface. The ovule primordia were staged as early stages 1-I, 1-II, 2-I and 2-II (Schneitz et al., 1995) (Figure 3.14-A). Ovules are initiated by periclinal divisions on the sub epidermal tissue of the

placenta (Hill and Lored, 1994). Stage 1-I is marked by a small bulge on the surface of the placenta. Stage 1-II is when the bulge forms a protrusion-like elongated finger-like structure. Stage 2-I is marked by the presence of MMC in the L2. Stage 2-II is when the inner integument initiation happens. Morphologically, all these landmarks were used for staging the young ovules. But, with the advancement of 3D digital models, we could now quantify several features of the primordia like the total number of cells in the primordia and in individual cell layers, the proximal distal (P-D) extension of the primordia etc. It was unknown if the primordia have distinct growth phases or pulses where it grows from one stage to another or if it undergoes a continuous growth.

To better understand the primordial development, total volume of the primordia was quantified for all primordia dataset, also total number of cells and the P-D extension (Figure 3.14- D, E, F). The values were ranked in an increasing order to evaluate the presence of any growth pulses. Results from overall primordia volume (size), total number of cells and P-D extension indicate that there is no growth pulse between different stages and ovules grow in a more continuous fashion (Figure 3.14- D, E, F). A steady and continuous rise in the values indicate that there are no major fluctuations happening to the primordia growth.

The primordia grow in a continuous manner making it challenging to define proper quantitative landmarks for separation of different stages. However, one can rank ovules based on the three features used here (size, total number of cells, P-D extension) and identify or compare ovules as young or old to each other. To further demonstrate the stage classification, a gene expression pattern could be used. Vijayan et al., 2021 demonstrates the WUS expression in young ovules. WUS expression was first observed in ovule primordia with a total number of 50 cells. The proposed criterion is to cluster all ovules with less than 50 total number of cells as stage 1-I. Stage 2-I is defined by the emergence of the large L2-derived MMC at the distal tip of the primordium (Schneitz et al., 1995), it was easy to identify MMC based on their position and size. After detailed visual and quantitative examination of the L2 cell volumes the proposed definition of a stage 2-I is when the MMC has a minimal cell volume of 335 μm^3. At stage 2-I, an average MMC volume was quantified 543 μm^3 (543.3 ± 120.6). Later this was used to define a stage 2-I ovule. A stage 1-II ovule never has a distinct enlarged distal cell, a maximum volume of L2 cells at stage 1-II was observed 297.6 μm^3 and visual inspection revealed that this cell wasn't at the distal end of L2. That makes it easy to cluster stage 1-II ovules as they have more than 50 cells in total and

no distal enlarged L2 cell with a volume more than 335 μm^3. Taking these considerations into account, ovule primordia grow to a volume of about 1.8 × 104 μm^3, or a cell number range of approximately 125–135 total cells, and to a height range of about 41– 43 µm when they enter stage 2-I. For each of the three parameters, a small number of ovules fell into the range of overlap: five for the PD extension and seven each for total volume of primordium and total cell number per primordium. These numbers account for 3.6% and 5.1% of the total of 138 scored ovules, respectively. The results indicate that ovule primordia reach a size threshold before they transition from stage 1-II to stage 2-I.

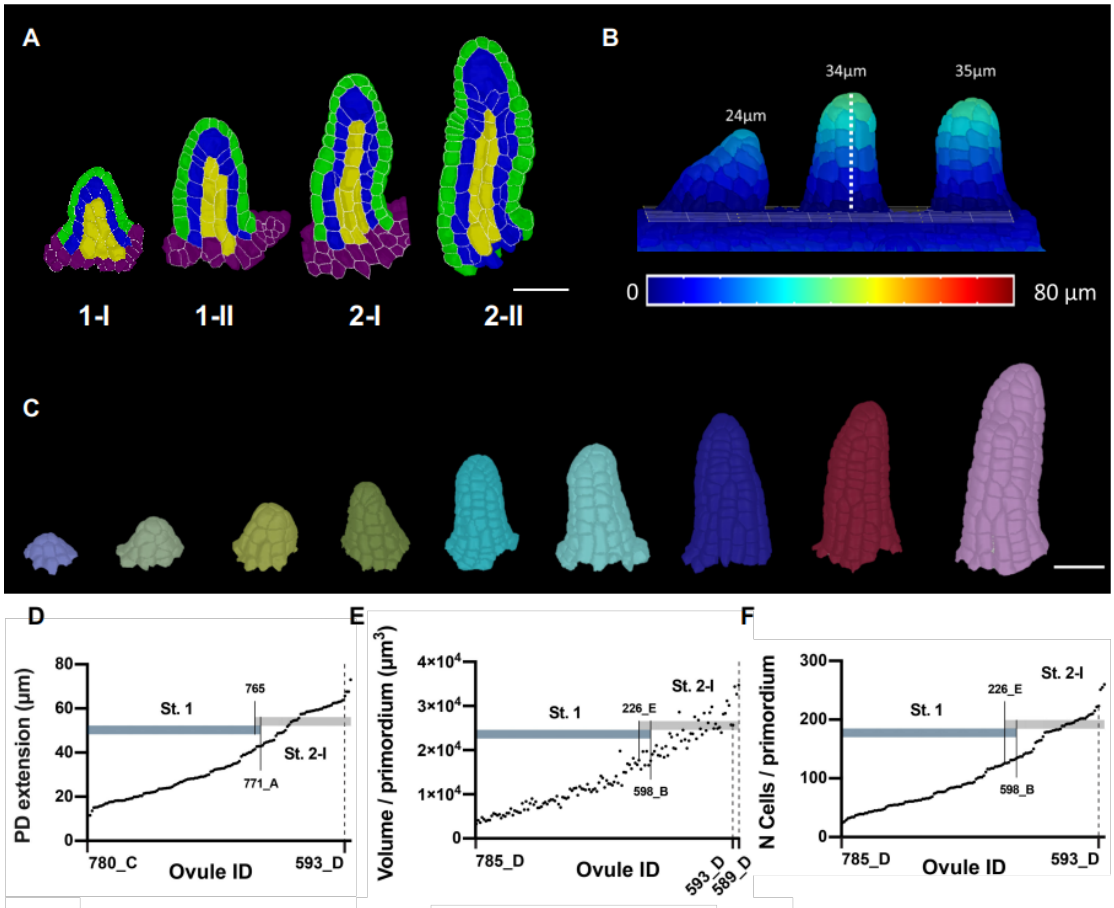

Figure 3.14 Ovule primordia grows in a continuous manner

2D mid sagittal section view of 3D digital ovules with L1, L2 L3 tissue labels in green, blue and yellow colors respectively (A). Figure Illustrating quantification of proximal-distal extension of three different ovules from the same placental surface (B). Heatmap indicates the P-D distance values quantified. 3D ovules of different P-D length extracted from the dataset to demonstrate continuous growth of primordia (C). Graph representing the continuous growth of primordia from the quantitative measures of P-D extension (D), overall size or volume of the primordia (E) and total number of cells per primordia (F). Scale bar 20µm.

3.2.5 Tissue specific analysis of ovule primordia

Overall size of an organ is influenced by two main factors, cell proliferation and cell expansion. Plot of continuous volume increase and cell increase (Figure 3.14 E, F) indicate that the organ size increases during growth, and it also involves cell proliferation as the total number of cells increases. But these graphs do not infer anything about individual cell volumes. Is there cell enlargement happening along with cell proliferation during the growth? To further investigate this, a quantitative study was performed by grouping ovules into stage 1-I, 1-II and 2-I with the criterions mentioned before.

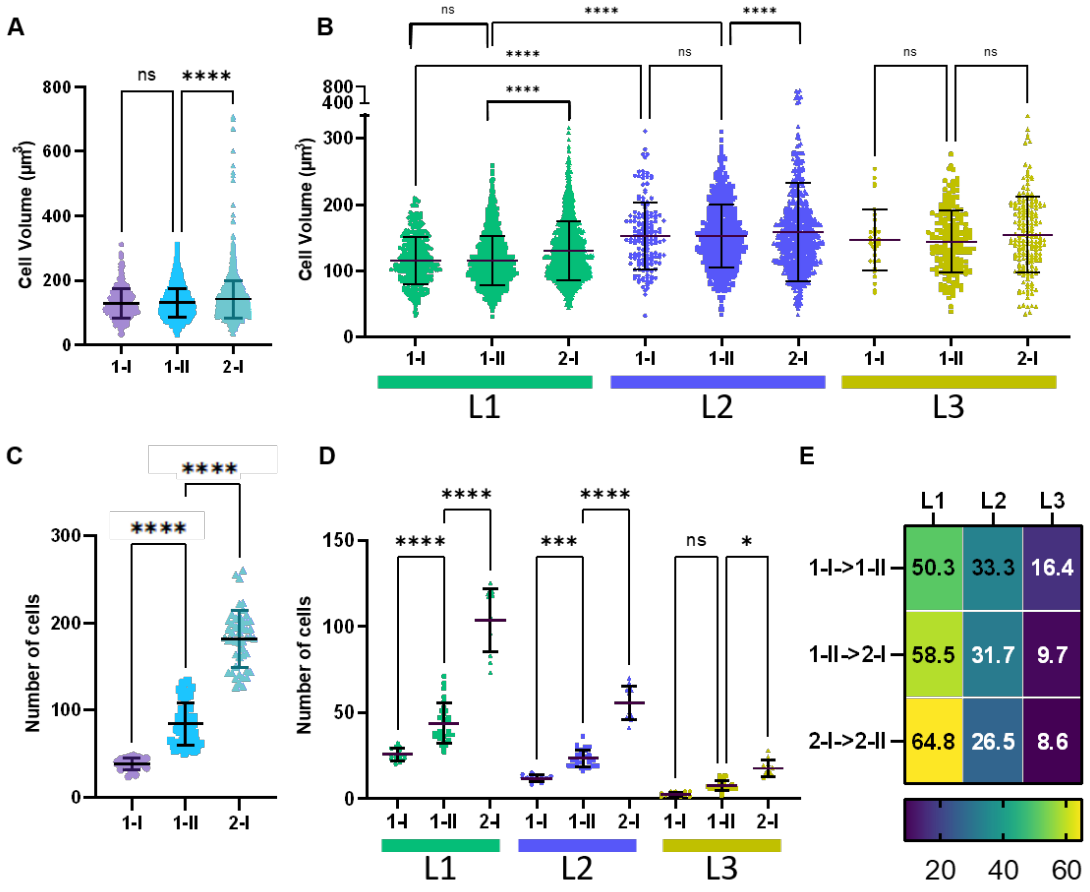

Figure 3.15 Tissue specific quantitative analysis of cell expansion and proliferation in ovule primordia
Quantification of overall mean cell volumes of stage 1-I, 1-II and 2-I (A). Quantification of cell volumes in L1, L2 and L3 layers from stage 1-I, 1-II and 2-I (B). Quantification of total number of cells per primordia at different stages (C). Quantification of number of cells in different tissue layers (D). Heatmap depicting the percentage of proliferation caused by individual tissues during growth from one stage to the next (E).

Results indicate that there is no statistically significant increase in cell volume from stage 1-I to 1-II, whereas there is a significant increase from stage 1-II to 2-I (Figure 3.15- A). Mean cell volume of primordia increases just 1.1 times from stage 1-I to 2-I (128.4 ± 44.6 to 141.3 ± 58.2 μm^3). The significant increase could be mainly because of the presence of an enlarged MMC cell in L2 during stage 2-I. To evaluate this hypothesis, similar quantitative analysis was performed by grouping the cells based on L1, L2 and L3 (Figure 3.15- B). As expected, results indicate that there is a significant increase in L2 cell volume from stage 1-II to 2-I. Interestingly, L1 cell volume was also significant between stage 1-II and 2-I. There was no significant increase in cell volumes of the L3 layer of cells. Overall, the results indicate that there is a small, but significant increase in mean cell volume between stage 1-II and 2-I and is caused because of cell volume increase in the L1 and L2 layer of cells. Additionally, cell volumes of L1, L2 and L3 at the same stage were compared to identify if there are any volumetric differences between cells of these tissues. Results indicate that L1 is composed of cells of significantly low volumes compared to L2. This is true for stages 1-I and 1-II (Figure 3.15- B).

Total volume increase of the organ or the organ size from one stage to the next is influenced by an increase in total number of cells by proliferation and or by an increase in individual cell volumes. Quantitative analysis of the total number of cells at different stages reveals that there is a significant increase in the number of cells and hence cell proliferation is happening (Figure 3.15- C). Further, analysis of the same with tissue layer resolutions indicate that L1 and L2 undergoes proliferation while the organ grows from stage 1-I to 2-I (Figure 3.15- D). L3 from stage 1-I to 1-II had no significant increase in cell numbers. Additionally, the results were analyzed as a form of heat percentages indicating which of the tissues are undergoing active divisions (Figure 3.15- E). The heat values indicate the percentage of cell proliferation caused by the tissue of interest compared to the overall cell proliferation. Essentially, the increase in number of cells was quantified between stages and with tissue resolution to calculate the percentages. As an example, a ratio of increase in number of L1 cells to the overall increase in the number of cells gives the percentage of proliferation caused by the divisions in L1 cells. Results indicate that the organ grows mainly by divisions in the L1 cell layer and partly by L2 cell layer. L3 is least involved in cell proliferation. Overall, L3 maintains the cell volume and undergoes less divisions compared to L1 and L2. Cell volume analysis also revealed that L1 cells are significantly lower in size than L2 layer cells. Overall, the results also infer that L1 is composed of actively dividing small

cells compared to L2 and L3 and hence L1 is composed of many cells compared to L2 and L3.

A summary of both cell volume analysis and proliferation indicates that most of the growth happening from stage 1-I to 2-I is by cell proliferation, while we can't exclude a small factor of growth caused by a significant increase in L1 and L2 cell volumes from stage 1-II to 2-I.

3.2.6 Synchrony and variability in ovule primordia development

Ovule development in Arabidopsis thaliana has been described to follow a stereotypical pattern (Robinson-Beers et al., 1992; Schneitz et al., 1995). However, it is unclear if ovules within a pistil develop in a synchronous fashion. Taking advantage of the 3D ovule atlas, an initial assessment of the regularity of primordia formation between ovules developing within a given pistil was visually identified (Figure 3.16-A). Spacing between primordia was observed to be not uniform (Figure 3.16-A). Quantitative analysis of the P-D extension of these primordia reveals that there is a maximum of about two times difference in these primordia from the same pistil fragment (Figure 3.16-B). Six out of the eight analyzed stage 1 primordia showed a comparable number of cells (140.5 ± 10.84, mean ± SD, ovules 1–5, 7) (Figure 3.16-C). However, two primordia exhibited a smaller number of cells with ovule #6 being composed of 91 and ovule #8 of 57 cells, respectively. Interestingly, the cell number of a primordium does not necessarily translate into its respective height or proximal-distal (PD) extension. For example, ovule #2, which is composed of 150 cells and thus of the second largest number of cells of the analyzed specimen, represents the second shortest of the eight primordia with a height of 26.5µm (Figure 3.16-C). Its comparably large number of cells relates to a wide base of the primordium. Taken together, this analysis indicates that ovule primordium formation within a pistil is relatively uniform, however, some fluctuation can be observed.

Additionally, the variability in ovules was tested by looking at ovules composed of relatively few cell differences. Eight ovules composed of a total number of cells 57 to 62 were extracted from the dataset and the total volume difference was quantified (Figure 3.16-D). These ovules are not necessarily from the same pistil fragment. Results illustrate that there is a maximum of 1.27 times difference in total volumes of these ovules (7787 ± 712.6, mean ± SD) (Figure 3.16-D). A violin plot of individual cell volumes from these eight ovules illustrates that there is no same pattern of cell

volumes (Figure 3.16-E). There is fluctuation in the pattern observed on different ovules. One can't exclude that the cells quantified here are at different cell cycle phases and the enlarged ones are the cells prior to division. Overall, the quantification indicates that there are differences in overall and individual cell volumes between ovules composed of the similar total number of cells.

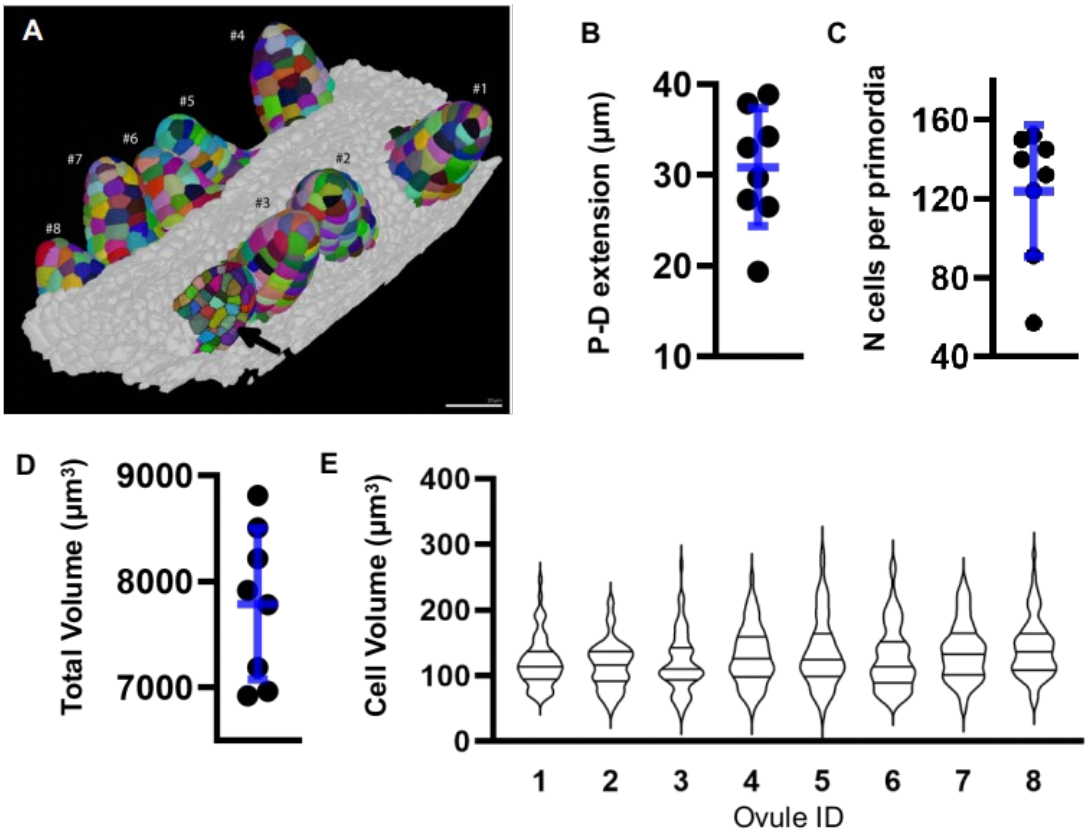

Figure 3.16 Synchrony and variability in ovule primordia development
Multiple ovule primordia from the same pistil fragment illustrating the fluctuation in synchronous development and spacing between primordia (A) (Figure modified from Wolny et al., 2020). Numbers indicate the ovule identity used to explain the differences. Plot illustrating P-D extension of the primordia (B). Plot illustrating the total number of cells per primordia (C). Plot depicting the total volumes of the ovule primordia used for variability study (D). Violin plot depicting the variability in individual cell volumes of these eight ovules (E). Scale bar 20µm.

3.2.7 First morphological manifestation of polarity in the young ovule

A mature ovule has a final orientation along the pistil axis where it orients its micropyle towards the apex of the pistil (Figure 3.17-A). This is termed gynapical-gynbasal orientation as it denotes the ovule orientation with the apical-basal axis of the gynoecium (Simon et al., 2012). The study here aims to understand when and how

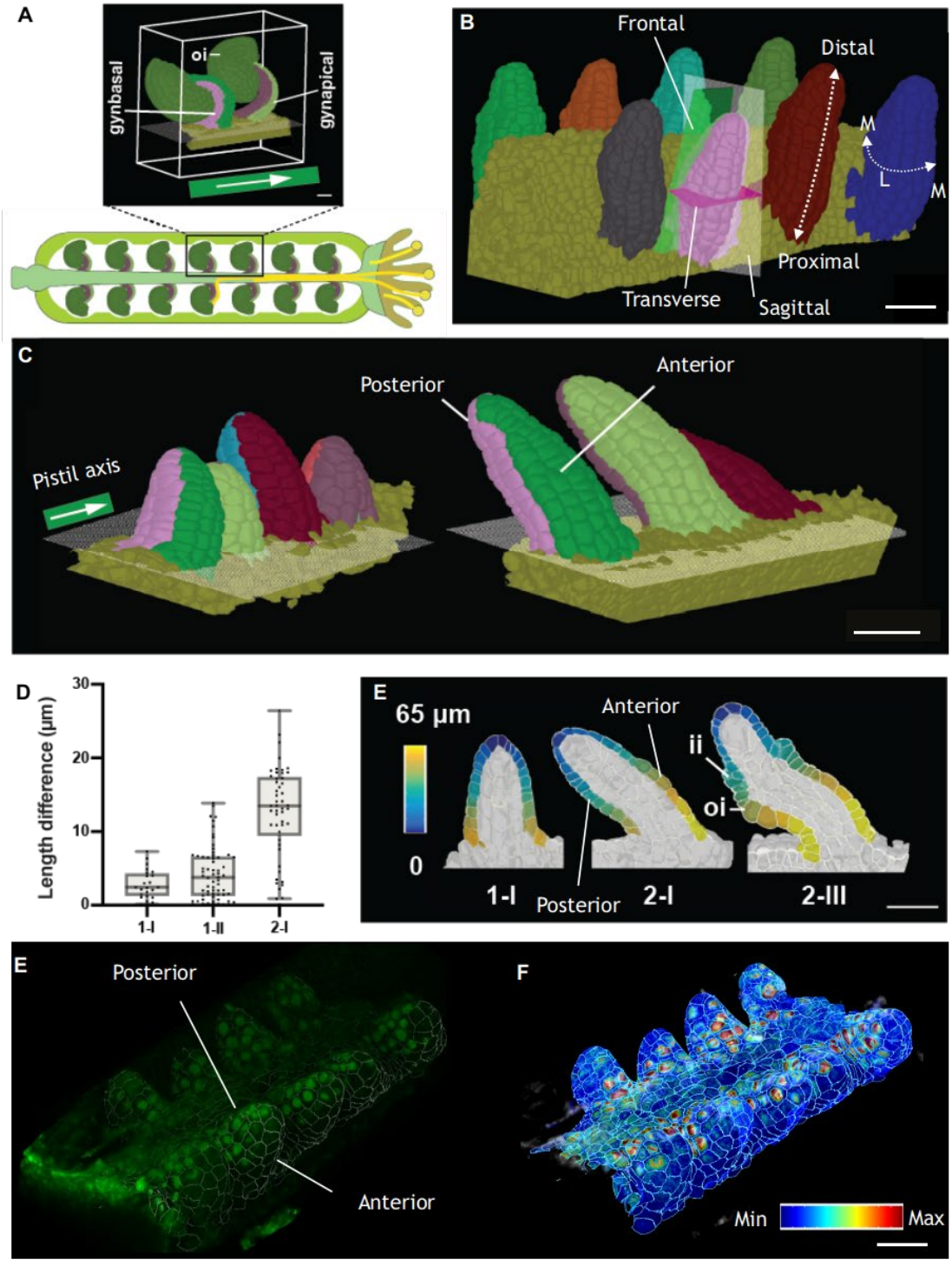

Figure 3.17 Ovule primordia slanting

Figure illustrating the final gynapical-gynbasal orientation of mature ovules (A). a 3D bounding box with ovules indicates the final orientation of the ovules inside the gynoecium with the micropyle facing the apex (stigma) of the gynoecium. The posterior and anterior sides of the ovules are oriented gynbasally and gynapically, respectively. The white arrow indicates the pistil axis. Figure illustrating the three primary planes of ovule primordia: sagittal, transverse and frontal (B). Proximal distal axis on the anterior side of the organ is marked, additionally the Medial (M) lateral (L) axis is marked on the rightmost ovule primordia by M, L, M. 3D meshes with multiple ovules from the

same carpel attached to the placenta showing unslanted ovules at stage 1-I and slanted ovules at stage 2-I (C). The 2D grid represents the surface of the placenta. Color labels depict the anterior and posterior cells, respectively. The pistil axis is indicated with an arrow. Plot depicting the extent of slanting, quantified by the difference in maximal length on the anterior and posterior sides of the ovule at stages 1-I, 1-II, and 2-I (D). Data points indicate individual ovules. Mean ± SD are represented as bars. 2D mid sagittal section views of 3D cell meshes from early ovules depicting the quantified distances on both halves of the organ (E). Stages are indicated. The heatmap on the surface cells of posterior and anterior halves depicts the quantified distance value between individual measured cells to the distal tip of primordia. Number of 3D digital ovules scored for slanting: 23 (stage 1-I), 49 (stage 2-I), 66 (stage 1-II). Expression pattern of KAN1:GFP in a young ovule primordia (F, G). Signal quantification of the KAN1:GFP from the surface segmented mesh. Heatmap is max min normalized according to signal intensity for simplicity. Number of ovule primordia scored for KAN1 signal- 9. Scale bars: 20 µm.

this orientation is set up during the development. 3D meshes of the early primordia is a good starting point for better understanding this patterning.

