Tesis Doctoral

Evolución y desarrollo del sistema apertural del polen en eudicotiledóneas basales: análisis del gen *INAPERTURATE POLLEN1* e identificación de genes candidatos determinantes de la morfología apertural

Ph.D. Thesis

Evolution and development of the pollen apertural system in basal eudicots: analysis of the *INAPERTURATE POLLEN1* gene and identification of candidate genes determining apertural morphology

UNIVERSIDAD
DE GRANADA

Ismael Mazuecos Aguilera

Granada, Diciembre 2021

Programa de Doctorado Biología Fundamental y de Sistemas

Editor: Universidad de Granada. Tesis Doctorales
Autor: Ismael Mazuecos Aguilera
ISBN: 978-84-1117-205-9
URI: http://hdl.handle.net/10481/72364

Evolución y desarrollo del sistema apertural del polen en eudicotiledóneas basales: análisis del gen *INAPERTURATE POLLEN1* e identificación de genes candidatos determinantes de la morfología apertural

Evolution and development of the pollen apertural system in basal eudicots: analysis of the *INAPERTURATE POLLEN1* gene and identification of candidate genes determining apertural morphology

Memoria presentada por el graduado Ismael Mazuecos Aguilera para optar al Grado de Doctor con Mención Internacional en Ciencias Biológicas por la Universidad de Granada. Esta memoria ha sido realizada bajo la dirección de Dr. Víctor N. Suárez Santiago, Profesor Titular de Botánica de la Universidad de Granada.

Fdo: **Víctor N. Suárez Santiago**
Director de tesis

Fdo: **Ismael Mazuecos Aguilera**
Doctorando

El doctorando / *The doctoral candidate* Ismael Mazuecos Aguilera y el director de la tesis / *and the thesis supervisor* Víctor Nazario Suárez Santiago

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

/

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor/s and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / *Place and date:*

Granada, 10 de Noviembre de 2021

Víctor Nazario Suárez Santiago	**Ismael Mazuecos Aguilera**
Director de la Tesis / *Thesis supervisor;*	Doctorando / *Doctoral candidate*:
Firma / Signed	Firma / Signed

Contents

Resumen general 10

Summary 13

Introduction 17

- Pollen grain 17
- Pollen wall structure 18
- Exine ontogeny 20
- Pollen grain aperture 21
- Pollen aperture evolution 22
- Pollen aperture formation 23
- Molecular player of aperture formation 24
- Papaveraceae family as a model system to study aperture pattern formation 30
- References 34

Objectives 39

Chapter 1 40

- Abstract 42
- Introduction 43
- Materials and Methods 45
- Results 51
- Discussion 60
- References 67
- Supporting Information 72

Chapter 2 86

- Abstract 87
- Introduction 88

Material and methods 90

Results 94

Discussion 98

References 106

Supporting Information 111

General discussion 118

Conclusions 125

Conclusiones 127

Resumen general | Summary

Resumen general

Las aperturas del grano de polen son áreas donde la pared se encuentra debilitada o ausente y juegan un papel principal en su supervivencia y éxito reproductivo. Una de las grandes incógnitas en el campo de la evolución y desarrollo es el determinismo genético de la enorme variedad de sistemas aperturales (forma, número y posición de las aperturas) existente en angiospermas (Furness and Rudall, 2004; Zhou and Dobritsa, 2019).

En los últimos años ha habido avances importantes con el descubrimiento de las proteínas INAPERTURATE POLLEN 1 (INP1), indispensable para la formación de las aperturas en *Arabidopsis thaliana*, *Zea mays* y *Oryza sativa* (Dobritsa and Coerper, 2012; Li *et al*., 2018; Zhang *et al*., 2020); D6 PROTEIN KINASE LIKE 3 (D6PKL3) implicada en la formación y el determinismo del número de aperturas en *Arabidopsis* (Lee *et al*., 2018); y miembros de la familia ELMOD implicados en el determinismo del número de aperturas (ELMOD_A y B) y potencialmente en la forma de las aperturas (ELMOD_E) (Zhou *et al*., 2021). Sin embargo se desconoce si *INP1*, a pesar de la diversidad de su secuencia, conserva su función en otros grupos taxonómicos y apenas se conocen los procesos implicados en la formación de los diferentes sistemas aperturales. Estudios sobre el determinismo genético de las aperturas en la familia Papaveraceae, por su gran variedad de sistemas aperturales y por su posición filogenética, uno de los linajes más basales de Eudicotiledóneas (Wang *et al*., 2009), ayudarían a la comprensión de estas incógnitas.

Esta tesis doctoral se desarrolla con el objetivo de estudiar la implicación de *INP1* en la formación de aperturas en *Eschscholzia californica*, miembro de Papaveraceae. Por otro lado investigamos genes que posiblemente intervengan en la formación de las aperturas y en el determinismo de su forma.

En el Capítulo 1 llevamos a cabo un estudio funcional de *EcINP1* en *Eschscholzia californica* y proponemos genes candidatos a estar implicados en la formación de las aperturas. Analizamos la expresión temporal y espacial de *EcINP1*, validamos su función generando mutantes mediante silenciamiento génico inducido por virus (VIGS) y comparamos la tasa de germinación y el transcriptoma de plantas salvajes con el de mutantes *inp1*. *EcINP1* presenta su máxima expresión en anteras en estadio de desarrollo del polen de tétradas y su papel como un factor esencial para la

formación de las aperturas es conservado. En *Eschscholzia californica* las aperturas son prescindibles para la germinación del polen. Además, encontramos 971 genes diferencialmente expresados (DEGs) entre plantas salvajes y *EcINP1*-mutantes, entre los que destacamos los que son potenciales candidatos a intervenir en el proceso de formación de las aperturas, como los homólogos en *E. californica* de *NUCLEOSOME ASSEMBLY PROTEIN 1* (*NAP1*), *D6 PROTEIN KINASE LIKE 3* (*D6PKL3*) y *PROTEIN KINASE ASSOCIATED WITH BRX* (*PAX*). Estos genes mostraron patrones de expresión coincidentes con los de *EcINP1*, que apoyan su implicación en la formación de las aperturas y/o posible interacción de sus productos con la proteína EcINP1.

En el Capítulo 2 identificamos genes que podrían estar implicados en el determinismo de la forma de las aperturas. Secuenciamos el transcriptoma de cuatro especies de Papaveraceae, de las que tres no está disponible su genoma de referencia, por lo que generamos un ensamblaje de referencia *de novo* para cada una de ellas. Filtramos y anotamos los ensamblajes *de novo* generados. Comparamos el transcriptoma de dos especies con polen colpado con el de otras dos con polen porado y analizamos los DEGs entre ambos grupos. Encontramos 531 DEGs entre los cuales no se encontraba el gen *ELMOD_E*, el único descrito hasta el momento como un posible factor determinante para la forma de las aperturas (Zhou *et al.,* 2021), por tanto parece que *ELMOD_E* no regula el cambio de colpo a poro en Papaveraceae. En cambio *INP1* si se encontraba diferencialmente expresado. Por otro lado, entre los DEGs encontramos genes implicados en procesos que posiblemente intervienen en la formación de las aperturas, como los relacionados con la síntesis o degradación de la calosa, destacando el factor de transcripción *DYSFUNCTIONAL TAPETUM 1* (*DYT1*) o con la organización de elementos del citoesqueleto.

Referencias

Dobritsa AA, Coerper D. 2012. The novel plant protein INAPERTURATE POLLEN1 marks distinct cellular domains and controls formation of apertures in the Arabidopsis pollen exine. *The Plant Cell* 24: 4452-4464.

Furness CA, Rudall PJ. 2004. Pollen aperture evolution–a crucial factor for eudicot success?. *Trends in plant science* **9**: 154-158.

Lee BH, Weber ZT, Zourelidou M, *et al.* 2018. Arabidopsis protein kinase D6PKL3 is involved in the formation of distinct plasma membrane aperture domains on the pollen surface. *The Plant Cell* **30**: 2038-2056.

Li P, Ben-Menni Schuler S, Reeder SH, Wang R, Suárez Santiago VN, Dobritsa AA. 2018. INP1 involvement in pollen aperture formation is evolutionarily conserved and may require species-specific partners. Journal of Experimental Botany 69: 983-996.

Wang W, Lu AM, Ren Y, Endress ME, Chen ZD. 2009. Phylogeny and classification of Ranunculales: evidence from four molecular loci and morphological data. *Perspectives in Plant Ecology, Evolution and Systematics* **11**: 81-110.

Zhang X, Zhao G, Tan Q. 2020. Rice pollen aperture formation is regulated by the interplay between OsINP1 and OsDAF1. *Nature Plants* **6:** 394-403.

Zhou Y, Dobritsa AA. 2019. Formation of aperture sites on the pollen surface as a model for development of distinct cellular domains. *Plant Science* **288**: 110222.

Zhou Y, Amom P, Reeder SH, Lee BH, Helton A, Dobritsa AA. 2021. Members of the ELMOD protein family specify formation of distinct aperture domains on the Arabidopsis pollen surface. *Elife* **10**: e71061.

Summary

The apertures of the pollen grain are areas where the pollen wall is weakened or absent and play a major role in their survival and reproductive success. One of the great mysteries in the field of evolution and development is the genetic determinism of the enormous variety of aperture systems (shape, number and position of apertures) in angiosperms (Furness and Rudall, 2004; Zhou and Dobritsa, 2019).

In recent years there have been important advances with the discovery of the INAPERTURATE POLLEN 1 (INP1) protein, essential for aperture formation in *Arabidopsis thaliana*, *Zea mays* and *Oryza sativa* (Dobritsa and Coerper, 2012; Li *et al.*, 2018; Zhang *et al.*, 2020); D6 PROTEIN KINASE LIKE 3 (D6PKL3) involved in the formation and determinism of aperture number in *Arabidopsis* (Lee *et al.*, 2018); and ELMOD family members involved in aperture number determinism (ELMOD_A and B) and potentially aperture shape (ELMOD_E) (Zhou *et al.*, 2021). However, whether *INP1*, despite its sequence diversity, conserves its function in other taxonomic groups is unknown, also the processes involved in the formation of the different apertural systems are poorly understood. Studies on the genetic determinism of apertures in the family Papaveraceae, due to its great variety of aperture systems and its phylogenetic position, one of the most basal lineages of eudicots (Wang *et al.*, 2009), would help to understand these questions.

This doctoral thesis is developed with the aim of studying the involvement of *INP1* in the formation of apertures in *Eschscholzia californica*, a member of Papaveraceae. On the other hand, we investigate genes that are possibly involved in the formation of the apertures and in the determinism of their shape.

In Chapter 1 we perform a functional study of *EcINP1* in *Eschscholzia californica* and propose candidate genes to be involved in the formation of the apertures. We analysed the temporal and spatial expression of *EcINP1*, validated its function by generating mutants through virus-induced gene silencing (VIGS) and compared the germination rate and transcriptome of wild-type plants with that of *inp1* mutants. *EcINP1* is maximally expressed in anthers at the tetrad stage of pollen development and its role as an essential factor for aperture formation is conserved. In *Eschscholzia californica* the apertures are dispensable for pollen germination. In addition, we found 971 differentially expressed genes (DEGs) between wild-type and

EcINP1-mutants, among which we highlight those that are potential candidates for involvement in the aperture formation process, such as the *E. californica* homologues of *NUCLEOSOME ASSEMBLY PROTEIN 1* (*NAP1*), *D6 PROTEIN KINASE LIKE 3* (*D6PKL3*) and *PROTEIN KINASE ASSOCIATED WITH BRX* (*PAX*). These genes showed expression patterns coincident with those of *EcINP1*, supporting their involvement in the formation of the apertures and/or possible interaction of their products with the EcINP1 protein.

In Chapter 2 we identified genes that could be involved in aperture shape determinism. We sequenced the transcriptome of four species of Papaveraceae, three of which do not have their reference genome available, so we generated a de novo reference assembly for each of them. We filter and annotate the generated de novo assemblies. We compared the transcriptome of two species with colpate pollen with that of two others with porate pollen and analysed the differentially expressed genes between the two groups. We found 531 DEGs among which *ELMOD_E*, the only gene described so far as a possible determinant of aperture shape (Zhou *et al.,* 2021), was not found, so *ELMOD_E* may not be regulating the change from colpo to pore in Papaveraceae. In contrast, *INP1* was found to be differentially expressed. On the other hand, among the DEGs we found genes implicated in processes possibly involved in aperture formation, such as genes related to the synthesis or degradation of callose, including the transcription factor *DYSFUNCTIONAL TAPETUM 1* (*DYT1*) or with the organisation of cytoskeletal elements.

References

Dobritsa AA, Coerper D. 2012. The novel plant protein INAPERTURATE POLLEN1 marks distinct cellular domains and controls formation of apertures in the Arabidopsis pollen exine. *The Plant Cell* 24: 4452-4464.

Furness CA, Rudall PJ. 2004. Pollen aperture evolution–a crucial factor for eudicot success?. *Trends in plant science* **9**: 154-158.

Lee BH, Weber ZT, Zourelidou M, *et al.* 2018. Arabidopsis protein kinase D6PKL3 is involved in the formation of distinct plasma membrane aperture domains on the pollen surface. *The Plant Cell* **30**: 2038-2056.

Li P, Ben-Menni Schuler S, Reeder SH, Wang R, Suárez Santiago VN, Dobritsa AA. 2018. INP1 involvement in pollen aperture formation is evolutionarily conserved and may require species-specific partners. Journal of Experimental Botany 69: 983-996.

Wang W, Lu AM, Ren Y, Endress ME, Chen ZD. 2009. Phylogeny and classification of Ranunculales: evidence from four molecular loci and morphological data. *Perspectives in Plant Ecology, Evolution and Systematics* **11**: 81-110.

Zhang X, Zhao G, Tan Q. 2020. Rice pollen aperture formation is regulated by the interplay between OsINP1 and OsDAF1. *Nature Plants* **6:** 394-403.

Zhou Y, Dobritsa AA. 2019. Formation of aperture sites on the pollen surface as a model for development of distinct cellular domains. *Plant Science* **288**: 110222.

Zhou Y, Amom P, Reeder SH, Lee BH, Helton A, Dobritsa AA. 2021. Members of the ELMOD protein family specify formation of distinct aperture domains on the Arabidopsis pollen surface. Elife 10: e71061.

Introduction

Introduction

1. Pollen grain

Angiosperms have been surprisingly successful in colonizing all terrestrial environments. Part of this success has been due to the simplification of structures that allows a better use of energy (Hackenberg and Twell, 2019). The male gametophyte of spermatophytes, and specially angiosperms, has undergone such a degree of reduction that its development occurs within the wall of the microspore, the pollen grain. Plant reproduction success depends largely on the viability of pollen until ovule fertilization. Therefore, the gametophyte must be protected against high temperatures, desiccation, ultraviolet light and microbial attack; this protective function is fulfilled by the wall of the pollen grain, a very complex cell wall. In addition, the pollen wall aids pollination, since more elaborate outer pollen walls are more attractive to pollinators. Also, depending on the structure and composition of the wall, which is very diverse, plants can distinguish their own pollen from others in order to avoid embryo lethality (Matamoro-Vidal *et al.,* 2016; Ariizumi and Toriyama, 2011). Due to its great morphological variety, the pollen grain has been widely used in palynological studies and in numerous applications, such as taxonomy (Blackmore *et al.,* 2007). The tips of pollen tubes are also very interesting as models to study the formation of different domains in the plasmatic membrane (PM) of plant cells (Zhou and Dobritsa, 2019).

The process of pollen grain formation is known as microsporogenesis, in which four haploid microspores are generated from one mother cell of the pollen grain by meiosis. Subsequently, within the pollen wall, the male gametophyte is developed, a process known as gametogenesis. During microsporogenesis and gametogenesis the pollen grain goes through different stages, which will be mentioned in some chapters of this thesis. These stages of pollen development were described by Owen and Makaroff (1995) and are summarized as follows:

- Pre-meiosis phase. In this initial phase, microspore mother cells (MMC) are located inside the anther surrounded by the tapetum, the innermost cell layer of the anther wall surrounding the locule where pollen develops. The tapetum provides different compounds to the MMC through plasmodesmata that connect

them. The MMC increases in size and callose begins to be deposited around them.

- Meiosis stage. The thickness of the callose layer increases as meiotic division occurs, after which the microsporocyte contains four haploid nuclei. In angiosperms there are two types of meiotic cytokinesis: simultaneous, in which cytokinesis and callose deposition do not occur until both meiotic nuclear divisions have been completed; and successive, in which callose cell walls are formed after meiosis I and meiosis II (Furness and Rudall 2004). After meiosis, each of the four nuclei is positioned at the periphery of the cell, with a tetrahedral distribution in simultaneous meiosis. Small amounts of callose are deposited between the microsporocyte nucleus.
- Tetrad stage. After cytokinesis, the four haploid microspores are bounded by the callose wall. Some pollen grain wall precursors are deposited in this stage.
- Young microspore stage. The individual cells of the tetrad separate and remain independent due to the degradation of callose. The pollen grain wall begins to be discernible and will continue to develop in the following phases.
- Vacuolated microspore stage. The cells increase in size and a large number of vacuoles form and merge into a single large vacuole that occupies a large part of the cell volume. This phase ends when mitotic division begins.
- Bicellular pollen stage. After mitotic division, the vegetative cell and generative cell are originated. The generative cell occupies a marginal position next to the cell wall. Lipids and starch accumulate in the cell interior.
- Mature pollen stage. This is the final phase of the pollen grain development process. At this stage the cytoplasmic content consists mainly of large amounts of starch and lipids and the pollen wall is completely formed.

2. Pollen wall structure

The structure of the pollen wall is well known and, despite the high pollen diversity, it is significantly similar among taxa being composed of two layers. The innermost layer, the intine, is located directly above the PM and is formed by hydrolytic enzymes, hydrophobic proteins, cellulose, hemicellulose, and pectic polymers, similar to the

primary walls of common plant cells (Ariizumi and Toriyama, 2011). The outermost, the exine, is one of the most complex cell walls in plants and consists of several layers as described below (Ariizumi and Toriyama, 2011). Finally, surrounding the exine and filling its cavities is the pollen coat, pollenkitt or tryphine, an oily substance made up of lipids, proteins, pigments and aromatic compounds (Wang and Dobritsa, 2018).

The exine is a highly resistant layer, which is mainly due to its composition based on sporopollenin, a biopolymer that is highly resistant to physical, chemical and biological degradation processes (Ariizumi and Toriyama, 2011). The high stability of sporopollenin to non-oxidative degradation is demonstrated by the morphological preservation of ancient pollen grains over many millions of years (Piffanelli *et al*., 1998). Sporopollenin is present in the wall of pollen grains and spores of land plants, and was indispensable for their colonization of the terrestrial environment (Wallace *et al*., 2011). Although sporopollenin composition is not entirely known, due to its high stability, it is composed of different biopolymers insoluble in aqueous and organic solvents derived from saturated precursors, such as long-chain fatty acids or long aliphatic chains (Ariizumi and Toriyama, 2011).

Although the variety of exine patterns is enormous depending on the species, its architecture and the layers that form it are similar for all of them. Two different systems have been used to describe exine stratification (Erdtman, 1952; Fægri, 1956; Figure 1). Following the terminology of Erdtman, the exine consists of an outer layer, the sexine, and an inner layer, the nexin. The sexine is usually species-specific and is formed by vertical rods that form the *baculae* or *columellae*, on top of which the *tectum* is arranged as a roof and finally in the outermost part are the *supratectal elements*, small protrusions with different shapes. The nexin is sometimes subdivided in an outer layer, *nexine I* or *foot layer*, and an inner layer, *nexine II*. In some species some of the sublayers may be missing, depending on the layers present and their composition and shape, the different patterns observable in the exine and characteristic of each taxon will be created (Ariizumi and Toriyama, 2011; Wang and Dobritsa, 2018).

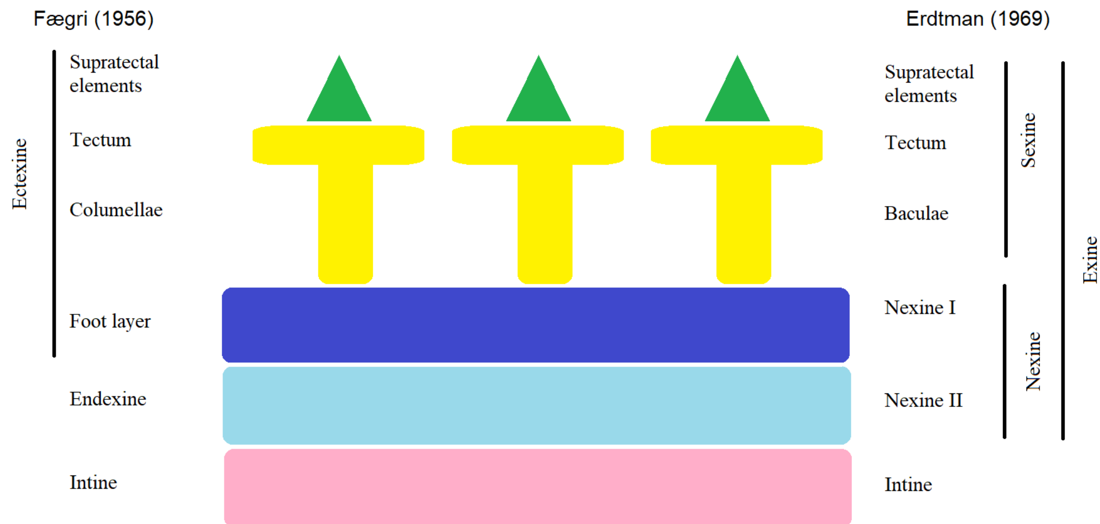


Figure 1. Pollen wall structure of angiosperm showing the two main terminologies used to describe its layers

3. Exine ontogeny

The exine begins its development at the tetrad stage, in which the four microspores are surrounded by the callose (β-1-3-glucan) wall that separates them from each other.