Ovule primordia can be looked at three primary planes: sagittal, transverse, and frontal plane (Figure 3.17-B). Frontal plane is identified by the plane on which adjacent ovule primordia can be found on the same plane. The sagittal plane is orthogonal to the frontal plane and is the plane mostly used for representation of mature ovules. Transverse plane is identified by the plane parallel to the placental surface and orthogonal to both the transverse and frontal plane. Upon careful examination of the meshes containing multiple ovules from the same pistil fragment, I found that ovule primordia at stage 1-II were positioned at a hitherto undescribed slant relative to the placenta surface (Figure 3.17-C). The frontal plane separates the slanted halves of the organ, a longer half, and a shorter half.

Even though the slant was morphologically cogent, a quantitative analysis was performed to better evaluate the extent of slant at different developmental stages. For quantitative analysis, the PD distance along the surface of the organ was extracted from the meshes using cell distance tools (Detailed method for this analysis is explained in materials and methods section 2.13). Essentially, the distances along the surface of the shorter and longer side of the slant was used for quantification. Results from quantitative analysis support qualitative morphological analysis. Slant was barely detected at stage 1-I, became more tangible during stage 1-II, and was prominent by stage 2-I (Figure 3.17-D, E).

Detailed examination of the orientation of slant illustrates that the primordia slanting is along an axis orthogonal to the apical-basal axis of the gynoecium, where the short side of the slant faces the developing septum. Since this axis showed a different spatial arrangement to the future gynapical-gynbasal orientation, I defined the short half of the primordium slant as the posterior and the opposite half as the anterior sides of the primordium. The three PD pattern elements are supported by gene expression patterns. For example, the homeobox genes WUSCHEL (WUS) and BELL1 (BEL1), are specifically expressed in the nucellus and chalaza, respectively (Reiser et al. 1995; Gross-Hardt et al. 2002; Vijayan et al. 2021). Slant formation coincides with anterior expression of PHABULOSA (PHB) during stage 1 (Sieber et al., 2004). In emerging lateral organs, members of the class III HD-ZIP and KANADI gene family are typically expressed in neighboring domains (Emery et al. 2003; Kerstetter et al. 2001; McConnell et al. 2001; Yadav et al. 2013). Spatial signal distribution of pKAN1::KAN1:2xGFP, a well-characterized reporter for KANADI1 (KAN1) expression (Caggiano et al. 2017) was found to be always at the posterior half of the ovule primordia (Figure 3.17-E, F). The results indicate that expression of pKAN1::KAN1:2xGFP serves as a marker for the posterior domain.

Slanting represents the first morphological manifestation of polarity in the young ovule primordium that separated the organ along the frontal plane. Overall, slanting has turned the organ from a more radial symmetric initial structure to a bilateral symmetric structure. Morphological examination of slanted ovules indicate that they don't align with the final orientation of the organ along the gynapical-gynbasal axis, instead, it was found that slant is along an axis perpendicular to the final orientation of the organ.

3.2.8 Post slanting reorientation of anterior-posterior axis

Since, the slant orients the organ orthogonal to the final gynapical-gynbasal orientation along the pistil axis, the hypothesis here is that the organ might undergo a turn of about 90 degrees from its slanted orientation such that the slant translates to the final gynapical-gynbasal orientation. Otherwise, the final orientation would be established with a different mechanism irrespective of the slant.

To further understand how the organ orient along the final gynapical-gynbasal axis, stages followed by slanting was carefully analyzed for any changes in the orientation of the central axis of the organ and any changes due to the integument initiation. Outer integument was found to be always initiated at the posterior side of the organ.

Following the outer integument initiation, the organ tends to have a slightly turned orientation towards the pistil apex (Figure 3.18-A). The alignment of sagittal plane is a good estimate of the orientation of the organ. Overall, it seems to be a series of events starting from early-stage ovules where they first form the slant that defines the anterior and posterior side of the organ. Further, outer integument initiation occurs at the posterior side and at the same time or afterwards, the ovule undergoes a slight turn that finally orients the organ along the gynapical-gynbasal axis.

To further clarify and evaluate this statement, a qualitative study was performed on 40 stage 2-III and 2-IV 3D digital ovules attached to placenta. Orientation of the sagittal plane relative to the pistil central axis was scored in every case (Figure 3.18-

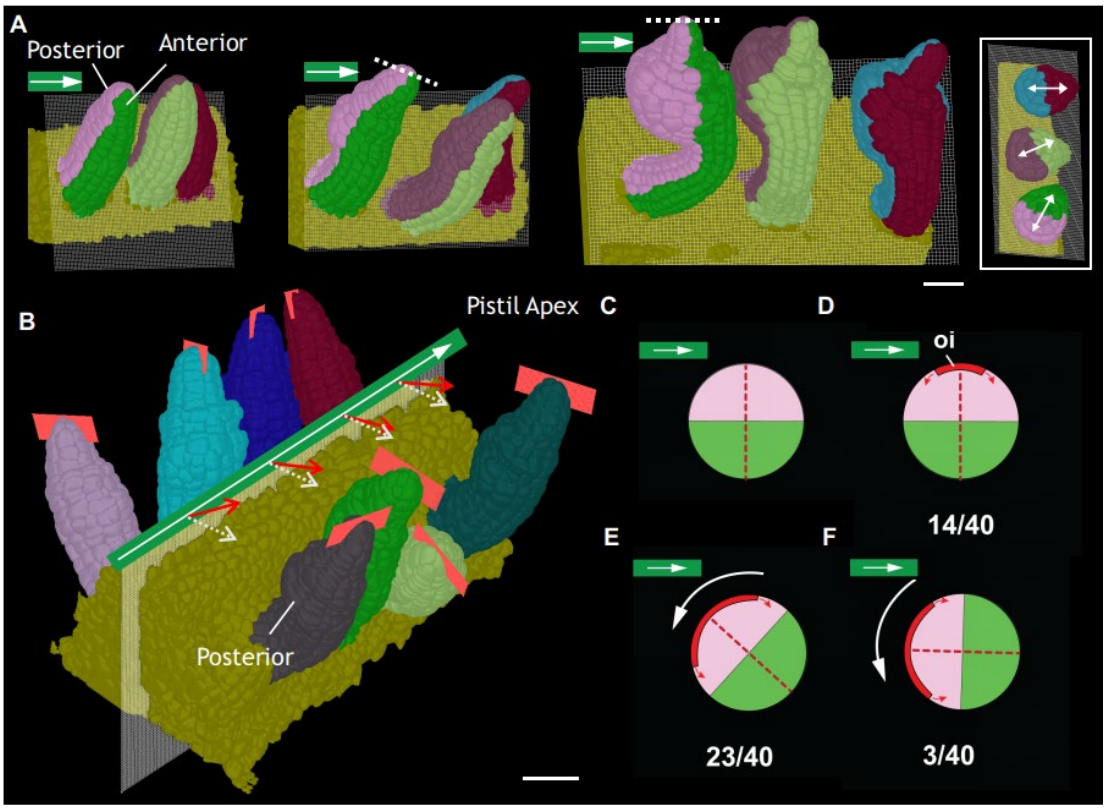

Figure 3.18 Reorienting the ovule central axis after integument initiation
Reorientation of the anterior-posterior axis relative to the apical-basal axis of the gynoecium after outer integument initiation (A). From left to right: 3D mesh of slanted ovules of early stage 2, stage 2-III ovules with a slight turn in the orientation and stage 2-IV/2-V ovules with a more pronounced turn in their orientation. Dotted lines indicate the direction of orientation of the organ with respect to the pistil axis (arrows in green box). Box highlights the top view of the rightmost image with arrows pointing the direction of orientation of the organ. Images represent multiple ovules from the same pistil fragment. Grid at the base of the organ represents the surface of the placenta. 3D mesh of a pistil fragment with multiple ovules illustrating the orientation of the ovule central axis after integument initiation (B). Rectangular red grid at the distal tip of primordia represents the orientation of the sagittal plane of the organ. Arrow on the

green bounding box points the pistil central axis to the apex. While dotted lines represent an ideal orthogonal direction from the pistil axis at which the slanting occurs at stage 1-II and later. Red arrows at an angle to the white dotted line illustrate the extent of anterior posterior orientation of the organ with respect to the perpendicular orientation of slant that occurs at a previous stage. Essentially, the degree of turn made by the organ. Cartoon depicting transverse sections of stage 2 ovules and summarizing the orientations of the anterior and posterior halves of the ovule during the turn along the PD funicular axis (C-F). The relative numbers of ovules per degree of turn are given as ratios to the overall number of organs scored. The dotted red line indicates the mid-sagittal plane. The horizontal white arrow on the green bounding box marks the pistil axis. Colors represent anterior and posterior half of the organ as marked in figure A. A red arc around the anterior half of the cartoon illustrates the extent of outer integument outgrowth. White arrows indicate the direction of turn in the orientation. Cartoon illustrating the perpendicular orientation of the sagittal plane to the pistil axis after slanting (C). Cartoon illustrating the early stage 2-III ovules where the outer integument has just initiated and that has an orientation similar to slanted ovules (D). Cartoon illustrating the advanced grown outer integument and the organ orientation with a slight angle parallel to the pistil central axis (E). Cartoon illustrating a more pronounced outer integument growth as in stage 2-IV ovules and with the orientation nearly parallel to the pistil axis (F). Number of 3D digital ovules scored for degree of turn: 8 (stage 2-IV), 32 (stage 2-III). Scale bars: 20 μm.

C-F). Figure 3.18-B illustrates an example analysis of the orientation of eight stage 2-III ovules from a pistil fragment. Results from the analysis of 40 ovules indicate that ovules exhibiting slightly more advanced outer integument undergoes an advanced turn of the central axis compared to ovules just after integument initiation. A cartoon model was framed to score the orientation of the organ and correlate it to the advancement in integument growth. Results suggest that most of the ovules at early integument initiation don't orient with a slight angle parallel to the pistil axis, rather they were observed more like the slanted ovules indicating that the turn in orientation might be triggered only after posterior initiation of outer integument at early stage 2-III (Figure 3.18-D). Most of the later stage 2-III and 2-IV ovules undergo a slight turn (Figure 3.18-E) supporting the above statement. None of the wild-type ovules were observed with a slight angle opposite to the final orientation of the organ along the gynapical-gynbasal. Among the few late stage 2-IV ovules with advanced outer integument growth, the ovule orientation was nearly parallel to the pistil axis indicating that the ovule at around late stage 2-IV already achieved the final orientation along the gynapical-gynbasal axis (Figure 3.18-F). Thus, the study proposes that the final gynapical-gynbasal orientation of the ovule is the result of a multi-step morphogenetic process involving the early establishment of an anterior-posterior axis oriented normally to the long axis of the gynoecium followed by outer integument initiation and a turn in the central axis of the organ.

3.2.9 Morphologically discernible polarity within funiculus and embryo sac

Slanting and the posterior initiation of the outer integument provide anatomically recognizable signs of anterior-posterior polarity in the ovule. Expression analysis of KANADI confirms the early establishment of anterior posterior polarity in ovules. Further exploration of the 3D dataset reveals two new findings regarding similar polarity in mature ovules (Vijayan et al. 2021). Firstly, the posterior placement of the phloem in the funiculus (Figure 3.19-A). The vascular tissue reporter line (pPD1::GFP, Bauby et al., 2007) confirms that it is located throughout on the posterior half of the funiculus. Secondly, the embryo sac nuclei were observed to follow a different positioning of its nuclei. A four nuclear embryo sac has two nuclei at the chalazal pole and two at the micropylar pole. The chalazal nuclei were aligned along the proximal-distal axis, whereas the micropylar nuclei were aligned along the anterior posterior axis of the organ (Figure 3.19-B, C). Overall, the anterior posterior polarity seems to be central in creating noise to the original radial symmetry of the organ and further establishing bilateral symmetry.

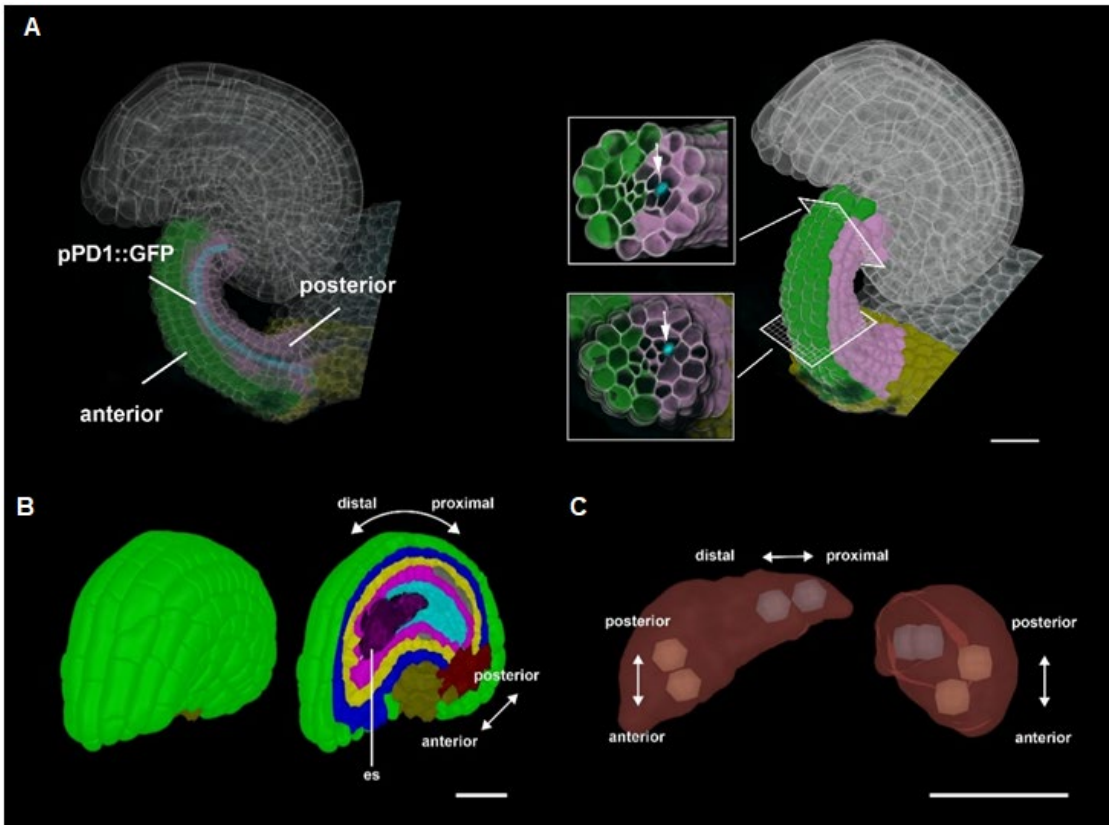

Figure 3.19 Morphologically discernible polarity within funiculus and embryo sac

Stage 3-IV wild-type ovule representing the polarity in vascular tissue (A). Expression pattern of the sieve element marker pPD1::GFP (blue signals) (Bauby et al., 2007).

The anterior-posterior polarity in the funiculus is indicated and marked by two colors. Note the posterior position of the pPD1::GFP signal. Number of 3D pPD1::GFP digital ovules scored: 8 (stage 3-IV). Stage 3-IV 3D digital wild-type ovule representing the polarity in embryo sac nuclei (B). The different polarities are indicated. Embryo sac extracted from the 3D mesh for visualization (C). Nuclei are marked by an artificial sphere on top of the image for better visualization of the faint embryo sac nuclei. The central vacuole is omitted. The two distal nuclei near the micropylar pole are usually located on top of each other (7/10 embryo sacs scored) and at a right angle to the PD axis. A slight tilt is frequently observed (5/7). The two proximal nuclei next to the chalaza are always arranged along the PD axis (10/10). Number of 3D embryo sacs scored: 10 (stage 3-IV). Abbreviation: es, embryo sac. Scale bars: 20 µm.

3.2.10 Cellular architecture of internal central region

With the advantage of the 3D cell meshes, it's now possible to extract individual cells or tissues from inside the organ and to make endless 2D sections of the planes of interest. This study aims to identify any polarity within any group of cells inside the organ. To define the internal regions, it's essential to better understand the outside structure and tissues in 3D. Mature ovule have a central proximal distal axis as described before, a medial lateral axis and a typical curved sagittal plane that separates the organ into two halves (Figure 3.20-A). A sagittal section image provides a good understanding of the overall structure in 2D but lacks the entire 3D information. A clipping plane through the transverse section of the organ gives a clear picture of the central region (Figure 3.20-B). Chalaza is the central region just above the funiculus. The proximal chalaza is surrounded by cells of the oi2 layer of the outer integument. Another clip section on the frontal plane just above the chalaza cuts through the four layers of integument and nucellus displaying the symmetric organization of integument tissues around the nucellus in 3D (Figure 3.20-B).

This study has identified chalaza or the internal central region with two groups of visually differentiated cells which was previously unreported. Previous genetic results as well as evolutionary considerations implied that the central region or chalaza can be subdivided into distal and proximal tiers flanked by the inner and outer integuments, respectively (Baker et al., 1997; Endress, 2011; Gasser and Skinner, 2019; Sieber et al., 2004). This study has further divided this into anterior and posterior chalaza based on morphological evidence. Anterior chalaza is here defined by the region mainly delineated anteriorly and medio-laterally by the oi1 at the proximal region and partly by the oi2 layers of outer integument at the distal region.

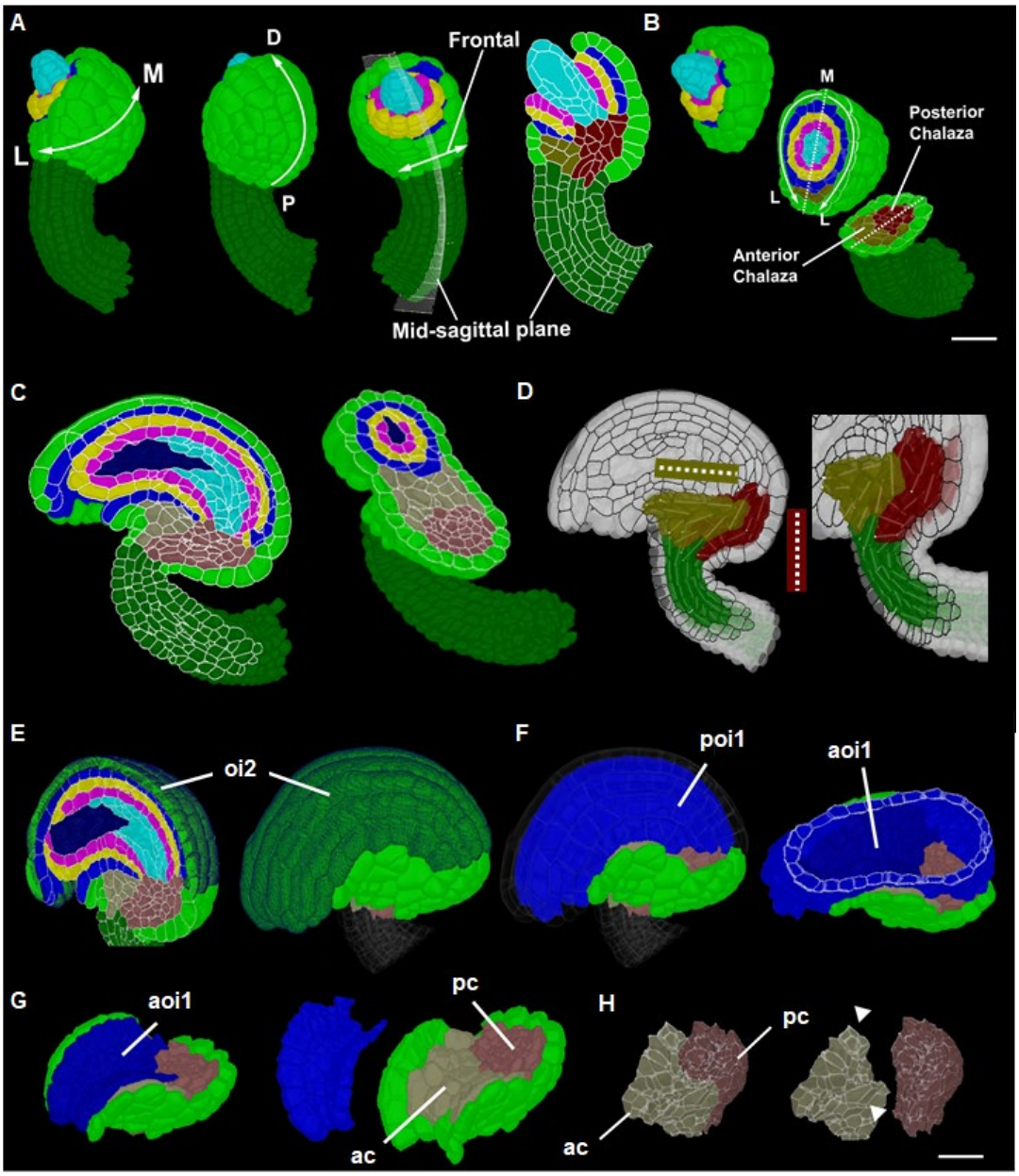

Figure 3.20 Anterior posterior polarity of central chalaza region Different views of a 3D stage 2-V ovule representing the outer polarities of the organ (A). From left to right: Medial lateral axis (M-L), proximal distal axis (P-D), curved mid sagittal section on the 3D frontal view and a 2D sagittal section image displaying the internal tissues in 2D. A 3D clip section image breaks open the 3D ovule to three pieces to display the transverse section with anterior and posterior chalaza (bottom image), frontal plane with symmetric integuments and nucellus tissue (middle image) and the distal region of the organ (top image) (B). 2D sagittal section and clip plane transverse section image of a mature stage 3-IV ovule representing the two tissue groups of the chalaza, anterior and posterior chalaza (C). 3D view of anterior and posterior chalaza with the white background representing the 2D sagittal plane with cell outlines, image on the right is a zoomed view from a posterior camera (D). White lines inside every chalazal cell represent the major axis of the cell extracted with principal component analysis in MGX. Alignment of the major axis indicates the orientation of the cell and essentially the tissue. White dotted line inside the bounding box of respective chalazal

tissue represents the visually observed general orientation of outer cells of chalaza. Progressive removal of outer tissues to visualize the central region of the ovule (E-H). 3D clip section view of a mature stage 3-IV ovule displaying the posterior oi2 cells (highlighted in dark green color) (E). Removing the posterior oi2 cells leaves the surface of oi1 exposed (F). Other inner tissues are removed to visualize the anterior and posterior oi1 layer of cells (anterior oi1 cells are highlighted with dark blue color). Anterior oi1 cells in contact with the anterior chalaza are extracted from the mesh for visualization of chalaza before and after separating the anterior oi1 cells (G). After separating the anterior oi1 cells which were in contact with anterior chalaza, the surface of anterior chalaza is first seen (right image). Anterior and posterior chalaza are extracted from the whole organ and separately visualized in contact with each other and separated (H). Arrowhead indicates the wing-like lateral extension of the anterior chalaza which were in contact with the tube-like oi1 cells. Abbreviation: L- lateral, M-Medial, P-Proximal, D-Distal, oi2-outer layer of outer integument, poi1-posterior oi1, aoi1-anterior oi1, ac-anterior chalaza, pc-posterior chalaza. Scale bars: 20 µm.