During the tetrad stage, each of the microspores produces a thin layer of primexine under the callose wall, which can act by trapping primexine subunits around microspores or act as a physical support for primexine assembly (Ariizumi and Toriyama, 2011; Wang and Dobritsa, 2018; Figure 2). Primexine is a glycoclayx-like fibrillar polysaccharide material, which acts as a scaffold, matrix, or template of sporopollenin. The genetic control of primexine formation is carried out by the sporophytic tissue, which has been demonstrated with some mutants defective in pyrimexin formation. While primexine is developed, the PM of each microspore begins to undulate and separate from the callose wall (Ariizumi and Toriyama, 2011; Wang and Dobritsa, 2018; Figure 2). These undulations influence the final configuration of the exine, since between the peaks of the undulations more primexine is deposited and in these areas the baculae will appear. Thus, the height of the probacuoles, the precursors of baculae, depends on the height that the primexin reaches in these areas (Wang and Dobritsa, 2018).

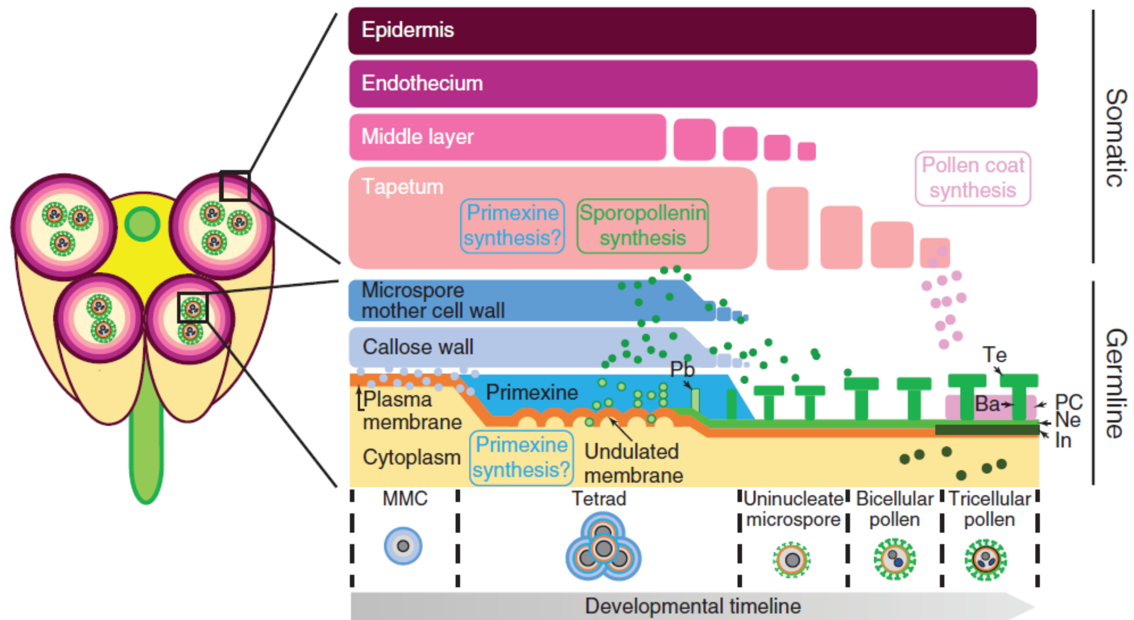


Figure 2. Events associated with the ontogeny of the pollen wall during its different stages of development. Ba, Baculae; In, Intine; MMC, Microspore mother cell; Ne, Nexine; Pb, Probaculae; PC, Pollen coat; Te, Tectum. Image taken from Wang and Dobritsa, 2018.

Later, as the callose wall disappears, sporopollenin precursors begin to deposit and polymerize. The sporopollenin precursor is produced from within the haploid microspore and from the tapetum (Ariizumi and Toriyama, 2011; Blackmore *et al.*, 2007; Figure 2). The cells of the tapetum also synthesize β-1,3-glucanase (callase), which degrades the callose wall, allowing microspores to remain free in the locule (Aboulela *et al.*, 2018). Sporopollenin will continue to form until the bicellular pollen stage, at which time the exine structure will be complete; also, at this stage, the intine begins to develop (Ariizumi and Toriyama 2011; Figure 2). Finally, before the mature pollen stage, the tryphine or pollen coat covering the exine is deposited (Blackmore *et al.*, 2007).

4. Pollen grain aperture

In certain areas of the pollen wall primexine is not deposited and therefore the exine is weakened or absent; these areas are known as apertures. The apertural areas, being covered only by the intine, offer less resistance to rupture by the pollen tube. Thus the apertures are usually the areas through which the pollen grain germinates. For example, in grasses, apertures are the only area through which the pollen tube can exit, as mutants lacking apertures do not germinate resulting in their sterility (Li *et al.*, 2018). However, in other species, such as *Arabidopsis* or *Eschscholzia californica*, apertures are

dispensable and their mutants lacking apertures are fertile (Dobritsa *et al*., 2018; Mazuecos-Aguilera *et al*., 2021). In addition, some species have inaperturate pollen and therefore do not need apertures to germinate. The apertures are also involved in the exchange of solutes and water with the environment and stigma (Wang and Dobritsa, 2018). On the other hand, when the pollen grain dehydrates and decreases in volume, the membrane of the aperture sites folds inward, so that the edges of each aperture are touching each other, closing up the aperture site. This process is called harmomegathy and allows the pollen grain to adapt to the change in volume without breaking the wall and avoiding water loss (Matamoro-Vidal *et al*., 2016; Wang and Dobritsa, 2018).

The pollen aperture system, defined by the shape, number, position orientation or size of the apertures, is species-specific, but can vary widely between species. According to their shape they can be: elongate furrow-like (colpate, if furrows placed elsewhere and sulcate if furrows across the middle of the distal face) or round, pore-like (porate). Depending on their number, they can be inaperturate (without any apertures), mono-aperturate (with one aperture), di-aperturate (with two apertures), tri-aperturate (with three apertures), or poly-aperturate (with more than three apertures). With respect to their position, apertures were classified in: polar apertures located at or toward the poles, equatorial apertures located on or at the equator and global apertures more or less uniformly scattered over the surface of the pollen grain (Sporne, 1972; Walker and Doyle, 1975). The enormous diversity of aperture patterns among species has been used in studies of taxonomy and phylogeny of flowering plants.

5. Pollen aperture evolution

Among angiosperms there are two basic aperture patterns: with a single aperture in the polar position (monosulcate), found in many species of basal groups and monocots; and and a derivative model with three apertures in the equatorial position (tricolpate), present in eudicots. Although many species have aperture patterns very different from these prototypes (Wang and Dobritsa, 2018; Furness and Rudall, 2004). Thus, from an evolutionary perspective, in the lineage leading to eudicots there was a change in apertural position that apparently favoured an increase in the number of apertures (Furness and Rudall, 2004). This increase in apertures may have provided a selective advantage by increasing the number of possible sites for pollen germination and harmomegathy process.

The change in apertural position has been related to the change in the pattern of meiotic cytokinesis, from successive in most of the basal angiosperms to simultaneous in eudicots (Furness and Rudall, 2004). This change in the apertural pattern may have been a key innovation in the success of eudicots (75% of angiosperms) and subsequent radiation (Blackmore *et al*., 1995). As with other characters, the most derived groups of angiosperms and especially the core eudicots have well defined and fixed apertural systems, while the basal groups usually show higher variability due to the lack of character fixation (Blackmore *et al*., 1995).

6. Pollen aperture formation

The different aperture patterns are determined during the early stages of microsporogenesis, hence the apertures are often fully visible in the late-tetrad stage (Nadot *et al*., 2006). For the formation of apertures PM domains must be created, which will attract specific proteins and lipids that will prevent exine deposition. Therefore, apertures have emerged as a model for studying the formation of distinct PM domains (Zhou and Dobritsa, 2019; Zhou *et al*., 2021). Many cellular processes associated with microsporogenesis have been proposed to be involved in the aperture formation, although this is not yet well understood.

As discussed above, there appears to be a relationship between the type of meiotic cytokinesis and the type of apertures developed. On the other hand, detailed observations of microsporogenesis in species with different aperture patterns show that apertures are defined according to the regions where cytokinesis is completed (Nadot *et al*., 2006). However, some *Arabidopsis* mutants defective in meiotic cytokinesis exhibit apertures, so cytokinesis would not be required for their formation (Wang and Dobritsa, 2018). Consequently, the type of cytokinesis and the last points of contact between sister microspores does not determine the type of apertural pattern in *Arabidopsis*. On the other hand mutants without meiotic nuclear divisions lack apertures. The absence of meiotic division could disrupt the formation of aperture domains in the PM that indicate where the apertures should form (Wang and Dobritsa, 2018).

As in exine formation, callose also appears to be involved in the formation of apertures. The apertures form at sites where the PM is in close contact with the callose wall. Also, additional callose deposits have been observed at the sites where apertures

will appear in different species and in some species, additional callose deposits and apertures coincide in shape (Albert *et al*., 2010; Prieu *et al.,* 2017).

Other factor that has been shown to play a crucial role in determining the aperture pattern is the presence of endoplasmic reticulum (ER) shields underneath of the PM in the regions where the apertures are formed (Blackmore *et al*., 2007; Nadot *et al*., 2006). Possibly the ER prevents the deposition of factors required for exine formation, such as any component of primexine (Wang and Dobritsa, 2018).

7. Molecular player of aperture formation

Many processes could be involved in aperture formation and there is also a great variety of aperture patterns. Therefore, many molecular players related to the creation of PM domains, deposition/degradation of callose, and prevention of exine deposition must be involved in aperture formation. Furthermore, the change from monosulcate to tricolpate apertural pattern evolved only once and was irreversible, so this change must have involved mutation of many genes (Furnes and Rudall, 2004). Forward genetic approaches have been used to identify genes involved in aperture formation. Using genetic screens the identification of different mutants with variations in aperture patterns has been possible (Wang and Dobritsa, 2018).

7.1. *INAPERTURATE POLLEN 1* and partners

The first discovered molecular player involved in aperture formation, *INAPERTURATE POLLEN 1* (*INP1*), was described by Dobritsa and Coerper (2012) in *Arabidopsis thaliana*, which has tricolpate pollen, the most typical of eudicots. This protein of unknown biochemical function is indispensable for aperture formation, since *inp1* knockout mutants lack apertures in their walls.

INP1 protein is synthesised by the sporophytic MMC before and during meiotic cytokinesis, and is distributed equally throughout the cell (Figure 3). Simultaneous MMC cytokinesis generates four microspores arranged in a tetrahedral tetrad. In each of these microspores INP1 protein assembles into three equidistant longitudinal punctate lines between the PM and the callose wall, where INP1 acts as a marker of the preaperture domains (Figure 3). INP1 causes formation of membrane ridges in the aperture domains to bring the PM in close contact with the callose wall preventing primexin formation and subsequent sporopollenin deposition in these regions. The

formation of these bridges is perhaps due to interaction of INP1 with callose or a callose-related protein, although this has not been confirmed. In support of this, INP1 requires the callose wall for its correct localisation, since in callose defective *CALLOSE SYNTHASE* (CALS5) mutants INP1 assembled into puncta at the abnormal sites (Dobritsa and Coerper, 2012; Dobritsa *et al*., 2018; Wang and Dobritsa, 2018).

During exine development INP1 keeps the specific membrane domains attached to the callose wall (Figure 3), but it is not responsible for specifying where the domains are created. The aperture domains are already pre-specified prior to the arrival of INP1, so INP1 does not determine the type, position or number of apertures, although it can influence the length of the apertures depending on their dosage. Thus, the apertures become shorter when *INP1* expression is reduced (Dobritsa *et al*., 2018, Dobritsa and Coerper, 2012; Reeder *et al*., 2016).

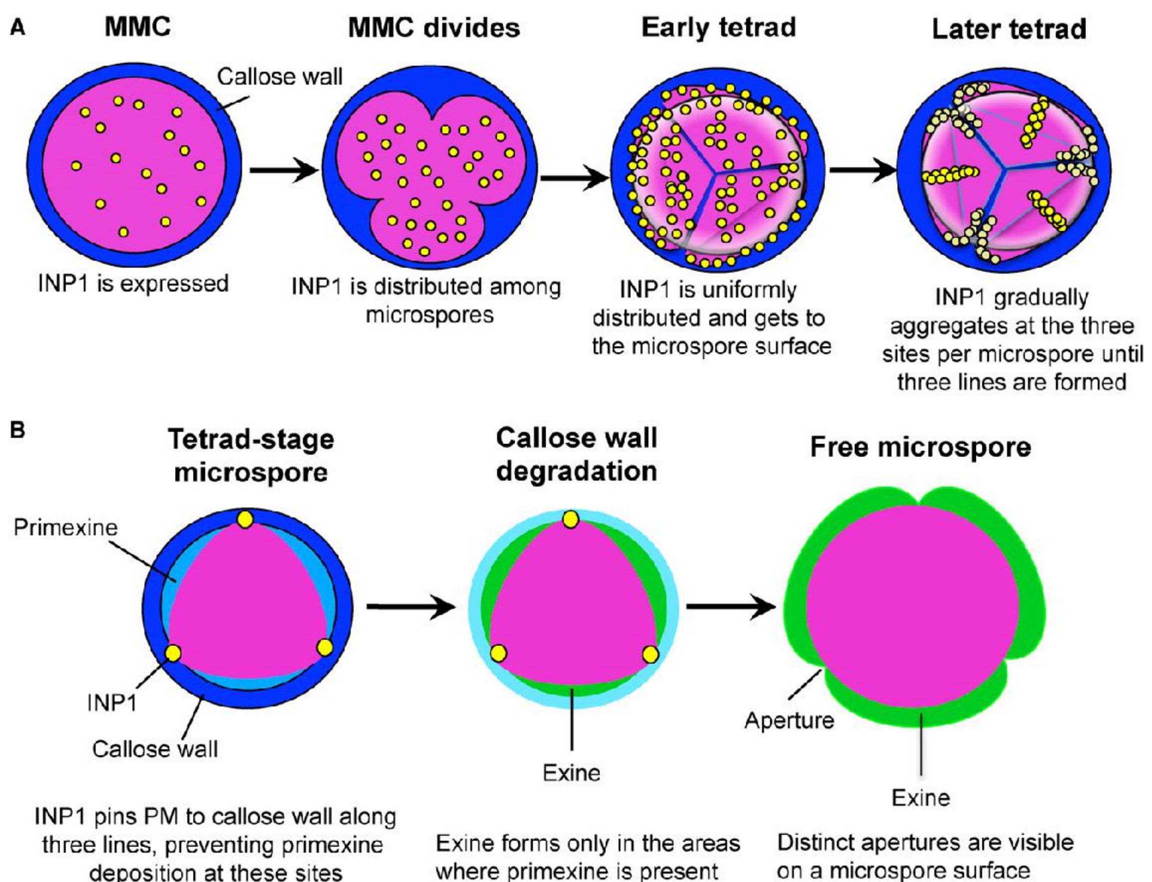


Figure 3. Model of INP1 distribution and assembly during aperture formation. Image taken from Dobritsa *et al.,* 2018.

Despite the great variety of apertural patterns found in the different taxonomic groups of angiosperms, INP1 function appears to be conserved. Thus, INP1 is indispensable for the formation of apertures in monocots such as maize and rice (Li *et al*., 2018; Zhang *et al*., 2020), which have a single polar pore surrounded by a ring-shaped annulus and covered by a lid-like operculum. Pollen grains of maize (*Zea mays*) and rice (*Oryza sativa*) mutants for *ZmINP1* and *OsINP1* have no apertures and are infertile because, unlike *Arabidopsis* mutants, they cannot germinate. (Li *et al*., 2018; Zhang *et al*., 2020).

Although INP1 function is conserved, its sequence is highly variable, especially among eudicots, which is consistent with the wide variety of apertural patterns found in eudicots. Despite the coincidence of the variety of aperture patterns and the variety of INP1 sequences in eudicots, INP1 is not responsible for determining the different aperture patterns; since when the apertures of *Arabidopsis* mutants were restored with *INP1* sequences from *Capsella rubella* (*CrINP1*) (another Brassicaceae), they had *Arabidopsis*-like morphology and not *Capsella*-like morphology (Li *et al*., 2018).

On the other hand, the restoration of apertures with *INP1* sequences from phylogenetically more distant species is not possible, because INP1 function also depends on interactions with co-evolved partners in their respective species (Li *et al*., 2018). A species-specific partner for INP1 in *Arabidopsis* and tomato is its paralogue INP2, and probably also in many other species in which orthologues of *INP2* have been found. INP2 resembles INP1 in its protein structure, expression patterns, evolutionary divergence trends, mutant phenotype and genetic interactions. In addition to having a similar function to INP1, INP2 physically interacts with INP1 and is involved in its orientation towards aperture domains and its permanence in these areas (Lee *et al*., 2021).

As with *INP1*, restoration of apertures in *Arabidopsis inp2* mutants with tomato *INP2* sequences (*SlINP2*) is not possible. However, when tomato *INP1* (*SlINP1*) and *SlINP2* are expressed together in *Arabidopsis*, they are able to restore apertures, indicating that INP1 and INP2 act as co-evolved species-specific partners, and form a functional module essential for the formation of pollen apertures. Furthermore, AtINP2 interacts with AtINP1 but not with SlINP1, just as SlINP2 interacts with SlINP1 but not with AtINP1, further confirming that *INP1* and *INP2* are functionally species-specific.

On the other hand, when *Arabidopsis* apertures are restored with *SlINP1* and *SlINP2*, they have *Arabidopsis*-like morphology and not tomato-like morphology, as happens when restored with the *CrINP1* sequence. Therefore neither *INP1* nor *INP2* define the position, number or morphology of the apertures, so each species will have variations in the upstream mechanisms or differences in other interactors of the INP proteins (Lee *et al*., 2021).

In rice, the orthologue of *AtINP1*, *OsINP1*, is indispensable for aperture formation and interacts with the lectin receptor-like kinase *DEFECTIVE IN APERTURE FORMATION1* (*OsDAF1*), which is essential for annulus formation. In this species, OsINP1 has not only conserved its function for aperture formation, but has diversified its role to recruit OsDAF1 to its correct subcellular location via direct protein interaction. However OsINP1 localizes and functions normally in the absence of OsDAF1 (Liang *et al*., 2020).

7.2. *D6 PROTEIN KINASE-LIKE3*

Another protein involved in aperture formation in *Arabidopsis* is the protein kinase D6 PROTEIN KINASE-LIKE3 (D6PKL3), which acts upstream of INP1 and INP2. D6PKL3 and the other members of the D6PK subfamily are involved in domain formation in different vegetative cells, although only D6PKL3 is involved in aperture formation (Lee *et al*., 2018). The *d6pkl3* mutants have abnormal apertures, which are partially or, in some cases, completely covered with exine. Thus, *d6pkl3* mutants have a similar but less drastic phenotype than *inp1* and *inp2* mutants, in which the phenotype is completely inaperturate. Therefore, the role of *D6PKL3* in the aperture formation process, although important, is less essential than that of INP proteins (Lee *et al*., 2018; Lee *et al*., 2021; Zhou and Dobritsa 2019).

Like INP1 and INP2, D6PKL3 localises to three PM domains in tetrad-stage microspore where the apertures will later be developed (Lee *et al*., 2018; Lee *et al*., 2021; Zhou and Dobritsa, 2019). However, D6PKL3 is assembled in these domains before the end of meiotic cytokinesis, so it is assembled before the assembly of INP1 and INP2 (Lee *et al*., 2018). Furthermore, INP1 needs D6PKL3 for recruitment to its correct location indicating that D6PKL3 acts upstream of INP1 and INP2 in the establishment of the aperture domains (Lee *et al*., 2018). On the other hand, D6PKL3 also requires INP proteins for its maintenance in these domains. Despite the reciprocal

dependence between INP proteins and D6PKL3, they do not interact directly and may interact via modification of the PM in these areas (Lee *et al.*, 2018). The indirect interaction of these proteins is due to the fact that D6PKL3 occupies a location at the cytoplasmic site of the membrane in the aperture domains while INP proteins assemble between the external part of the membrane and the callose wall. Thus, unlike INP1, D6PKL3 does not depend on callose for its proper localisation. These different localisations indicate that INP proteins and D6PKL3 play different roles in aperture formation (Lee *et al.*, 2018; Lee *et al.*, 2021; Zhou and Dobritsa, 2019).

D6PKL3 requires its kinase activity for correct function, indicating that it may act by phosphorylating another unknown aperture factor. In addition, D6PKL3 also requires a phosphoinositide-interacting polybasic motif for its localisation in the aperture domains. In connection with this, two phosphoinositide species, phosphatidylinositol- 4-phosphate [PI(4)P] and phosphatidylinositol-4,5- bisphosphate [PI(4,5)P2], have been found in the aperture domains of the PM, which interact in-vitro with D6PKL3. Phosphoinositides are involved in polarity generation, thus these two phosphoinositides species could act by recruiting proteins required for aperture specification and formation, such as INP1 and/or INP2 (Lee *et al.*, 2018; Zhou and Dobritsa, 2019).

On the other hand, D6PKL3 could be modulating the number of apertures, since in transgenic plants in which *D6PKL3* is overexpressed, pollen grains appear with double parallel apertures or with more than three apertures (four or six). In these pollen grains INP1 is also recruited to the newly formed apertures, which further confirms that D6PKL3 acts upstream of INP1 and recruits it to the aperture domains. Although overexpression of *D6PKL3* generates pollen with more than three apertures, D6PKL3 is not the only factor that specifies where the apertures form, since *d6pkl3* mutants have apertures that are deficient but located at wild-type positions (Lee *et al.*, 2018; Zhou and Dobritsa, 2019; Lee *et al.*, 2021).

7.3. *MACARON and ELMOD_A*

Two members of the ELMOD protein family, MACARON/ELMOD_B (MCR) and ELMOD_A, which act upstream of INP1, INP2 and D6PKL3, are able to determine the position of apertures, although maintaining the equidistant placement, as well as the

number of apertures (Zhou *et al.*, 2021). In mammals, ELMOD proteins act as non-canonical GTPase activating proteins (GAPs), while in plants its function is unknown.