Posterior chalaza was defined as bounded medio-laterally and posteriorly by the proximal oi2 layers of cells and distally by the anterior oi1 layer of cells. The central difference in these two groups of cells is in their tissue shape. The anterior chalaza has a wing-like lateral extension probably to accommodate the tube-like oi1 layer of cells on top of this tissue (Figure 3.20-H). There is no typical posterior chalazal cell and typical anterior chalazal cells because they have differences relative to where they are in space. Nevertheless, most of the outside visible anterior chalazal cells seem to have their long axis along the anterior posterior axis of the organ, whereas most of the outside visible posterior chalazal cells seem to have their long axis parallel to the PD axis of the organ (Figure 3.20-D). This study has only approached this in a qualitative manner, a detailed quantitative analysis is relevant in this case when spatial information is also available. Essentially, the qualitative analysis indicates that not all cells follow the same pattern, the extent of pattern depends on the spatial position of cells, hence an analysis including spatial coordinates are more relevant here than just a quantitative analysis of all cells within the tissue group. Overall, the internal central region of the ovule is different from the outside integument cells in terms of not following a filar and layer wise or bundled arrangement, but rather what appears to be a random packing of cells when looking at them with naked eyes.

Table 2. Total number of cell, total volume and cell density of anterior chalaza at different stages. Number of 3D digital ovules scored: 10 (stages 2-III- 3-II, 3-IV, 3-VI), 11 (stages 3-III, 3-V). Values represent mean ± SD.

Stage	Mean number of anterior cells	Mean anterior tissue volume (μm^3)	Anterior Cell density (Cells per 1000 μm^3)
2-IV	9.20 ± 2.25	1470.00 ± 343.30	6.26
2-V	13.80 ± 5.98	3125.00 ± 1309.00	4.42
3-I	32.50 ± 8.64	8787.00 ± 2266.00	3.70
3-II	50.10 ± 8.21	14918.00 ± 1743.00	3.36
3-III	50.82 ± 12.89	16167.00 ± 3522.00	3.14
3-IV	51.50 ± 10.63	17898.00 ± 3693.00	2.88
3-V	60.73 ± 13.48	20455.00 ± 4104.00	2.97
3-VI	59.50 ± 6.10	25508.00 ± 4013.00	2.33

Table 3. Total number of cell, total volume and cell density of posterior chalaza at different stages. Number of 3D digital ovules scored: 10 (stages 2-III- 3-II, 3-IV, 3-VI), 11 (stages 3-III, 3-V). Values represent mean ± SD.

Stage	Mean number of posterior cells	Mean posterior tissue volume (μm^3)	Posterior Cell density (Cells per 1000 μm^3)
2-IV	48.40 ± 9.00	5283.00 ± 1087.00	9.16
2-V	71.20 ± 10.43	8543.00 ± 1074.00	8.33
3-I	86.40 ± 20.42	10431.00 ± 2499.00	8.28
3-II	108.00 ± 19.34	13772.00 ± 2750.00	7.84
3-III	121.40 ± 23.92	15043.00 ± 3163.00	8.07
3-IV	138.00 ± 18.29	17646.00 ± 2786.00	7.82
3-V	148.50 ± 27.47	18671.00 ± 3728.00	7.95
3-VI	209.00 ± 31.46	27675.00 ± 4636.00	7.55

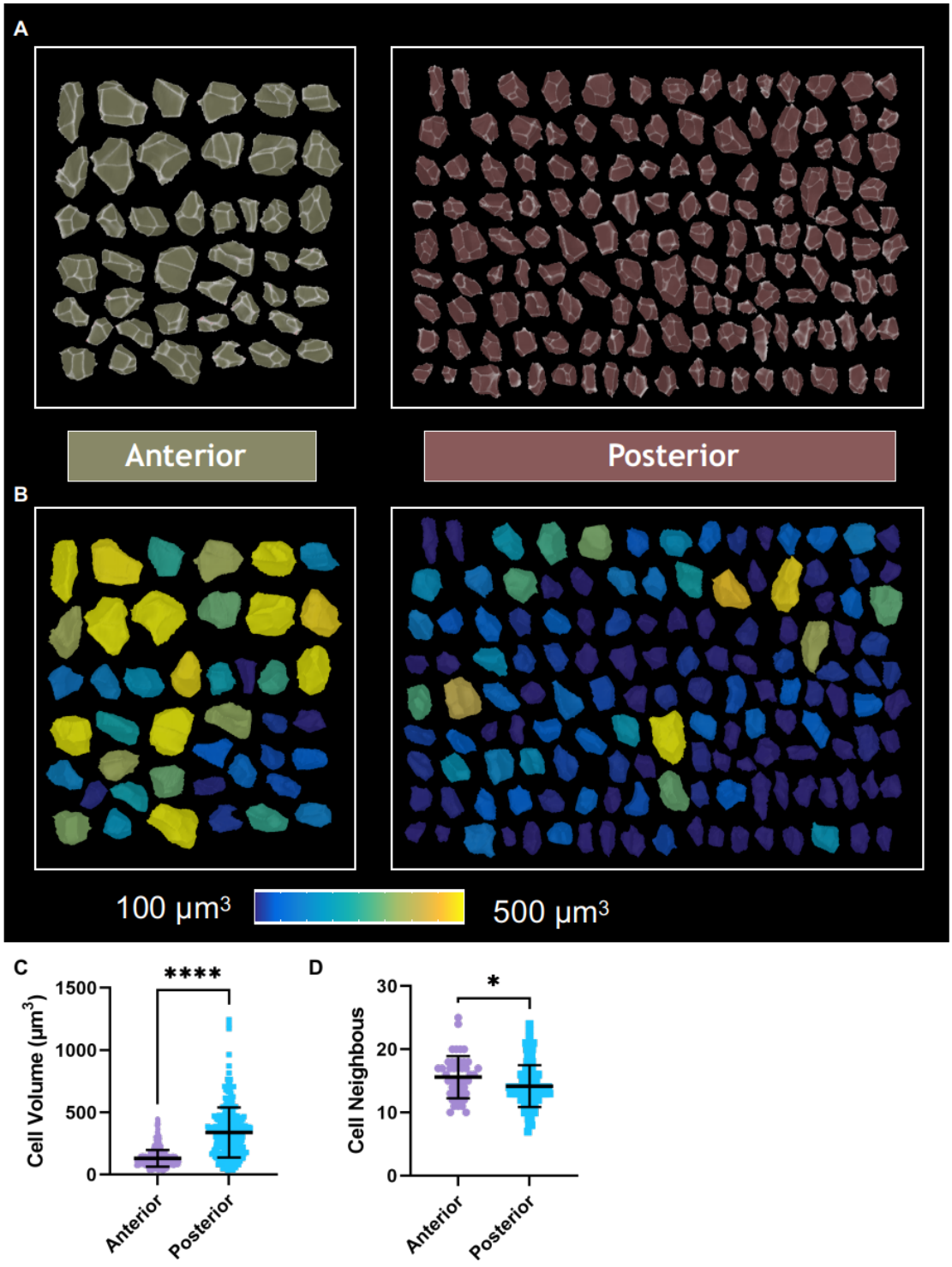

Figure 3.21 Cell shape, size, and neighbors of anterior and posterior chalaza 3D cells of anterior and posterior chalaza isolated from the tissue group (A). White border lines are the PlantSeg boundary predictions projected on the surface of cell meshes for visualization of cell faces shared with neighboring cells. Cell color represents the tissue groups. Heatmap of cell volumes on the isolated cells (B). Quantitative analysis of cell volumes and cell neighbors of stage 3-IV anterior and posterior chalaza (C, D) Number of 3D digital ovules scored for cell volume and neighbor analysis: 10 stage 3-IV ovules.

The packing of these cells is observed to follow a pattern that probably have a major motif on filling the internal space and holding the structure intact while the integument tissues along with embryo sac are outgrowing. Anterior chalaza is composed of bigger cells compared to posterior chalaza (Figure 3.21-B, C**)**. The mean cell volume of anterior chalaza is about 2.5 times higher than the posterior chalaza. For the same reason, there are differences in the packing of these cells. Posterior chalaza is a densely packed tissue compared to anterior chalaza. Table 2 summarizes the mean number of anterior cells, anterior tissue volume and their cell density calculated by estimating how many cells are present per 1000 μm^3 volume and same for the posterior chalaza. Values indicate that there is 2.7 times increase in cell density on the posterior chalaza at stage 3-IV compared to anterior chalaza. Any functional relevance of the packing of these cells has to be further explored and is beyond the scope of this study.

Cell shapes of these two groups of chalazal cells also seem to have visual differences (Figure 3.21-A). Cell size difference is one main factor that can also influence the shape of the cell. Bigger cells have large outside surface area. When it's surrounded by smaller cells, the outside surface area of the cell is shared with several cells whereas it would be only a few cell neighbors when it's surrounded by large cells. Essentially, the number of cell neighbor's influences the shape of the cell by adding different faces to the surface of the cell and such complex cell shapes could generate high level packing of cells in 3D.

3D morphological analysis of these isolated single cells from the tissue groups creates a notion that the anterior chalaza has more complex cell shapes than the posterior chalaza (Figure 3.21-A). Using MGX 3D cell meshes, a measure of number of cell neighbors or cell degree for a given cell was quantified (Figure 3.21-D). Results indicate that posterior chalazal cells have a similar number of cell neighbors compared to anterior chalazal cells. This correlates to posterior chalaza composed of small cells and surrounded by other small cells, whereas the anterior chalaza is composed of bigger cells. Every neighbor of a cell shares a surface with the cell of interest, creating individual cell faces. Essentially, the number of cell neighbor is an indirect measure of how many faces a cell has. A combination of cell volumes and cell neighbors is hence an estimate of how complex is the shape of a cell. Overall, the results support the initial notion that the posterior chalaza have complex cell shapes compared to anterior chalaza. Other quantitative measures of cell shapes like convexity, compactness, solidity, anisotropy elongation etc. were also performed on these tissue groups (results

not shown), but the complex shapes of these cells suggest that it might require any new tools to define its precise shapes.

3.2.11 Differential growth of integument tissues

Wild type Arabidopsis ovule development involves coordinated growth of different tissues. Integuments are one of the major tissues in the ovule, two layers of the outer integument (oi1 and oi2) and two layers of the inner integument (ii1 and ii2). Integuments grow around the nucellus containing embryo sac in an asymmetric fashion to form the final curved shape for the ovule. Stage 2-III of ovule development is characterized by a slightly protruded oi on the posterior side of the organ (Figure 3.22 A). Inner integuments are already initiated at stage 2-II. One major finding by looking at the 2D images is that there is differential growth of integument layers apart from the differential initiation of the two integument tissues. The outgrowing outer integument grows at a higher rate than the inner integument. Eventually, the outer integument two layers grow past the inner integument at around early-stage 3-I. This is an interesting morphological event that might have importance in the ovule curvature.

Analysis of this differential growth by looking at the tissue labelled section reveals different events. First, slow-growing but already initiated inner integument at stage 2-III undergoes a minimum outgrow before the outer integument gets initiated. The P-D extension of the inner integument is hence higher than the oi at stage 2-III. Once the outer integument starts to outgrow, there is a small phase where the oi is growing, but not able to completely extend till the distal tip of the ii as the ii started to grow already at an early stage and is also continuously growing. At the next step, the oi outgrows the ii, oi then continues to grow further, probably until a rough outer curved structure is made. Further events include the ii enclosing the nucellus completely and both the ii and nucellus growing within the structure or space made by the outer integument.

In summary, the sequence of events can be classified into three main phases. Phase one, where the oi is approaching the ii. Phase two, where the oi has outgrown ii and phase three where the ii has enclosed the nucellus completely.

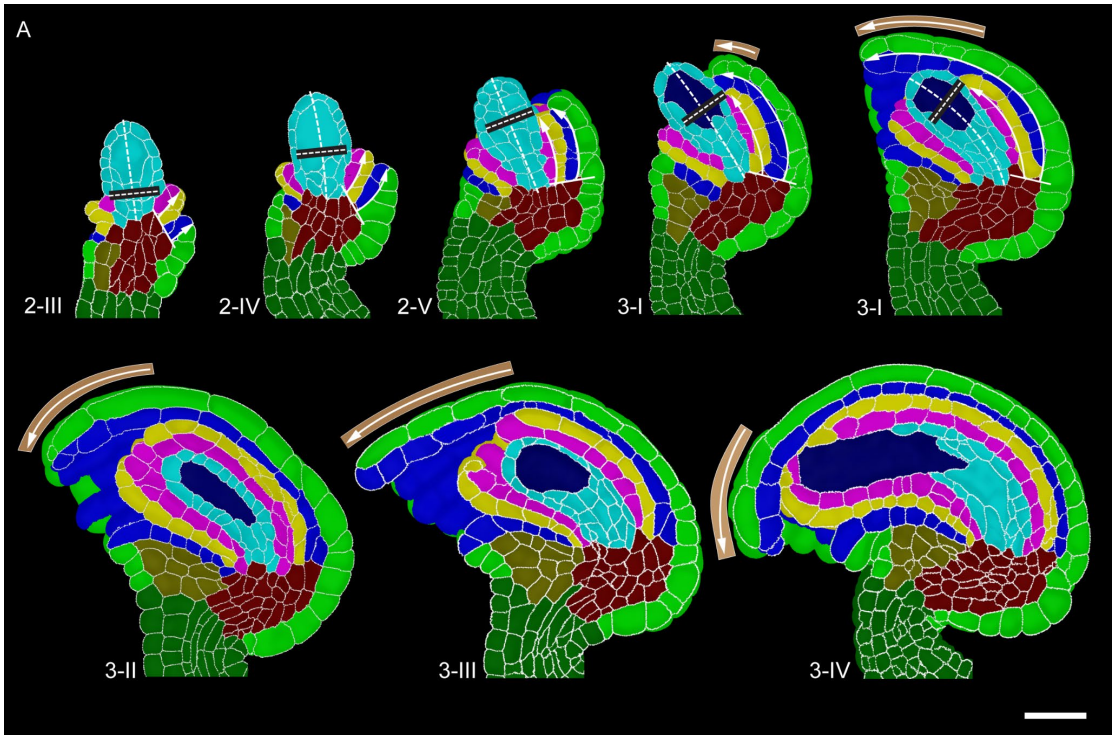

Figure 3.22 Differential growth of integument tissues.
The figure highlights the changes in the mid-sagittal section during different developmental stages of Wild type Arabidopsis ovules. Different stages are indicated. White solid line with arrows from stages 2-III to 3-I indicate the extent of outgrowth of inner and outer integuments. The white solid line at the distal end of the posterior side indicates the extent of outgrowth of the outer integument ahead of the inner integument at the indicated stages. Dotted line on black background inside the nucellus represents the extent of posterior inner integuments compared to the overall PD extension of the nucellus. The dotted lines are just to understand the different stages based on the extent of integuments or nucellus. Scale Bar 20µm.

3.2.12 Quantitative analysis of integument growth and proliferation patterns

A quantitative analysis was performed on the dataset exported from the meshes. Total tissue volume and the total number of cells per integument tissue layers were separately analyzed from stages 2-III to 3-VI. Results of tissue volume indicate that all four layers of integuments increase in volume during growth (Figure 3.23 A). Oi2 seems to have a higher increase in tissue volume compared to inner layers of tissues. Results of total number of cells per tissue indicate that there is active cell proliferation happening in all four layers of integuments (Figure 3.23 C). At stage 2-IV, the total number of cells and tissue volume of ii2 is slightly higher than the other tissues (Figure 3.23 B, D). This represents the early phase of integument outgrowth where the oi is still approaching the ii, a higher tissue volume and cell number of the ii2 indicates that ii2 is volumetrically huge than oi at this stage of development, it might be also growing at a higher rate. From stage 2-IV to 3-I, there is a comparatively higher increase in the

growth and proliferation of the oi1 layer of cells (Figure 3.23 A, B, C, D). Essentially, the oi1 growth and proliferation outcompete the ii2. This reflects the phase where the oi approaches and outgrows the ii and that the oi grows at a higher rate than the ii. From Stage 3-IV to 3-VI, results indicate that oi2 continues its volume increase as previous stages compared to oi1 (Figure 3.23 A). Interestingly, the cell proliferation pattern is different, oi1 displays a slightly higher cell proliferation compared to the oi2 layer, essentially, the outer integument inner layer has more cells than the outer layer at this phase. It might also indicate that the oi2 cells are undergoing more cell expansion at this growth phase as they increase in volume, but less in cell proliferation compared to the oi1 layer of cells.

To better understand the growth and proliferation of these tissues in a comparative manner, growth and proliferation normalized to the overall integument growth and proliferation was calculated. Detailed explanations about the growth and proliferation rate calculation can be found under the materials and methods section 2.15. Values above one indicates the preferential growth or proliferation of this tissue during the stage transition (Figure 3.23-E). Values below one indicates a lower rate with respect to overall rate of growth or proliferation of integuments. Interestingly, one can correlate these graphs to few events of the differential growth of integument tissues. Growth and proliferation heatmap of ii2 infers three main phases, firstly the initial hype in the growth and proliferation just after initiation at stage 2-III to 2-IV transition, then a sudden decrease from stage 2-IV to 2-V and an increase thereafter to equal growth with respect to overall growth of integuments. Interestingly, oi1 also has some growth and proliferation dynamics at the same growth phase. oi1 increases in growth and proliferation to maximum at stage 2-IV to 3-I transition when the ii2 decreases in growth and proliferation. Results suggest that ii have a growth phase with reduced growth and proliferation, at the same phase, oi have a higher growth and proliferation. This supports the initial observation that the outer integument outgrows and grow past the inner integument. It's still unclear if this is essential for forming the final shape of the organ.

Cell proliferation patterns indicate that oi2 always undergo less proliferation compared to overall proliferation of integument cells. Whereas oi2 is undergoing almost equal growth with respect to overall growth of integuments. Interestingly, oi2 has a higher growth and proliferation from stage 2-IV to 2-V transition. This suggests that the oi1 and oi2 have a hype at the same point of development that might be

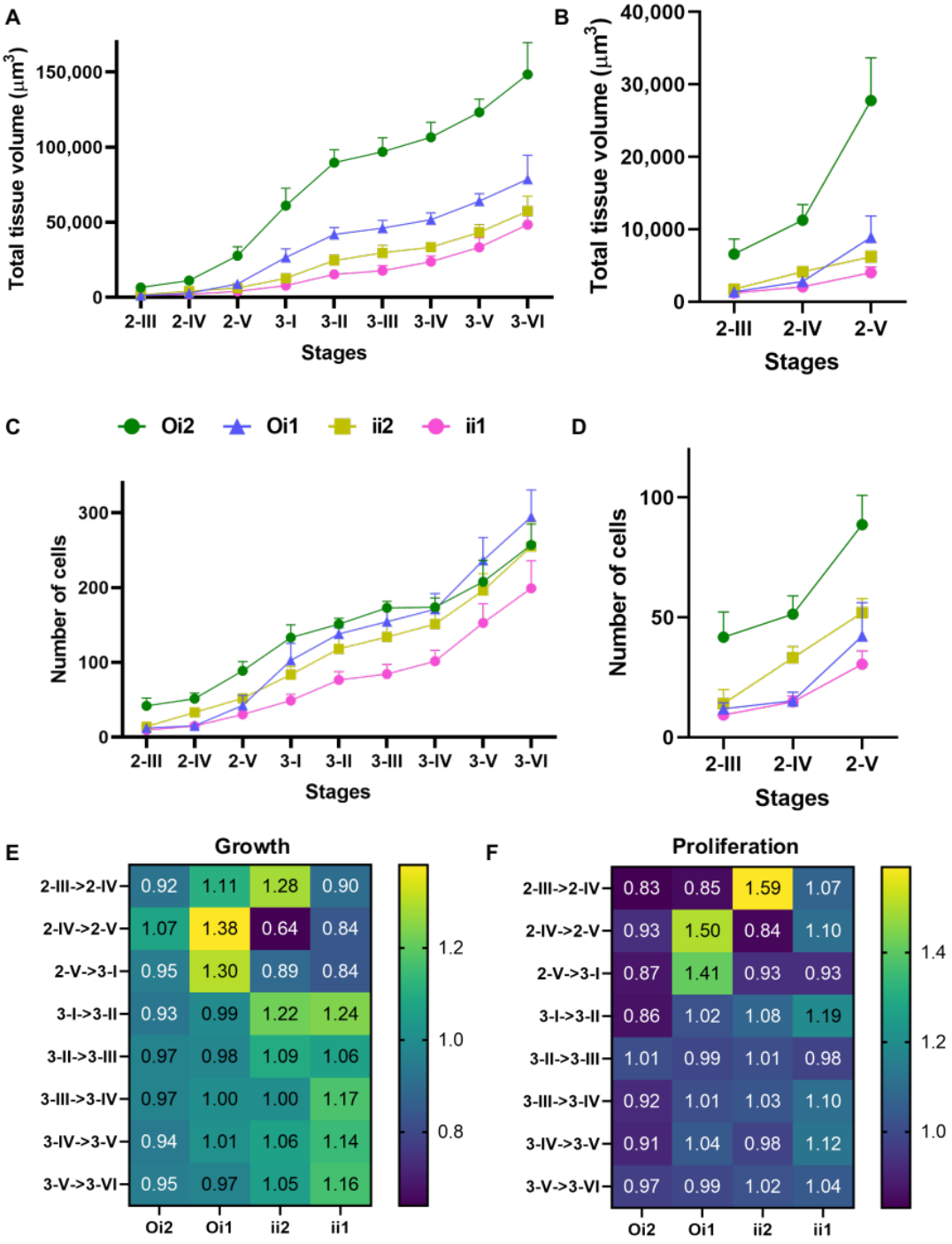

Figure 3.23 Quantitative analysis of integument growth and proliferation patterns

Graph representing the total tissue volume and total number of cells of integument tissues at different stages of development (A-D). A zoomed view of early stages from figure A and C (B, D). Colors represent different tissues indicated in the legends (A-D). Heatmap of tissue growth and proliferation patterns normalized to overall growth and proliferation of integument tissues respectively (E, D). Transition of different stages are indicated. Value = 1, represents growth or proliferation in the tissue happening equally to the overall growth or proliferation of the four layer of integuments. Value <1 represents proliferation or growth performance of the tissue low compared to overall growth or proliferation.

essential to grow past the inner integument. At a later stage, mainly at the transition between stage 3-II to 3-III one can also notice a phase where all the integument layers grow or proliferate at the same rate. During the stage transition 3-III to 3-VI, most of the cell proliferation occurs in the ii1 layer. The heat map calculations take into consideration the entire length of oi2 (includes the proximal region), hence the normalized heat values doesn't reflect the growth or proliferation contribution of oi2 to overall growth of integuments, but other three layers can be compared. The heat values (Figure 3.23 E, F) of oi2 has to be ignored. This analysis also excludes the radial growth factor. Outer layer being outside the radial structure, have to grow more to reach to same point in radial vector as to inner tissues. Unfortunately, there isn't a precise way to calculate this ratio of radial growth for integuments.

3.2.13 Friction due to contact on the surfaces

Integuments are initiated from the chalaza, and they grow around the nucellus, forming two layers of outer and two layers of inner integuments in wild type Arabidopsis. The two layers of outer integuments share a common wall between them, similarly, the two layers of inner integument share a common wall between them keeping them connected, but the inner and outer integument are separate tissues, and they don't share any common wall between them or in another way, they aren't physically joined. But due to the positioning and growth of these tissues one on top of the other, the inner surface of oi1 and the outer surface of ii2 are in direct contact. Essentially there is no shared wall between the inner and outer integument, but the surfaces are in direct contact that leads to friction on contact surfaces, otherwise called contact friction (Figure 3.24). Similar contact friction also exists between the inner integument and nucellus.

Integuments undergo differential initiation and outgrowth during development. Contact friction comes to play mainly during differential outgrowth. An imaginary smooth surface between the outer and inner integument would allow uncontrolled slide over between them when one tissue has a higher growth rate, whereas a rough surface would provide friction that limits slide over or that enables a controlled slide over depending on the friction at the surface. Overall, these results validate that there is an unshared wall between inner and outer integument and that could act as a system where the slide over could be controlled by providing controlled friction at the surfaces between the tissues. To further understand if the surface between the inner and outer integuments are smooth or rough, one requires a detailed analysis of the cell surfaces from the 3D mesh.