ELMOD_A and MCR are paralogues, sharing 86% sequence identity, and their mutants show pollen with a single ring-shaped pollen aperture, composed of two equidistantly placed longitudinal apertures that overlap, resulting in the loss of one aperture (Zhou *et al.*, 2021). However, MCR plays a more prominent role than ELMOD_A; since the reduction of the number of apertures in *elmod_a* mutants requires the loss of a copy of *MCR*, whereas *mcr* mutants show the same phenotype although both copies of *ELMOD_A* are present. Thus, in wild-type plants MCR can specify three normal apertures in the absence of ELMOD_A but not vice versa. Also, when these genes are overexpressed, the number of apertures increases to six in the case of *MCR* and to four in *ELMOD_A* (Zhou *et al.*, 2021).

In addition, in *mcr* mutants INP1 and D6PKL3 proteins are located in the ring-shaped membrane domain, indicating that MCR acts upstream of them and pre-marks the aperture domains where they should act. Unlike the other aperture factors, MCR and ELMOD_A do not accumulate in the aperture membrane domains and remain dispersed in the cytoplasm (Zhou *et al.*, 2021).

On the other hand, a third member of the ELMOD family, ELMOD_E, has an ability to influence the morphology of aperture domains. The ELMOD_E overexpression at the tetrad stage in *Arabidopsis* mutants using the *MCR* promoter, produces pollen with round apertures instead of narrow longitudinal furrows apertures. However, *elmod_e* mutants produce normal pollen, because its expression is very low in *Arabidopsis*, and ELMOD_E probably competes for the ligands of MCR and ELMOD_A. Nevertheless, *ELMOD_E* could play a role in determining the shape of apertures in *Arabidopsis* and other species, thus contributing to the diversity of aperture patterns in nature (Zhou *et al.*, 2021).

7.4. *DOUGHNUT* and future perspectives

Finally, another mutant of *Arabidopsis* called *doughnut* (*dnt*) presents changes in aperture number, morphology, and positions. Pollen of the *dnt* mutant shows two round, holelike apertures with an internal deposition of exine, and its position changes from the equator to the poles of the pollen grain. However, the gene or genes involved in this

mutation have not yet been discovered. In this mutant, INP1 is located at the poles, so the protein(s) that cause this mutation acts genetically upstream of INP1 (Zhou and Dobritsa, 2019).

Despite the aperture factors discovered, their function and interaction mechanisms need to be better understood, in addition to the search for many other molecular players that might be involved in the formation of membrane domains, in the avoidance of primexine deposition, in aperture shape determinism or in the interaction of the known aperture factors.

8. Papaveraceae family as a model system to study aperture pattern formation

The family Papaveraceae Juss. includes 44 genera and around 790 species (Zhang *et al*., 2008). Papaveraceae belongs to the order Ranunculales, the most basal lineage of eudicots (Hoot *et al*., 2015); and within Ranunculales, Papaveraceae and Eupteleaceae K. Wihl. have the most basal position with respect to the rest of families. Papaveraceae is composed of four subfamilies: Fumarioideae Eaton, Hypocoideae (Dum.) Hutchinson, Papaveroideae Eaton, and Pteridophylloideae Murb. (Thorne, 1992; Mabberley, 2008). The main morphological features, common to all Papaveraceae species, are related to floral characters; since all of its taxa have two sepals, petals often four, in two whorls, gynoecium syncarpous (paracarpous) and bicarpellate or polycarpellate, placentation parietal (Wang *et al.,* 2009).

As with other characters, the more derived groups of angiosperms and especially eudicots (core eudicots) have well-defined and fixed apertural systems, whereas the basal groups have more variability due to the lack of character fixation. Thus, the order Ranunculales, the most basal lineage of eudicots, and particularly the family Papaveraceae, shows great pollen diversity. Therefore, the phylogenetic position of Papaveraceae together with its pollen and aperture diversity makes this family of great interest for comparative studies in the field of evolution and development. Information on the role of genes controlling pollen aperture formation in Papaveraceae may be crucial to confirm the degree of conservation of the gene function of genes discovered in model species. In addition, the great pollen diversity within the same group would

facilitate the discovery, by comparative studies, of molecular players that determine the different aperture patterns.

For the development of this thesis we chose the species described below. These were selected for comparative analysis due to their taxonomic category and apertural system. With the exception of *Eschscholzia californica*, protocols for transformation have not been described in these species and their genome is not available.

Eschscholzia californica

California poppy, *Eschscholzia californica* Cham. is a species belonging to the subfamily Papaveroideae, tribe Eschscholtzieae Baill. This plant is native to the west coast of North America, and is used globally as an ornamental plant for its beautiful flower (Becker et., al 2005). The pollen grains are mostly spheroidal in shape and have a 5-7-zonocolpate apertural system (PalDat, 2017; Figure 4a). The California poppy is also used as a medicinal herb, as it produces several types of alkaloids, so it has also been used to understand the molecular basis of alkaloid chemistry (Carlson *et al*., 2006; Hori *et al*., 2018).

Eschscholzia californica has certain characteristics that make it interesting as a model species (Becker et., al 2005; Carlson *et al*., 2006; Hori *et al*., 2017; Wege *et al*., 2007).

- It is easy to grow, it is a qualitative long-day plant and their generation time is fast, around 3 months. They also produce a large amount of seeds per plant, which is interesting in studies that need to obtain successive generations.
- It is diploid and its genome size is relatively small (502 Mb, 1,115 Mbp) approximately 6.5 times the size of *A. thaliana* and smaller than the genome of *Papaver somniferum* (3,724 Mbp), another well-studied Papaveraceae. The genome of *Eshscholzia californica* has been sequenced. On the other hand, a large number of expressed sequence tags have been sequenced by the Floral Genome Project.
- A protocol for carrying out virus-induced gene silencing (VIGS) has been developed. This method, although transient, is faster to obtain functional data by

down-regulating gene expression. For localising the expression of certain genes, in situ hybridization protocols have also been developed.

- Unlike *Papaver somniferum*, California poppy does not require the government research permits needed to work.
- Shoot, inflorescence and flower morphogenesis has been studied in detail.
- Each plant produces a large number of flowers and each flower has between 18 and 34 stamens with a large pollen production. This facilitates the study of flower and pollen grain development. In addition, the flowers are larger than those of *Arabidopsis* and other model plants, with a bud diameter of about 5.5 mm before anthesis, which facilitates handling.

These characteristics make *Eschscholzia californica* a good model plant together with the already used core eudicots: *Antirrhinum majus* (snapdragon), *Arabidopsis thaliana*, *Lycopersicon esculentum* (tomato), *Nicotiana tabacum* (tobacco), *Petunia_hybrida* and *Pisum sativum* (pea); and monocots: *Oryza sativa* (rice), *Zea mays* (corn).

Dactylicapnos torulosa

Dactylicapnos torulosa (Hook.f. and T. Thomson) is a anual plant belonging to the subfamily Fumarioideae that grows in Southwest China. This plant is used in traditional Chinese medicine due to its alkaloid production, so it has also been used to determine its alkaloid composition (Rucker *et al.*, 1994). Its pollen grains have 6 colpos. This plant is easy to cultivate, diploid, fast-growing and produces a large number of flowers (Frisby and Hind, 2008; Figure 4b).

Fumaria bracteosa

Fumaria bracteosa (Pomel) also belongs to the subfamily Fumarioideae and, like the previous species, has also been used as a medicinal plant due to its alkaloid content (Halim *et al.*, 1986; Lidén, 1993). Regarding its apertural system, the pollen is tetra- to hexa-porate (Shamso and Toshiyuki, 2012). This anual plant also is diploid and shows rapid growth and flowering, although germination is more complicated (Figure 4c).

Roemeria refracta

Roemeria refracta DC. Belonging to the subfamily Papaveroideae. It also contains alkaloids and has therefore been studied extensively (Gozler *et al*., 1990). Its pollen grains have six pores (Layka, 1975; Figure 4d). *Roemeria refracta* also is diploid and has a fast generation time.

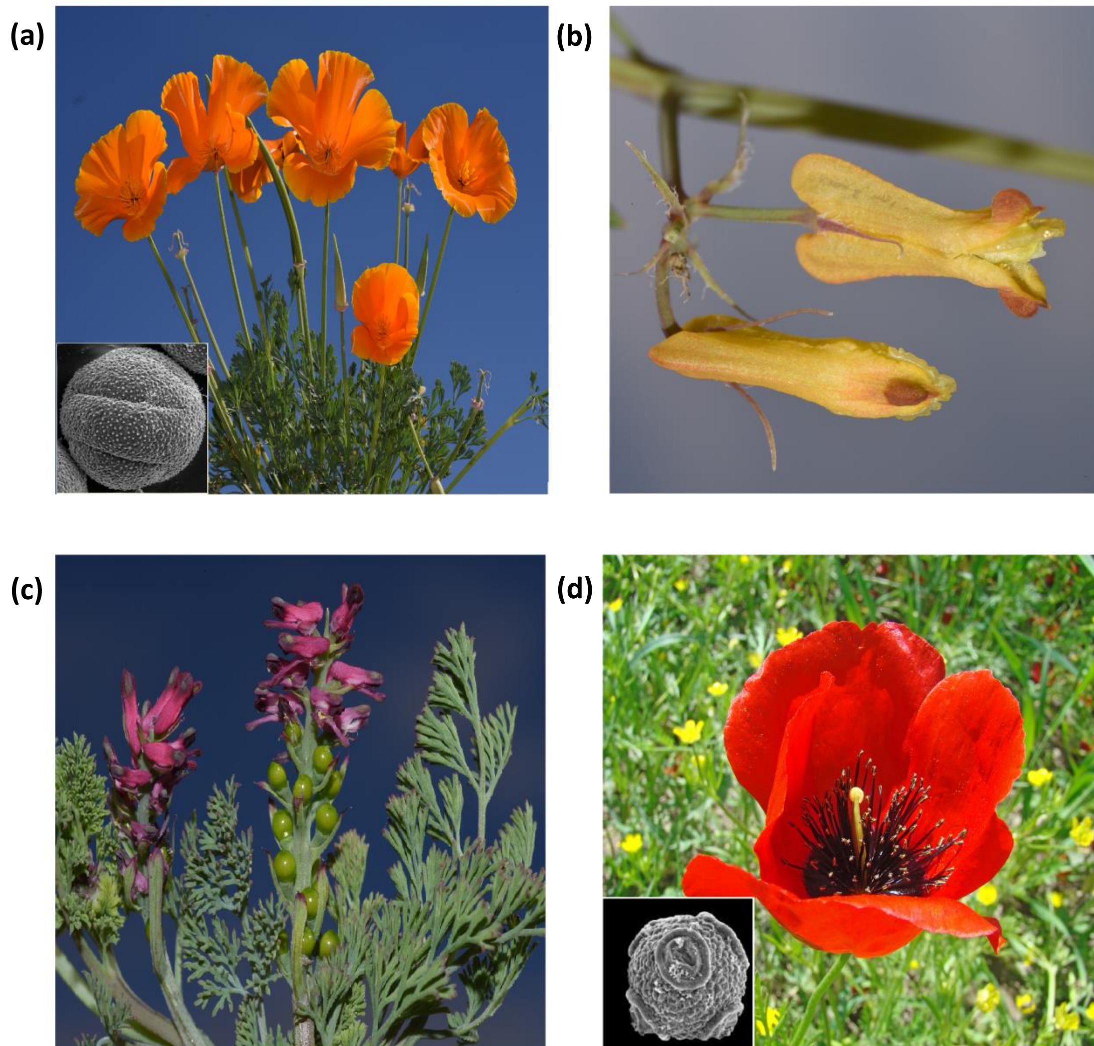


Figure 4. Species used for the development of the thesis and their pollens. **(a)** *Eschscholzia californica* (image taken by Gabriel Blanca) and scanning electron microscopy (SEM) image of its pollen grain. **(b)** *Dactylicapnos torulosa* (image taken by Gabriel Blanca). **(c)** *Fumaria bracteosa* (image taken by Gabriel Blanca). **(d)** *Roemeria refracta* (image from iNaturalist; iNaturalist.org; taken by Katunchik) and SEM image of its pollen grain (image taken from Geran and Sharifnia, 2008).

References

Aboulela M, Nakagawa T, Oshima A, Nishimura K, Tanaka Y. 2018. The Arabidopsis COPII components, AtSEC23A and AtSEC23D, are essential for pollen wall development and exine patterning. *Journal of experimental botany* **69:** 1615-1633.

Albert B, Nadot S, Dreyer L, Ressayre A. 2010. The influence of tetrad shape and intersporal callose wall formation on pollen aperture pattern ontogeny in two eudicot species. *Annals of Botany* **106**: 557-564.

Ariizumi T, Toriyama K. 2011. Genetic regulation of sporopollenin synthesis and pollen exine development. *Annual review of plant biology* **62**: 437-460.

Becker A, Gleissberg S, Smyth DR. 2005. Floral and vegetative morphogenesis in California poppy (Eschscholzia californica Cham.). *International Journal of Plant Sciences* **166**: 537-555.

Blackmore S, Stafford P, Persson V. 1995. Palynology and systematics of Ranunculiflorae. In: Jensen U., Kadereit J.W. (eds.) Systematics and evolution of the Ranunculiflorae. *Plant Syst. Evol. [Suppl.]* **9**:71-82.

Blackmore S, Wortley AH, Skvarla JJ, Rowley JR. 2007. Pollen wall development in flowering plants. *New Phytologist* **174**: 483-498.

Carlson JE, Leebens-Mack JH, Wall PK, *et al.* 2006. EST database for early flower development in California poppy (Eschscholzia californica Cham., Papaveraceae) tags over 6000 genes from a basal eudicot. *Plant Molecular Biology* **62**: 351-369.

Dobritsa AA, Coerper D. 2012. The novel plant protein INAPERTURATE POLLEN1 marks distinct cellular domains and controls formation of apertures in the Arabidopsis pollen exine. *The Plant Cell* **24**: 4452-4464.

Dobritsa AA, Kirkpatrick AB, Reeder SH, Li P, Owen, HA. 2018. Pollen aperture factor INP1 acts late in aperture formation by excluding specific membrane domains from exine deposition. *Plant Physiology* **176**: 326-339.

Erdtman, G. 1952. Pollen morphology and plant taxonomy. *Geologiska Föreningens i Stockholm Förhandlingar* **74**: 526–527.

Fægri, K. 1956. Recent trends in palynology. *The Botanical Review* **22**: 639–664.

Frisby S, Hind N. 2008. 625. DACTYLICAPNOS TORULOSA: Papaveraceae. *Curtis's Botanical Magazine* 230-236.

Furness CA, Rudall PJ. 2004. Pollen aperture evolution–a crucial factor for eudicot success?. *Trends in plant science* **9**: 154-158.

Geran A, Sharifnia F. 2008. Micro-macromorphological studies of the genus Glaucium (Papaveraceae) in Iran.

Gözler B, Öziç P, Freyer AJ, Shamma M. 1990. Morphinandienone alkaloids from Roemeria refracta. *Journal of Natural Products **53***: 986-988.

Halim AF, Salama OM, Amer MMA. 1986. Alkaloids of Fumaria bracteosa. *Planta medica* **52**: 414-414.

Hackenberg D, Twell D. 2019. The evolution and patterning of male gametophyte development. *Current topics in developmental biology* **131**: 257-298.

Hoot SB, Wefferling KM., Wulff JA. 2015. Phylogeny and character evolution of Papaveraceae sl (Ranunculales). *Systematic botany* **40**: 474-488.

Hori K, Yamada Y, Purwanto R, *et al.* 2018. Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. *Plant and Cell Physiology* **59**: 222-233.

Layka S. 1975. Les caractères de l'endexine chez les Papavéracées. *Bulletin de la Société Botanique de France* **122(sup1):** 103-107.

Lee BH, Wang R, Moberg IM, *et al.* 2021. A species-specific functional module controls formation of pollen apertures. *Nature Plants* **7**: 966-978.

Lee BH, Weber ZT, Zourelidou M, *et al.* 2018. Arabidopsis protein kinase D6PKL3 is involved in the formation of distinct plasma membrane aperture domains on the pollen surface. *The Plant Cell* **30**: 2038-2056.

Li P, Ben-Menni Schuler S, Reeder SH, Wang R, Suárez Santiago VN, Dobritsa AA. 2018. INP1 involvement in pollen aperture formation is evolutionarily conserved and may require species-specific partners. *Journal of Experimental Botany* **69**: 983-996.

Lidén M. 1993. Fumariaceae. In *Flowering Plants· Dicotyledons* (pp. 310-318). Springer, Berlin, Heidelberg.

Mabberley DJ. 2008. Mabberley's plant-book. A protable dictionary of plants, their classification and uses. Cambridge, Cambridge University Press.

Matamoro-Vidal A, Raquin C, Brisset F, *et al.* 2016. Links between morphology and function of the pollen wall: an experimental approach. Botanical journal of the Linnean Society **180**: 478-490.

Mazuecos-Aguilera I, Romero-García AT, Klodová B, *et al.* 2021. The role of INAPERTURATE POLLEN1 as a pollen aperture factor is conserved in the basal eudicot Eschscholzia californica (Papaveraceae). *Frontiers in Plant Science* **1405**.

Nadot S, Forchioni A, Penet L, Sannier J, Ressayre A. 2006. Links between early pollen development and aperture pattern in monocots. *Protoplasma* **228**: 55-64.

Owen HA, Makaroff CA. 1995. Ultrastructure of microsporogenesis and microgametogenesis in Arabidopsis thaliana (L.) Heynh. ecotype Wassilewskija (Brassicaceae). *Protoplasma* **185**: 7-21.

PalDat 2017. PalDat – A Palynological Database, www.paldat.org (accessed 2021)

Piffanelli P, Ross JH, Murphy DJ. 1998. Biogenesis and function of the lipidic structures of pollen grains. *Sexual plant reproduction* **11:** 65-80.

Prieu C, Toghranegar Z, Matamoro-Vidal A, Nadot S, Albert B. 2017b. Additional callose deposits are located at the future apertural regions in sulcate, ulcerate, porate, colporate, colpate and syncolpate pollen grains. *Botanical Journal of the Linnean Society* **183**: 271-279.

Reeder SH, Lee BH, Fox R, Dobritsa AA. 2016. A ploidy-sensitive mechanism regulates aperture formation on the Arabidopsis pollen surface and guides localization of the aperture factor INP1. *PLoS Genetics* **12**: e1006060.

Rücker G, Breitmaier E, Zhang GL, Mayer R. 1994. Alkaloids from Dactylicapnos torulosa. *Phytochemistry* **36**: 519-523.

Shamso E, Toshiyuki F. 2012. The Pollen Flora of Faiyum, Egypt I-Archichlamydeae. *Taeckholmia* **32**: 1-40.

Sporne KR. 1972. Some observations on the evolution of pollen types in dicotyledons. *New Phytologist* **71**: 181-185.

Suárez-Santiago VN, Fernández-Fernández MC, Pérez-Gutiérrez MA, Schuler SBM, Salinas-Bonillo MJ, Romero-García AT. 2018. Morphological and ultrastructural diversity and character evolution of the pollen in the tribe Chelidonieae (Papaveraceae). *Review of Palaeobotany and Palynology* **258**: 83-97.

Thorne RF. 1992. Classification and geography of the flowering plants. *Botanical review* **58**: 225-348

Walker JA, Doyle JA. 1975. The bases of angiosperm phylogeny: palynology. *Annals of the Missouri Botanical Garden* **62**: 664–723.

Wallace S, Fleming A, Wellman CH, Beerling DJ. 2011. Evolutionary development of the plant spore and pollen wall. *AoB plants* **2011**.

Wang R, Dobritsa AA. 2018. Exine and aperture patterns on the pollen surface: their formation and roles in plant reproduction. *Annual Plant Reviews Online* 589-628.

Wang W, Lu AM, Ren Y, Endress ME, Chen ZD. 2009. Phylogeny and classification of Ranunculales: evidence from four molecular loci and morphological data. *Perspectives in Plant Ecology, Evolution and Systematics* **11**: 81-110.

Wege S, Scholz A, Gleissberg S, Becker A. 2007. Highly efficient virus-induced gene silencing (VIGS) in California poppy (Eschscholzia californica): an evaluation of VIGS as a strategy to obtain functional data from non-model plants. *Annals of Botany* **100**: 641-649.

Zhang ML, Su ZY, Lidén M, Grey-Wilson C. Papaveraceae. In: Wu Z, Raven PH, Hong D. 2008. Flora of China, Volume 7: Menispermaceae through Capparaceae. Science Press.

Zhang X, Zhao G, Tan Q, *et al*. 2020. Rice pollen aperture formation is regulated by the interplay between OsINP1 and OsDAF1. *Nature plants* **6**: 394-403.

Zhang X, Zhao G, Tan Q. 2020. Rice pollen aperture formation is regulated by the interplay between OsINP1 and OsDAF1. *Nature Plants* **6:** 394-403.

Zhou Y, Amom P, Reeder SH, Lee BH, Helton A, Dobritsa AA. 2021. Members of the ELMOD protein family specify formation of distinct aperture domains on the Arabidopsis pollen surface. *Elife* **10**: e71061

Zhou Y, Dobritsa AA. 2019. Formation of aperture sites on the pollen surface as a model for development of distinct cellular domains. *Plant Science* **288**: 110222.

Objectives

Objectives

The ultimate goal of this thesis is to understand the role of genes involved in aperture formation and in the determination of the different aperture systems of the pollen grain in the family Papaveraceae. Information about the role of aperture development control genes in this family is crucial to an understanding of angiosperms as a whole, given its position between core eudicots on the one hand and basal angiosperms and monocots on the other. To this end, we set out the following specific objectives:

1. Analysis of the involvement of the *INAPERTURATE POLLEN 1* gene in the formation of pollen grain apertures in *Eschscholzia californica*. This objective will be achieved by sequence characterisation, analysis of the spatio-temporal expression pattern and validation of the function by gene silencing.