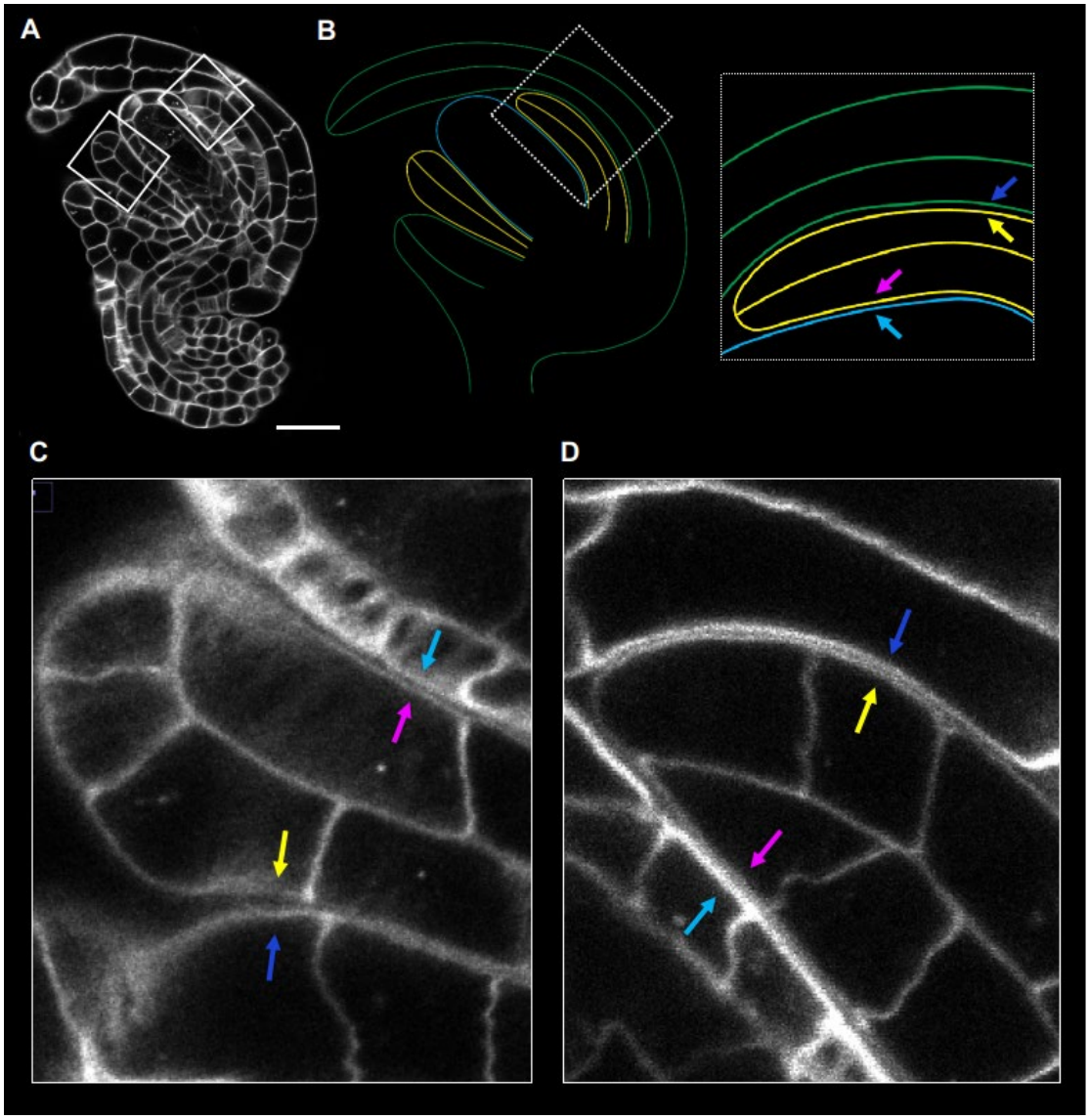

Figure 3.24 Contact friction in mature Arabidopsis ovules
A sagittal section of a stage 3-II ovule (A). A cartoon of integument and nucellus from the same section in A (B). Colors: Yellow represents inner integument, green represents outer integument and cyan represents nucellus. Right image: A zoomed view from the bounding box in the cartoon highlighting the surfaces of interest for contact friction. A zoomed view displaying the two highlighted regions in figure A (C, D). Anterior region (C) and posterior region (D). Colored arrows indicate the gaps between the indicated tissue layers indicating that they don't have a shared cell wall. Essentially, between inner integument and outer integument and between nucellus and inner integument. Scale Bar 20µm.

3.2.14 Cell Morphological changes due to contact friction

Having said about contact friction between the outer and inner integument, a detailed investigation of these tissue surfaces indicates that there are cellular morphological

changes at surfaces that are in contact between the inner and outer integument. *ino-4* mutant is a good candidate to visualize these morphological changes as they have only partially grown outer integument (Fig. 3.25). The outer integument partially covers the inner integument, the surface of the inner integument can be separated into two regions. Proximal region covered and the distal region uncovered by the outer integument. The tissue outer surface is smooth when it is uncovered by the outer integument, whereas the inner integument has a rough surface with small faces when they are in contact with the outer integument. Essentially, the contact friction introduces different faces on an otherwise smooth surface corresponding to cell faces of the inner layer of the outer integument with which the contact between the surfaces occurs.

Overall, the results indicate that the surfaces are rough upon contact between the inner and outer integument, which suggest that for the tissues to differentially outgrow, they have to surpass the contact friction at the rough surface. Alternatively, the outer layer must drag the underlying layer. The tissue does not have the freedom to free float on the surface. Additionally, the rough surface that is observed might also have another advantage in holding the outer integument tight to the inner integument compared to an otherwise smooth surface, overall making the organ tightly packed. Despite the unshared wall between the inner and outer integument, one couldn't find a loose connected cell junction or gap at junctions indicating that the contact friction enables tight packing of these two tissues which don't have a shared wall. The smooth outer surfaces of the distal ii2 uncovered by outer integument indicates that the contact friction imposes cell morphological changes of the tissues after contact.

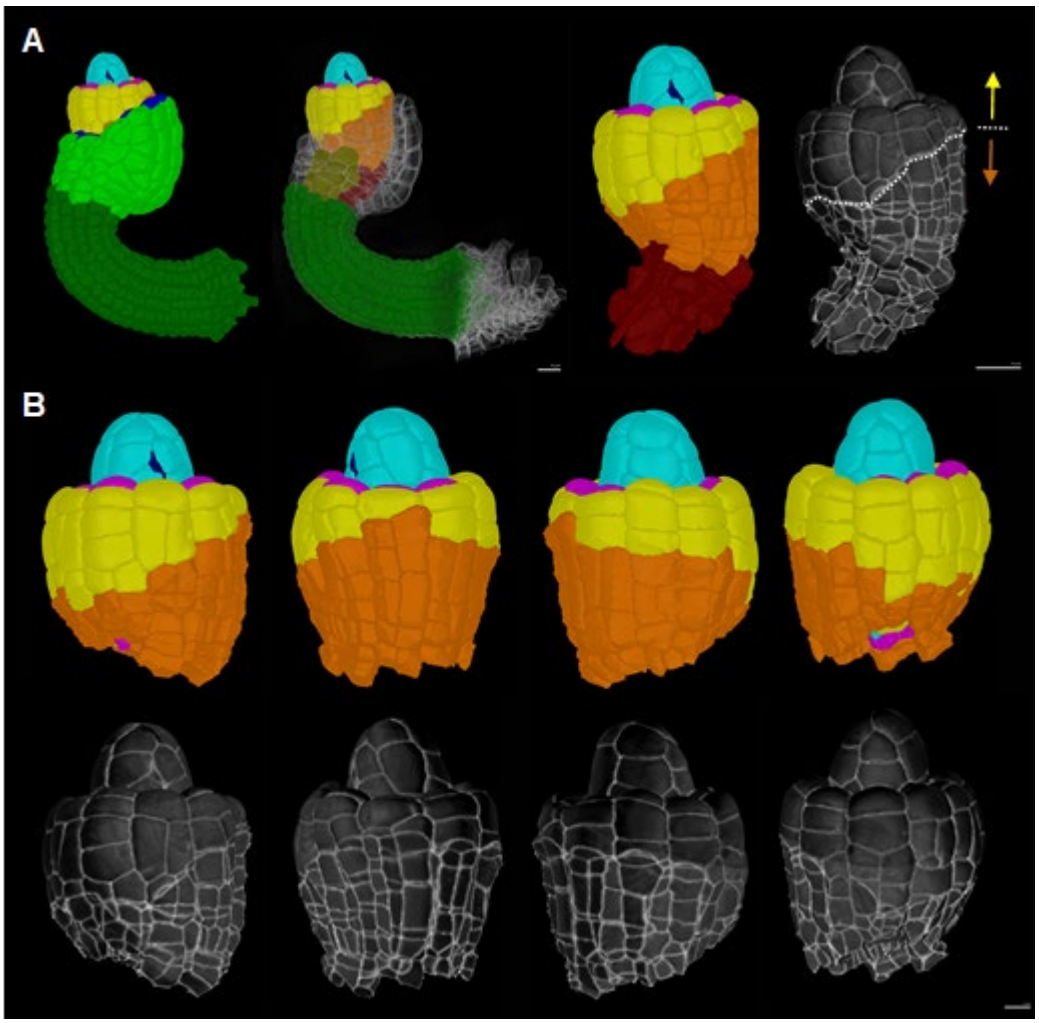

Figure 3.25 Cell morphological changes due to contact friction
Figure representing a 3D *ino-4* mutant ovule (A). From left to right; first: A 3D *ino-4* ovule with tissues labelled in the mesh; second: the outer integument tissues are removed from the mesh view, but the cell boundary predictions image stack is overlaid to display the two regions of ii2 tissue; third: a zoomed view of the ovule in the same orientation as before; fourth: same zoomed image but displaying the projected cell boundary signal onto the 3D mesh. Same *ino-4* ovule in different rotations to display the morphological changes due to the surface of ii2 covered by oi1 (B). The surface of ii2 covered by oi1 is labelled in orange color for visualization. The uncovered cells are marked by yellow color. Signal projection onto the 3D mesh view displayed the small faces of cells made by surface contact. Scale Bar 20µm.

3.2.15 Morphological changes in OI1 during differential outgrowth of integuments

At a mature stage of ovule development, the differential outgrowth of integument tissues results in geometric changes in the oi1 layer of the outer integument. Once the outer integument grows past the inner integument, it leaves behind the inner layer of the outer integument with no internal tissue layers touching them, essentially cells

growing in space. The oi1 layer can hence be separated into two regions, a proximal region where it touches the inner integument and a distal region where its inner surface isn't in contact with any tissues, otherwise called exposed cells of oi1 (Figure 3.26 A, B). The morphology of oi1 tissue looks close to a bent cylinder, but upon estimating the diameter of the tissue, results indicate that it resembles a conical cylinder with the radius becoming small as it tends to be distally located (Figure 3.26 B). This indicates that as soon as the outer integument outgrows the inner integument, they shrink such that they form only a small opening called the micropyle. A hypothesis here is that the outgrown outer integument additionally blocks the further P-D extension of the inner integument by performing the shrink.

A heatmap of cell depth of the oi1 layer indicates that it has non-uniform cell depth along the P-D axis. Interestingly, the exposed cells seem to have a higher cell depth compared to proximal cells of oi1. Additionally, one interesting observation was that some cells of the oi1 layer which are in contact with the distal-most ii2 cells have very low cell depth (Figure 3.26 C). The 2D sagittal section image also indicates that the oi1 cells at this position are shrunk and there might be forces between the ii2 distal cells and these oi1 cells. The reduced cell depth could be a result of these forces. It seems to be also supporting the previous hypothesis that radial shrink in the outgrown oi1 tissue blocks the inner integument outgrowth.

Heatmap of cell volume indicates that distal exposed cells are higher in volume (Figure 3.26 D). Heatmap of cell flatness defined by the ratio of mid to min cell axis indicates that the oi2 cells at the central posterior region are highly flat (Figure 3.26 E). Essentially, the highly flat regions are those where the inner integument distal tip is or was in contact with the oi1 layer during the differential outgrowth of integuments. A quantitative analysis was also performed on the exposed and unexposed cells at different developmental stages (Figure 3.26 F, G) to understand differences in cell volume and depth of these cells. Results indicate that exposed cells gain in volume and depth from stage 3-I onwards. There is a further significant increase in cell depth and volume of exposed cells from stages 3-I to 3-II. However, from 3-II to 3-III, there is no significant increase in the exposed cell geometry, but they remain higher in volume and depth compared to non-exposed cells. Overall, this is a great example of tissue growth and how different a tissue could grow when in the absence of a neighboring tissue or cells.

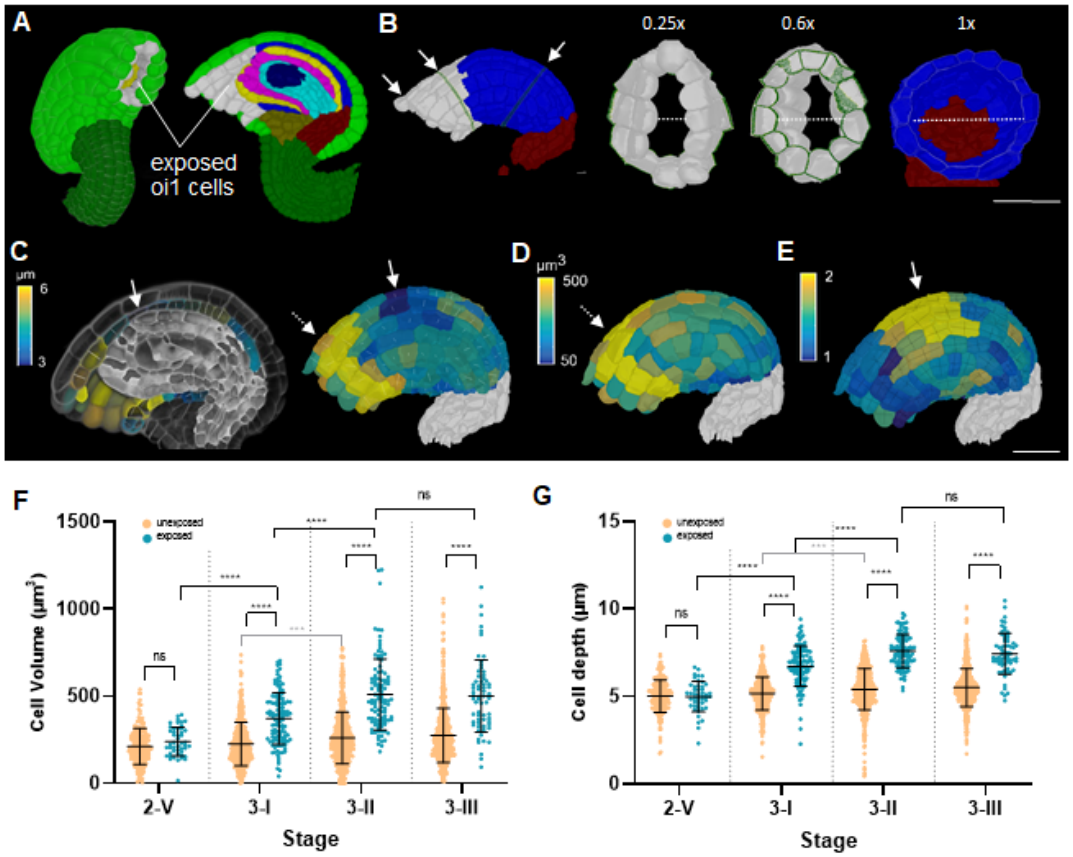

Figure 3.26 Morphological changes in OI1 during differential outgrowth
3D outside view and a sagittal section view of a mature ovule at stage 3-III displaying the exposed cells of oi1 in white color (A). 3D view of oi1 still attached to chalaza clearly displaying the two regions in oi1 (B). Arrows indicate the three regions from left to right on a transverse section displaying the radius of the conical cylinder-like oi1 tissue at the section. Heatmap of oi1 cell depth at the sagittal section view overlaid with the cell boundary prediction to indicate the squeezed region of the oi1 that is in contact with the distal ii2 cells. The right image represents the same heatmap of cell depth, but with a 3D view of only the oi1 cells showing the heat colors. Arrows indicate the squeezed region with low cell depth and the distal region with high cell depth. Solid white line inside the cell represents the cell depth axis. Heatmap of cell volume of the oi1 tissue after removing the outside tissues for better visualization (D). Heatmap of measure of flatness of cell is displayed (E). Solid black lines inside the cell represent the max, mid and mix axis of cells. Graph representing the cell volumes and cell depth of exposed and unexposed oi1 cells from different developmental stages (F, G). Scale Bar 20µm.

3.2.16 Proximal distal growth gradient

Integument tissues form layered arrangements of cells, one tissue on top of the other. Moreover, within a tissue, there are filar or bundled arrangements of cells. Cell files can be long and continuous from the proximal to the distal end. Sometimes they are not continuous throughout, but they still follow a filar arrangement of at least a few cells along the proximal-distal (P-D) axis indicating the divisions are mainly along the P-D

axis (Figure 3.27- A, B). A heatmap of cell volume reveals that there are distally located enlarged cells in the oi2 layer of cells. Proximal cells are relatively small in size, there seems to have a gradient increase (Figure 3.27- A). Similarly, a heatmap cell volume of inner layers also displays a higher volume at the distal end, but at a different scale compared to the oi2 (Figure 3.27- B). Further quantitative analysis has to be performed to better understand the pattern, one requires spatial coordinates to map the cell volume along the central organ axis.

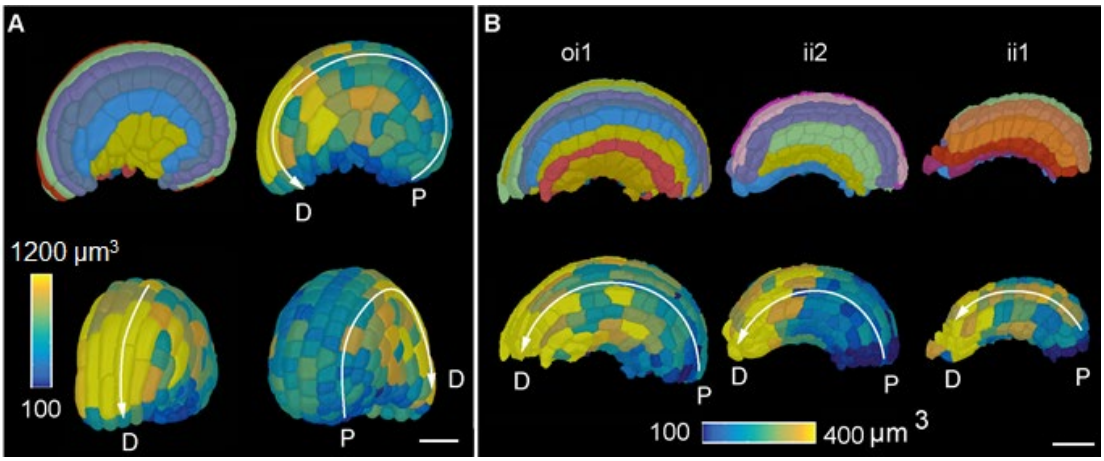

Figure 3.27 Proximal distal growth gradient in Arabidopsis integuments
3D surface view of oi2 from multiple viewing angles (A). Proximal distal axis is labeled, heatmap represents 3D cell volume, cell files are labelled in the top left image. 3D surface view of oi1, ii2 and ii1 layer representing the heatmap of cell volumes in the bottom row and the filar arrangement of cells in the top row (B). Scale Bar 20µm.

3.3 Spatial coordinate system for 3D Arabidopsis ovules using 3DCoordX

3.3.1 Curvature-induced complications in the assignment of axial position

Based on the structure of the tissue under study, different analytic strategies for identifying the position of individual cells are employed. iRoCs and 3DCellAtlas are two well-known approaches for annotating positions to plant organs. These tools work well for straight structures or slightly bent structures (Figure 3.28 A). For example, an Arabidopsis root or hypocotyl is a simple structure to annotate positions along its central axis. When it comes to highly curved structures of interest, this method fails to index axial positions in space (Figure 3.28 B). These approaches result in cells of the same indexed position having different absolute axial distances to a common reference (Figure 3.28 B).

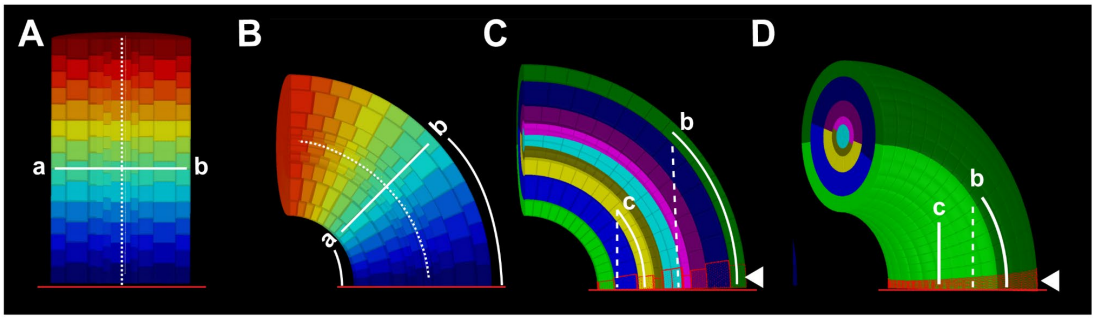

Figure 3.28. Axial cell distance determination in curved tissue.
Section through an artificial template of a tube-like and straight tissue consisting of multiple concentric cell layers (A). The heatmap indicates distance from the origin (red line at bottom). The dashed line outlines the central axis. Note that the two cells (a,b) at the same cell index position also show the same absolute axial distance to the origin (B) Same structure as in (A) but curved. Note that cells a and b differ in their axial distances to the origin. (C) Same structure as in (B). The separate cell layers are distinguished by their different colors. Two cells in different layers are highlighted (c, b). Dashed lines indicate shortest distances to the origin ignoring tissue layers. Sold lines mark the shortest distances to the origin that are restricted to tissue layers. Confining the shortest distance to a given layer reduces axial distance errors. The red line at the bottom highlights the origin. The arrowhead marks the origin cells outlined in red. (D) 3D representation of (C) revealing how the AP boundary further minimizes the axial distance error for a cell in the posterior half of the structure.

Thus, a different strategy was devised in this study to minimize such axial distance errors when assigning positional annotation to cells in for example the slanted primordium or the strongly curved funiculus and integuments. The new approach is based on tissue restriction that includes distinction of radial layers, subdivision into two

AP or ML domains. The coordinate system would have an origin or common reference at a favorable end of the structure or inside the structure of interest. Further, distances could be mapped to cells based on the shortest path through the connected tissue and the coordinate origin. Essentially, the approach annotates positional coordinates as distance from the origin to the cell through the tissue connections of the cell of interest. For example, a cell within the yellow layer at the anterior region has its distance from the origin as its shortest path from the origin to this cell through the yellow layer of cells at the anterior position (Figure 3.28 C, D). The coordinate origin can be a small point like bezier ring that fits into the tip of a cone shaped structure or a wide bezier ring that can be approximated as an origin for the integuments. As a next step, these principles must be applied on real templates of ovule primordia and integuments to have spatial coordinates for cells.

3.3.2 Assignment of PD position to individual cells in 3D ovule primordia

To apply this method of tissue restricted spatial coordinates in 3D digital ovule primordia, the dataset was first annotated with radial and AP domains (Figure 3.29 A). The method was implemented in MGX as a package called “3DCoordX” referring to 3D cell distance coordinates. 3DCoordX annotates PD positions of individual cells based on its position from the origin of the coordinates. The coordinate origin in case of an ovule primordia was placed at the distal tip of the organ, further different parameters like cell neighbor’s and its tissue connectivity is automatically determined and is used for providing an accurate coordinate value.

Positioning the coordinate origin (Bezier ring) at the distal tip correlates with a biologically relevant auxin maximum at the tip as inferred from the expression of the auxin response reporter pDR5::GFP (Benková et al. 2003), the spatial signal of the auxin sensor R2D2 (Liao et al. 2015) , and the finding that polar auxin transport mediated by PINFORMED1 (PIN1) is required for ovule primordium formation (Bencivenga et al. 2012; Galbiati et al. 2013).

The coordinate value in this case is the shortest distance from the origin of the coordinates to the cell through its respective tissue label (Figure 3.29 B). For individual tissues to first have a seed to propagate the coordinate values of other cells within the tissue, a group of origin cells are first defined with a parameter referring to its direct 3D distance to the coordinate origin. Once the origin cells have their coordinate values, other cells within the tissue find its shortest path between its cell centroids to the

nearest direct cell centroid and get the entire stretch to the coordinate origin as a value of distance, similarly origin cells are provided to all tissues and a possibility for adding an additionally selected cell as origin cell is included as a parameter. Alternatively, 3DCoordX also provides the value of the cell index from the coordinate origin (Figure 3.29 C). Thus, following radial and AP tissue annotation and placing a bezier point like ring at the distal end of the ovule primordia, 3DCoordX automatically determines the PD position of each cell, either in terms of cell index or absolute or relative distance to an origin. A detailed method is described as a user manual (See under section 2.19).

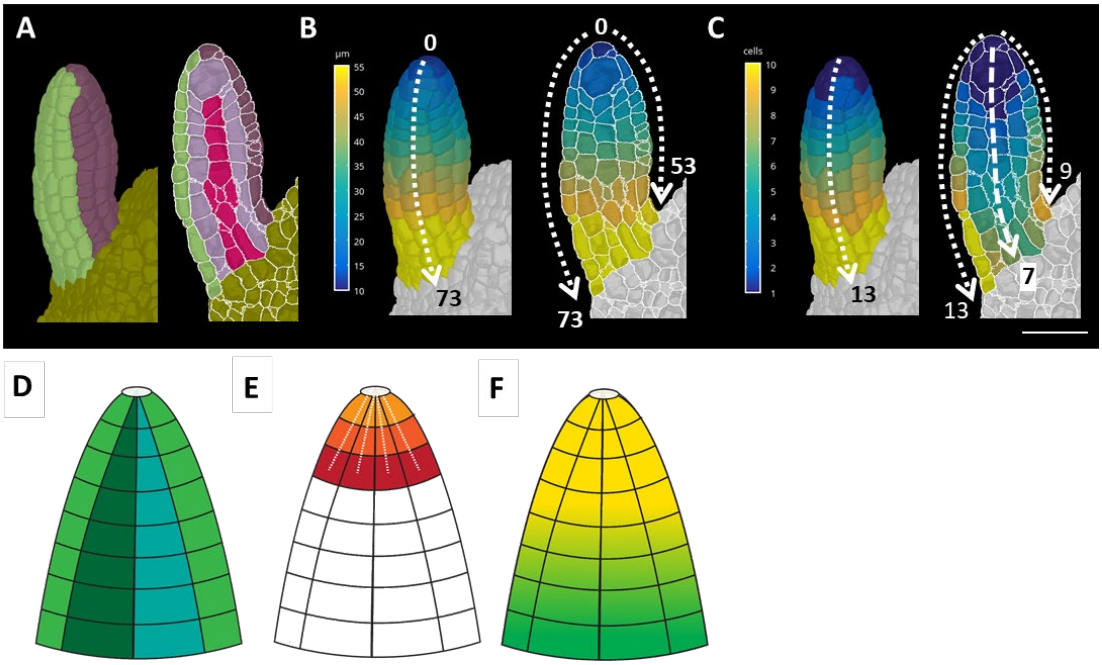

Figure 3.29 Principle of 3DCoordX organ coordinates and its application in ovule primordia

Anterior-Posterior L1 tissue labelled Arabidopsis ovule primordia (A). Left image is a 3D side view, the right image is the sagittal section displaying the interior tissue layers. Distance coordinate map for the same ovule in A (B). Heatmap indicates distance in µm from the origin. Distances along the anterior and posterior sides are indicated. Cell coordinate map for the same ovule in A (C). Heat values indicate the number of cells an individual cell of interest is away from the origin of the coordinates. Cells along the anterior and posterior sides are indicated. Scale bar 20 µm. Cartoon illustrating the principle of 3DCoordX organ coordinate annotation in a cone-shaped structure (D-F). As a first step, the coordinate origin (small ring which is a point-like origin) is placed at the distal tip of a cell type annotated structure. Colors represent the tissue groups (A). In a second step, direct distance measurement finds cells within the provided limit and annotates coordinates values to those cells as the shortest path to the cells from the origin (B). This direct distance measurement does not take different tissue types into account. The white dotted line represents the direct connection from the coordinate origin to these cells. In a third step, these direct cells act as seeds for the coordinate values for other cells within their tissue groups (C). Essentially, for the rest of the cells in the organ, their distance is restricted within the tissue groups labeled in (D) and their

shortest distance to any of the direct cells within their cell type is measured using the Measure 3DCoordX process. The gradient from yellow to green in figure F represents the ideal coordinates to be obtained. All these three steps are performed in just one click using the process: Mesh/CoordX/Measure 3DCoordX

In case of the mature funiculus, essentially a curved cylinder, a larger Bezier ring is placed at its proximal end. Origin cells are then defined by their close user-specified distance to the Bezier ring in 3D (usually 5-15 µm) (Figure 3.28 C,D). Importantly, distances to other cells from the origin cell confined to a given tissue layer and may not cross the AP boundary. The restriction to the tissue layers a cell belongs to removes a large part of the axial distance error as the shortest path through the tissue layers cannot extend through interior tissues (Figure 3.28 C). On top of this restriction, prohibiting the shortest path from passing the AP boundary minimizes the error further (Figure 3.28 D). It should be noted that with this approach small axial distance errors still occur within the anterior or posterior domains depending on the number of laterally arranged cell files within these areas. The remaining errors are typically in the range of a few microns but can be eliminated when taking individual cell files into account. The software can do so, however, the procedure involves cumbersome manual annotation of all cell files for each cell layer. Moreover, the gain of resolution is minimal.

3.3.3 A coordinate system for integuments

Integuments have a more complex architecture compared to the structures discussed above. Outer integument grows into a hood-like shell whereas the inner integument resembles more like a curved cylinder, forming four layers of integument tissues one on top of the other. Integuments can be also subdivided into medial and lateral cells based on the cell position close to posterior central midline and anterior central midline respectively. Overall, the morphological complexity of the shape of integuments makes it challenging to apply any spatial coordinates on them. However, the 3DCoordX principle used for ovule primordia could be also applied on integuments. In an initial step, the integumentary tissues are labelled manually, and its medial and lateral cells need to be defined. Cells can be grouped into medial and lateral groups of cells if one has an existing medial lateral coordinate.

The medio-lateral (ML) and proximal distal (PD) coordinates of all integument cells are established in this study by a semi-automatic manner. For medial lateral coordinates, the user first selects a file of midline cells in each cell layer at the posterior half of the organ (Figure 3.30 A).

Next, the process “Cell distance” is applied which calculates how many cells separate a given cell from the midline (Figure 3.30 B). Cells along theML axis were further subdivided medial and lateral domains, and this is taken as the final tissue label for organ coordinates. In the following step, PD distance coordinates are assigned for all integument cells. A Bezier ring is first placed at the proximal end of the inner side of the outer integument (next to oi1 layer) facing the ii2 layer (Fig. 3.30 D). This circular origin is in the same plane as the ring-shaped expression pattern of the pCUC3::CUC3:CFP reporter which marks the proximal base of the ii1 and ii1’ layers of the inner integument, respectively (Figure 3.30 E). Members of the CUC gene family are generally required for primordium initiation and organ boundary formation (Aida et al. 1997; Ishida et al. 2000; Takada et al. 2001; Breuil-Broyer et al. 2004; Sieber et al. 2007; Galbiati et al. 2013; Gonçalves et al. 2015). Origin cells are then defined by their close user-specified distance to the Bezier ring in 3D (about 5-15 µm). As a direct result of the placement of the Bezier ring, cells of the oi2 layer that are in direct contact with the subepidermal proximal chalaza are assigned a negative value for the PD position (Figure 3.30 F). This feature can be used to separately cluster and analyze those cells. Finally, PD distance coordinates of the integument cells are obtained by searching for the shortest path through the cell centroids to the centroid of an origin cell using the tool “3D Cell Distance Coordinates” (Figure 3.30 F, G). The search is again restricted to a given tissue layer and may not cross the ML boundary. Taken together, the tools “Cell Distance” and “3D Cell Distance Coordinates” assign ML and PD positions for all cells of the integuments. A detailed method for 3DCoordX is described as a user manual (See under section 2.19).

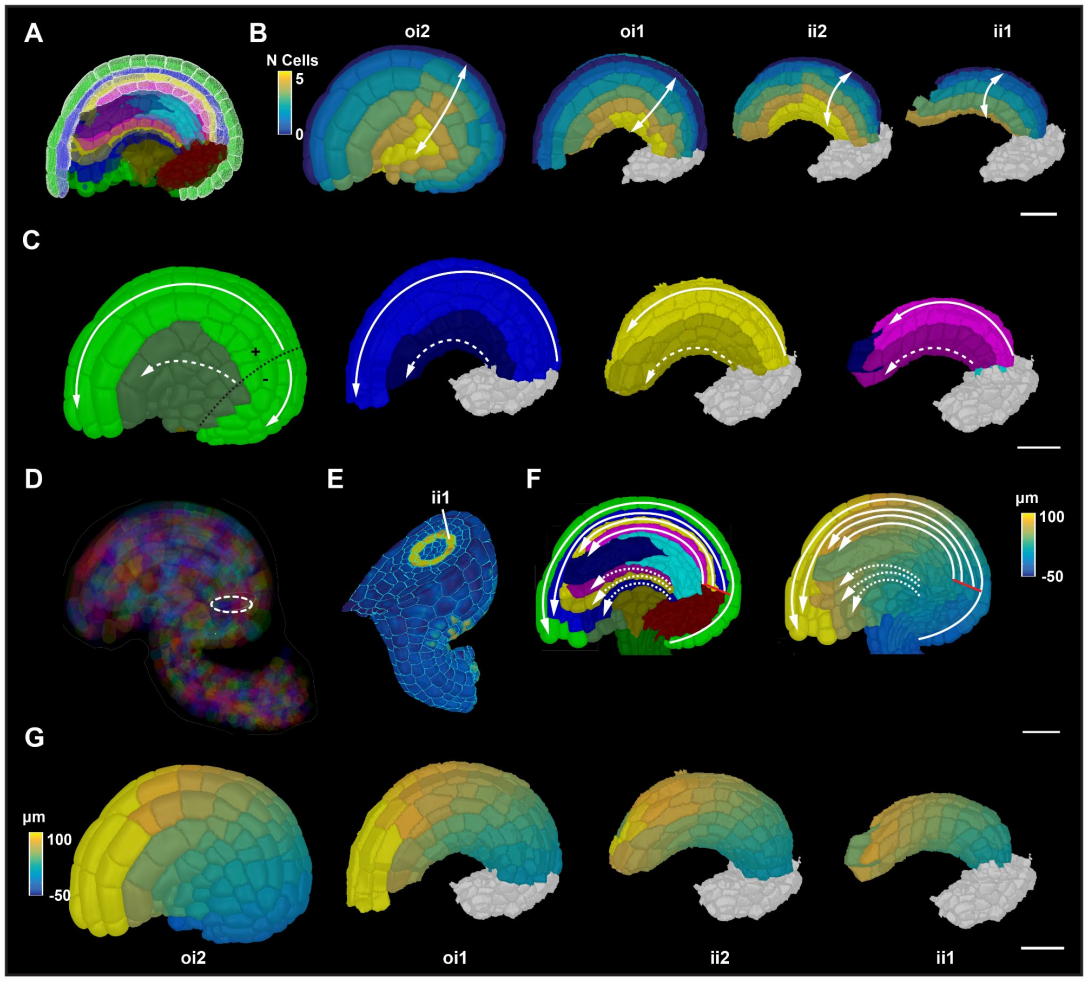

Figure 3.30. Integument coordinate system using 3DCoordX toolbox.
Mid-sagittal section highlighting the selected medial cells on the posterior side of the four layers of integument tissues for medio-lateral coordinate annotation (A). Colors represent tissue annotations similar to Figure 3.12. (B) Heatmap of medio-lateral cell coordinates. Heat values indicate the lateral position in terms of the number of cells from the median file of cells. Different integument tissues are extracted from the 3D mesh to display the medio-lateral coordinates at their tissue surface. (C) 3D surface view of integument tissues similar to (B). Medial and lateral cells are distinguished. Solid white line represents the tissue restricted coordinate direction along the medial group of cells. White dashed line represents the tissue restricted coordinate direction along the lateral group of cells. Black dotted line on oi2 represents the coordinate origin projected on the surface which separates the proximal oi2 cells with negative coordinate values (D) Semi-transparent view of a mature 3D ovule displaying the coordinate origin as a ring inside the organ. (E) 3D clipping view of a transverse section of an ovule highlighting the ring-like expression of the pCUC3::CUC3:CFP reporter in yellow. (F) Left panel: sagittal section of a mature ovule displaying the coordinate directions of the medial and lateral group of cells in solid and white lines, respectively. Solid red line indicates the origin of the coordinate system. Right panel: Sagittal section displaying the heatmap of distance coordinates. Solid red line indicates the origin of the coordinate system. (G) 3D surface view of integument tissues similar to (B) displaying the distance coordinates at the surface of internal tissues. Scale bars: 20 µm.

3.3.4 Spatial mitotic distribution in Ovule Primordia

Taking advantage of the ovule primordia coordinate system, mitotic cells were mapped to the cell such that their spatial arrangement could be studied. Cells that exhibit mitotic figures were manually selected and labelled on the 3D mesh (Figure 3.31 A). A total of 52 mitotic cells were found from 52 ovule primordia. Radial analysis indicates that 33 mitotic cells are in L1, 18 in the L2, and 1 in the L3. This result is in line with the observed differences in cell numbers between the three layers. Mitotic cells from the L1 were further looked at if they are in the anterior or posterior side of the organ. Results indicate that more cell divisions occur in the L1 anterior domain (23 L1 anterior mitotic cells and 10 L1 posterior mitotic cells).

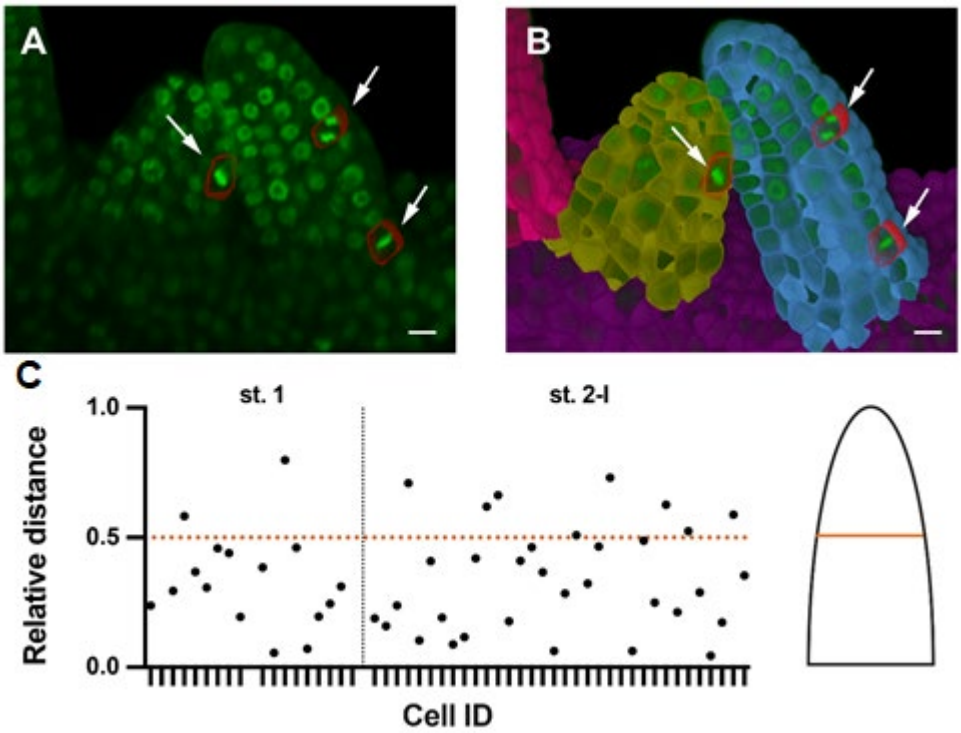

Figure 3.31 Mitotic patterns in the ovule primordium.
Different stages of wild-type (Col-0) early ovule primordia are analyzed. (A) 3D frontal plane view of ovule primordia displaying the To-PRO-3 nuclear stain. Cells undergoing mitosis are outlined in red and marked by arrows. (B) Same section in (A) with an overlay of 3D labelled meshes. Arrows indicate mitotic figures. Plot showing the relative distance along the proximal-distal axis of mitotic cells in a primordium. Relative distance was calculated by the following formula: relative disance = 1 - distance coordinate/mean length. Scale bar 20 µm.

To the end, the spatial mitotic distribution song the P-D axis of the organs as quantified. Results indicate that about 80 to 85 percent of scored mitotic cells were in the proximal half of the developing ovule primordia (Figure 3.31 C). Taken together with radial analysis of cell numbers and volumes (Section 3.2.2, Figure 3.15), the data indicate

that primordium outgrowth is preferentially driven by an increase in cell number, not cell volume (Figure 3.15 A-E, Figure 3.31 C). In addition, they suggest unequal spatial distribution of mitoses between cell layers and along the AP and PD axes. A higher number of mitoses in the anterior domain could explain primordium slanting. The data further indicate that a cell proliferation zone located in the bottom half of the developing primordium contributes significantly to its outgrowth.

3.3.5 Funiculus curvature correlates with differences in cell number and cell volume along the AP and PD axes

This study aims to understand any spatial pattern of cells along the central axis of the funiculus using 14 3D digital ovules of stage 3-IV (Figure 3.32 A-E). By this stage growth of the funiculus has ceased (Vijayan et al. 2021). Again, like the ovule primordia, the funiculus tissue was subdivided into radial layers and with anterior or posterior subdivision. Patterns along the L1 anterior and posterior domain indicate that the maximal extension and number of cells along the anterior side is higher compared to the posterior side.

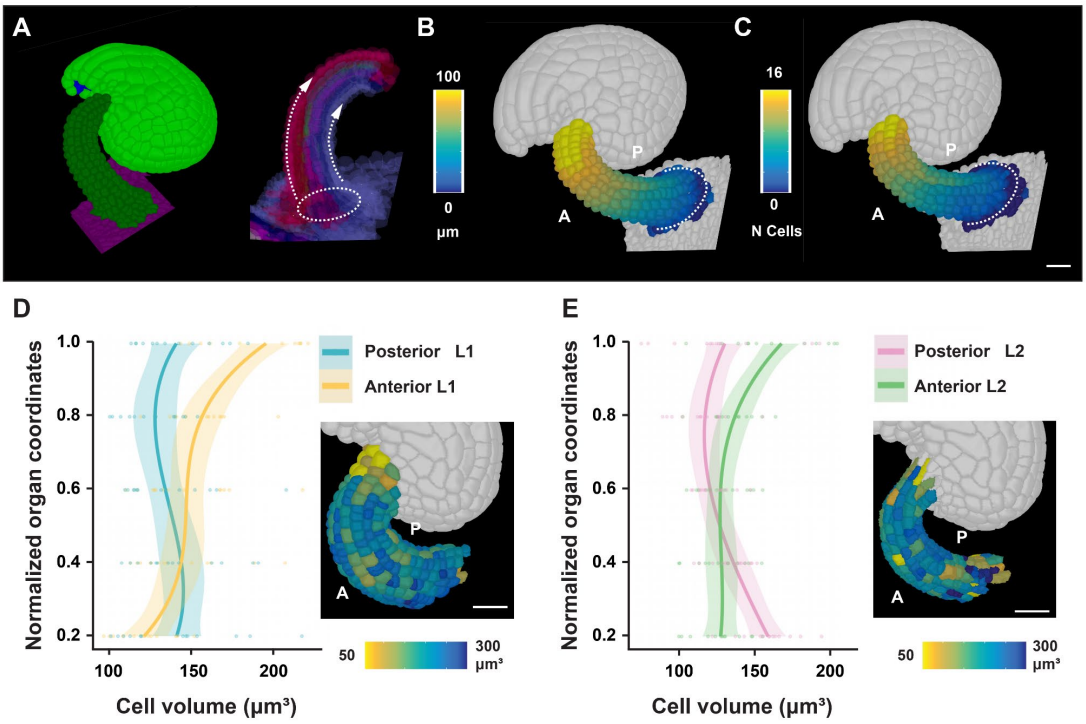

Figure 3.32. Cellular features of funiculus curvature.
Wild-type (Col-0) ovules of stage 3-V are analyzed. (A) Left panel: Tilted side-view of a 3D cell mesh. Right panel: Semitransparent 3D mesh view of funiculus extracted from the 3D mesh of the organ. The Bezier ring serving as origin is placed at the proximal base of the funiculus. The dashed arrow lines

indicate the coordinate direction along the anterior and posterior midlines. (B, C) Same specimen as in (A). The anterior and posterior sides are marked. (B) Heatmap of cell numbers along the PD axis of the funiculus. (C) Heatmap of cell distances along the PD axis of the funiculus. White dotted line indicates the coordinate origin as a ring. (D) Graph depicting cell volumes of anterior and posterior L1 cells in relation to the relative PD position. The inset in the bottom right corner shows a heat map of cell volume in the L1 of the funiculus. 14 3D digital ovules were analyzed. Number of cells xx ≤ n ≤ yy (E) Similar graph as in (D) revealing cell volumes of anterior and posterior L2 cells. Scale bars: 20 µm

Results of cell volume along the organ coordinates indicate that there is a gradient along the proximal-distal axis of the anterior L1 and L2 cell volumes. Additionally, the anterior distal-most cells feature nearly 1.5 to 2 times the volume of cells located at the anterior proximal end (Figure 3.32 D, E). The posterior side of the organ was not observed with any similar volume increase. In summary, the data suggest that a combination of differential cell proliferation along the AP axis and unequal cell growth along the PD axis of the anterior domain contributes to the curvature of the funiculus.

3.3.6 Proximal-distal growth gradient in Arabidopsis integuments

Genetic data indicate that asymmetric growth of the outer integument contributes significantly to the anatropous shape of the ovule (Baker et al. 1997; Schneitz et al. 1997; Vijayan et al. 2021). With the advancement of 3D digital ovule atlas (Vijayan et al. 2021), 3D quantitative cellular features can be studied and visualized as heatmaps and such features could be analyzed with tissue clusters from the dataset (Figure 3.33, A, B). Several of the analyses with tissue clustering gives a good idea of how the tissue growth happens. Additionally, the heatmap visualization enabled the first understanding by visual quantitative values in space. Further data exploration that aims at comparing multiple samples or different tissues from the same sample was impossible without a proper spatial coordinate system. The current coordinate system allows to perform such analysis of tissues and understand the changes in the same platform. Analysis of 3D cell volumes of integument tissues from a single ovule along the proximal-distal axis reveals that the cell volumes of outer layers, especially the oi2 layer is higher compared to inner layers and they undergo a gradient increase from proximal to the distal end (Figure 3.33, C). Proximal cells exhibited small cell volumes while, except for small cells at the tip of the integument, distal cells were characterized by larger sizes.

The gradient increase is observed in all four layers of integuments, but the slope of the gradient is reduced when it comes to inner layers. Additionally, as a proof of concept for mapping a single cell feature from multiple samples, cell length to cell position along the PD axis in the medial domain of the oi1 layer across five different specimens was compared (Figure 3.33, D). Results indicate that cell length increased along the PD axis. Essentially, the data identified patterns from different tissues inside an organ that can be qualitatively looked at the same platform with spatial resolution. Results suggest that preferential cell elongation along the PD axis is an important factor underlying differential growth of the outer integument and ovule curvature.