2. Identification of possible new genes involved in aperture formation by comparative transcriptome analysis of wild-type *Eschscholzia californica* vs *inp1* transient mutant *Eschscholzia californica*.

3. Identification of possible new genes involved in aperture shape determinism by comparative transcriptome analysis of two pairs of Papaveraceae species belonging to two different subfamilies and where within each pair the species has colpate pollen or porate pollen.

Chapter 1

The Role of INAPERTURATE POLLEN1 as a Pollen Aperture Factor Is Conserved in the Basal Eudicot *Eschscholzia californica* (Papaveraceae)

Ismael Mazuecos-Aguilera, Ana Teresa Romero-García, Božena Klodová, David Honys, María C. Fernández-Fernández, Samira Ben-Menni Schuler, Anna A. Dobritsa and Víctor N. Suárez-Santiago

Frontiers in Plant Science, 1405 (2021).

A previous versión to the published one is here presented as copyright requirement for the publication of this thesis by the Library of the University of Granada. Please see the final version here.

Abstract

Pollen grains show an enormous variety of aperture systems. What genes are involved in the aperture formation pathway and how conserved this pathway is in angiosperms remains largely unknown. *INAPERTURATE POLLEN1* (*INP1*) encodes a protein of unknown function, essential for aperture formation in *Arabidopsis*, rice and maize. Yet, because INP1 sequences are quite divergent, it is unclear if their function is conserved across angiosperms. Here, we conducted a functional study of the *INP1* ortholog from the basal eudicot *Eschscholzia californica* (*EcINP1*) using expression analyses, virus-induced gene silencing, pollen germination assay, and transcriptomics. We found that *EcINP1* expression peaks at the tetrad stage of pollen development, consistent with its role in aperture formation, which occurs at that stage, and showed, via gene silencing, that the role of *INP1* as an important aperture factor extends to basal eudicots. Using germination assays, we demonstrated that, in *Eschscholzia*, apertures are dispensable for pollen germination. Our comparative transcriptome analysis of wildtype and silenced plants identified over 900 differentially expressed genes, many of them potential candidates for the aperture pathway. Our study substantiates the importance of *INP1* homologs for aperture formation across angiosperms and opens up new avenues for functional studies of other aperture candidate genes.

Keywords: *Eschscholzia californica*, *INAPERTURATE POLLEN1*, Papaveraceae, pollen, pollen aperture, RNA-seq, transcriptome analysis, VIGS

1. Introduction

Pollen, the male gametophyte of spermatophytes, is surrounded by a robust, sporopollenin-based pollen wall, called exine, which isolates and protects it from the external environment (Ariizumi and Toriyama, 2011). Although most of the pollen surface is covered by exine, in many plants certain regions of the pollen surface receive little to no exine deposition. These regions, known as pollen apertures, represent some of the most characteristic and well-defined elements of the pollen surface (Furness and Rudall, 2004; Zhou and Dobritsa, 2019). Apertures often serve as the sites for pollen tube exit (Heslop-Harrison and Heslop-Harrison, 1985; Edlund *et al.*, 2004; Edlund *et al.*, 2016). They are also involved in the exchange of water and solutes with the medium and allow the rigid exine to adjust to changes in pollen volume due to dehydration/rehydration during pollination (Heslop-Harrison, 1979).

Numerous studies have been carried out to characterise the large diversity of aperture patterns in angiosperms (e.g. Wodehouse, 1935; Blackmore *et al.*, 1995; Furness and Rudall, 2004; Matamoro-Vidal *et al.*, 2016; Pérez-Gutiérrez *et al.*, 2015; Wortley *et al.*, 2015). Although the number, morphology and position of apertures often vary among species, these attributes usually remain stable at the intraspecific level, suggesting a genetic control of aperture patterning which changed multiple times during the evolution of flowering plants. However, few studies so far have probed the molecular mechanisms of aperture formation.

Genetic screening in *Arabidopsis thaliana* has made it possible to identify mutants defective in aperture formation (Dobritsa *et al.,* 2011; Lee *et al.*, 2018; Plourde *et al.*, 2019). One of the genes discovered in these screens, *INAPERTURATE POLLEN1* (*INP1*), encodes a protein of unknown function that is essential for aperture formation (Dobritsa and Coerper, 2012). In the *Arabidopsis inp1* mutant, pollen completely lacks apertures, producing inaperturate phenotype.

During pollen development, *Arabidopsis INP1* (*AtINP1*) first becomes expressed in pollen mother cells, where its protein is evenly distributed in the cytoplasm. Later, at the post-meiotic tetrad stage, INP1 accumulates at the specific plasma-membrane (PM) regions in each microspore (the predecessor of the pollen grain) and assembles into punctate lines at the sites where apertures will be formed (Dobritsa and Coerper 2012; Dobritsa *et al.*, 2018). *AtINP1* reaches peak expression at the tetrad stage, and after the release of microspores from tetrads its expression quickly disappears. At the aperture

PM domains, AtINP1 appears to localize at the interface between the PM and the callose wall surrounding the tetrad of microspores, where it may act as a bridge, keeping the aperture domains of each microspore attached to the callose wall to prevent sporopollenin deposition at these regions (Dobritsa *et al.*, 2018). It is not known whether AtINP1 interacts with other proteins to form these bridges (Zhou and Dobritsa, 2019).

Although both aperture patterns and the INP1 protein sequences are quite divergent across angiosperms, the involvement of INP1 proteins in the process of aperture formation seems to be conserved, as indicated by the loss of apertures in the *inp1* mutants of maize and rice (Li *et al.*, 2018; Zhang *et al.*, 2020). Normally, these two monocot species have the ulcerate aperture pattern characteristic of the Poaceae family, with a single pore-like polar aperture, very different from the typical eudicot tricolpate pattern of *Arabidopsis* pollen, with its three meridional furrows. Also, while the INP1 protein sequences are very similar between maize and rice, they share only ~35% sequence identity with AtINP1 (Dobritsa and Coerper, 2012; Li *et al.*, 2018). It was, therefore, surprising that these quite divergent proteins are all involved in the same process. Like AtINP1, the rice INP1 (OsINP1) localizes to the aperture PM domains in microspores, assembling into a single ring-like structure that pre-marks the position of a single pore-like aperture (Zhang *et al.*, 2020). Interestingly, although *inp1* pollen in *Arabidopsis* is fertile, *inp1* pollen in rice and maize loses its ability to germinate pollen tubes and becomes completely sterile, demonstrating that apertures in these species, but not in *Arabidopsis*, are essential for pollen fertility.

Although the INP1 homologues from several species are all involved in the same process, they cannot readily substitute for each other. The sequence divergence among INP1 homologues was suggested to be responsible for the fact that only proteins from closely related species from Brassicaceae were able to restore apertures in the *Arabidopsis inp1* mutants during the interspecific complementation experiments, while the orthologues from Solanaceae, Papaveraceae, and Poaceae were unable to do it (Dobritsa and Coerper, 2012; Li *et al.*, 2018). Based on these results, the need for species-specific partners to cooperate with INP1 in the control of aperture formation was postulated (Li *et al.*, 2018). However, a functional analysis of the INP1 homologues from more plant lineages is necessary to ascertain that the role of INP1 as an aperture factor is conserved across angiosperms.

In this study, we investigated the functional conservation of the INP1 orthologue EcINP1 from *Eschscholzia californica* (California poppy), a species of Papaveraceae family. This is an early divergent family within the order Ranunculales, the lineage that first diverged from the large group of eudicots (Wang *et al*., 2009; Hoot *et al*., 2015). At the base of the eudicot clade there was a transition in the pollen grain aperture system, from pollen with one polar aperture –typical of basal angiosperms and monocots– to pollen with three apertures in equatorial positions, which may have been a key innovation involved in the success and diversification of eudicots (Furness and Rudall, 2004). Because of the phylogenetic position of Ranunculales, between monocots and the core eudicots, and its high pollen aperture system diversity (Blackmore *et al.*, 1995; Pérez-Gutiérrez *et al*., 2016; Zhang *et al*., 2017), the Ranunculales species provide the opportunity to study the conservation of genetic mechanisms involved in the aperture system and its evolution. Unlike pollen from the species in which INP1 was previously studied, pollen of *E. californica* develops between five and seven colpate apertures. AtINP1 and EcINP1 share only ~44% sequence identity, and it was previously shown that EcINP1 cannot substitute for AtINP1 in *Arabidopsis* (Li *et al*., 2018). However, it remained unclear if this was due to the EcINP1's dependence on additional factors from *E. californica* or because EcINP1 is no longer involved in aperture formation. Here, we 1) analysed patterns of *EcINP1* expression to demonstrate that it is also expressed in anthers during the tetrad stage of pollen development, 2) inactivated *EcINP1* with virus-induced gene silencing (VIGS) to reveal that the *INP1* involvement in aperture formation is conserved in *E. californica*, 3) tested whether pollen germination in this species requires the presence of apertures, and 4) identified genes whose expression was affected by the silencing of *EcINP1*, which might represent candidates involved in the aperture formation mechanisms.

2. Materials and Methods

2.1. Plant material

The seeds of *Eschscholzia californica* Cham. were purchased from the company Seedaholic (Cloghbrack, Clonbur, Galway, Ireland). They were sown in pots (9x9x9 cm^3) with universal substrate and vermiculite mixed in a 3:1 ratio and kept in a greenhouse at a temperature range between 26 ºC to 14 ºC under a light/dark cycle of

16/8 h. Each pot was fertilized once at the beginning of the experiment and watered every day.

2.2. DNA/RNA extraction, PCR amplification of *EcINP1* analysis of intraspecific variability, and phylogenetic analysis

Genomic DNA (gDNA) was isolated from fresh leaves using the NucleoSpin® Plant II Kit (Macherey-Nagel GmbH and Co., Germany), and total RNA was extracted using the NucleoSpin® RNA Plant kit (Macherey-Nagel GmbH and Co., Germany), following the manufacturer's instructions. 1 µg of RNA was reverse-transcribed to cDNA using SuperScript III Reverse Transcriptase (Invitrogen, Karlsruhe, Germany) and oligo(dT)18 (ThermoFisher Scientific, USA).

The full-length coding sequence of *EcINP1* was published in Li *et al*. (2018; ENA accession number LT840341). To study intraspecific variation of *EcINP1*, PCR amplification was carried out on both gDNA and cDNA from bulked individuals (5 individuals for gDNA and 5 for cDNA) using the primers EcaINP1-F and EcaINP1-R (Li *et al*., 2018; Supplementary Table 1). PCR products were cloned using the Strataclon blunt PCR cloning kit (Agilent Technologies, USA), and 17 cDNA and 10 gDNA clones were sequenced. In addition, we included four sequences from the RNA-seq data generated in our laboratory, the sequence published in Li *et al*. (2018), one complete sequence from the 1000 Plants Project (1KP; Wickett *et al*., 2014), and the genomic sequence retrieved from the *Eschscholzia* Genome Database (Hori *et al*., 2018). The nucleotide sequences were aligned in Bioedit (Hall, 2004) and, after intron removal from the gDNA sequences, the observed genetic distances (*p*-distance) were estimated with MEGA X (Kumar *et al*., 2018).

For phylogenetic analyses, nucleotide sequences from other Papaveraceae and different angiosperm species were obtained by BLAST searches from NCBI, 1KP, Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html), and Phytometasyn (Xiao *et al*., 2013; https://bioinformatics.tugraz.at/phytometasyn/) databases (Supplementary Table 2). Sequences from the same species with DNA identity of >95% were regarded as possible alleles, so only one sequence was considered. After translating the nucleotide sequences, all the protein sequences were aligned using the Clustal algorithm in Bioedit and adjusted manually. The aligned sequences were used to generate a Maximum Likelihood (ML) tree. The ML analysis was performed with the program

PhyML v3.0 (Guindon and Gascuel, 2003) through the web platform "ATCG: Montpellier Bioinformatics Platform", applying the substitution model JTT automatically selected by Smart Model Selection in PhyML (Lefort *et al*., 2017) and the Akaike Information Criterion (AIC; Akaike 1974). The search for the optimal tree was carried out using the subtree-pruning and regrafting algorithm from five random starting trees generated by the parsimony algorithm (Guindon *et al*., 2010). Branch support was assessed using the approximated likelihood ratio test (aLRT) with the Shimodaira-Hasegawa-like test interpretation (Anisimova and Gascuel, 2006; Guindon *et al*., 2010).

A ML phylogenetic analysis was also performed to establish the homology relationships of two genes of interest, identified by transcriptome analyses (see below), related to the AGC1 group kinases from *Arabidopsis*. This analysis was based on the protein sequences of the kinase domain of all AGC1 kinases from *Arabidopsis* and those sequences obtained by BLAST from the genomes of *E. californica* and *Papaver somniferum*. The sequences of the AGC3-group kinases from *Arabidopsis* were used as an outgroup. Sequence alignment and ML analysis were conducted as described above.

2.3. Semi-quantitative reverse transcription-polymerase chain reaction (semi-qRT-PCR) and quantitative RT-PCR (qRT-PCR)

The spatial pattern of expression of the *EcINP1* gene was tested by semi-qRT-PCR on vegetative organs, floral organs of flowers at anthesis, and immature and mature fruits. The temporal expression pattern in anthers at different stages of pollen development was tested by qRT-PCR. Stages of pollen development were determined by optical microscopy (Olympus-CX31, Japan) using glycerin gelatine with basic fuchsin (50 ml of glycerin, 7 g of gelatine, 1 g of phenol, a few crystals of basic fuchsin, and 42 ml of distilled water) pre-warmed to 30-35 ºC before staining pollen grains. RNA extraction and cDNA synthesis were performed as described above and cDNA pools were diluted to 50 ng/µl for subsequent semi-qRT-PCR and qRT-PCR analyses. qRT-PCR was also used to test for reduction in the *EcINP1* expression in the VIGS experiments, to validate the expression levels of eight genes of interest selected with the help of transcriptome analyses, and to analyse the temporal expression pattern of three of these eight genes throughout pollen ontogeny. *ACTIN* served as the reference gene for calculating the relative expression intensities in both semi-qRT-PCR and qRT-PCR analyses, using the $2^{\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). qRT-PCR was performed using the

FastGene IC Green 2x qPCR mix (NIPPON Genetics, Tokyo, Japan) according to the manufacturer's instructions, and the qTower 2.2 real-time PCR thermocycler (Analytik Jena, Germany). Gene-specific primers used for semi-qRT-PCR and qRT-PCR reactions were designed using the software Primer3 (Untergasser *et al*., 2012; Supplementary Table 1). All experiments were repeated with three biological and three technical replicates.

2.4. Western blot

Anthers of wild-type plants were collected when pollen grains were at the pre-meiosis, tetrad, microspore and mature stages. For protein extraction, the anthers were crushed in the extraction buffer (50 mM Tris-HCl, 9 M Urea, 1% (v/v) Triton X-100, pH 7.4), and the insoluble material was removed by centrifugation at 10,000 g for 10 min at 4 °C. The protein extracts were subjected to sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE). Protein profiles were determined by means of Stain-free technology using a Gel DocTM EZ System (Bio-Rad; Supplementary Figure 1), and were used as a loading control as described by Welinder and Ekblad (2011) and using AlphaView software (Cell Biosciences, Santa Clara, CA) for protein quantification. Proteins were electroblotted from the gel onto a PVDF membrane in a Semi-Dry Transfer Cell (Bio-Rad), and the membrane was blocked for 1 h at room temperature in Tris-buffered saline (TBS) buffer, pH 7.4 containing 5% (w/v) defatted milk and 0.1% (v/v) Tween-20. Immunodetection of EcINP1 was carried out by incubation with a polyclonal antibody SAN-ESK generated against the epitope ESKQEILKTVEKDLMVEIEE (synthesised by Davids Biotechnologie GmbH, Regensburg, Germany) diluted 1:500 in TBS buffer (pH 7.4) containing 0.3% (v/v) Tween-20, overnight at 4 °C. An HRP-conjugated anti-chicken IgY (Abcam, ref. ab97135, Cambridge, UK), diluted 1:2000, served as the secondary antibody. Protein bands were visualized in a C-Digit scanner (LI-COR Biotechnology, United States).

2.5. Virus induced gene-silencing

The tobacco rattle virus (TRV)-based system (Liu *et al*., 2002) was used for VIGS experiments. A pTRV2:*EcINP1* construct was made by cloning a 480-bp fragment of the *EcINP1* coding region, amplified with the Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific, USA) and primers EcaINP1-F2-EcoRI and EcaINP1-R2-BamHI (Supplementary Table 1), between the *Eco*RI and *Bam*HI restriction sites.

pTRV2:*EcINP1* as well as negative- and positive-control constructs (respectively, pTRV2:empty, without any insert, and pTRV2:*PDS*, with a fragment of the phytoene desaturase gene) were transformed into *Agrobacterium tumefaciens* strain pGV3101. Three-week-old plants were inoculated with *Agrobacterium* in the hypocotyl following Tekleyohans *et al*. (2013) and left at 4 ºC overnight, after which they were transferred to the greenhouse at a temperature range between 15-24 ºC. 78 plants were infected with pTRV2:*EcINP1*, 15 with pTRV2:*PDS*, and 30 with pTRV2:empty. In addition, 50 non-treated plants were grown to observe the wild-type phenotype.

Pollen phenotypes were observed by optical microscopy in the fully open flowers. For 55 plants, pollen was collected from three flowers and three anthers per flower and stained with glycerin gelatine with basic fuchsin. Phenotypes were classified as follows: wild-type, with normal 5-7-colpate apertures; inaperturate, without apertures; and affected, with abnormal apertures that were shorter or shallower than normal. In addition, anthers from three *EcINP1*-silenced plants and three wild-type plants were fixed according to Fernández *et al*. (1992) and pollen was observed with a scanning electron microscope (model SMT; Zeiss) at the Centro de Instrumentación Científica (University of Granada). Pollen was also stained with auramine O as previously described (Reeder *et al*., 2016) and observed with confocal microscopy (Nikon A1+, Japan).

In order to test the effectiveness of VIGS, the expression levels of *EcINP1* were measured by qRT-PCR on the first bud (2-3 mm; pollen at the tetrad/free-microspore stage) of four silenced (one with the affected phenotype and three with the inaperturate phenotype) and three non-silenced (pTRV2:empty; wild-type phenotype) plants, as described above.

2.6. *In vitro* pollen germination

Pollen grains were cultured for 2 h on plates with 5 ml of liquid germination medium (15 g sucrose, 30 mg $Ca(NO_3)_2$, 20 mg $MgSO_4$, 10 mg KNO_3, and 10 mg H_3BO_3 in 100 ml of distilled water) at 25 ºC under moist conditions and in the dark. For both wild-type and inaperturate pollen from three wild-type plants and two VIGS plants, we counted 100 pollen grains from three anthers per flower and three flowers per plant. The pollen count was carried out using an inverted optical microscope (ZEISS Axio Scope A1, Germany) and, after checking the normality of the data, an analysis of variance

(ANOVA) was carried out using IBM SPSS statistics (SPSS Inc., IBM Company, 2020).

2.7. Transcriptome sequencing, RNA-seq data analysis and annotation

RNA from anthers with pollen at the free-microspore stage was extracted from three inaperturate and three wild-type plants as described above. The extracts were sent to Macrogen Inc. (Korea), where libraries were prepared using the Illumina TruSeq Stranded Total RNA Sample Preparation Kit with Ribo-Zero Plant and sequenced on an Illumina Hiseq 2500 platform as 150 nucleotides paired-ends. Raw data was generated using the Illumina package bcl2fastq.

Raw single reads (in FASTQ format) were subjected to sequence quality control using FastQC v0.11.8 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adaptors and low-quality sequences were removed from the data set using TrimGalore v0.6.4 (https://github.com/FelixKrueger/TrimGalore) and a minimum quality score of 20. The library's characteristics and trimming efficiency were checked and reads aligned to the genome of *Eschscholzia californica* (ftp://ftp.kazusa.or.jp/pub/eschscholzia/ECA_r1.0.cds.fa.gz) using the software HISAT2 v2.1.0 (https://ccb.jhu.edu/software/hisat2/; Pertea *et al.*, 2016). Transcripts were assembled and quantified using the software StringTie v2.0 (https://ccb.jhu.edu/software/stringtie/), with the merge option (Pertea *et al.*, 2016). Analysis of differentially expressed genes (DEGs) was performed with the DESeq2 R package (v1.24.0) using the coverage produced by StringTie. DEGs were annotated by BLASTX search against the genome of *Papaver somniferum* (https://www.ncbi.nlm.nih.gov/assembly/GCF_003573695.1) and the SwissProt Database (https://data.broadinstitute.org/Trinity/Trinotate_v3_RESOURCES/uniprot_sprot.pep.gz), using default parameters and extracting only the top hit for each sequence. To assign a function to each DEG, annotated DEGs were further annotated with Gene Ontology (GO, Ashburner *et al*., 2000) and Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa and Goto, 2000) databases using the Blast2GO v5.2.5 application (Götz *et al*., 2008) and the GhostKoala mapping tool (Kanehisa *et al*., 2016), respectively. To predict putative plant transcription factors (TF), coding sequences were aligned to TF domains from the Plant Transcription Factor Database (PlantTFDB, Zhang *et al.,* 2011).

To validate the gene expression results, qRT-PCR experiments were performed for eight DEGs. These genes were selected because their predicted functions make them interesting candidates for a role in pollen aperture formation. The primer sequences used are listed in Supplementary Table 1.

2.8. Data Availability

All raw sequences for transcriptome are available in the European Nucleotide Archive (ENA) under the accession numbers ERS6376182-ERS6376184 for wild-type samples and ERS6376185-ERS6376187 for VIGS-treated samples.