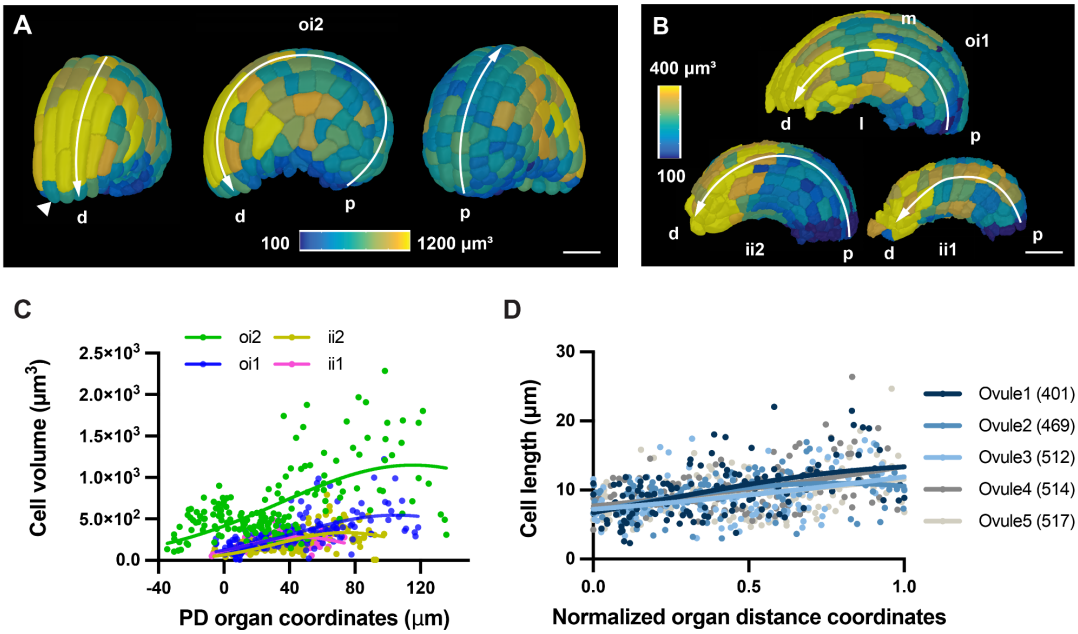

Figure 3.33. 3D spatial analysis of integument cell geometry
Wild-type (Col-0) ovules of stage 3-IV are analyzed. (A) Heatmap of oi2 cell volumes. Panels depict a tilted frontal view (left), side view (center), and tilted back view (right). The PD axis is indicated. The arrowhead marks a small tip cell. (B) Side views of the same specimen as in (A) showing the 3D surface view of internal tissues. Heatmaps of the oi1, ii2, and ii1 layers, respectively. (C) Graph showing cell volume in relation to PD distance for the four integument layers of a representative ovule. The respective nonlinear Gaussian regression curves are indicated. 95 ≤ n ≤ 172. (D) Cell lengths in relation to normalized PD distance. Cells in the median oi1 layer of five ovules were analyzed. The respective nonlinear Gaussian regression curves are indicated. 126 ≤ n ≤ 133. Note the increase in cell length towards the distal end. Scale bars: (A, B) 20 µm. Abbreviations: p, proximal; d, distal; m, medial; l, lateral.

3.4 3D Nuclei atlas within the context of cells

After all the advancements in microscopic imaging of cell boundaries and nuclei stains, this study demonstrated the power of deep learning algorithms in image segmentation of cells and nuclei. With the application of 3DcoordX, organ coordinates can be also obtained. Combining all these features, a 3D digital atlas can be now looked at where the cellular features can be linked to the features from the nuclei and they can be studied at the same time. For example, one can investigate the nuclear size, shape and other features and compare over the size and shape of the cell in which the nucleus of interest is located. This is quite an advancement in 3D analysis of cellular whole organs. Additionally, since the cells and nuclei are now linked, tissue labels from the cells can be now applied on nuclei and tissue specific nuclei analysis could be also performed (Figure 3.34 D). The organ coordinates obtained from the cells can be also performed with just the nuclei or the nuclei can be now looked at spatial resolution with 3DCoordX. Overall, this study has increased the possibility to extract hidden information from a complex 3D microscopic image of cell boundary and nuclei of the whole organ.

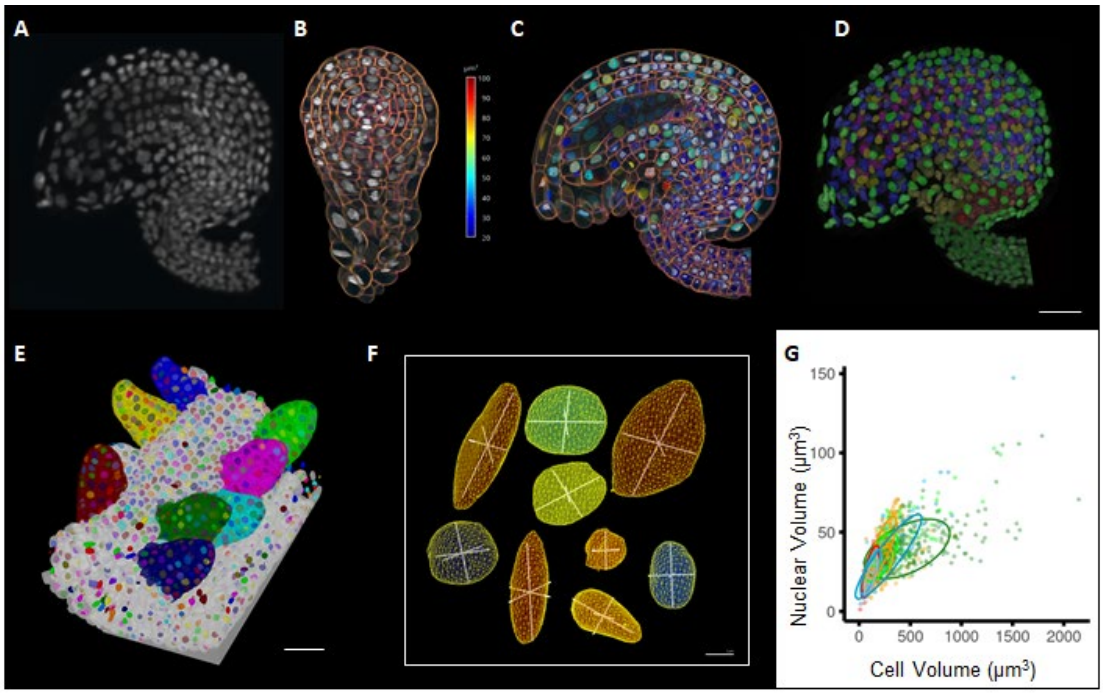

Figure 3.34 3D Nuclei atlas within the context of cells
Figure demonstrating the application of cell segmentation and nuclei segmentation from microscopic images. Mid sagittal section of a mature Arabidopsis ovule displaying raw microscopic image of nuclei stain (A). Transverse section of the same ovule displaying the overlay of cell wall stain and nuclei raw images (B). Figure displaying

the 3D segmented nuclei with the heatmap displaying the nuclei volumes in 3D (C). Figure displaying the application of cell-based tissue labels to nuclei mesh, essentially the tissue labels from cell segmentation is transferred to nuclei for visualizing the nuclei from different tissues that in turn allows the extraction and analysis of nuclei from different tissues separately (D). 3D view of multiple ovule primordia displaying cell and nuclei segmentation that can be now used to extract information regarding the nuclei size and cell size or its ratio in different ovules of the same stage (E). Figure demonstrating the different shapes of nuclei in Arabidopsis ovules (F). These nuclei linked to cells now allows to look at the cell shapes when the nuclei are shaped differently. Figure demonstrating the differences in cell volume and nuclei volume from all cells in a mature Arabidopsis ovule (G). Ellipse represents different tissues within the organ.

4 Discussion

4.1 Method development for generating 3D Arabidopsis ovule atlas with cellular and tissue resolution

This study aims to understand qualitative and quantitative cellular patterns from microscopic instance segmented images of Arabidopsis ovules. The prerequisite for generating a precise 3D digital model of an organ is very high. It starts from optimized microscopy with a good signal to noise ratio, high image resolution in the xy and z axis of imaging, image processing and instance cell segmentation. The requirements are mainly to obtain a good image that can be instance cell segmented

hours of manual correction, tissue annotation, and finally, the procedure repeated for all developmental stages with a good number of samples. Essentially, the whole procedure requires inputs for improvements from multiple angles that could make the task easy and results reliable. This section of the study has reliably come up with several improvements for the whole atlas generation for Arabidopsis ovules which is now also applicable for other organs.

4.1.1 Improvements in sample preparation and microscopic imaging for instance cell segmentation

This study has identified that tissue clearing is essential to attain a high-quality atlas of organs with cellular resolution. A novel protocol for rapid, reliable, and simple imaging of fixed and cleared Arabidopsis ovules is out from this study (Tofanelli, Vijayan et al., 2019). The strategy combines two recently depicted methods for clearing and cell wall staining, ClearSee and SR2200 staining, and incorporates several modifications and additional improvements for sample preparation that highly improved the final results of microscopic images and that now allows instance cell segmentation of such images when acquired at a high resolution. Also, the methods using fixed and cleared samples stained with a cell wall stain are compatible with an FP-based reporter line that can be now imaged along with cell outlines from the cell wall stain reliably. Additionally, this high-performance cell wall stain eliminated the need for a cell outline reporter line for performing a whole organ cell segmentation, this further expands the project to examine several genotypes and mutant phenotypes without generating a cell outline reporter line in its background.

Nuclear staining with TO-PRO-3 is another important adaptation in the above-mentioned method. TO-PRO-3 allows detection of all nuclei in Arabidopsis ovules, even those of the developing gametophytes that allow identifying mature ovule developmental stages based on gametophytic nuclei. Moreover, it is compatible with SR2200, the emission spectra of which do not overlap. Overall, the nuclei stain can be imaged without any hassle and allows to confirm the presence of a cell. This is especially useful when the cell outlines are weak and it's hard to distinguish between cells in any regions.

Stains have reduced tissue penetration, additionally, nuclei stain had a rapid loss of signal intensity when applied with the original ClearSee method for sample preparation, this study has made a slight modification in staining solution (staining in PBS solution) and sample mounting in Vectashield that retains the nuclei stain signal and also allows deeper stain penetration. This slight modification was critical as it now allowed better sample preservation for long-term observation up to six months in current understanding. Multiple numbers of ovules can be images from the same microscopic slide with full 3D resolution and without significant phototoxicity to the sample due to the adapted sample mounting in Vectaschield. Overall, Vectaschield mounting has allowed more freedom on time and helped in a rapid increase in samples that can be imaged or stored for future imaging.

Deep imaging was another challenge as light scatters in deeper tissue layers. Here, the issue was solved using just a confocal microscope and without going into a more advanced multiphoton microscope that allows deeper tissue penetration. The mentioned depth of the tissues is in the range of 50-80 µm. This study uses a 63× glycerol objective with a numerical aperture of 1.30, sample mounting in Vectashield makes maximum compatibility with the glycerol immersion and objective as it has the refractive index similar to glycerol. A consequence of z scanning, the signal intensity reduces as one goes into deeper tissue layers. This was homogenized with an automated z compensation in the microscope that increased the laser power and detector gain in a gradient fashion when the scanning was performed along the z direction. These improvements are initially developed for Arabidopsis ovules, but they are also applicable for other plant organs. Overall, this study has made critical changes to the sample preparation using the ClearSee method that valuably added to several improvements of the final microscopic image.

Digital 3D models with cellular resolution are crucial for the quantitative analysis of organ morphogenesis. These images generated by improved methods are instance cell segmentable as they were acquired with high image resolution. The demands for initial image resolution were kept very high because there weren't any existing machine learning tools and the tools used required high image quality as it was just based on ITK watershed. Overall, the existing tools require higher-level human input in correcting image segmentation, and it is not ideal when the whole procedure needs to be applied on hundreds of samples from various developmental stages. Further work must be done to improve the rapid, and ideally automated, generation of 3D digital ovule models, for example by improving image analysis of ovules and the segmentation algorithm, through deep learning approaches.

4.1.2 PlantSeg deep learning-based image improvement for precise instance 3D cell segmentation

PlantSeg is a simple, powerful, and versatile tool for plant cell segmentation that came to form based on the images provided from this study. PlantSeg uses the latest development in machine learning and computer science to perform the image improvement task, to start the whole pipeline from scratch, computer scientists require ground truth images or segmentation of the provided raw images for training purposes. Training is performed to create a model that could ideally generate reliable predictions of new raw input images. This study provided hand-curated high-resolution images, which are resampled with several factors of downsampling to perform PlantSeg training of different resolutions.

PlantSeg resulted in current best and automated instance cell segmentation of provided images. It went down from almost 30 minutes to 5 minutes of human input needed to correct a segmentation from a mature Arabidopsis ovule containing about 2000 cells in 3D. Nevertheless, the segmentation results produced by PlantSeg on new ovule datasets are not fully perfect and still require some human proof-reading to reach 100% accuracy. PlantSeg advancement has also generated images of cell boundary predictions, this is now a good representation of the raw images for visualization purposes that helped in better understanding the morphology of the structure with less background noise.

Thus, this study has provided a new benchmark dataset of large number of high-quality images of Arabidopsis ovules with gold standard ground truth labels for machine

learning-based image segmentation that is applied for PlantSeg training. Additionally, the training was also performed on reduced dimensions of 2x and 3x which now allows to segment images from even low resolution. Essentially for previously unaccepted image quality standards for 3D cell segmentation, PlantSeg now offers the possibility to perform 3D cell segmentation in those images in a good manner.

4.1.3 Deep learning-based 3D nuclei instance segmentation using a novel training dataset for faint images of nuclei

The image dataset of Arabidopsis ovules includes a nuclei image stack which was initially used to identify gametophytic nuclei stages and for validating the presence of nuclei when it comes to proofreading the cell segmentation. Perfect instance segmentation for such faint nuclei channels is nearly impossible with any possible machine learning tools. This study has broken the conventions and generated a new training dataset that provides ground truth images for those faint nuclei channels using the nuclei segmentation of a bright nuclei reporter as a pseudo segmentation. This approach has greatly improved the results of nuclei segmentation. Now the To-PRO3 faint nuclei stain images can be segmented, nuclei sizes and shapes can be quantified from those segmentations. Signal intensity of the nuclei stain and any reporter activity can be also quantified in 3D with these nuclei segmentation.

Additionally, the first steps for further improvement in cell segmentation based on nuclei domain knowledge were also applied in this case here. This would require a further improvement in the computer scripts to develop further to a higher level of precision. This study did not perform a scored comparison of the results of nuclei segmentation and the methods for improving cell segmentation based on nuclei domain knowledge, but a qualitative and morphological analysis was only performed on the results.

4.1.4 Whole organ tissue annotation in 3D digital atlas

A 3D digital organ with full cellular resolution has an unprecedented analytical capability. Examination of 3D digital ovules revealed new insights that could have only been obtained by studying the morphogenesis of the organ in 3D with cellular resolution imaging and cell segmentation. Tissues were labelled from early to late stages of development. The 3D tissue annotation of different tissues in the organ resulted in gaining few previously unobserved morphological features of the organ

including defining different regions of the organ with proper cellular resolution. Additionally, this has added a valuable clustering feature of the organ, without tissue annotation the possibility of the study was limited to understanding the total number of cells in the organ and other morphological features. Now with tissue annotations, there is a possibility for tissue specific quantification and analysis of cellular features. Surface rendering the internal tissues was another advanced tissue structure analysis that was possible after tissue annotations to the digital models. This now allows a better understanding of the shapes of internal tissues by itself and how this is different in different genotypes and at different developmental stages.

4.2 Stage-specific 3D digital ovules atlas with cellular and tissue resolution

Several characteristic features of the organ were identified based on the tissue labels annotated to the 3D digital images after cell segmentation. Also, a quantitative analysis was performed on growth and proliferation of tissues at different stages. This study has provided a benchmark with a lot of possible information on wild-type Arabidopsis ovule development. A few of the highlights are mentioned below.

4.2.1 Growth patterns forming ovule primordia

The dataset of Arabidopsis ovules comprises early to late stages of development. Global analysis of early development reveals that the ovules follow a continuous growth or continuous increase in cell number and overall volume. This confirms that there are no specific growth pulses or phases of early development, but rather a continuous growth and proliferation of cells. Additionally, the PD extension of the primordia was also quantified for all these ovule primordia to obtain continuous growth. Tissue specific analysis reveals further details on what tissue is composed of the most number of cells and their relative growth and proliferation rates. Results indicate that L1 is composed of more cells and L1 undergoes more proliferation. This is also essential as L1 forms the outer surface of the organ, for layered cone shaped structure, the surface area of the outer layer has to be higher than inner layers, thus the L1 is composed of a large number of cells that makes sense. But interestingly, L1 is composed of small cells. Overall, these tissue specific analysis indicate that the pattern in different tissue layers are not the same, cellular divisions and growth in different tissues within ovule primordia are different.

Ovule synchrony was also looked at by quantifying the number of cells, total volume, PD extension and cell volumes from ovules within a pistil fragment. Variations do occur within ovules of the same pistil; the analysis reveals that these ovules don't have the same characteristic developmental stage and a slight variation is to be expected. But here the variation is quantified with cellular resolution, allowing the possibility to characterize and compare the synchrony in any mutant genotype of interest.

4.2.2 Slanting and polarity in ovule primordia formation

The results provide a new understanding of the establishment of ovule polarity. It was previously observed that ovules are oriented relative to the long axis of the gynoecium (Simon et al., 2012). The gynapical (micropylar) side is pointing to the stigma and the gynbasal (chalazal) side toward the base of the gynoecium. It remained unclear, however, when this polarity became established and whether it could be recognized throughout the internal tissues of the ovule. Results indicate that the final orientation of the ovule does not correspond to how the polarity is established. The emergence of a slant in the ovule primordium represents the earliest morphological manifestation of polarity in the ovule. Importantly, this early anterior-posterior-axis was initiated at a right angle to the long axis of the gynoecium. Anatomical and marker gene expression data further suggest that the anterior-posterior polarity is maintained during subsequent development and throughout all major tissues along the PD axis of the ovule. This is indicated by the polar distribution of the distal two nuclei of the four-nuclear embryo sac, the morphological differences between the anterior and posterior chalaza in the proximal central region, and the posterior placement of the phloem within the funiculus (Vijayan et al., 2021). The gynbasal-gynapical orientation of the ovule is eventually achieved by a turn in the funiculus. Thus, the alignment of the ovule with the apical-basal axis of the gynoecium does not directly correspond to the establishment of anterior-posterior polarity in the primordium but rather depends on a subsequent morphogenetic process in the funiculus. A basic molecular framework underlying primordium outgrowth and integument formation has been established (Chaudhary et al., 2018; Cucinotta et al., 2014; Gasser and Skinner, 2019). However, it is not known what regulates anterior-posterior polarity, slanting, and funiculus twisting. The mechanism may involve PHB and cues from the replum-septum boundary that is positioned in the medial plane of the gynoecium (Reyes-Olalde and de Folter, 2019; Roeder and Yanofsky, 2006).

4.2.3 Polarity in the internal central region of Arabidopsis ovule

The internal central region of chalaza was less studied previously mainly because of no 3D data availability. The cellular architecture of the internal central region indicated that it's composed of two different groups of cells. Anterior and posterior chalaza. These two groups of cells display differences in the shape and size of cells. They display morphological differences in cellular packing compared to other cells in the organ which is visible by a qualitative analysis. The shape formed by these two tissue groups is very complex. A hypothesis here is that the cells of anterior chalaza are composed of loosely packed large cells, whereas the cells of posterior chalaza might be composed of tightly packed small cells. This hypothesis is mainly derived by looking at the shapes and sizes and how they are packed in 3D. Unfortunately, less tools are developed to study the cellular packing of these tissues and shapes of cells formed by these two tissues.

4.2.4 Mapping characteristic cellular feature to ovule curvature formation

This study has identified differential growth of integument tissue layers as a highlight feature of Arabidopsis ovule development. A detailed morphological analysis of the integument growth at the sagittal section revealed characteristic events during integument outgrowth and curvature. Results strongly suggest that differential initiation and outgrowth have a major role in forming the curved shape of the ovule. A major event in this differential outgrowth is when the outer integument two layers outcross the inner integument tissues. One hypothesis is that once the outer integument outcrosses the inner integument, it has gained control in the development of the inner integument. In an exaggerated manner, the outer integument might be able a stop to the PD growth and extension of the inner integument and allow the inner integument to only follow the outer integument. This was considered as the outer integument was found to be encroaching the area where the inner integument has to eventually grow out. This area is filled with cells of adaxial oi (oi1) with expanded cells in the depth axis, called exposed cells of oi1 in this study.

Additionally, there are contact frictions between the inner and outer integument. The outer two layers grow together and they would ideally be able to slide over the inner integument surface as they does not share a wall between them. Neverthelss, the walls between inner and outer integuments are in direct contact. But the extent of contact and the resulting friction might be also able to control the amount

of sliding over between the inner and outer integument. Control of contact friction might be another mechanism behind ovule development and curvature.

There are several features displayed by the integuments during its differential outgrowth. Contact friction is one advanced feature of this organ. The morphology of cell surfaces changes when they are covered by cells. This study shows a prime example of outer integument outcrossing the inner integument and the cell shapes of the surface cells of the inner integument are altered. The otherwise smooth surface of ii2 is now rough with faces of cells of oi1 projected on its surface due to contact between these two layers as the oi outcrosses the ii.

A global and tissue-specific analysis of integument tissues was performed. Quantitative analysis of integument growth and proliferation revealed different patterns in integument tissues. The growth and proliferation rates of integument tissues correlate with the different phases of integument outgrowth. The early phase of integument outgrowth where there is a pulse in oi growth that leads to the oi outcrossing the ii is clearly visualized by the growth and proliferation rates of oi1. In contrast, the values from oi2 don't say much. This could be because all the oi2 cells are included in this analysis. An analysis excluding the proximal oi2 cells covered by the chalazal cells might have made more sense. This can be now performed by just clustering the cells with positive coordinate values in the 3DCoordX attribute. Due to time constraints, this is not included in this study.

Results also indicate that the inner integument undergoes high growth and proliferation at its early growth, this can be taken as a mechanism for ii to already attain sufficient growth before the oi starts to grow. Essentially, if the oi starts to initiate at the same time to ii, the shape of the organ might have been different. This can be a debatable statement as it's not supported by experimental evidence. Current knowledge is that before the oi outgrows, the ii has to perform some sufficient growth that allows oi to grow on the surface of ii along the PD axis, further there is preferential growth of oi that leads them to outcross ii, then there is a phase where all these tissues grow at almost the same rate. Quantification strongly suggests that there are growth pulses in oi cells that help them to grow faster than ii and outcross. Ovule curvature and these growth pulses can be related, but any statement doesn't not stand without proper experimental evidence. Nevertheless, the results suggest that oi outcrossing ii might be essential to form the final curved structure. Also, ii initiations before oi and outgrow already at a

higher rate until oi starts to outgrow might be an essential feature of the organ to form the curved shape.

Differential growth and proliferation rates say much about tissue growth, having cellular level information adds more details to the quantifications. Results from surface rendered heatmaps of cell volumes of integument layers indicate that there is a prominent gradient increase in cell volumes of integument cells. Essentially, the cells at the distal end are composed of big cells compared to the cells at the proximal end. Correlating this to the differential growth indicated further about the organ structure formation. Ovule integument cells are attached to one end, growth is possible along the curved axis and growth pushes the cells to more distal positions. Essentially the growth is favourable at the distal end compared to the proximal end. Considering differential growth of integuments as an essential feature of the organ, for the organ to attain differential growth, the easiest mechanism would be to pattern the cells at the distal end with higher cell volumes, this is what we observe as a polarity or gradient increase in cell volume along the proximal-distal axis of the organ. Considering if the organ decided to produce larger cells at the proximal pole, this would require higher-level coordination and cells might end up protruding from the surface, instead of when the distal cells are enlarged, only a few cells distal to the cell of interest have to be taken care of. Overall the pattern of cell enlargement in integument cells might be an additive mechanism to generate differential growth of integument.

Differential outgrowth resulted in the differential cellular property to distal exposed cells of oi1. This is another observed feature that can be also correlated to the final shape formation. Quantification of cellular features of these cells indicate that they expand in depth. This is an interesting morphological feature and can be correlated as a competition between oi and ii. Once the oi outcrosses the ii, these exposed cells appear or in other terms, the cells are exposed and they display such characteristic higher cell depth. This can be a simple event irrelevant to the shape of the organ, but on the other hand it can also be a highly relevant feature. A hypothesis here is that the higher depth cells might be able to make a stop to the ii outgrowth by making physical contact with the distal growing end of ii. This makes ii to a point that it can now grow by only following the oi. Also, if the curvature of the organ is looked at, this feature is almost at the last stage to form the final bending. The problem can be also looked at differently. If the oi1 cells might have not increased in-depth, the inner integument would be still able to grow slightly following the shape made by oi, but the ii would have

been able to grow with few constrictions in its path. Having said this, the shape of the ovule might have been more straight at the distal end. Overall it's still a debatable feature, but a valuable pattern found from 3D digital ovule images. Additionally, the oi1 shape is featured as a cylinder with its closing end smaller than the opening end. Shapewise, this indicates that the oi1 shrinks the ii cells as it comes to the distal end. In other words, the oi1 physically occupies the space for ii to outgrow, resulting in the ii to grow slowly or to produce smaller cells or push back hard the oi1 cells. This is now direct evidence of interaction between these tissue layers during outgrowth.