3. Results

3.1 Intraspecific variability and phylogeny of *EcINP1*

INP1 is a novel protein whose structure and domain organization are not well understood. Previously, AtINP1 was divided into five regions, based on their evolutionary conservation and position relative to the single predicted domain in this protein, the domain of unknown function DOG1 (Li *et al*., 2018). To better understand the relative importance of different INP1 regions, we explored the extent of the intraspecific variation of the EcINP1 sequences. To this end, we amplified and analysed the gDNA and cDNA sequences from multiple plants of *E. californica*, as well as retrieved information from sequence databases. In total, we obtained 34 different *EcINP1* sequences: 31 by us (17 from cDNA, 10 from gDNA, and four from our RNA-seq data) and three from databases (two transcriptomic and one genomic sequences). After *in silico* translating these sequences, we found that five cDNA sequences had substitutions resulting in stop codons, and those were removed from the alignment and not analysed further. The combination of sequences obtained from cDNA and gDNA made it possible to determine the location of an intron near the beginning of the coding sequence (as in *INP1* orthologues from several other species), as well as to discover a possible second intron, which was found in four cases near the very end of the coding sequence, suggesting the existence of a second possible transcript isoform (Supplementary Figure 2).

The intraspecific variation of *EcINP1* was low. The mean observed genetic distance between sequences was 0.92%, with a total of 34 variable nucleotide positions. In 11 sequences, an insertion of one triplet (AAT) was present at the beginning of the

acidic domain (one of the protein regions proposed for AtINP1 by Li et al., 2018). All pairwise comparisons were always at more than 97.5% sequence identity. At the protein level, the number of variable positions was 17, distributed within the protein regions as follows: 3 N-terminal; 5 DOG1; 3 acidic; 2 middle; 4 C-terminal. The sequences with the extra triplet in the acidic region contained an extra asparagine, and one of the transcripts included three additional amino acidic residues at the end of the C-terminal domain (Supplementary Figure 2, sequence 5).

Curiously, in our intraspecific library of *EcINP1* we uncovered a sequence with significant homology to multiple dispersed *EcINP1* fragments that also included 255 bp of the ubiquitin-like domain-containing CTD phosphatase (Supplementary Figure 3, *EcUBCL1*). Portions of this sequence correspond to fragments from the *5'UTR*, the N-terminal region, the DOG1 domain, the acidic region, as well as portions of the C-terminal region and the 3'UTR of *EcINP1* (Supplementary Figure 3). This sequence was also found by BLAST searches in the GenBank (JG611242), the 1KP (TUHA-2032598, NJKC-2028667, RKGT-2014503, EVOD-2107709), and the *Eschscholzia* Genome (Eca_sc001433.1_421821..424882) databases. We performed a VIGS experiment to silence this sequence, using a 493 bp region that included the ubiquitin-like domain-containing CTD phosphatase region (Supplementary Figure 3; see Supplementary Table 1 for primer sequences), but did not observe any abnormal phenotypes in the treated plants (data not shown), so the functional significance of this chimeric sequence remains unknown.

To verify the relationships of the *EcINP1* gene with *INP1* homologues from other plants, we conducted a phylogenetic analysis using sequences from other members of Papaveraceae and from other angiosperm species. The ML tree shows that EcINP1 is grouped with the rest of INP1 sequences from Papaveraceae, which all fall into the Ranunculales clade, together with sequences from Ranunculaceae and Berberidaceae (Figure 1). In general, the INP1 tree follows a pattern according to the angiosperm phylogeny.

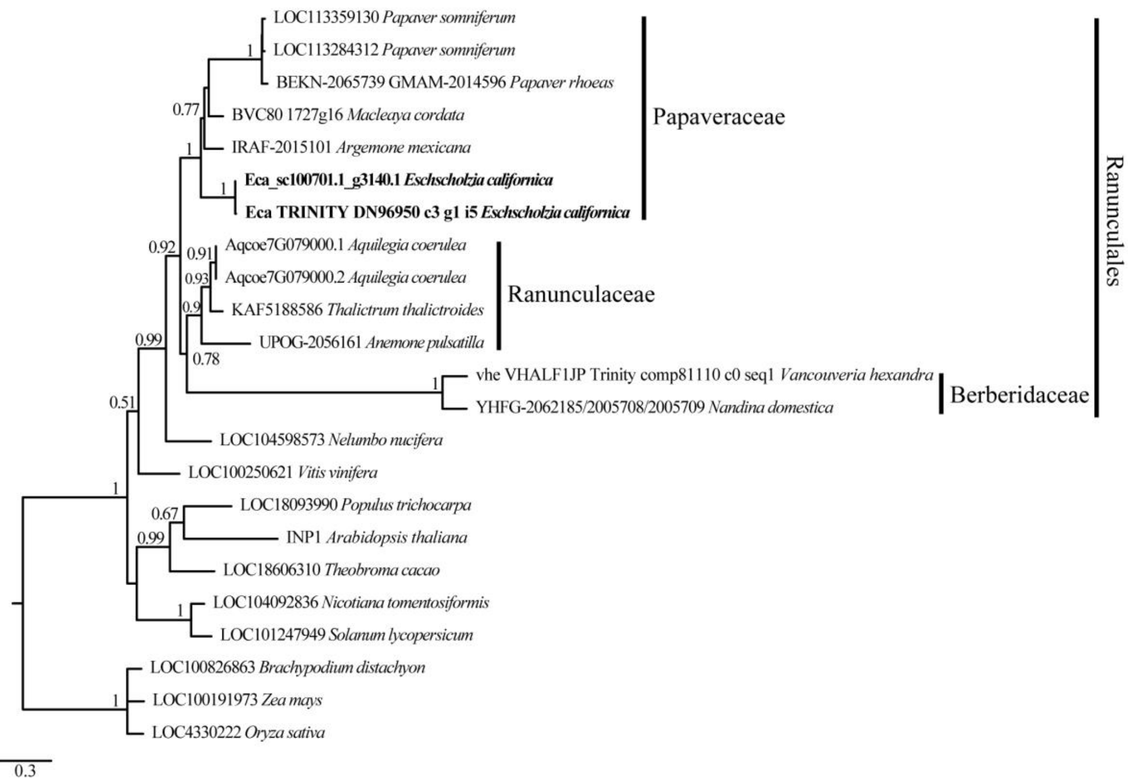


Figure 1. Maximum likelihood phylogenetic tree of INP1-like protein sequences showing EcINP1 homology relationships of *Eschscholzia californica*. The name of each sequence is composed of the locus identifier followed by the species name. Sequence accession numbers can be found in Table S2. The EcINP1 sequences representing two protein isoforms are in bold type. The families to which the selected sequences from the order Ranunculales belong are indicated on the right.

3.2. EcINP1 is expressed in anthers during pollen aperture development

To test where EcINP1 is expressed, we performed a semi-qRT-PCR on multiple plant organs. *EcINP1* transcripts were found in all tested vegetative, floral, and fruit organs. The highest signal intensity was observed in stamens, followed by the gynoecium (Figure 2A).

We then used qRT-PCR to assess the pattern of temporal expression of *EcINP1* in anthers during different stages of pollen ontogeny (Figure 2B). The *EcINP1* expression starts at the end of the microsporogenesis process, during the pre-mother cell stage. It reaches maximal levels during the tetrad stage, and essentially disappears during the microspore stage. We have also developed an antibody against an epitope in the middle region of EcINP1. Western blot analysis on protein extracts from anthers at different stages revealed a main band with a molecular weight corresponding to the ~35 kDa expected for EcINP1, and a secondary band that might represent protein dimers.

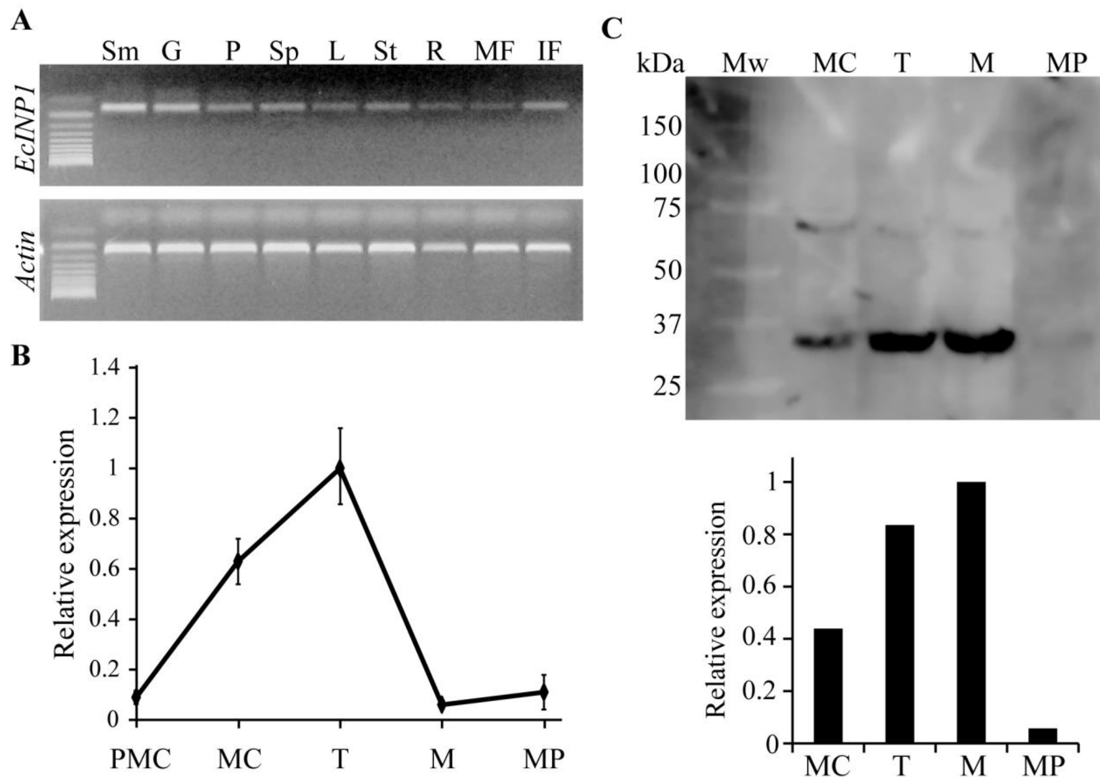


Figure 2. Expression patterns of *EcINP1* of *Eschscholzia californica*. **(a)** Expression analysis of *EcINP1*, based on semi-quantitative RT-PCR, in different plant organs. Sm, stamen; G, gynoecium; P, petal; Sp, sepal; L, leaf; St, stem; R, root; MF, mature fruit; IF, immature fruit. **(b)** qRT-PCR-based expression analysis of *EcINP1* in anthers at different pollen development stages. Actin was used as a normalization control. **(c)** Detection of EcINP1 protein levels by Western blot immunoassay in anthers at different pollen development stages. The bar plot shows the relative expression of EcINP1 (normalized to total proteins detected with a Stain-free technology gel as a loading control and quantified with the AlphaView software). PMC, pre-mother cell; MC, mother cell; T, tetrad; M, microspore; MP, mature pollen.

This analysis showed that EcINP1 protein was highly expressed at the tetrad stage and was still present at high concentration at the free-microspore stage (Figure 2C), suggesting that the protein may persist longer than the transcript. Taken together, these results indicate that EcINP1 is produced in developing pollen at the stages concurrent to aperture development.

3.3. *EcINP1* is required for the formation of pollen apertures

To test whether the INP1 involvement in aperture formation is conserved in *E. californica*, we attempted to silence *EcINP1* using the tobacco rattle virus (TRV)-based VIGS system (Liu *et al*., 2002). To this end, we infected plants with the pTRV2:*EcINP1* construct containing a 480-bp fragment of the *EcINP1* coding region.

The silencing of *EcINP1* did not affect vegetative growth, branching, or leaf morphology in any of the 78 plants treated with pTRV2:*EcINP1*, compared to controls (untreated plants and plants treated with the empty vector). 55 of the 78 pTRV2:*EcINP1* plants flowered and did not show any variation in the floral-organ phenotypes.

Yet, when we examined their pollen, we found that their aperture phenotypes showed significant abnormalities. We observed pollen from 139 flowers (34 plants produced three flowers, 16 plants – two flowers, and 5 plants – one flower), collecting three anthers per flower, so, in total, pollen from 417 anthers was analysed. Within each anther, all pollen presented the same phenotype, but, occasionally, the phenotypes differed between pollen grains from different anthers of the same flower. 33.3% of the observed anthers in pTRV2:*EcINP1* plants presented pollen with some aperture defect (Table 1), compared to 0% in wild-type plants or in those infected with the empty vector. In pTRV2:*EcINP1* plants, 15.3% of anthers had pollen that completely lacked apertures (inaperturate phenotype), while 18% of anthers showed pollen with apertures that were shorter or shallower than normal (affected phenotype) (Table 1; Figure 3A). This result provides strong evidence that, like its homologues from *Arabidopsis*, rice and maize, EcINP1 is involved in the formation of pollen apertures.

Table 1. Overview of the observed phenotypes of *EcINP1* VIGS in *Eschscholzia californica*

pTRV2:*EcINP1* inoculated plants	78
Flowered plants	55
Observed anthers	417
VIGS phenotype pollen anthers	139 (33,3%)
Inaperturate pollen anthers	64 (15,3%)
Affected pollen anthers	75 (18%)
Aperturate pollen anthers	278 (66,6%)
pTRV2:empty inoculated plants	30
Flowered plants	24
Observed anthers	180
Aperturate pollen anthers	180 (100%)
Wild plants	50
Flowered plants	36
Observed anthers	246
Aperturate pollen anthers	246 (100%)

Since we tracked the developmental order of the flowers collected from each plant, we noticed that the occurrence of anthers with inaperturate pollen often correlated with the order of flower development, decreasing from the oldest flower (the second in development, where the first bud was used to extract RNA) to the youngest flower (the fourth) in the same silenced plant (Figure 3B). This was likely due to the decrease of the silencing effect as flowering progressed (Wege *et al*., 2007).

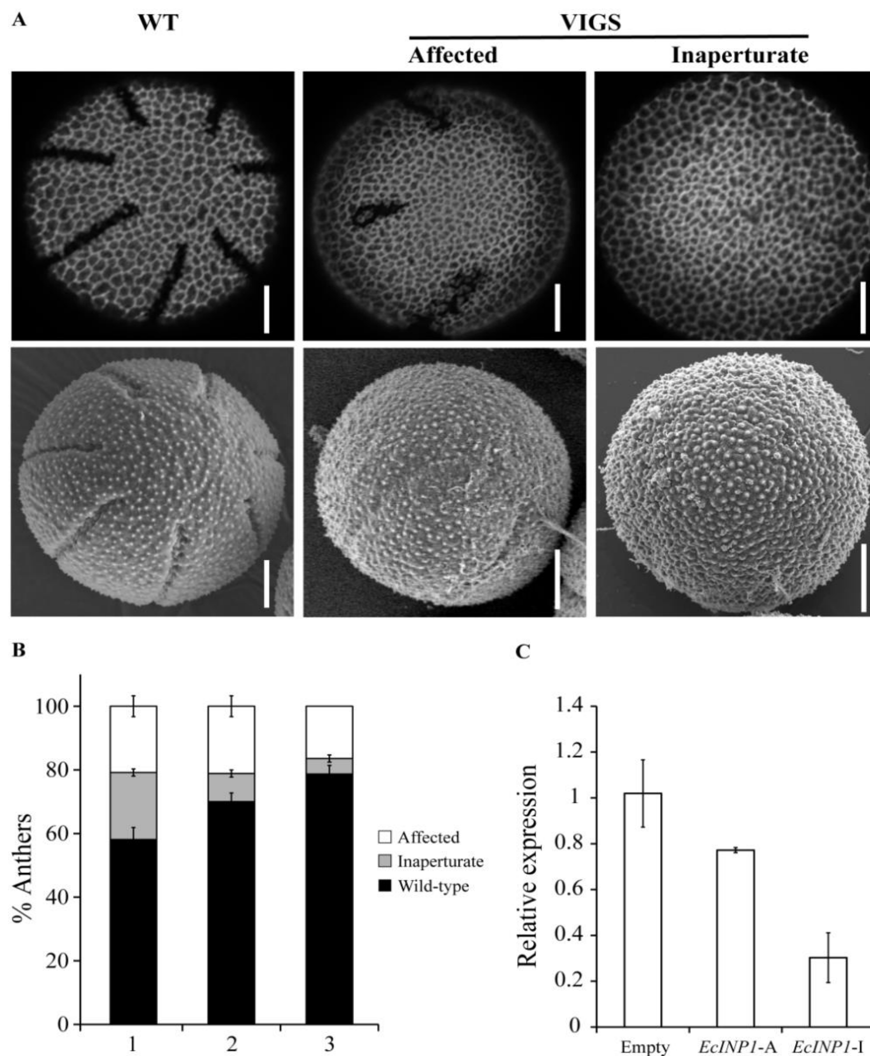
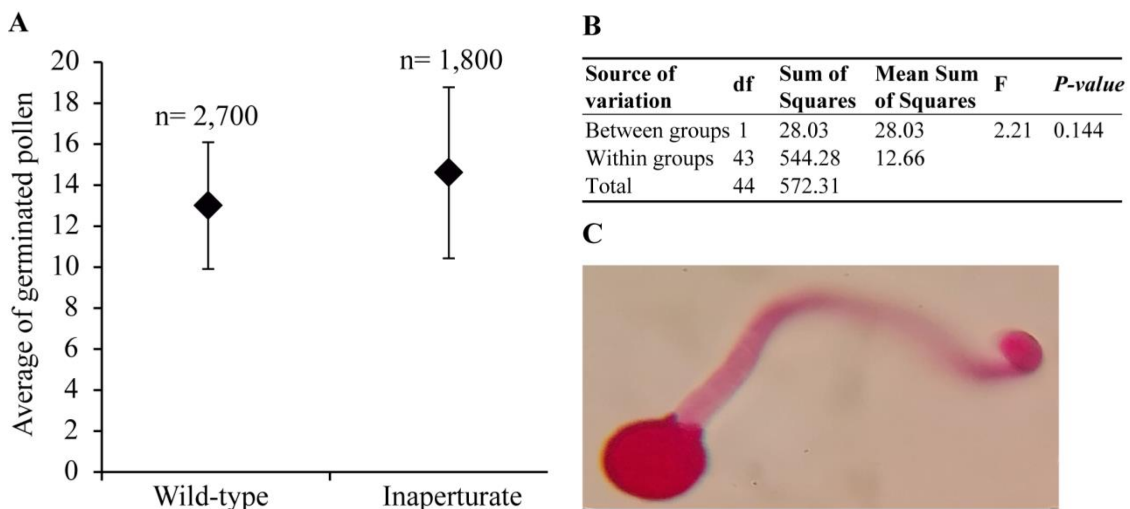


Figure 3. Results of *EcINP1* silencing in *Eschscholzia californica* by VIGS. **(a)** Images showing the observed pollen phenotypes: normal aperturate wild-type (WT), on the left; affected, with shorter and/or shallower apertures, in the middle; inaperturate, on the right. Top images were taken with a confocal microscope and bottom images with a scanning electron microscope. Scale bars = 0.5 □m **(b)** Stacked bar plot showing the distribution of observed phenotypes (black: wild-type, grey: inaperturate, white: affected) according to the developmental order of the three flowers (first = (1), second = (2), and third = (3)) sampled for each plant treated with pTRV2:*EcINP1*. Quantification of phenotypes is shown as the percentage of anthers with pollen of each phenotype, since within each anther all pollen grains showed the same phenotype. **(c)** Comparison of *EcINP1* expression, by qRT-PCR analysis of the first flower bud, in plants treated with pTRV2:empty (n=3) and pTRV2:*EcINP1* (n = 4). Student's *t* test *P*-value = 0.047.

qRT-PCR performed on young buds confirmed the downregulation of *EcINP1* in the pTRV2:*EcINP1* plants, with an approximately two-fold reduction in the *EcINP1* levels in these plants compared to the plants infected with the empty vector (Figure 3C).

3.4. Apertures in *E. californica* are dispensable for pollen germination

As mentioned earlier, pollen apertures are usually thought to serve as the sites for pollen tube exit during germination. Consistent with this notion, inaperturate pollen grains in rice and maize lose their ability to germinate (Li *et al*., 2018; Zhang *et al*., 2020). However, in *Arabidopsis*, *inp1* pollen still shows normal fertility in the absence of apertures (Dobritsa *et al*., 2011; Albert *et al*., 2018). Therefore, using an *in vitro* germination test, we tested whether apertures were essential or dispensable for pollen tube germination in *E. californica*. No significant difference was found between the germination of wild-type pollen and the inaperturate pollen from the silenced plants (Figures 4A,B). Thus, the presence of apertures is not essential for pollen germination, and pollen tubes in *E. californica*, like those in *Arabidopsis*, are capable of emerging directly through the pollen wall (Figure 4C).



Source of variation	df	Sum of Squares	Mean Sum of Squares	F	*P-value*
Between groups	1	28.03	28.03	2.21	0.144
Within groups	43	544.28	12.66		
Total	44	572.31			

Figure 4. Results of the *Eschscholzia californica* pollen germination test. **(a)** Scatter plot of the mean number of germinated pollen grains per each 100 pollen grains counted, for wild-type and inaperturate pollen. n, total number of pollen grains counted. **(b)** Summary table of the ANOVA analysis. **(c)** Optic microscope image of a germinated inaperturate pollen, stained with basic fuchsin, with the pollen tube emerging through the pollen wall.

3.5. Transcriptome in the *EcINP1*-silenced plants changes significantly

To determine if silencing of *EcINP1* could affect expression of any genes, we performed RNA-seq analysis of the anthers from the pTRV2:*EcINP1* VIGS plants and compared them with the results for the wild-type plants. Sequencing, using the Illumina platform with Phred quality score, yielded an average of 50,566,911 high-quality (HQ) reads per sample (Supplementary Table 3), after initial quality filtering. On average, 53.53% of reads were aligned to the *E. californica* reference genome and assembled into 34,729 contigs with an average length of 1,006 bp (Supplementary Table 3).

The hierarchical clustering, heatmap and principal component analyses showed that there were significant differences between the wild-type and VIGS samples (Figures 5A-C). The filtering of the genes with the $\log_2$-fold change of >2 and with the *p*-value and *p*-adj of < 0.05 identified 971 differentially expressed genes (DEGs), of which 488 were upregulated and 483 were downregulated. As expected, *EcINP1* itself was one of the DEGs, showing downregulation in the VIGS samples with a $\log_2$-fold change of 2.49, consistent with the result of the qRT-PCR. Out of the 971 DEGs, 211 were annotated through BLASTX with the genome of *Papaver somniferum*, 677 with SwissProt Database, and the remaining 83 did not give hits in any database (Supplementary Table 4).

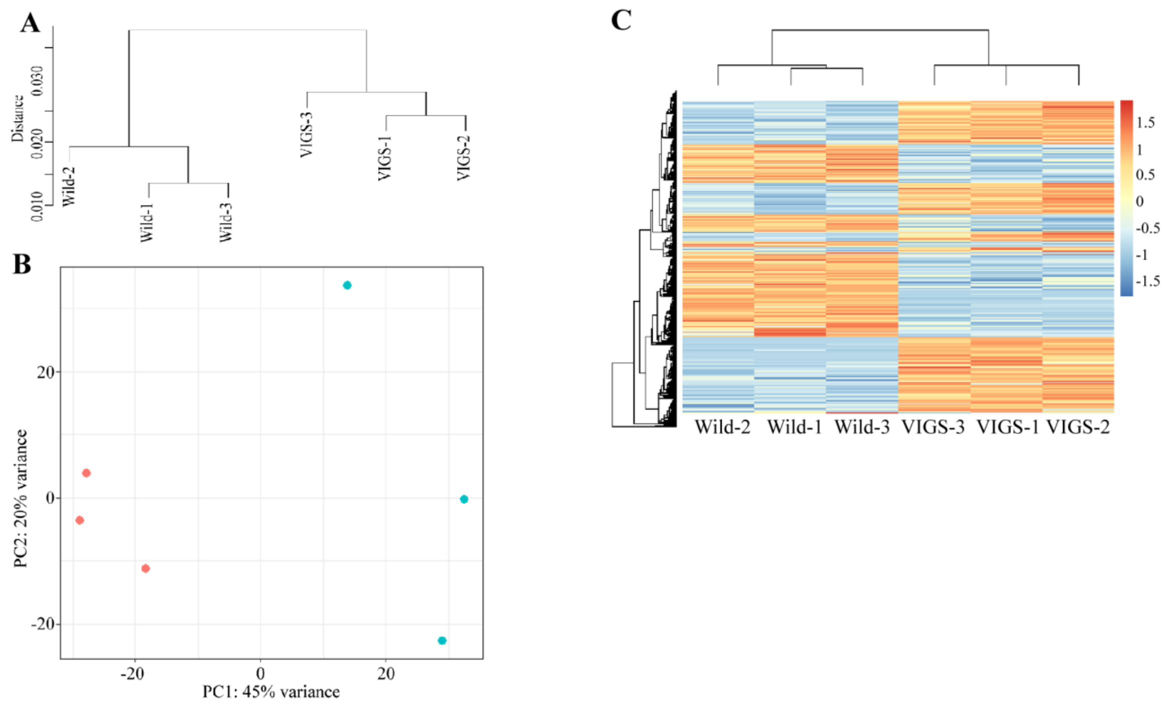


Figure 5. Analysis of differentially expressed genes (DEGs) between transcriptomes of wild-type and VIGS-treated plants of *E. californica*. **(a)** Hierarchical clustering shows dissimilarity among the transcriptome samples; distance is calculated by Pearson correlation coefficient. **(b)** Principal component analysis of the transcriptome samples. Red points represent wild-type plants; blue points represent pTRV2:*EcINP1*-treated plants. **(c)** Heatmap of transcriptomes of wild-type and pTRV2:*EcINP1*-treated plants. Heatmap scale bars indicate $\log_2$fold changes.

As transcription factors (TFs) play important roles in plant development, we used the Plant Transcription Factor Database (PlantTFDB) to align DEGs with TF domains. 30 DEGs showed high homology to 16 known TF families, among which the NAC family, with 8 genes, was the most represented (Supplementary Table 4; Supplementary Figure 4A). With respect to functional characterization, 277 (28.5%) of 971 DEGs were classified using GhostKoala into different functional categories, with proteins involved in processing of genetic information being the most common category, followed by proteins involved in carbohydrate metabolism and in signalling and cellular processes, all processes that could be relevant to aperture formation (Supplementary Figure 4B; Supplementary Table 4). We further describe the identity of some of the most interesting DEGs in the Discussion. For eight of those genes, their differential expression was confirmed with qRT-PCR (Figure 6).

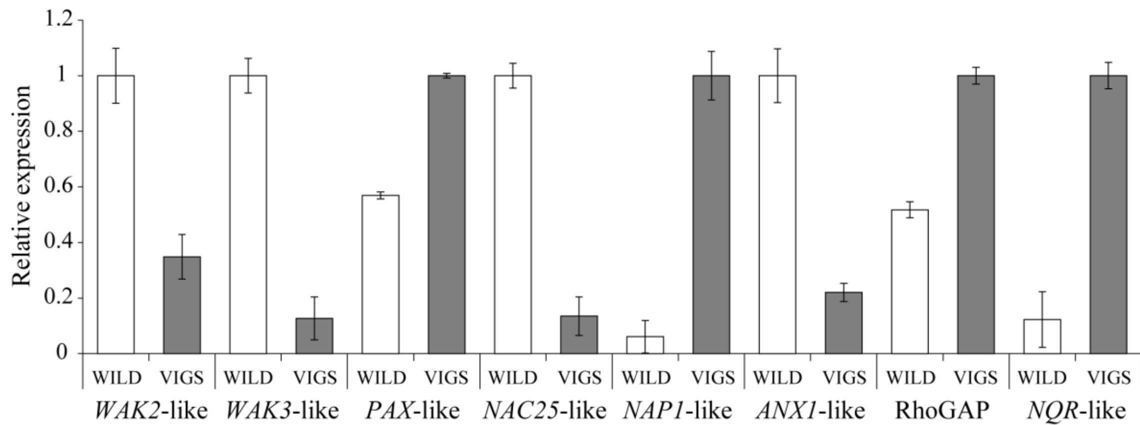


Figure 6. Results of the qRT-PCR analysis to confirm the differential expression of eight DEGs detected through the transcriptome analysis in *Eschscholzia californica*. *WAK2-like*, Eca_sc194563.1_g0010.1; *WAK3-like*, Eca_sc000051.1_g1410.1; *PAX-like*, Eca_sc095603.1_g0010.1; *NAC25-like*, Eca_sc194734.1_g0710.1; *NAP1-like*, Eca_sc004324.1_g2710.1; *ANX1-like*, Eca_sc194691.1_g0370.1; Rho GTPase-activating protein (*RhoGAP*), Eca_sc009796.1_g0090.1; NADPH:*quinone oxidoreductase* (*NQR)-like*, Eca_sc194522.1_g0180.1.

3.6. Phylogenetic and expression analyses of three interesting DEGs

Among the DEGs, we found two upregulated transcripts of the AGCVIII protein serine-threonine kinases from the AGC1 group (in the nomenclature of Galván-Ampudia and Offringa, 2007), putatively annotated as D6 PROTEIN KINASE-LIKE1 (D6PKL1; scaffold Eca_sc004639.1_g0030.1) and D6 PROTEIN KINASE-LIKE2 (D6PKL2; Eca_sc095603.1_g0010.1). In *Arabidopsis*, the protein D6 PROTEIN KINASE-LIKE3 (D6PKL3) belonging to the same group of kinases is involved in aperture formation (Lee *et al*., 2018). To clarify the homology relationships of the two D6PKL genes found in our DEG set, we carried out phylogenetic analysis on sequences of the AGC1 group kinases from *Arabidopsis*, *E. californica* and *Papaver somniferum*. The DEG sequence Eca_sc095603.1_g0010.1 was not included in the analysis due to its incompleteness in the database. However, it appears to be identical to Eca_sc000153.1_g1520.1, which was used instead.

This analysis showed two distinct clades within the D6PK family (Supplementary Figure 5), suggesting a duplication event and a posterior divergence before the eudicot diversification from their common ancestor, with one strongly supported clade, including the *Arabidopsis* D6PK, D6PKL1, D6PKL2 and their orthologues, and a more weakly supported clade, which included D6PKL3 and the rest of the *E. californica* and *P. somniferum* sequences. Eca_sc004639.1_g0030.1, affected

by the *EcINP1* silencing, belongs to this second clade, falling into a subclade unique to the Papaveraceae. The Papaveraceae sequences in the first clade are clearly related to D6PKL2, suggesting that an ancestral D6PKL2 sequence diverged from the lineage leading to D6PK-D6PKL1 in the ancestor of the eudicots. In *Arabidopsis*, D6PKL3 is the most diverged sequence in the D6PK family. In *E. californica*, the DEG sequence Eca_sc004639.1_g0030.1 is also clearly differentiated from the other members of the family (Supplementary Figure 5). The second D6PK family-related DEG, Eca_sc000153.1_g1520.1, is clearly linked to a close relative of the D6PK family, the *Arabidopsis* gene At2g44830, encoding PROTEIN KINASE ASSOCIATED WITH BRX (PAX) (Supplementary Figure 5).

Another DEG we considered interesting was the homologue of *NUCLEOSOME ASSEMBLY PROTEIN 1* (*NAP1*). In dioecious wild grapevine *Vitis vinifera* subsp. *sylvestris*, this gene, together with the *INP1* homologue, is part of the sexual locus whose members are functional/present in the male but not the female genome (Badouin *et al*., 2020). To find out if NAP1 and the two AGC1 kinases (D6PK-like and PAX-like) could be acting simultaneously with INP1, we analyzed their transcript expression at different stages of pollen development. The results showed that all three genes reach maximal expression levels during the tetrad stage (Supplementary Figure 6), mimicking the expression pattern of *EcINP1* (Figure 2B).

4. Discussion

The molecular mechanisms involved in the formation of pollen apertures remain largely unknown. Studies in this area have so far been mostly carried out in two model species, the core eudicot *A. thaliana* (Brassicaceae, with tricolpate pollen) and the monocot *O. sativa* (Poaceae, with ulcerate pollen). This study is the first functional characterization of an *INP1*-like gene in a basal eudicot, *E. californica* (Papaveraceae, Ranunculales, with penta- to heptacolpate pollen), and therefore, in the evolutionary context, it complements our understanding of the role of this gene in the diversification of the angiosperm apertural system. Moreover, this is the first study in which a comparative transcriptome analysis was performed for identifying potential candidate genes involved in aperture formation, calling for further functional studies.

Our functional study showed that *EcINP1* is required for the formation of apertures in *E. californica*, similar to its homologues in *Arabidopsis,* maize and rice.

This result extends the conserved role of the gene to basal eudicots and supports the hypothesis that INP1 was involved in the aperture formation in the monocot-eudicot common ancestor (Li *et al*., 2018). In the interspecies complementation experiments conducted by Li *et al*. (2018), EcINP1 failed to restore apertures in the *Arabidopsis inp1* mutants. Our results now demonstrate the functionality of EcINP1 and support sequence divergence as the cause of this failure.

At the intraspecific level, we have detected a single copy of *EcINP1* in *E. californica*, as deduced from BLAST searches in the *Eschscholzia* genome database and from the high identity among the observed variants of the gene (> 97.5%). One of the most variable regions of the gene and protein is the very end of the C-terminus, where we even detected a second intron and a shift in the reading frame resulting in extra amino acids (Supplementary Figure 2). Li *et al*. (2018) showed that this region was dispensable for the formation of the punctate INP1 lines and apertures in *Arabidopsis*, and that it was poorly conserved among the INP1 homologues from other species. This evidence suggests that the low functional significance of the C-terminal region leads to relaxation of the selective pressure acting on it, allowing its divergence at the sequence level even within a species. At the interspecific level, this region is highly variable and had to be excluded from the phylogenetic analysis due to alignment ambiguity. In addition, a 30-bp region located in the acidic region was also excluded because of its high variability. This region, together with the DOG1 domain and the middle region, is part of the INP1 central region, essential for the function and stability of the protein and containing amino acids critical for species-specific interactions (Li *et al*., 2018). However, whether the acidic region itself is involved in species-specific interactions is not known. At the intraspecific level, we have detected moderate variation in the acidic region, with three amino acid substitutions and an indel of an asparagine. We have also detected intraspecific changes in other important areas of the central portion of EcINP1 (five in the DOG1 domain and two in the middle region). Further studies will be required to understand whether these changes in the regions critical for species-specific interactions could be related to the observed variability of the pollen aperture system in *E. californica* (e.g., in the number of apertures or their length).

The gene expression pattern of *INP1* homologues during microsporogenesis also seems to be conserved across angiosperms. Similar to its counterparts in *Arabidopsis* and rice (Dobritsa and Coeper, 2012; Zhang *et al*., 2020), *EcINP1* begins expression in

pollen mother cells, reaches the maximum level at the early tetrad stage, and, after the release of microspores, its transcript practically disappears. In *Arabidopsis*, the AtINP1 protein signal disappears soon after the release of microspores from the tetrad, suggesting rapid degradation of the protein (Dobritsa and Coerper, 2012; Dobritsa *et al*., 2018). The EcINP1 protein, unlike AtINP1, reaches its maximum concentration during the free microspore period (Figure 2C), so degradation of the protein does not occur until after that stage. A delay in the degradation of OsINP1 has also been documented in rice and could be associated with the functional diversification of that protein, which, together with the lectin receptor-like kinase (RLK) DEFECTIVE IN APERTURE FORMATION1 (OsDAF1) with which it interacts, is involved in the formation of the pore annulus (Zhang *et al*., 2020). Thus, a delay in the EcINP1 degradation beyond aperture formation may suggest a possibility of its functional diversification.

Our pollen germination assay showed that apertures in *E. californica*, like those in *Arabidopsis*, are not essential for the exit of pollen tubes, and that pollen tubes can break through exine in inaperturate pollen grains. This is different from grasses, in which inaperturate mutants fail to germinate pollen tubes and show complete male sterility (Li *et al*., 2018; Zhang *et al*., 2020). Differences in exine morphology (thickness and tectum sculpture; Li *et al*., 2018) as well as in physiology of pollen and stigma (Edlund *et al*., 2016) have been proposed as possible causes for differences in dependence of pollen tubes on the presence of apertures. Inaperturate sterile pollen in dioecious species, produced by female flowers and acting as a reward or attractant for pollinators, is a character that has independently evolved at least six times among eudicots (Furness, 2007). Recently, Badouin *et al*. (2020) found that in female flowers of *Vitis vinifera* subsp. *sylvestris* the *INP1* homologue has an 8-bp deletion in the DOG1 domain, which results in a premature stop codon and is probably responsible for the absence of apertures. However, they did not study whether the lack of apertures in this pollen is sufficient to cause sterility. In this taxon, *INP1* is located within the sex locus, where four other genes, present in male plants, are missing in the female plants and could also be the candidates for pollen sterility. Interestingly, we discovered one of these four genes, *NUCLEOSOME ASSEMBLY PROTEIN 1* (*NAP1*), as part of the DEG set in the *EcINP1*-silenced plants, suggesting that a possibility of its interaction with *INP1* should be examined. Consistent with this notion, we found that *EcNAP1*-like is most strongly expressed during the tetrad stage of pollen ontogeny (Supplementary

Figure 6A), when *EcINP1* expression is also maximal and when the process of aperture development begins.

INP1 is an essential factor for the development of apertures, but it is not the main factor defining aperture number, positions, and morphology (Reeder *et al*., 2016; Dobritsa *et al*., 2018; Zhou and Dobritsa, 2019). Still, in *Arabidopsis*, it has been observed that there is a relationship between the INP1 levels and aperture length, with the lower transcript levels correlating with shorter apertures (Dobritsa and Coerper, 2012). In *E. californica*, this relationship also appears to exist, since older VIGS-silenced plants, in which *EcINP1* was likely inactivated only partially, often produced pollen with apertures that were shorter or shallower than normal (Figure 3B). However, in no case were the changes in number, position or shape of apertures observed.

Although accumulating evidence indicates the need for species-specific partners to cooperate with INP1 in different species to control the formation of apertures (Li *et al*., 2018; Zhou and Dobritsa, 2019; Zhang *et al*., 2020), these other molecular players remain largely unknown. Our transcriptome analysis presents a basis for identifying DEGs that may represent some candidates with which EcINP1 interacts. Among the 971 DEGs, we found two belonging to the AGC1 kinases: one related to D6PKL3 of *Arabidopsis* and the other to PAX. In *Arabidopsis*, the four membrane-associated kinases of the D6PK family directly regulate the PIN-FORMED (PIN)-mediated auxin transport required for phototropic responses (Zourelidou *et al*., 2009). However, D6PKL3 is also involved in pollen aperture formation (Lee *et al*., 2018). D6PKL3 appears to act upstream of INP1, possibly specifying domains in the PM to indicate the sites where INP1 must attach. At the same time, INP1 also seems to control D6PKL3 localization at the aperture domains (Lee *et al*., 2018). Similar to the D6PK-family proteins, PAX regulates the activity of PIN1 in developing protophloem sieve elements (Marhava *et al*., 2018). Marhava *et al*. (2020) have shown that PAX and its partner BREVIS RADIX (BRX) influence the local abundance of PIN1 by recruiting phosphatidylinositol-4-phosphate 5-kinases (PIP5Ks) to partition the PM into distinct domains. In the *Arabidopsis* aperture formation process a possible link between AGC1 kinases and phosphatidylinositol lipids in the aperture PM domains was also proposed (Lee *et al*. 2018). The upregulation of the transcripts of these AGC1 kinases in the *E. californica* VIGS-silenced plants suggests their direct or indirect interaction with EcINP1. The temporal expression pattern of these kinases, coincident with that of

EcINP1 (Supplementary Figures 6B,C; Figure 2B), provides further evidence pointing towards the connection between these genes in determining the formation of apertures. Thus, further studies will be necessary to test possible roles of D6PK-like and PAX-like in the establishment of the PM domains of future apertures and EcINP1 polarisation.

Among the annotated DEGs, we also found several RLKs. In rice, the above-mentioned lectin RLK OsDAF1 was recently shown to be involved in the formation of pollen apertures (Zhang *et al*., 2020). Interestingly, one of the upregulated DEGs was annotated as a G-type lectin S-receptor-like serine/threonine-protein kinase (homologous to the *Arabidopsis* At2g19130). Also, among the downregulated genes we found homologues of two *WALL-ASSOCIATED RECEPTOR KINASES* (*WAK*), *WAK2* and *WAK3* (Figure 6), encoding cell wall-associated RLKs (He *et al*., 1999). In *Oryza sativa,* a WAK-RLK gene *DEFECT IN EARLY EMBRYO SAC1* (*OsDEES1*) was shown to have some effect on pollen viability and pollen tube growth (Kohorn, 2001; Wang *et al*., 2012).

In addition, ten other RLKs were found in our DEG set. Among them was a downregulated homologue of the pollen-specific RLK *ANXUR1* (*ANX1*) (Figure 6). In *Arabidopsis*, ANX1, together with its homologue ANX2, controls cell wall integrity and pollen tube rupture by regulating pollen-expressed NADPH oxidases, as well as exocytosis and secretion of cell wall materials (Miyazaki *et al*., 2009; Boisson-Dernier *et al*., 2009; Fehér and Lajkó, 2015). Related to the possible role of ANX1 in aperture formation, we also found an upregulated homologue of *NADPH:quinone oxidoreductase* (*NQR*) (Figure 6), proposed to be a part of the ANX pathway (Fehér and Lajkó, 2015). Also, we found an upregulated DEG coding for the homologue of the leucine-rich repeat RLK PXY-LIKE1 (PXL1), very closely related to PXY which helps to maintain cell polarity required for the orientation of cell division during vascular development (Fisher and Turner, 2007). Since cell polarity in developing microspores likely plays a role in the establishment of aperture domains (Lee *et al*., 2018; Zhou and Dobritsa, 2019), polarity-related proteins could be good candidates for functional testing in the future. Additionally, among the upregulated DEGs, we found an uncharacterized Rho GTPase-activating protein (RhoGAP) (Figure 6). Rho GTPases, acting as molecular switches that cycle between the inactive cytosolic GDP-bound state and the active membrane-bound GTP-state, have been implicated in the control of cell

polarity, cellular domain formation, and cytoskeletal organization (Craddock *et al*., 2012), all of which could be important for the formation of aperture domains.

So far, nothing is known about the factors regulating expression of genes involved in pollen aperture formation. Among the annotated DEGs, we found 30 that correspond to transcription factors. Eight belong to the NAC family, whose members regulate many developmental processes in plants. One of them, downregulated in the VIGS plants (Figure 6), is annotated as a homologue of *NAC TRANSCRIPTION FACTOR 25* or *TAPNAC*, known to be expressed in the tapetum of the *Arabidopsis* anthers (Alvarado *et al*., 2011).

In summary, our study extends INP1 involvement in aperture formation to basal eudicots. This functional conservation is quite remarkable, given the low protein conservation of INP1 and the large variations in aperture patterns across angiosperms. There are many questions to be answered about the aperture pathway. Characterising the function of known genes in species from relevant angiosperm groups will allow a better assessment of the functional conservation of these genes in the phylogenetic scale of flowering plants. Of particular importance would be the identification of new factors involved in aperture formation. Characterisation of proteins interacting with INP1 and/or D6PKL3, as well as genes responsible for abnormal aperture systems, such as the recently discovered *macaron* and *doughnut* mutants (Plourde *et al.,* 2019), or the regulatory genetic network, would help to better understand the process of aperture formation and its evolution in angiosperms.

References

Akaike H. 1974. A new look at the statistical model identification. *IEEE transactions on automatic control* **19**: 716-723.

Albert B, Ressayre A, Dillmann C, Carlson AL, Swanson RJ, Gouyon PH, *et al.* 2018. Effect of aperture number on pollen germination, survival and reproductive success in *Arabidopsis thaliana*. *Annals of Botany* **121:** 733-740.