The ovule curvature is a complex morphogenesis, this study has tried to identify the key features of the organ that could have direct or indirect link to forming the curved shape. It could be also that some of the features are a resultant of different characteristic pattern established in the organ. To answer this in a clear manner, different mutant genotypes have to be established with differences in each of the feature mentioned in this study. For example, what happens in a mutant which doesn't make distal enlarged cells? what happens when the outer integument gets initiated at the same time to inner integument? What happens if the inner integument prolongs to grow at a higher rate than outer integument? Validating the importance of all these characteristic features is time consuming and out of the scope of the current study.

4.3 Spatial coordinate system for 3D Arabidopsis ovules

Analysis of the organ with tissue resolution revealed several features of the organ. Most of these features follows differences along the central axis of the organ. Spatial patterns can be morphologically studied by visual analysis of the 3D digital organs, but they lack any quantitative analysis that can be linked to the spatial resolution. This study has come with a toolbox 3DCoordX, that allows annotation and analysis of 3D digital organ coordinates. It's a user-interactive open-source toolbox that assigns to each cell 3D positional information relative to the major axes of the organ or tissue under study.

Essentially, the toolbox is an advancement for a spatial resolution to the dataset, the novelty of this toolbox is that it can be applied on simple and also complex-shaped organs. Arabidopsis ovules form a highly curved shape at maturity, 3DCoordX allows tissue-restricted coordinate annotation to these cells positioned at the curved central proximal-distal axis. In the end, a quantitative feature of the organ can be mapped to

the spatial coordinate value of the cell, and a high-end spatial and tissue-specific quantitative analysis can be performed.

With the 3DCoordX coordinates, one can now map multiple specimens or can understand variations within a sample with spatial resolution. 3DCoordX toolbox is a simple approach using tissue annotations to build spatial coordinates that follow the tissue patterns. Coordinates have an origin, the origin used here is a 2D ring. The ring can be as small as a point or can we a wide origin. For ovule primordia, the point like ring is used to form coordinate origin. The approach is to first annotate direct cells from the origin. Further distances to cells within individual cells and direct cells are measured, thus for a selected cell, its distance coordinate is its distance to its nearest direct cell plus the distance coordinate of the direct cell, indirectly providing the distance between the cell to the ring origin. These direct cells are applicable since the measurements are performed in a tissue-restricted manner. Direct cells act as seeds to individual tissues. Essentially, the coordinate values follow tissue restriction. This novel method for coordinate annotation provides accurate coordinates for simple and even complex plant organs. The same principle is used for integuments. Integuments follow a curved structure, its coordinate origin is placed inside the organ where the integuments are initiated. Following the anterior and posterior tissue annotations within the tissue layer, the 3DCoordX will search for the distance between the direct cells and all cells within the individual tissues and provide the added coordinate value between the cell and the coordinate origin. Additionally, the package also allows defining medial-lateral coordinates for such complex organs. The same principle can be applied to other organs like Utricularia, which is a curved leaf, Marchantia, which is an archegonia. Overall, the principle is now widely applicable in other organs and spatial resolution is now not a complicated task for 3D digital organs.

3DCoordX also includes an automated tissue annotation approach for simple layered tissue structure. It takes a segmented image stack and the respective 3D mesh as input and automatically clusters mesh cells into tissue layers. This is done in multiple steps combined to a single click process. Firstly the surface cells are figured out by their outside exposed surface area or if the segmented cell has voxels exposed to air. Once the surface cells are clustered as L1, they are used to define the inner tissue layers in terms of distances from L1. This is user-friendly as it doesn’t use surface mesh to do the clustering. Surface mesh generation is critical and is often time-consuming. This approach is fast and reliable for simple layered organs.

Having the 3DCoordX values for ovules, different findings are now defined in a quantitative manner. Mitotic cells in ovule primordia are now mapped to the coordinates, this allows to define of the spatial pattern of mitotic cells from multiple samples of ovule primordia and come with a general pattern of mitotic cells in early ovules. Funiculus is a curved structure within the ovule, 3DCoordX can be also applied in the funiculus to understand the spatial patterns in the anterior and posterior half of the organs and along the proximal-distal axis. The differential growth of integuments is now defined with spatial resolution. Detailed analysis indicates that there are severe changes in cellular features along the central curved axis of the organ and many of them might have influence in shaping the structure to a curved form at maturity.

Although these patterns are now defined with spatial resolution, the science behind the patterns is still unclear. If the observed pattern is a resultant pattern or the pattern is inherent? Do these patterns really have an influence on the final organ shape or it's a passive observable pattern that doesn't have a major role in shaping the structure? A strong pattern like differential growth of integuments is observed, but what happens when the differential growth of integument tissues is altered? Several of such questions are unanswered in this study due to time constraints and limitations.

4.4 3D Nuclei atlas within the context of cells

This section includes a final refinement in the tools and techniques used in the study to make the best use of them. 3D cell segmentation and nuclei segmentations are available from this study. Tissue annotations and 3DCoordX spatial coordinates are also available. All this information can be used at the same platform to perform a combined analysis of cells and nuclei within the organ. Tools are finally developed for linking cells to nuclei within the 3D meshes in MGX. The result is that one can now quantify cellular patterns and nuclei patterns at the same platform. An example is one can look at the ratio of cell volume and nuclei volumes in different tissues and with spatial resolution. One can understand the position of nuclei within the cell and ask questions like if the shapes of nuclei follow a pattern that is also visible from the cell segmentation? Tissue annotations from cells are now applicable for nuclei, one can extract nuclei from different tissues and perform any global analysis. Also, nuclei from different stages of development and genotypes can be now looked at with spatial and tissue resolution. Overall, the final package for 3D nuclei atlas provides an extended

version of the most possible information to be extracted from the 3D cell segmentation and nuclei segmentation of the microscopic images.

5 Conclusion

This study has taken ovules as a model system to understand deed into 3D cellular patterns shaping an organ. The study initially focused on developing proper tools and techniques for 3D cell outline image analysis. Several advanced methods were developed as part of the study that have valuable helped the plant science community. Starts with advancements in microscopic imaging, machine learning-based PlantSeg pipeline for accurate image segmentation, and 3DCoordX, a toolbox for coordinate annotation in simple and complex plant organs. This study made use of the developed methods to further understand the organ with fine details.

Several features of the organ were identified from this study. A few highlights include firstly the slanting in ovule primordia. Slanting forms the first observable polarity in the organ that is further translated into the anterior-posterior organ axis, posterior initiation of integuments and finally orienting along the pistil axis with the micropyle pointed towards the pistil apex, called the gynapical-gynbasal orientation of the organ at maturity. Integuments are fascinating structure that grows around the organ. This study has identified differential growth of integuments and related events and features of the organ. Exposed cells, distal enlargement of integument cells, and contact friction are the currently observed highlights that can be used to explain the differential growth of the organs and how that translates to the final shape of the organ. Quantitative analysis of growth and proliferation in integument tissues resulted in advanced interpretation of the cellular patterns and features observed.

Finally, 3DcoordX, a spatial coordinate system for annotation and analysis of 3D organ morphogenesis was setup for this complex shaped organ which can be also applied on various other organs. 3DcoordX allowed extensive analysis that includes spatial coordinates to the system which was previously missing. Spatial resolution helped to define integuments proximal-distal growth gradient in a quantitative manner. Identified that mitotic cells appear at proximal 65% of the organ body in early development by mapping mitotic cells to organ coordinates from multiple sample normalization. 3DCoordX is an advanced toolbox that it can be applied on various simple and complicated shaped structures. Overall, this study has extensively defined the cellular organization in wild type Arabidopsis ovule development with fine details that now stands as a benchmark dataset for understanding various features of the organ at

different developmental stages and how that correlates with the tissue and organ shapes at their respective stages of development.

6. References

1. Adams DC, Rohlf FJ, Slice DE. 2004. Geometric morphometrics: Ten years of progress following the ‘revolution.’ Italian Journal of Zoology 71:5–16. doi:10.1080/11250000409356545

2. Adams Dean, Rohlf James, E Slice Dennis. A field comes of age: geometric morphometrics in the 21st century. 2013. Italian Journal of Mammalogy. https://doi.org/10.4404/hystrix-24.1-6283

3. Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M. 1997. Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. Plant Cell 9:841–857. doi:10.1105/tpc.9.6.841

4. Aoyagi Y, Kawakami R, Osanai H, Hibi T, Nemoto T. 2015. A rapid optical clearing protocol using 2,2’-thiodiethanol for microscopic observation of fixed mouse brain. PLoS ONE 10:e0116280. doi:10.1371/journal.pone.0116280

5. Autran D, Baroux C. 2020. Developmental constraints modulate reproductive fate and plasticity within the Arabidopsis ovule primordium. BioRxiv. doi:10.1101/2020.07.30.226670

6. Bailoni A, Pape C, Wolf S, Beier T, Kreshuk A, Hamprecht FA. 2019. A generalized framework for agglomerative clustering of signed graphs applied to instance segmentation. arxiv.org. https://arxiv.org/abs/1906.11713

7. Balasubramanian S, Schneitz K. 2000. NOZZLE regulates proximal-distal pattern formation, cell proliferation and early sporogenesis during ovule development in Arabidopsis thaliana. Development 127:4227–4238.

8. Barbier de Reuille P, Routier-Kierzkowska A-L, Kierzkowski D, Bassel GW, Schüpbach T, Tauriello G, Bajpai N, Strauss S, Weber A, Kiss A, Burian A, Hofhuis H, Sapala A, Lipowczan M, Heimlicher MB, Robinson S, Bayer EM, Basler K, Koumoutsakos P, Roeder AHK, Smith RS. 2015. MorphoGraphX: A platform for quantifying morphogenesis in 4D. eLife 4:05864. doi:10.7554/eLife.05864

9. Bassel GW, Stamm P, Mosca G, Barbier de Reuille P, Gibbs DJ, Winter R, Janka A, Holdsworth MJ, Smith RS. 2014. Mechanical constraints imposed by 3D cellular geometry and arrangement modulate growth patterns in the Arabidopsis embryo. Proc Natl Acad Sci USA 111:8685–8690. doi:10.1073/pnas.1404616111

10. Bencivenga S, Simonini S, Benková E, Colombo L. 2012. The transcription factors BEL1 and SPL are required for cytokinin and auxin signaling during ovule development in Arabidopsis. Plant Cell 24:2886–2897. doi:10.1105/tpc.112.100164

11. Benková E, Michniewicz M, Sauer M, Teichmann T, Seifertová D, Jürgens G, Friml J. 2003. Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115:591–602. doi:10.1016/s0092-8674(03)00924-3

12. 12. Berleth T, Jürgens G. 1993. The role of the monopteros gene in organising the basal body region of the Arabidopsis embryo | Development. Development 118:575–587

13. Bowman JL, Smyth DR, Meyerowitz EM. 1991. Genetic interactions among floral homeotic genes of Arabidopsis. Development 112:1–20. doi:10.1242/dev.112.1.1

14. Breuil-Broyer S, Morel P, de Almeida-Engler J, Coustham V, Negrutiu I, Trehin C. High-resolution boundary analysis during Arabidopsis thaliana flower development. Plant Journal. 2004;38: 182–192

15. Chaudhary A, Gao J, Schneitz K. 2018. The genetic control of ovule developmentReference Module in Life Sciences. Elsevier. doi:10.1016/B978-0-12-809633-8.20737-1

16. Chevalier D, Batoux M, Fulton L, Pfister K, Yadav RK, Schellenberg M, Schneitz K. 2005. STRUBBELIG defines a receptor kinase-mediated signaling pathway regulating organ development in Arabidopsis. Proc Natl Acad Sci USA 102:9074–9079. doi:10.1073/pnas.0503526102

17. Christensen CA, King EJ, Jordan JR, Drews GN. 1997. Megagametogenesis in Arabidopsis wild type and the Gf mutant. Sex Plant Reprod 10:49–64. doi:10.1007/s004970050067

18. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O (2016) 3d u-net: learning dense volumetric segmentation from sparse annotation International Conference on Medical Image Computing and Computer-Assisted Intervention. arXiv:1606.06650

19. Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16:735–743. doi:10.1046/j.1365-313x.1998.00343.x

20. Coen E, Kennaway R, Whitewoods C. 2017. On genes and form. Development 144:4203–4213. doi:10.1242/dev.151910

21. Coen E, Rebocho AB. 2016. Resolving conflicts: modeling genetic control of plant morphogenesis. Dev Cell 38:579–583. doi:10.1016/j.devcel.2016.09.006

22. Coen E, Rolland-Lagan A-G, Matthews M, Bangham JA, Prusinkiewicz P. 2004. The genetics of geometry. Proc Natl Acad Sci USA 101:4728–4735. doi:10.1073/pnas.0306308101

23. Coen ES, Meyerowitz EM. 1991. The war of the whorls: genetic interactions controlling flower development. Nature 353:31–37. doi:10.1038/353031a0

24. Cucinotta M, Colombo L, Roig-Villanova I. 2014. Ovule development, a new model for lateral organ formation. Front Plant Sci 5:117. doi:10.3389/fpls.2014.00117

25. Drews GN, Lee D, Christensen CA. 1998. Genetic analysis of female gametophyte development and function. Plant Cell 10:5–17.

26. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, Liu Y, Topol E, Dean J, Socher R. 2021. Deep learning-enabled medical computer vision. npj Digital Med 4:5. doi:10.1038/s41746-020-00376-2

27. Fernandez R, Das P, Mirabet V, Moscardi E, Traas J, Verdeil J-L, Malandain G, Godin C. 2010. Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. Nat Methods 7:547–553. doi:10.1038/nmeth.1472

28. Galbiati F, Sinha Roy D, Simonini S, Cucinotta M, Ceccato L, Cuesta C, Simaskova M, Benkova E, Kamiuchi Y, Aida M, Weijers D, Simon R, Masiero S, Colombo L. 2013. An integrative model of the control of ovule primordia formation. Plant J 76:446–455. doi:10.1111/tpj.12309

29. Gasser CS, Broadhvest J, Hauser BA. 1998. Genetic analysis of ovule development. Annu Rev Plant Physiol Plant Mol Biol 49:1–24. doi:10.1146/annurev.arplant.49.1.1

30. Gasser CS, Skinner DJ. 2019. Development and evolution of the unique ovules of flowering plants. Curr Top Dev Biol 131:373–399. doi:10.1016/bs.ctdb.2018.10.007

31. Galbiati F, Sinha Roy D, Simonini S, Cucinotta M, Ceccato L, Cuesta C, et al. An integrative model of the control of ovule primordia formation. Plant J. 2013;76: 446–455

32. Gibson WT, Veldhuis JH, Rubinstein B, Cartwright HN, Perrimon N, Brodland GW, Nagpal R, Gibson MC. 2011. Control of the mitotic cleavage plane by local epithelial topology. Cell 144:427–438. doi:10.1016/j.cell.2010.12.035

33. Gonçalves B, Hasson A, Belcram K, Cortizo M, Morin H, Nikovics K, et al. A conserved role for CUP-SHAPED COTYLEDON genes during ovule development. Plant J. 2015;83: 732–742.

34. Grossniklaus U, Schneitz K. 1998. The molecular and genetic basis of ovule and megagametophyte development. Seminars in Cell and Developmental Biology 9:227–238.

35. Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB. 1998. Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. Science 280:446–450. doi:10.1126/science.280.5362.446

36. Gupta A, Harrison PJ, Wieslander H, Pielawski N, Kartasalo K, Partel G, Solorzano L, Suveer A, Klemm AH, Spjuth O, Sintorn I-M, Wählby C. 2019. Deep learning in image cytometry: A review. Cytometry A 95:366–380. doi:10.1002/cyto.a.23701

37. Hafiz AM, Bhat GM. 2020. A survey on instance segmentation: state of the art. Int J Multimed Inf Retr 9:171–189. doi:10.1007/s13735-020-00195-x

38. Hallou A, Yevick HG, Dumitrascu B, Uhlmann V. 2021. Deep learning for bioimage analysis in developmental biology. Development 148. doi:10.1242/dev.199616

39. Hasegawa J, Sakamoto Y, Nakagami S, Aida M, Sawa S, Matsunaga S. 2016. Three-dimensional imaging of plant organs using a simple and rapid transparency technique. Plant Cell Physiol 57:462–472. doi:10.1093/pcp/pcw027

40. Hejnowicz Z. 2014. Trajectories of principal directions of growth, natural coordinate system in growing plant organ. Acta Soc Bot Pol 53:29–42. doi:10.5586/asbp.1984.004

41. Hernandez-Lagana E, Mosca G, Mendocilla Sato E, Pires N, Frey A, Giraldo-Fonseca A, Grossniklaus U, Hamant O, Godin C, Boudaoud A, Grimanelli D,

42. Horňáková A, Lange J-H, Andres B (2017) Analysis and Optimization of Graph Decompositions by Lifted Multicuts Proceedings of the 34th International Conference on Machine Learning. pp. 1539–1548. doi.org/10.48550/arXiv.1503.03791

43. Hoffman DP, Slavitt I, Fitzpatrick CA. 2021. The promise and peril of deep learning in microscopy. Nat Methods 18:131–132. doi:10.1038/s41592-020-01035-w

44. Hong L, Dumond M, Tsugawa S, Sapala A, Routier-Kierzkowska A-L, Zhou Y, Chen C, Kiss A, Zhu M, Hamant O, Smith RS, Komatsuzaki T, Li C-B, Boudaoud A, Roeder AHK. 2016. Variable Cell Growth Yields Reproducible OrganDevelopment through Spatiotemporal Averaging. Dev Cell 38:15–32. doi:10.1016/j.devcel.2016.06.016

45. Hong L, Dumond M, Zhu M, Tsugawa S, Li C-B, Boudaoud A, Hamant O, Roeder AHK. 2018. Heterogeneity and robustness in plant morphogenesis: from cells to organs. Annu Rev Plant Biol 69:469–495. doi:10.1146/annurev-arplant-042817-040517

46. Huang C, Wang Z, Quinn D, Suresh S, Hsia KJ. 2018. Differential growth and shape formation in plant organs. Proc Natl Acad Sci USA 115:12359–12364. doi:10.1073/pnas.1811296115

47. Hülskamp M, Misŕa S, Jürgens G. 1994. Genetic dissection of trichome cell development in Arabidopsis. Cell 76:555–566. doi:10.1016/0092-8674(94)90118-x

48. Ishida T, Aida M, Takada S, Tasaka M. Involvement of CUP-SHAPED COTYLEDON genes in gynoecium and ovule development in Arabidopsis thaliana. Plant Cell Physiol. 2000;41: 60–67.

49. Jenik PD, Irish VF. 2000. Regulation of cell proliferation patterns by homeotic genes during Arabidopsis floral development. Development 127:1267–1276.

50. Jürgens G, Mayer U, Ramon A. TR, Berleth T, Miséra S. 1991. Genetic analysis of pattern formation in the Arabidopsis embryo. Development 113:27–38. doi:10.1242/dev.113.Supplement_1.27

51. Kokkinos I. 2015. Pushing the boundaries of boundary detection using deep learning. arXiv. https://arxiv.org/abs/1511.07386

52. Kierzkowski D, Runions A, Vuolo F, Strauss S, Lymbouridou R, Routier-Kierzkowska A-L, Wilson-Sánchez D, Jenke H, Galinha C, Mosca G, Zhang Z, Canales C, Dello Ioio R, Huijser P, Smith RS, Tsiantis M. 2019. A Growth-Based Framework for Leaf Shape Development and Diversity. Cell 177:1405-1418.e17. doi:10.1016/j.cell.2019.05.011

53. Koncz C, Schell J. 1986. The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel Agrobacterium binary vector. Mol Gen Genet 204:383–396.

55. Kurihara D, Mizuta Y, Sato Y, Higashiyama T. 2015. ClearSee: a rapid optical clearing reagent for whole-plant fluorescence imaging. Development 142:4168–4179. doi:10.1242/dev.127613

56. Kwiatkowska D. 2008. Flowering and apical meristem growth dynamics. J Exp Bot 59:187–201. doi:10.1093/jxb/erm290

57. LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. Nature 521:436–444. doi:10.1038/nature14539

58. Liao C-Y, Smet W, Brunoud G, Yoshida S, Vernoux T, Weijers D. 2015. Reporters for sensitive and quantitative measurement of auxin response. Nat Methods 12:207–10, 2 p following 210. doi:10.1038/nmeth.3279

59. Lei Tao, Wang Risheng, Wan Yong , Zhang Bingtao, Meng Hongying, K Nandi Asoke. 2020. Medical Image Segmentation Using Deep Learning: A Survey. https://arxiv.org/abs/2009.13120. Machine Learning | The MIT Press. n.d. https://mitpress.mit.edu/books/machine-learning-1

60. Mansfield SG, Briarty LG, Erni S. 1991. Early embryogenesis in Arabidopsis thaliana. .The mature embryo sac. Can J Bot 69:447–460.

61. Masson PH, Tasaka M, Morita MT, Guan C, Chen R, Boonsirichai K. 2002. Arabidopsis thaliana: A Model for the Study of Root and Shoot Gravitropism. Arabidopsis Book 1:e0043. doi:10.1199/tab.0043

62. Meijering E. 2020. A bird's-eye view of deep learning in bioimage analysis. Comput Struct Biotechnol J 18:2312–2325. doi:10.1016/j.csbj.2020.08.003

63. Meyerowitz EM. 1997. Genetic control of cell division patterns in developing plants. Cell 88:299–308. doi:10.1016/s0092-8674(00)81868-1

64. Mitteroeckera Philipp, Gunzb Philipp, Windhager Sonja, Schaefer Katrin. 2013. A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. Italian Journal of Mammalogy. https://doi.org/10.4404/hystrix-24.1-6369

65. Modrusan Z, Reiser L, Feldmann KA, Fischer RL, Haughn GW. 1994. Homeotic rransformation of ovules into carpel-like structures in Arabidopsis. Plant Cell 6:333–349. doi:10.1105/tpc.6.3.333

66. Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D. 2019. Deep learning for cellular image analysis. Nat Methods 16:1233–1246. doi:10.1038/s41592-019-0403-1

67. Montenegro-Johnson T, Strauss S, Jackson MDB, Walker L, Smith RS, Bassel GW. 2019. 3DCellAtlas Meristem: a tool for the global cellular annotation of shoot apical meristems. Plant Methods 15:33. doi:10.1186/s13007-019-0413-0

68. Montenegro-Johnson TD, Stamm P, Strauss S, Topham AT, Tsagris M, Wood ATA, Smith RS, Bassel GW. 2015. Digital Single-Cell Analysis of Plant Organ Development Using 3DCellAtlas. Plant Cell 27:1018–1033. doi:10.1105/tpc.15.00175

70. Moreno N, Bougourd S, Haseloff J, Feijó JA. 2006. Imaging Plant Cells In: Pawley JB, editor. Handbook Of Biological Confocal Microscopy. Boston, MA: Springer US. pp. 769–787. doi:10.1007/978-0-387-45524-2_44

71. Musielak TJ, Schenkel L, Kolb M, Henschen A, Bayer M. 2015. A simple and versatile cell wall staining protocol to study plant reproduction. Plant Reprod 28:161–169. doi:10.1007/s00497-015-0267-1

72. Musielak TJ, Slane D, Liebig C, Bayer M. 2016. A versatile optical clearing protocol for deep tissue imaging of fluorescent proteins in Arabidopsis thaliana. PLoS ONE 11:e0161107. doi:10.1371/journal.pone.0161107

73. Murphy, K. P. (2012). Probabilistic Machine Learning: An Introduction . Cambridge, USA: MIT Press.

74. Nakajima K. 2018. Be my baby: patterning toward plant germ cells. Curr Opin Plant Biol 41:110–115. doi:10.1016/j.pbi.2017.11.002

75. Organ geometry channels reproductive cell fate in the Arabidopsis ovule primordium | eLife. n.d. https://elifesciences.org/articles/66031

76. Palmer WM, Martin AP, Flynn JR, Reed SL, White RG, Furbank RT, Grof CPL. 2015. PEA-CLARITY: 3D molecular imaging of whole plant organs. Sci Rep 5:13492. doi:10.1038/srep13492

77. Pape C, Matskevych A, Wolny A, Hennies J, Mizzon G, Louveaux M, Musser J, Maizel A, Arendt D, Kreshuk A. 2019. Leveraging domain knowledge to improve microscopy image segmentation with lifted multicuts. Front Comput Sci 1. doi:10.3389/fcomp.2019.00006

78. Refahi Y, Zardilis A, Michelin G, Wightman R, Leggio B, Legrand J, Faure E, Vachez L, Armezzani A, Risson A-E, Zhao F, Das P, Prunet N, Meyerowitz EM, Godin C, Malandain G, Jönsson H, Traas J. 2021. A multiscale analysis of early flower development in Arabidopsis provides an integrated view of molecular regulation and growth control. Dev Cell 56:540-556.e8. doi:10.1016/j.devcel.2021.01.019

79. Robinson-Beers K, Pruitt RE, Gasser CS. 1992. Ovule development in wild-type Arabidopsis and two female-sterile mutants. Plant Cell 4:1237–1249. doi:10.1105/tpc.4.10.1237

80. Roeder AHK, Tarr PT, Tobin C, Zhang X, Chickarmane V, Cunha A, Meyerowitz EM. 2011. Computational morphodynamics of plants: integrating development over space and time. Nat Rev Mol Cell Biol 12:265–273. doi:10.1038/nrm3079

81. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241

82. Sapala A, Runions A, Routier-Kierzkowska A-L, Das Gupta M, Hong L, Hofhuis H, Verger S, Mosca G, Li C-B, Hay A, Hamant O, Roeder AH, Tsiantis M, Prusinkiewicz P, Smith RS. 2018. Why plants make puzzle cells, and how their shape emerges. eLife 7. doi:10.7554/eLife.32794

83. Schiefthaler U, Balasubramanian S, Sieber P, Chevalier D, Wisman E, Schneitz K. 1999. Molecular analysis of NOZZLE, a gene involved in pattern formation and early sporogenesis during sex organ development in Arabidopsis thaliana. Proc Natl Acad Sci USA 96:11664–11669. doi:10.1073/pnas.96.20.11664

84. Schmidt A, Schmid MW, Grossniklaus U. 2015. Plant germline formation: common concepts and developmental flexibility in sexual and asexual reproduction. Development 142:229–241. doi:10.1242/dev.102103

85. Schmidt T, Pasternak T, Liu K, Blein T, Aubry-Hivet D, Dovzhenko A, Duerr J, Teale W, Ditengou FA, Burkhardt H, Ronneberger O, Palme K. 2014. The iRoCS Toolbox--3D analysis of the plant root apical meristem at cellular resolution. Plant J 77:806–814. doi:10.1111/tpj.12429

86. Schmidt U, Weigert M, Broaddus C, Myers G. 2018. Cell Detection with Star-Convex Polygons In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II, Lecture Notes in Computer Science. Cham: Springer International Publishing. pp. 265–273. doi:10.1007/978-3-030-00934-2_30

87. Schneitz K, Baker SC, Gasser CS, Redweik A. 1998. Pattern formation and growth during floral organogenesis: HUELLENLOS and AINTEGUMENTA are required for the formation of the proximal region of the ovule primordium in Arabidopsis thaliana. Development 125:2555–2563.

88. Schneitz K, Hülskamp M, Kopczak SD, Pruitt RE. 1997. Dissection of sexual organ ontogenesis: a genetic analysis of ovule development in Arabidopsis thaliana. Development 124:1367–1376.

89. Schneitz K, Hülskamp M, Pruitt RE. 1995. Wild-type ovule development in Arabidopsis thaliana: a light microscope study of cleared whole-mount tissue. Plant J 7:731–749. doi:10.1046/j.1365-313X.1995.07050731.x

90. Shelhamer E, Long J, Darrell T. 2017. Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell 39:640–651. doi:10.1109/TPAMI.2016.2572683

91. Sieber P, Wellmer F, Gheyselinck J, Riechmann JL, Meyerowitz EM. Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. Development. 2007;134: 1051–1060.

92. Silveira SR, Le Gloanec C, Gómez-Felipe A, Routier-Kierzkowska A-L, Kierzkowski D. 2021. Live-imaging provides an atlas of cellular growth dynamics in the stamen. Plant Physiol. doi:10.1093/plphys/kiab363

93. Slice DE. 2007. Geometric morphometrics. Annu Rev Anthropol 36:261–281. doi:10.1146/annurev.anthro.34.081804.120613

94. Stegmaier J, Amat F, Lemon WC, McDole K, Wan Y, Teodoro G, Mikut R, Keller PJ. 2016. Real-time three-dimensional cell segmentation in large-scale microscopy data of developing embryos. Dev Cell 36:225–240. doi:10.1016/j.devcel.2015.12.028

95. Strauss S, Sapala A, Kierzkowski D, Smith RS. 2019. Quantifying Plant Growth and Cell Proliferation with MorphoGraphX. Methods Mol Biol 1992:269–290. doi:10.1007/978-1-4939-9469-4_18

96. Soeren Strauss, Adam Runions, Brendan Lane, Dennis Eschweiler, Namrata Bajpai, Nicola Trozzi, Anne-Lise Routier-Kierzkowska, Saiko Yoshida, Sylvia Rodrigues da Silveira, Athul Vijayan, Rachele Tofanelli, Mateusz Majda, Emillie Echevin, Constance Le Gloanec, Hana Bertrand-Rakusova, Milad Adibi, Kay Schneitz, George Bassel, Daniel Kierzkowski, Johannes Stegmaier, Miltos Tsiantis, Richard S. Smith. (2022) Using positional information to provide context for biological image analysis with MorphoGraphX 2.0. eLife. doi.org/10.7554/eLife.72601

97. Stringer C, Michaelos M, Pachitariu M. 2020. Cellpose: a generalist algorithm for cellular segmentation. BioRxiv. doi:10.1101/2020.02.02.931238

98. Sui L, Alt S, Weigert M, Dye N, Eaton S, Jug F, Myers EW, Jülicher F, Salbreux G, Dahmann C. 2018. Differential lateral and basal tension drive folding of Drosophila wing discs through two distinct mechanisms. Nat Commun 9:4620. doi:10.1038/s41467-018-06497-3

99. Takada S, Hibara K, Ishida T, Tasaka M. The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. Development. 2001;128: 1127–1135.

100. Telfer A, Bollman KM, Poethig RS. 1997. Phase change and the regulation of trichome distribution in Arabidopsis thaliana. Development 124:645–654.

101. Tofanelli R, Vijayan A, Scholz S, Schneitz K. 2019. Protocol for rapid clearing and staining of fixed Arabidopsis ovules for improved imaging by confocal laser scanning microscopy. Plant Methods 15:120. doi:10.1186/s13007-019-0505-x

102. Troadec M-B, Pagès J-C, the Scientific Committee of the High Council for biotechnology, France. 2019. Where are we with unintended effects in genome editing applications from DNA to phenotype: focus on plant applications. Transgenic Res 28:125–133. doi:10.1007/s11248-019-00146-1

103. Truernit E, Bauby H, Dubreucq B, Grandjean O, Runions J, Barthélémy J, Palauqui J-C. 2008. High-resolution whole-mount imaging of three-dimensional tissue organization and gene expression enables the study of Phloem development and structure in Arabidopsis. Plant Cell 20:1494–1503. doi:10.1105/tpc.107.056069

104. Truernit E, Haseloff J. 2008. Arabidopsis thaliana outer ovule integument morphogenesis: ectopic expression of KNAT1 reveals a compensation mechanism. BMC Plant Biol 8:35. doi:10.1186/1471-2229-8-35

105. Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers.

106. Cell Detection with Star-convex Polygons. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Granada, Spain, September 2018

107. Vaddepalli P, Fulton L, Batoux M, Yadav RK, Schneitz K. 2011. Structure function analysis of STRUBBELIG, an Arabidopsis atypical receptor-like kinase1 involved in tissue morphogenesis. PLoS ONE 6:e19730. doi:10.1371 journal.pone.0019730

108. Vicar T, Balvan J, Jaros J, Jug F, Kolar R, Masarik M, Gumulec J. 2019. Cell segmentation methods for label-free contrast microscopy: review and comprehensive comparison. BMC Bioinformatics 20:360. doi:10.1186/s12859-19-2880-8

109. Vijayan A, Tofanelli R, Strauss S, Cerrone L, Wolny A, Strohmeier J, Kreshuk A, Hamprecht FA, Smith RS, Schneitz K. 2021. A digital 3D reference atlas reveals cellular growth patterns shaping the Arabidopsis ovule. eLife 10. doi:10.7554/eLife.63262

110. Vijayan, A., Strauss, S., Tofanelli, R., Mody, T.A., Lee, K., Tsiantis, M., Smith, R.S., Schneitz, K., 2022. The annotation and analysis of complex 3D plant organs using 3DCoordX. Plant Physiology. doi.org/10.1093/plphys/kiac145

111. Villanueva JM, Broadhvest J, Hauser BA, Meister RJ, Schneitz K, Gasser CS. 1999. INNER NO OUTER regulates abaxial- adaxial patterning in Arabidopsis ovules. Genes Dev 13:3160–3169.

112. Villoutreix P. 2021. What machine learning can do for developmental biology. Development 148. doi:10.1242/dev.188474

113. Wang S, Yang DM, Rong R, Zhan X, Xiao G. 2019. Pathology image analysis using segmentation deep learning algorithms. Am J Pathol 189:1686–1698. doi:10.1016/j.ajpath.2019.05.007

114. Weigert M, Schmidt U, Boothe T, Müller A, Dibrov A, Jain A, Wilhelm B, Schmidt D, Broaddus C, Culley S, Rocha-Martins M, Segovia-Miranda F, Norden C, Henriques R, Zerial M, Solimena M, Rink J, Tomancak P, Royer L, Jug F, Myers EW. 2018. Content-aware image restoration: pushing the limits of fluorescence microscopy. Nat Methods 15:1090–1097. doi:10.1038/s41592-018-0216-7

115. Weigert M, Schmidt U, Haase R, Sugawara K, Myers G. 2020. Star-convex polyhedra for 3D object detection and segmentation in microscopy2020 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE. pp. 3655–3662. doi:10.1109/WACV45572.2020.9093435

116. Willis L, Refahi Y, Wightman R, Landrein B, Teles J, Huang KC, Meyerowitz EM, Jönsson H. 2016. Cell size and growth regulation in the Arabidopsis thaliana apical stem cell niche. Proc Natl Acad Sci USA 113:E8238–E8246. doi:10.1073/pnas.1616768113

117. Wolny, A., Cerrone, L., Vijayan, A., Tofanelli, R., Barro, A.V., Louveaux, M., Wenzl, C., Strauss, S., Wilson-Sánchez, D., Lymbouridou, R., Steigleder, S.S., Pape, C., Bailoni, A., Duran-Nebreda, S., Bassel, G.W., Lohmann, J.U., Tsiantis, M., Hamprecht, F.A., Schneitz, K., Maizel, A., Kreshuk, A., 2020. Accurate and versatile 3D segmentation of plant tissues at cellular resolution. eLife 9. doi:10.7554/eLife.57613

118. Xie S, Tu Z (2015) Holistically-nested edge detection. arXiv.http://arxiv.org/abs/1504.06375

119. Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, Aylward S, Metaxas D, Whitaker R. 2002. Engineering and algorithm design for an image processing Api: a technical report on ITK--the Insight Toolkit. Stud Health Technol Inform 85:586–592.

120. Yoshida S, Barbier de Reuille P, Lane B, Bassel GW, Prusinkiewicz P, Smith RS, Weijers D. 2014. Genetic control of plant development by overriding a geometric division rule. Dev Cell 29:75–87. doi:10.1016/j.devcel.2014.02.002

7 Supplementary Data

7.1 Supplementary table 1- Arabidopsis ovule developmental stages

(Table adapted from Schneitz et al., 1995)

Ovule stages	Ovule development		Flower development	
	Embryo sac and endosperm development	Sporophytic tissue development	Gynoecium development	Floral landmarks of corresponding floral stages
1. Early phase of ovule development				
1-I		Protrusions arise	Open cylinder	Locules appear in long stamens (floral stage 8)
1-II		Protrusions elongated	Cylinder constricted at apex	Petal primordia stalked at base (floral stage 9)
2. Megasporogenesis				
2-I	Megaspore mother cell enlarges		Cylinder closed	Petals level with short stamens (floral stage 10)
2-II		Inner integument initiates		
2-III		Outer integument initiates, chalazal nucellus divides		
2-IV	Meiosis	Chalazal nucellus enlarges	First immature papillar cells which do not cover all of the stigma yet, lateral vasculature visible as lighter strip	Filaments start to elongate, anthers green (floral stage 11)

2-V	Tetrad formation	Integuments extend toward the apex of the nucellus		
3. Megametogenesis				
3-I	Degenerating tetrad with mono-nuclear embryo sac	Outer integument envelops the nucellus and the inner integument, funiculus and the nucellus curve		
3-II	Two-nuclear embryo sac	Outer integument surrounds nucellus, micropylar end pointing more than 90° away from funiculus, further differential growth of integuments	Papillar cells small and covering all the stigma, style recognizable, valves almost visible as distinct structures	Petals level with long stamens, anthers turn yellow (flower stage 12)
3-III	Vacuole appears	Micropylar end points away about 90° from funiculus	Papillar cells grow longer	
3-IV	Four-nuclear embryo sac	Inner integument surrounds nucellus, endothelium differentiates	Valves more pronounced	
3-V	Eight-nuclear embryo sac, cellularization			
3-IV	Central cell nuclei fuse, antipodal cells degenerate	Additional cell layer initiated in inner integument	Well-extended papillar cells, prominent style and valves	
4. Postfertilization development				

7.2 Supplementary table 2 - Cells type labels and abbreviations of tissues.

Cell type	Morphological definition of cell type	Standardized cell type label
Radial organization of early-stage ovule tissue		
L1	Outermost cell layer (epidermis).	L1
L2	First sub-epidermal cell layer.	L2
L3	Innermost cell layer.	L3
MMC	Single large L2 cell in distal end of nucellus, eventually undergoing meiosis, volume ≥ 350 µm^3.	mmc
Tissues organization of late-stage ovule		
Nucellus	Distal region, harbors the mmc, proximally delineated by adaxial inner integument .	nu
Chalaza	Central region, flanked by the two integuments, distal end marked by adaxial inner integument, proximal end marked by abaxial outer integument, does not include epidermis.	ch
Anterior chalaza	Group chalazal cells, positioned at the anterior side, underlying the epidermal cells forming the frontal base of the hood-like structure generated by the epidermis-derived outer integument.	ac
Posterior chalaza	Group of proximal chalazal cells, positioned opposite to the cells of the anterior chalaza.	pc

Funiculus	Proximal region, stalk-like structure, carrying the vascular strand, distal end marked by chalaza, proximal end marked by placenta.	fu
Embryo sac	Haploid female gametophyte encompasses all stages up to but not including fertilization.	es
Integument tissues		
Abaxial outer integument	Outermost single cell layer of outer integument, entirely made of epidermal cells, proximal end marked by chalaza.	oi2
Adaxial outer integument	Innermost (dorsal) single cell layer of outer integument, proximal end marked by chalaza.	oi1
Abaxial inner integument	Outer (ventral) single cell layer of inner integument, entirely made of epidermal cells, proximal end marked by chalaza.	ii2
Adaxial inner integument	Inner (dorsal) single cell layer of inner integument, entirely made of epidermal cells, proximal end marked by chalaza.	ii1
Parenchymatic inner integument	Cell layer derived from adaxial inner integument, proximal end marked by chalaza.	ii1’
Integument coordinate system annotation		
Anterior	Anterior region of the integument layer.	Ant
Posterior	Posterior region of the integument layer.	Post
Proximal	Proximal half of the integument layer. Relative PD distance values form 0 to 50.	P
Distal	Distal half of the integument layer. Relative PD distance values form 50 to 100.	D

Negative oi2	Region of the outer integument placed behind the Bezier ring and characterized by relative PD distance negative values.	Neg oi2
Medio-lateral oi2	Medio-lateral region of the abaxial outer integument.	MED-LAT
Frontal oi2	Frontal region of the abaxial outer integument.	FRONT
Cell cycle		
M-phase	Mitotic figures (metaphase, anaphase).	M
Interphase	Cells in G0, S, or G2 phase.	I

8 Appendix

8.1 Broad applicability of 3DCoordX spatial coordinates in other organs

The approach for an organ-intrinsic coordinate system, 3DCoordX was useful beyond the Arabidopsis ovule and could be of value to provide spatial context to cellular growth patterns in different organs of various plant species. Several 3D digital plants can be analysed using this method. As an example, the cup-shaped trap of the aquatic carnivorous plant Utricularia gibba is used here. The traps represent a highly curved 3D leaf form (Płachno et al. 2015; Reifenrath et al. 2006). Quantitative growth analysis at the tissue level combined with computer modeling indicated that the complex shape transformations occurring during trap development are associated with differential rates and orientations of growth (Lee et al. 2019; Whitewoods et al. 2020). However, a quantitative analysis of cellular parameters has not yet been performed. To obtain first insight into the cellular basis of the growth patterns shaping the Utricularia trap a 3D digital representation with cellular resolution of an intermediate-stage trap collected 6 days after initiation was generated from the available dataset (Lee et al. 2019). By this stage an invagination in the near-spherical young trap had occurred, followed by the formation of further folds and tissue broadening, and resulting in the emergence of the interior trap door and threshold (Fig. 8.1 A-C). Various tissues were manually labelled, including the the abaxial and adaxial cells of the wall, the threshold, and the trap door, and distinguished between medial and lateral domains of the adaxial and abaxial wall, respectively. To define an origin of the distance coordinate system, an ellipsoid Bezier ring was placed at the boundary between the base of the threshold and the wall of the trap (Fig. 8.1 D). A basic spatial analysis of the patterns of cell volumes were performed (Fig. 8.1 E-G). Results indicate that cell volume varied along the measured distances. For example, a sudden increase in cell volume was observed in an interval from 240 µm to 320 µm for cells of the abaxial wall (Fig. 8.1 F). This region precedes a prominent kink in the abaxial wall (Fig. 8.1 E). By contrast, cell volumes of the adaxial wall dropped towards the end. The decrease in cell volume was likely associated with the tapering of the adaxial wall that could be observed in this area. Volumes of threshold cells positioned within a range of 80 to about 150 µm from the origin showed relatively small volumes in comparison to the cells flanking this interval (Fig. 8.1 G). The 80-150 µm zone corresponded to a section of the threshold which was only moderately curved and provided a large surface exposed to the lumen of the trap (Fig. 8.1 E). Taken together, the data revealed spatial differences in cell

volume for all three examined tissues of this specimen and support the notion that differential cell growth contributes to the morphogenesis of the Utricularia trap.

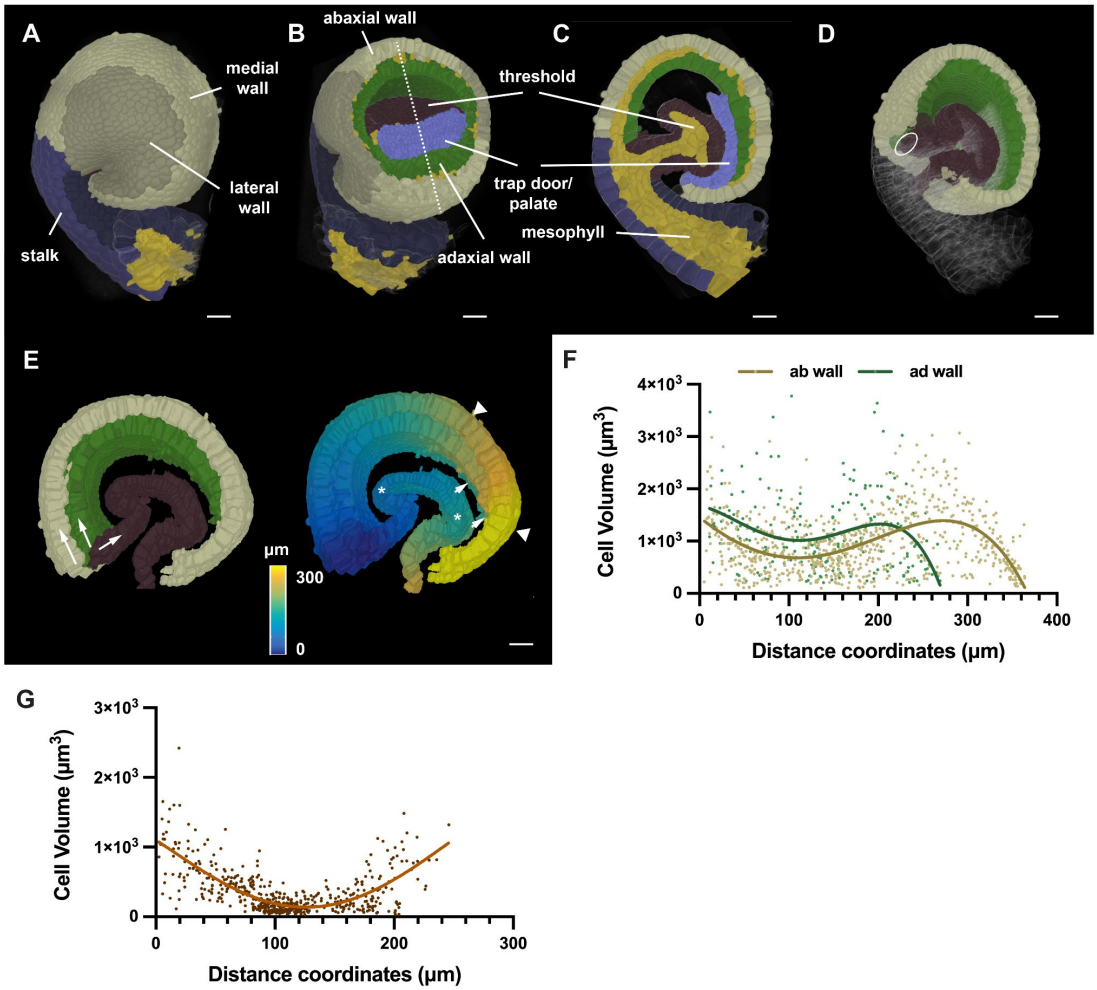

Figure 8.1. 3D digital Utricularia trap. A specimen 6 days after initiation is shown. (A) Side view of the 3D cell mesh with annotation of various tissues. (B) Tilted view of (A) with part of the wall removed by a tangential clipping plane. The dashed line indicates the mid-sagittal section shown in (C). (C) Mid-sagittal section. (D) Slanted 3D view of (A) with half of the trap cut off at the mid-sagittal plane shown in (B). The trap door was removed. The position of the Bezier ellipsoid is indicated. (E) Left panel: arrows indicate the direction of the distance coordinates through the epidermis of the abaxial and adaxial tissue of the wall and the threshold, respectively. Right panel: Heat map indicating distances. Wall and threshold are treated separately. Triangles mark the 240-320 µm interval of the abaxial wall. Arrows highlight the tapering end of the adaxial wall. Asterisks indicate the 80-150 µm intervall of the threshold. (F) Graph displaying cell volume of epidermal trap cells in relation to their position. Values for the abaxial and adaxial wall are superimposed. The respective nonlinear regression curves with fourth order polynomial fitting are indicated. ab wall: n = 786, ad wall: n = 231. (G). Graph displaying cell volume of threshold cells in relation to their position. The line marks a nonlinear regression curve. n = 533. Scale bars: 20 µm.

9 Acknowledgements

I would like to sincerely thank Prof. Dr. Kay Schneitz for providing me the opportunity to work on this fascinating project. Furthermore, for his excellent supervision, support and motivation in many ways, as well as the input through scientific discussions and suggestions throughout the project. Kay has given me the opportunity to work independently on an exciting topic and lay the foundation for my career. He has been instrumental in improving my scientific presentation and writing skills. Many thanks!

Furthermore, I would like to thank Prof. Dr. Frank Johannes and Prof. Dr. Alexis Maizel for taking over the thesis advisory committee. I would like to thank the entire Schneitz group and former members for all their help, scientific and technical support. Special thanks to Rachele for kickstarting and co-leading the project along with me. Thanks to Tejasvinee for taking over the project responsibly. I would also like to thank the CALM facility and Prof. Ramon Torres for excellent introduction and support to microscopes. I would also like to thank all members of the FOR2581 plant morphodynamics forschergruppe for their scientific collaboration and support. Special thanks to Soeren, Lorenzo, Adrian, Prof. Fred A Hamprecht, Prof. Alexis Maizel and Dr. Anna Kreshuk. I would like to also thank my mentor Prof. Caroline Gutjahr for all advice and support. Special thanks go to Susana, our administrative secretary for helping me with the German language related issues and bureaucracy. I thank Rutuja Lele for proofreading the study and Lukas Baumgarten for helping with German translation of some of the thesis materials. I would like to thank the bachelor and masters' students who worked with me. I would like to thank all the helping hands for the sources of materials I have used for the study that includes Arabidopsis seeds and reporter constructs. Finally, I would like to thank my family and friends for their support and constant encouragement through all these years.