Alvarado VY, Tag A, Thomas TL. 2011. A cis regulatory element in the TAPNAC promoter directs tapetal gene expression. *Plant Molecular Biology* **75**: 129-139.

Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology* **55**: 539-552.

Ariizumi T, Toriyama K. 2011. Genetic regulation of sporopollenin synthesis and pollen exine development. *Annual Review of Plant Biology* **62**: 437-460.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al*. 2000. Gene ontology: tool for the unification of biology. *Nature genetics* **25**: 25-29.

Badouin H, Velt A, Gindraud F, Flutre T, Dumas V, Vautrin S, *et al*. 2020. The wild grape genome sequence provides insights into the transition from dioecy to hermaphroditism during grape domestication. *Genome Biology* **21**: 1-24.

Blackmore S, Stafford P, Persson V. 1995. "Systematics and evolution of the Ranunculiflorae," in Plant Systematics and Evolution, ed. U. Jensen, J.W. Kadereit (Viena. Springer), 71-82.

Boisson-Dernier A, Roy S, Kritsas K, Grobei MA, Jaciubek M, Schroeder JI, *et al.* 2009. Disruption of the pollen-expressed FERONIA homologs ANXUR1 and ANXUR2 triggers pollen tube discharge. *Development* **136**: 3279-3288.

Craddock C, Lavagi I, Yang Z. 2012. New insights into Rho signaling from plant ROP/Rac GTPases. *Trends in Cell Biology* **22**: 492-501.

Dobritsa AA, Coerper D. 2012. The novel plant protein INAPERTURATE POLLEN1 marks distinct cellular domains and controls formation of apertures in the Arabidopsis pollen exine. *The Plant Cell* **24**: 4452-4464.

Dobritsa AA, Geanconteri A, Shrestha J, Carlson A, Kooyers N, Coerper D, *et al*. 2011. A large-scale genetic screen in Arabidopsis to identify genes involved in pollen exine production. *Plant Physiology* **157**: 947-970.

Dobritsa AA, Kirkpatrick AB, Reeder SH, Li P, Owen HA. 2018. Pollen aperture factor INP1 acts late in aperture formation by excluding specific membrane domains from exine deposition. *Plant Physiology* **176**: 326-339.

Edlund AF, Swanson R, Preus D. 2004. Pollen and stigma structure and function: the role of diversity in pollination. *The Plant Cell* **16(suppl 1)**: S84-S97.

Edlund AF, Zheng Q, Lowe N, Kuseryk S, Ainsworth KL, Lyles RH, *et al*. 2016. Pollen from Arabidopsis thaliana and other Brassicaceae are functionally omniaperturate. *American Journal of Botany* **103**: 1006-1019.

Fehér A, Lajkó DB. 2015. Signals fly when kinases meet Rho-of-plants (ROP) small G-proteins. *Plant Science* **237**: 93-107.

Fernández MC, Romero García AT, Rodríguez García MI. 1992. Aperture structure, development and function in Lycopersicum esculentum Miller (Solanaceae) pollen grain. *Review of Palaeobotany and Palynology* **72**: 41-48.

Fisher K, Turner S. 2007. PXY, a receptor-like kinase essential for maintaining polarity during plant vascular-tissue development. *Current Biology* **17**: 1061-1066.

Furness CA, Rudall PJ. 2004. Pollen aperture evolution – a crucial factor for eudicot success? *Trends Plant Science* **9**: 154-158.

Furness CA. 2007. Why does some pollen lack apertures? A review of inaperturate pollen in eudicots. *Botanical Journal of the Linnean Society* **155**: 29-48.

Galván-Ampudia CS, Offringa R. 2007. Plant evolution: AGC kinases tell the auxin tale. *Trends in Plant Science* **12**: 541-547.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307-321.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**: 696-704.

Hall T. 2004. BioEdit version 7.0. 0. http://mbio.ncsu.edu/BioEdit/bioedit.

He ZH, Cheeseman I, He D, Kohorn BD. 1999. A cluster of five cell wall-associated receptor kinase genes, Wak1–5, are expressed in specific organs of Arabidopsis. *Plant Molecular Biology* **39**: 1189-1196.

Heslop-Harrison J, Heslop-Harrison Y. 1985. Surfaces and secretions in the pollen–stigma interaction: a brief review. *Journal of Cell Science* (**Supplement 2**): 287-300.

Heslop-Harrison J. 1979. Pollen walls as adaptive systems. *Annals of the Missouri Botanical Garden* **66**: 813-829.

Hoot SB, Wefferling KM, Wulff JA. 2015. Phylogeny and character evolution of Papaveraceae sl (Ranunculales). *Systematic Botany* **40**: 474-488.

Hori K, Yamada Y, Purwanto R, Minakuchi Y, Toyoda A, Hirakawa H, *et al*. 2018. Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. *Plant and Cell Physiology* **59**: 222-233.

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**: 27-30.

Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology* **428**: 726-731.

Kohorn BD. 2001. WAKs; cell wall associated kinases. *Current Opinion in Cell Biology* **13**: 529-533.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**: 1547-1549.

Lee BH, Weber ZT, Zourelidou M, Hofmeister BT, Schmitz RJ, Schwechheimer C, *et al.* 2018. Arabidopsis protein kinase D6PKL3 is involved in the formation of distinct plasma membrane aperture domains on the pollen surface. *The Plant Cell* **30**: 2038-2056.

Lefort V, Longueville JE, Gascuel O. 2017. SMS: smart model selection in PhyML. *Molecular Biology and Evolution* **34**: 2422-2424.

Li P, Ben-Menni Schuler S, Reeder SH, Wang R, Suárez Santiago VN, Dobritsa AA. 2018. INP1 involvement in pollen aperture formation is evolutionarily conserved and may require species-specific partners. *Journal of Experimental Botany* **69**: 983-996.

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

Marhava P, Bassukas AEL, Zourelidou M, Kolb M, Moret B, Fastner A, *et al*. 2018. A molecular rheostat adjusts auxin flux to promote root protophloem differentiation. *Nature* **558**: 297-300.

Marhava P, Fandino ACA, Koh SW, Jelínková A, Kolb M, Janacek DP, *et al*. 2020. Plasma membrane domain patterning and self-reinforcing polarity in Arabidopsis. *Developmental Cell* **52**: 223-235.

Matamoro-Vidal A, Prieu C, Furness CA, Albert B, Gouyon PH. 2016. Evolutionary stasis in pollen morphogenesis due to natural selection. *New Phytologist* **209**: 376-394.

Miyazaki S, Murata T, Sakurai-Ozato N, Kubo M, Demura T, Fukuda H, *et al*. 2009. ANXUR1 and 2, sister genes to FERONIA/SIRENE, are male factors for coordinated fertilization. *Current Biology* **19**: 1327-1331.

Pérez-Gutiérrez MA, Suárez-Santiago VN, Fernández MC, Salinas-Bonillo MJ, Romero-García AT. 2015. Pollen morphology and post-tetrad wall development in the subfamily Fumarioideae (Papaveraceae). *Review of Palaeobotany and Palynology* **222**: 33-47.

Pérez-Gutiérrez MA, Fernández MC, Salinas-Bonillo MJ, Suárez-Santiago VN, Ben-Menni Schuler S, Romero-García AT. 2016. Comparative exine development from the post-tetrad stage in the early-divergent lineages of Ranunculales: the genera Euptelea and Pteridophyllum. *Journal of Plant Research* **129**: 1085-1096.

Plourde SM, Amom P, Tan M, Dawes AT, Dobritsa AA. 2019. Changes in morphogen kinetics and pollen grain size are potential mechanisms of aberrant pollen aperture patterning in previously observed and novel mutants of Arabidopsis thaliana. *PLoS Computational Biology* **15**: e1006800.

Reeder SH, Lee BH, Fox R, Dobritsa AA. 2016. A ploidy-sensitive mechanism regulates aperture formation on the Arabidopsis pollen surface and guides localization of the aperture factor INP1. *PLoS Genetics* **12**: e1006060.

Tekleyohans DG, Lange S, Becker A. 2013. "Virus-induced gene silencing of the alkaloid-producing basal eudicot model plant Eschscholzia californica (California Poppy)," in Virus-Induced Gene Silencing, ed. A. Becker (Totowa, NJ: Humana Press): 83-98.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, *et al*. 2012. Primer3 – new capabilities and interfaces. *Nucleic Acids Research* **40**: e115-e115.

Wang H, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, *et al*. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences* **106**: 3853-3858.

Wang N, Huang HJ, Ren ST, Li JJ, Sun Y, Sun DY, *et al*. 2012. The rice wall-associated receptor-like kinase gene OsDEES1 plays a role in female gametophyte development. *Plant Physiology* **160**: 696-707.

Wege S, Scholz A, Gleissberg S, Becker A. 2007. Highly efficient virus-induced gene silencing (VIGS) in California poppy (Eschscholzia californica): an evaluation of VIGS as a strategy to obtain functional data from non-model plants. *Annals of Botany* **100**: 641-649.

Welinder C, Ekblad L. 2011. Coomassie staining as loading control in Western blot analysis. *Journal of Proteome Research* **10**: 1416-1419.

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, *et al*. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences* **111**: E4859-E4868.

Wodehouse RP. 1935. Pollen grains. Their structure, identification and significance in science and medicine. London: Mcgraw-Hill Book Company, Inc.

Wortley AH, Wang H, Lu L, Li DZ, Blackmore S. 2015. Evolution of angiosperm pollen. 1. Introduction. *Annals of the Missouri Botanical Garden* **100**: 177-226.

Xiao M, Zhang Y, Chen X, Lee EJ, Barber CJ, Chakrabarty R, *et al*. 2013. Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *Journal of Biotechnology* **166**: 122-134.

Zhang H, Jin J, Tang L, Zhao Y, Gu X, Gao G, *et al*. 2011. PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. *Nucleic acids research* **39**: D1114-D1117.

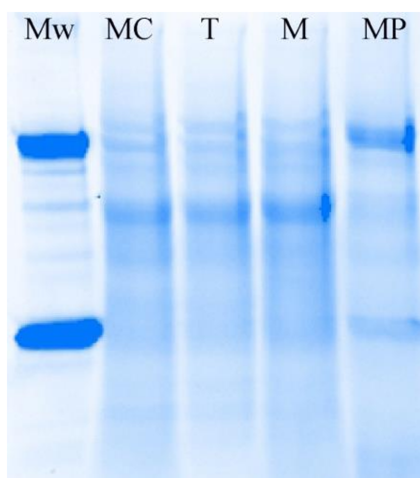
Zhang MY, Lu L, Wortley AH, Wang H, Li DZ, Blackmore S. 2017. Evolution of Angiosperm Pollen: 4. Basal Eudicots. *Annals of the Missouri Botanical Garden* **102**: 141-182.

Zhang X, Zhao G, Tan Q, Yuan H, Betts N, Zhu L, *et al*. 2020. Rice pollen aperture formation is regulated by the interplay between OsINP1 and OsDAF1. *Nature Plants* **6**: 394-403.

Zhou Y, Dobritsa AA. 2019. Formation of aperture sites on the pollen surface as a model for development of distinct cellular domains. *Plant Science* **288**: 110222.

Zourelidou M, Müller I, Willige BC, Nill C, Jikumaru Y, Li H, *et al.* 2009. The polarly localized D6 PROTEIN KINASE is required for efficient auxin transport in *Arabidopsis thaliana*. *Development* **136**: 627-636.

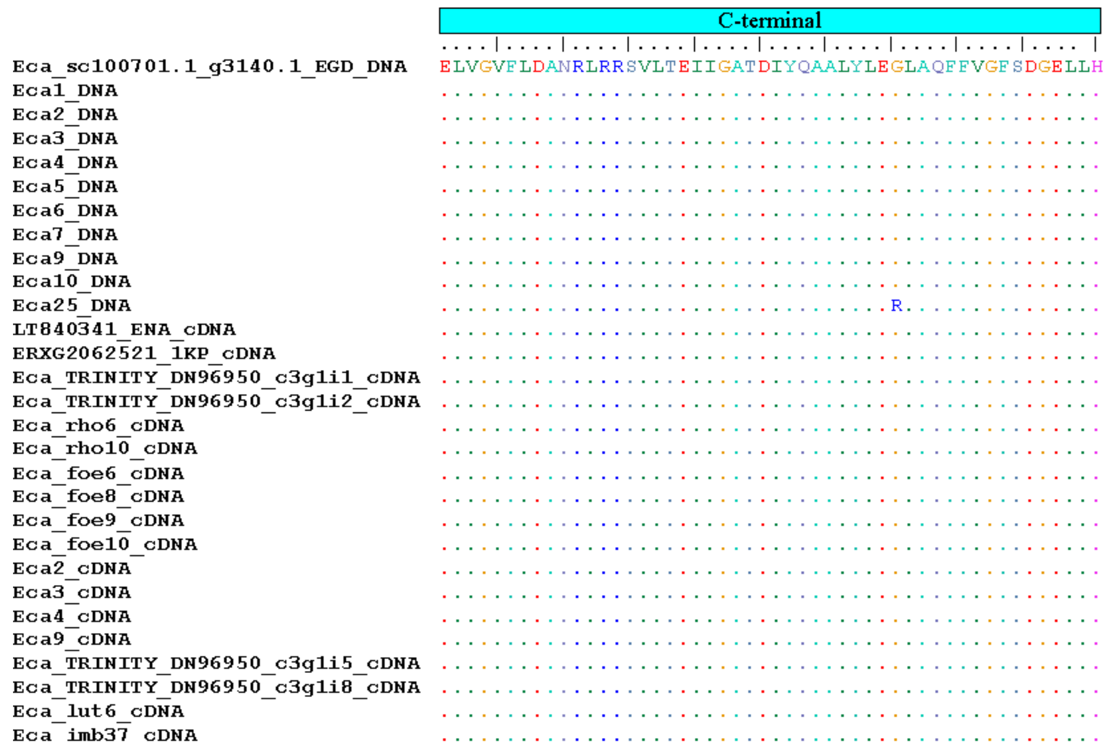
Supporting Information

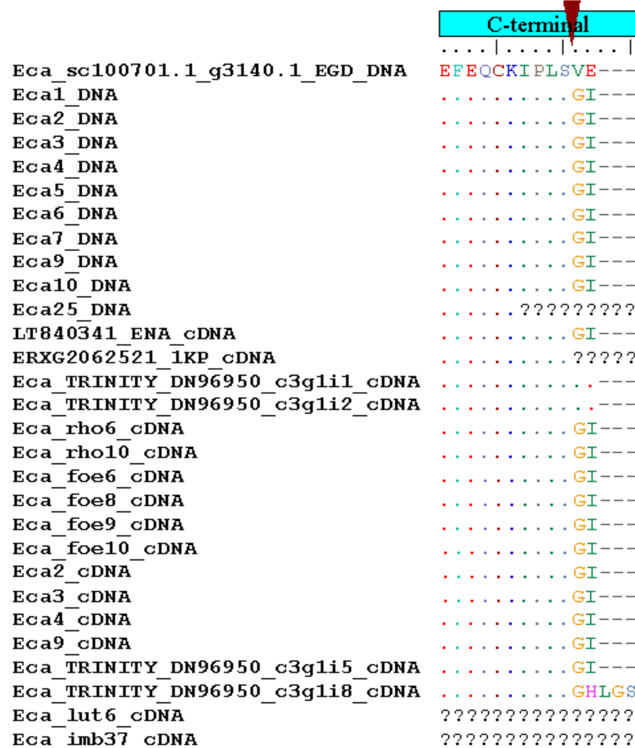


Supplementary Figure 1. Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) of *Eschscholzia californica* proteins visualized by Stain-free technology using a Gel DocTM EZ System. MC, mother cell; T, tetrad; M, microspore; MP, mature pollen.

```
                                    N-terminal                    DOG1
Eca_sc100701.1_g3140.1_EGD_DNA     MIKAAARFGRKKSTRLFKDFYLEWIETLKTNLLPLLRRSILVSSSNQLST
Eca1_DNA                           .....................V............................
Eca2_DNA                           .....................V............................
Eca3_DNA                           .....................V............................
Eca4_DNA                           ..................................................
Eca5_DNA                           .....................V............................
Eca6_DNA                           ..................................................
Eca7_DNA                           ..................................................
Eca9_DNA                           ..................................................
Eca10_DNA                          ..................................................
Eca25_DNA                          ??????????????????...V............................
LT840341_ENA_cDNA                  .....................V............................
ERXG2062521_1KP_cDNA               ..................................................
Eca_TRINITY_DN96950_c3g1i1_cDNA    ..................................................
Eca_TRINITY_DN96950_c3g1i2_cDNA    ..................................................
Eca_rho6_cDNA                      ..................................................
Eca_rho10_cDNA                     ..................................................
Eca_foe6_cDNA                      ..............................D...................
Eca_foe8_cDNA                      ..................................................
Eca_foe9_cDNA                      ..................................................
Eca_foe10_cDNA                     ......................G...........................
Eca2_cDNA                          ..................................................
Eca3_cDNA                          ..............................D...................
Eca4_cDNA                          ..................................................
Eca9_cDNA                          ..................................................
Eca_TRINITY_DN96950_c3g1i5_cDNA    ..................................................
Eca_TRINITY_DN96950_c3g1i8_cDNA    ..................................................
Eca_lut6_cDNA                      ?????????????????????????.........................
Eca_imb37_cDNA                     ???????????????.......G.....R.....................

                                                 DOG1
Eca_sc100701.1_g3140.1_EGD_DNA     HVQMIQHHFQNYYLTLDLAASEDVSQLLFPIWRNSLEKPFLWVGDFHPNL
Eca1_DNA                           ..................................................
Eca2_DNA                           ..................................................
Eca3_DNA                           ..................................................
Eca4_DNA                           ..................................................
Eca5_DNA                           ..................................P...............
Eca6_DNA                           ..................................P...............
Eca7_DNA                           ..................................P...............
Eca9_DNA                           ..................................P...............
Eca10_DNA                          ..................................P...............
Eca25_DNA                          ..................................................
LT840341_ENA_cDNA                  ..................................................
ERXG2062521_1KP_cDNA               ..........................I.......................
Eca_TRINITY_DN96950_c3g1i1_cDNA    ..........................I.......................
Eca_TRINITY_DN96950_c3g1i2_cDNA    ..........................I.......................
Eca_rho6_cDNA                      ....................................G.............
Eca_rho10_cDNA                     ....................................G.............
Eca_foe6_cDNA                      ..................................................
Eca_foe8_cDNA                      ..........................I.......................
Eca_foe9_cDNA                      ..........................I.......................
Eca_foe10_cDNA                     ..........................I.......................
Eca2_cDNA                          ..........................I.......................
Eca3_cDNA                          ..................................................
Eca4_cDNA                          ..........................I.......................
Eca9_cDNA                          ....................................G.............
Eca_TRINITY_DN96950_c3g1i5_cDNA    ..........................I.......................
Eca_TRINITY_DN96950_c3g1i8_cDNA    ..........................I.......................
Eca_lut6_cDNA                      ....................................G.............
Eca_imb37_cDNA                     ..........................I.....C.................
```

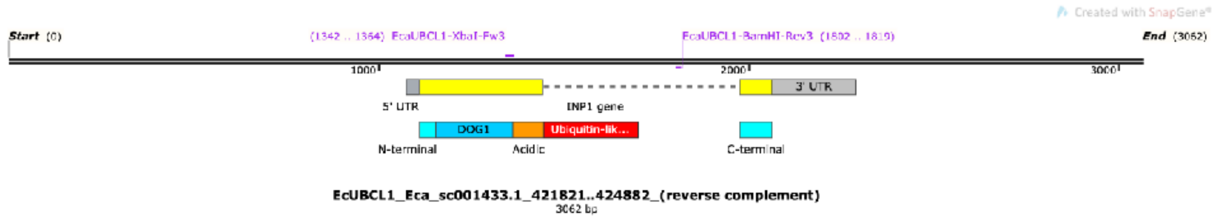
```
                                   |  DOG1  |            Acidic                      |Middle|
                                   ....|....|....|....|....|....|....|....|....|....|
Eca_sc100701.1_g3140.1_EGD_DNA     FTNLLRSFLNNN-SSDEEINDTHIEKSSNFPLVWKYPSKNLMNKIEQIEC
Eca1_DNA                           ............-......DN.............................
Eca2_DNA                           ............-......D..............................
Eca3_DNA                           ............-......D..............................
Eca4_DNA                           ............-......D..............................
Eca5_DNA                           ............-......DN.............................
Eca6_DNA                           ............-......D..............................
Eca7_DNA                           ............-......DN.............................
Eca9_DNA                           ............-......DN.............................
Eca10_DNA                          ............-......DN.............................
Eca25_DNA                          ............-.....................................
LT840341_ENA_cDNA                  ............-......DN.............................
ERXG2062521_1KP_cDNA               ............N......D..............................
Eca_TRINITY_DN96950_c3g1i1_cDNA    ............N......D..............................
Eca_TRINITY_DN96950_c3g1i2_cDNA    ............N......D..............................
Eca_rho6_cDNA                      ............-.....................................
Eca_rho10_cDNA                     ............-.....................................
Eca_foe6_cDNA                      ............-.....................................
Eca_foe8_cDNA                      ............N......D..................E...........
Eca_foe9_cDNA                      ............N......D..................E...........
Eca_foe10_cDNA                     ............N......D..................E...........
Eca2_cDNA                          ............N......D..............................
Eca3_cDNA                          ............-.....................................
Eca4_cDNA                          ............N......D..................E...........
Eca9_cDNA                          ............-.....................................
Eca_TRINITY_DN96950_c3g1i5_cDNA    ............N......D..............................
Eca_TRINITY_DN96950_c3g1i8_cDNA    ............N......D..............................
Eca_lut6_cDNA                      ............-.....................................
Eca_imb37_cDNA                     ............N......D..................E...........

                                   |                   Middle                  |C-terminal|
                                   ....|....|....|....|....|....|....|....|....|....|
Eca_sc100701.1_g3140.1_EGD_DNA     GLRSMVPTLVTRYRKSQSKFLDKCGLNWINCESKQEILKTVEKDLMVEIE
Eca1_DNA                           ................H.................................
Eca2_DNA                           ..................................................
Eca3_DNA                           ..................................................
Eca4_DNA                           ..................................................
Eca5_DNA                           ................H.................................
Eca6_DNA                           ..................................................
Eca7_DNA                           ................H.................................
Eca9_DNA                           ................H.................................
Eca10_DNA                          ................H.................................
Eca25_DNA                          ...............A................................T.
LT840341_ENA_cDNA                  ................H.................................
ERXG2062521_1KP_cDNA               ..................................................
Eca_TRINITY_DN96950_c3g1i1_cDNA    ..................................................
Eca_TRINITY_DN96950_c3g1i2_cDNA    ..................................................
Eca_rho6_cDNA                      ..................................................
Eca_rho10_cDNA                     ..................................................
Eca_foe6_cDNA                      ..................................................
Eca_foe8_cDNA                      ..................................................
Eca_foe9_cDNA                      ..................................................
Eca_foe10_cDNA                     ..................................................
Eca2_cDNA                          ..................................................
Eca3_cDNA                          ..................................................
Eca4_cDNA                          ..................................................
Eca9_cDNA                          ..................................................
Eca_TRINITY_DN96950_c3g1i5_cDNA    ..................................................
Eca_TRINITY_DN96950_c3g1i8_cDNA    ..................................................
Eca_lut6_cDNA                      ..................................................
Eca_imb37_cDNA                     ..................................................
```





Supplementary Figure 2. Protein alignment of 29 sequences of *Eschscholzia californica* INP1 (EcINP1). The alignment comprises 265 positions. The subdivisions of EcINP1 are shown according to Li *et al*. (2018). The possition of the introns is marked with a red triangle at the top of the alignment. 26 sequences were generated by us and three were downloaded from databases (indicated after the locus name: EGD, *Eschscholzia* Genome Database; ENA, European Nucleotide Archive; 1KP, 1000 Plants Project). 11 sequences come from the translation of gDNA and 18 from cDNA (indicated at the end of the name : DNA *vs* cDNA). As a reference sequence we have taken the sequence of the scaffold Eca_sc100701.1_g3140.1 obtained from the *Eschscholzia* Genome database. The named sequences starting with Eca_TRINITY come from transcriptomes obtained by us.

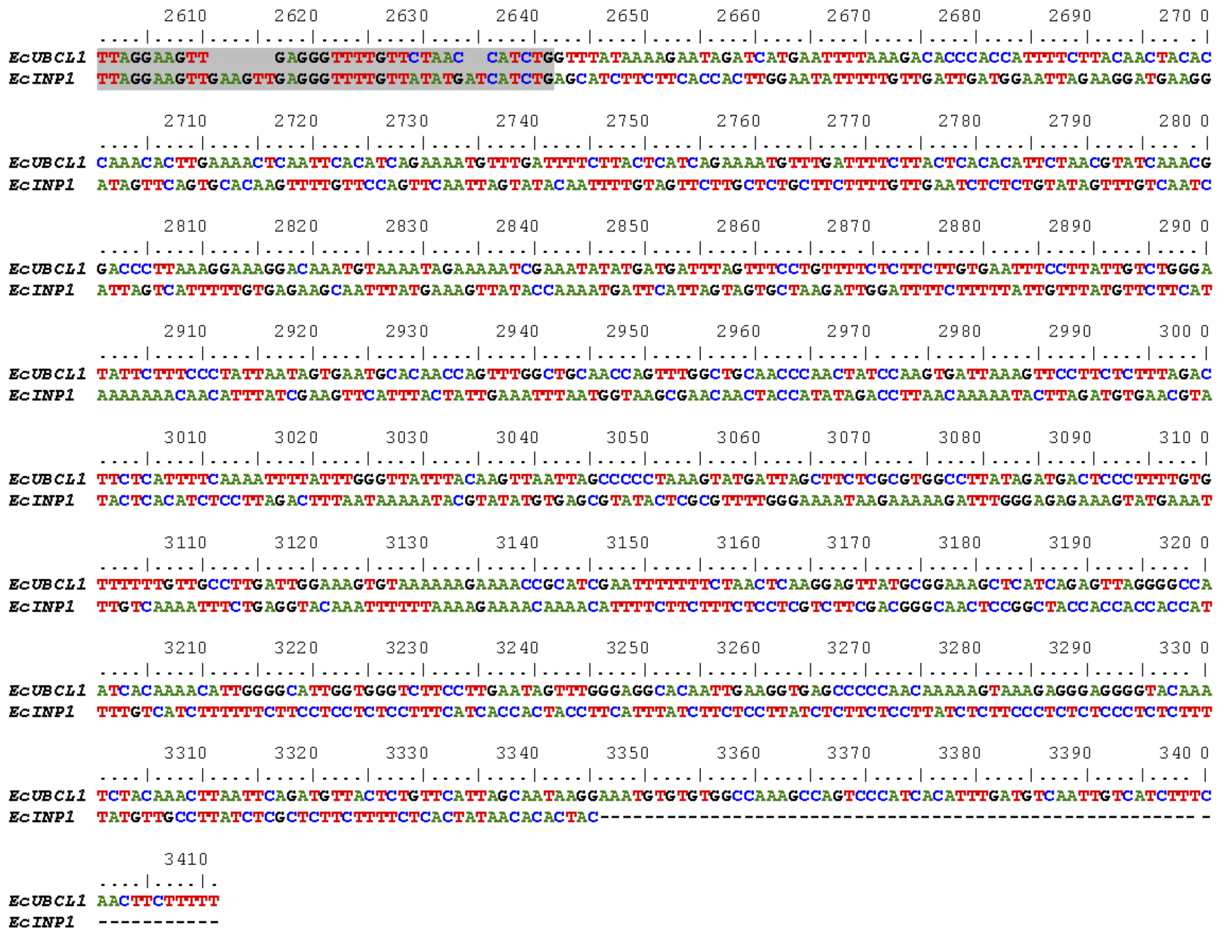
A



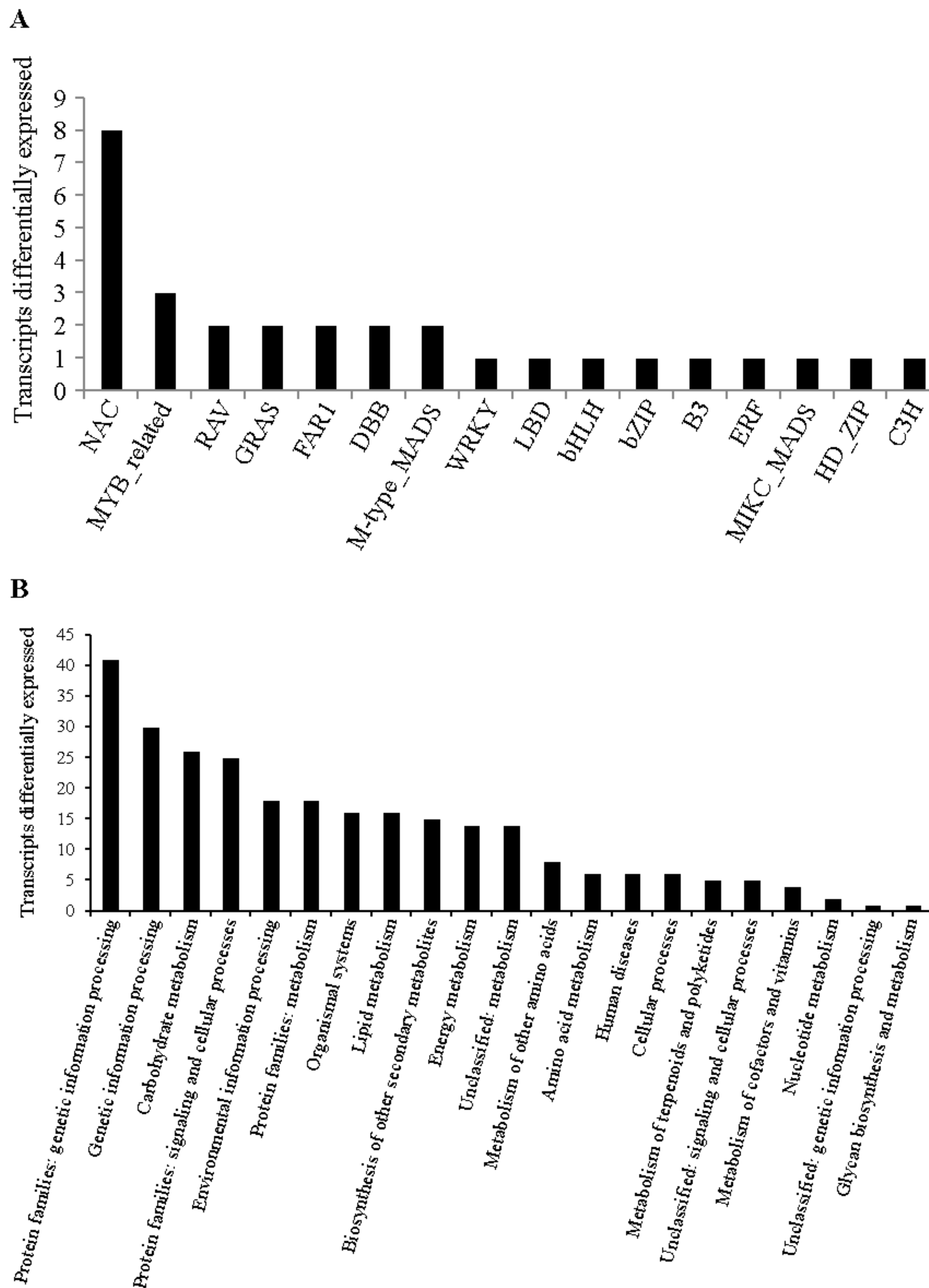
B

```
EcUBCL1_Eca_sc001433
EcINP1_Eca_sc100701

                 10        20        30        40        50        60        70        80        90       100
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  CAGCAAGTAGCAAAATTATAAATTCAGCTGACGGTTGCTTCTTTTAGCAAGAGTCTCAATCAGTCTCACTTAGCACACCTTAACCAGGTACTTTTATCTG
EcINP1   ----------------------------------------------------------------------------------------------------

                110       120       130       140       150       160       170       180       190       200
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  CTTTCTAAAAATAACTAATAAATTTCAATAACGTAGCGTCGACGATATTGATTGTGCCATCAATGCTTCGAGGGGGACTTTGGGCAAAGACTTAGTGAAT
EcINP1   ----------------------------------------------------------------------------------------------------

                210       220       230       240       250       260       270       280       290       300
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  GATTGTACCAATTCCTAAGAGAATAAGTTGTGTTTTAACTATCCAATGTTAATAAATGGCTTTATGAGCGAGGTGTAATTTGTTGTATGAAAAATATTGC
EcINP1   -----------------------------------------------------TCGGCTTTGATGAATTTTGAATTGTGATCAAGCATAACCAGTCCAAT

                310       320       330       340       350       360       370       380       390       400
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  GAATCAAATTGTCGAAATTTATTGTTTGAGTTTGGTTTAATCTTCATAAGATTCATTTAGACCTCCAAATATGGTGTTAACATTTGATTAGTTAGACTGC
EcINP1   TCCTATCTTCTGTTAAGAGAAAGAGAAAGATGCCTCGACATTAATCTTATAGCGGAGCGGACGTGGGGGTCCGTCTTTCCCGCTTTCAAGTATTACTGTA

                410       420       430       440       450       460       470       480       490       500
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  ACCCACATCTTCTAGTTATCATATCATATAGCATAGATAGAGATCAAAAGATGAAATGAGCACATAACAACTACAAAACGTAGTATGGCCTTGACACGTC
EcINP1   GCTAATCTTGAGAGATTTAGTCGTCTTGCGCTTCATGTGTAAAATTAAAAGAGTTAATTGTCTTTATATTGTCCTTCGAATTTAGCTGGTTATGTTGCAG

                510       520       530       540       550       560       570       580       590       600
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  AAGAGCATGAAATAAGCTCATAAACCATCATGAAAACAAAGACCTAGAAACGTATCCTGATGATCATGACACTAAATTCTACCTGCCATCAGGAACTAGT
EcINP1   ATAGATCCGAAGACGAGGCTTTCTACATATACCAAACACTCATGGCCAACAAGAAGAGATTAATAGCAGCGTTATCGTTATCTAAGTGTTTCCACGAGCA

                610       620       630       640       650       660       670       680       690       700
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  TTGTTTTTTAATATACATATTTAGAAGGTTAACTATCAACTCTACGACTGCAGTCCATTACCACCCCCGATGGTAGTTAATAGATTATGAAACTGGAGTG
EcINP1   TATCCATTGGGTGAGGTCGTCATTTGATCTGAAAATATTTCAATAAAACGTTAAACTTCTAGCAAATTGGACTTCTTATCTCACGTTCATTATCTCACAT

                710       720       730       740       750       760       770       780       790       800
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  ATTATTAGAGTCTTCTGCACCTTAATTTATAATATTGAGGTGACAGAAAGTAGAAACATATCAAGGAGATTAAGGGTCTGTTTGATACGTGAAAAAAGTG
EcINP1   TATCAAAGCTTCGTTGATGTTAAGACACTAGCGATCTTCTCCTTCTTCGTGACAGACCAGATCTAGTATCAAAAAAAAAAAACATCGATATCATCATCGC
```

```
             810       820       830       840       850       860       870       880       890       900
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  GGTAAGAAAACTGATTTTCTTAGGAAAATCCGTTTTCTGGTGTTTGGTTGGTGTTTATGAGTAAGAAAACCAAGTTCTTATAGAAAAAAAAGTGGGAAAA
EcINP1   TTTAGACACCTCAGTCCGAAACCAACAAGTACGAGATTTTGTTCCTGAGGGAGGCTCCACTAGCATACCAAAAAACCACAATGGTTATTAAGCTTGAGCG

             910       920       930       940       950       960       970       980       990      1000
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  AATGGTGTTTTATTTGCCCACAAAAAAGGTGGGAAAACCTAATTCCCACCCTCTCTCAAAAACGCGCTCTCTCTCCTCTTCTAGTCTTCCGTGTTTCTCT
EcINP1   TACAGGTATAACCCACCACGGGGACAATTTCTGGGAGTAATGTGGGACTATCACACCAATGGAGGAGACCATTCAGACCAGCCAGTCCCGATGGCAGCCC

            1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  CTCAAACTCTCATCTCTCTATTTCCATGGTTGAAGCTCCATCTCCATTTCTCATGGCTCCAAGAAAACTCTAGTTTGTAGAACTCCTTCAACAGCAATAT
EcINP1   TTATAAAAGTAACAGAAATTCAATTTTATCTTAAAACTTCCATGCTTGCTTGACAAAAACCCTAGATTGATCCTCGGTAGAACTCCTTCAACAGCAATAT

            1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  CAAAAGCAT-------------------------------------------------------------------------------------------
EcINP1   CAAAAGCATCTTTCTTCTATAACTATCCAAACTCTTCACAGAGCTCTCTCTCTCTCTCTAGAAATCCTGATCAAGAACAAGAACATTTTTCTCCCTTCAAG

            1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  ----------------------------------------------------------------------------------------------------
EcINP1   ATGATCAAAGCTGCAggtacatctctctctctctcattactttctctctacattcttcaagatagattcaaattttaatcatctctaatggttaat
         |---N-terminal---

            1310      1320      1330      1340      1350      1360      1370      1380      1390      1400
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  ---------------------------------------------CTTTTCAAAGATTTCTACGTGGAATGGATTGAAACCCTAAAAACCACTTTACTCC
EcINP1   tcctttttacttaacaGCTCGATTTGGTCGCAAGAAAAGCACTCGACTTTTCAAAGATTTCTACTTAGAATGGATTGAAACCCTAAAAACCAATTTACTCC
                        ------------------------------N-terminal------------------------------|---DOG1---

            1410      1420      1430      1440      1450      1460      1470      1480      1490      1500
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  CACTACTACGT---TTCAGTTCTGGTTTCTTCTTCAAACC--------------------ATGATTCAACACCATTTTCAAAACTATTACCTAACCCTAGA
EcINP1   CATTACTACGTCGTTCAATTTTGGTTTCTTCTTCAAACCAACTTTCTACCCATGTCCAAATGATTCAACACCATTTTCAAAACTATTACCTAACCCTAGA
         ---------------------------------------------DOG1---------------------------------------------

            1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  TCTTGCAGCTTCCGAAGATGTTTCACCCTAATCTCTTCAC--------TAATCTTCTTGAAAACCCCCTTTTTTTGGGTTGGCCATTTTCACCCTAATCT
EcINP1   TCTTGCAGCTTCTGAAGATGTTTCACAACTTCTTTTTCCCAATTTGGCGTAATTCCCTTGAAAAAACCC-TTTCTTTGGGTTGGTGATTTTCACCCTAATCT
         ---------------------------------------------DOG1---------------------------------------------

            1610      1620      1630      1640      1650      1660      1670      1680      1690      1700
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  CTTCACTAATCTCCTTGGATCATTTCTCAATAATAATAATAATAGTTCAGATGAAGAAATTTATGATAACCATATCGAAAAATCATCAAATTTCCATATG
EcINP1   CTTCACTAATCTCCTTAGATCATTTCTCAATAATAA------TAGTTCAGATGAAGAAATTTAATGATACCCATATTGAAAAAATCATCAAATTTCCCTC--
         -----------DOG1-----------|----------------------------Acidic----------------------------

            1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  GATTTGTTTGGTAAAGCATGAAATGGGTCGAATTGAAGATGGGGCAGCTCGGAGTACTCAGCAATCCTAACTACAAAATAACTGCTCTGCTTGACCACT
EcINP1   ---TTGTTTGGAAA--------------------------------------------------------------------------------------
         ----Acidic----

            1810      1820      1830      1840      1850      1860      1870      1880      1890      1900
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  TAGCCATGATTACAGTTCAATCTGATCACCGAGGCATCTTTGATTGCAAACTAGTGATCAAGAGTTAGTAATGCTCACCAACTATCTGCTTGAAATAGCT
EcINP1   ----------------------------------------------------------------------------------------------------

            1910      1920      1930      1940      1950      1960      1970      1980      1990      2000
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  GAGCTCAATTATATGAGTAAACTCGATCACAAGCATTGGGAGTCATATGTATCCAAGCAGCGAAGGCACCACTGAACATATATGAGAAATGTCGGCCGAG
EcINP1   ----------------------------------------------------------------------TACCCATCAAAGAATTTGATGAACAAAATTG
                                                                               ---------Acidic---------|

            2010      2020      2030      2040      2050      2060      2070      2080      2090      2100
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  CTCAGTCCTCTGCTGAGTTGTTATCTTATAAAATGGAGCTGACGAAGAGATCAAATCGAGCCGAAGTCGGCTTACGGGTACTTCTAGTATTTAATGTTAA
EcINP1   AACAGATTGAATGTGGGTTAAGATCAATGGTTCCTACCCTTGTTACTCGTTATAGGAAATCACAATCAAAATTTCTTGATAAATGTGGTTTGAATTGGAT
         ---------------------------------------------Middle---------------------------------------------

            2110      2120      2130      2140      2150      2160      2170      2180      2190      2200
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  TATATTATGAAATCTTTGACGCCAGAATGATTGGATGAGTGAAGAGGAATCAACAGAGTTGTAGTCATTTTAAACTTTAGAGTTTAGACTGTCATTTTAT
EcINP1   TAATTGTGAAAGTAAACAAGAGATTTTGAAGACAGTAGAAAAAGATTTAATGGTTGAGATTGAAGAACTTGTTGGTGTTTTTCTTGATGCTAATAGATTA
         ------------------------Middle------------------------|-------------------C-terminal-------------

            2210      2220      2230      2240      2250      2260      2270      2280      2290      2300
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  ACAATTGGGTTTGATTCTAATCAAGGGTGTTATCACTTTTGT-------------TTTATCTCCAAGGGCTTGCTCAATTTTTTGTTGGGTTTAGTGATG
EcINP1   AGAAGAAGTGTTTTAACCGAGATTATCGGTGCTACCGATATCTATCAGGCAGCTCTTTATCTCGAAGGGCTTGCTCAATTTTTTGTTGGGTTTAGTGATG
         ---------------------------------------------C-terminal---------------------------------------------

            2310      2320      2330      2340      2350      2360      2370      2380      2390      2400
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  GTGAATTACTACATGAATTTGAACATTGCAAGATTCCATTA                              CTTCTTGTTTTTGCTAAATCTAC
EcINP1   GTGAATTACTACATGAATTTGAACAATGCAAGATTCCATTATCAGTTGAGTAATTCTTAACTTGGGTATTTCAATTACTTCTTGTTCTTGCTAAATCTAT
         --------------------C-terminal--------------------

            2410      2420      2430      2440      2450      2460      2470      2480      2490      2500
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  AAGGAATTCAATTTTTAGGAAGTTCAAGATCTTTATTAAT   CATGTGGATATCTGATGTTTAA                     TAGTAGTTTTATTT
EcINP1   AAGGAATTCAGTTTTTAGGAAGTTCAAGATCTTTAATAATCATCATTTGGATATCTGATGTTCTTGCTGATTTTGATTTCATGGGTTAGTAGTTTTATTT

            2510      2520      2530      2540      2550      2560      2570      2580      2590      2600
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
EcUBCL1  GAATATGTGAACTTGTAATTCAACTTTCTTGTGTTTTTTAGATTTTGGAATTCTGAATT AGGTTTTGATTTGATTCAAAATT CACATTTACAGGGCAA
EcINP1   GAATATCTGAACTTGTAATTCAACATTCTTGTGTTTATTAAATTTTGCAATTCTGAATTTAGGTTTTGATTTGATTCAAATTTTCTCATTTACAGGGCAT
```

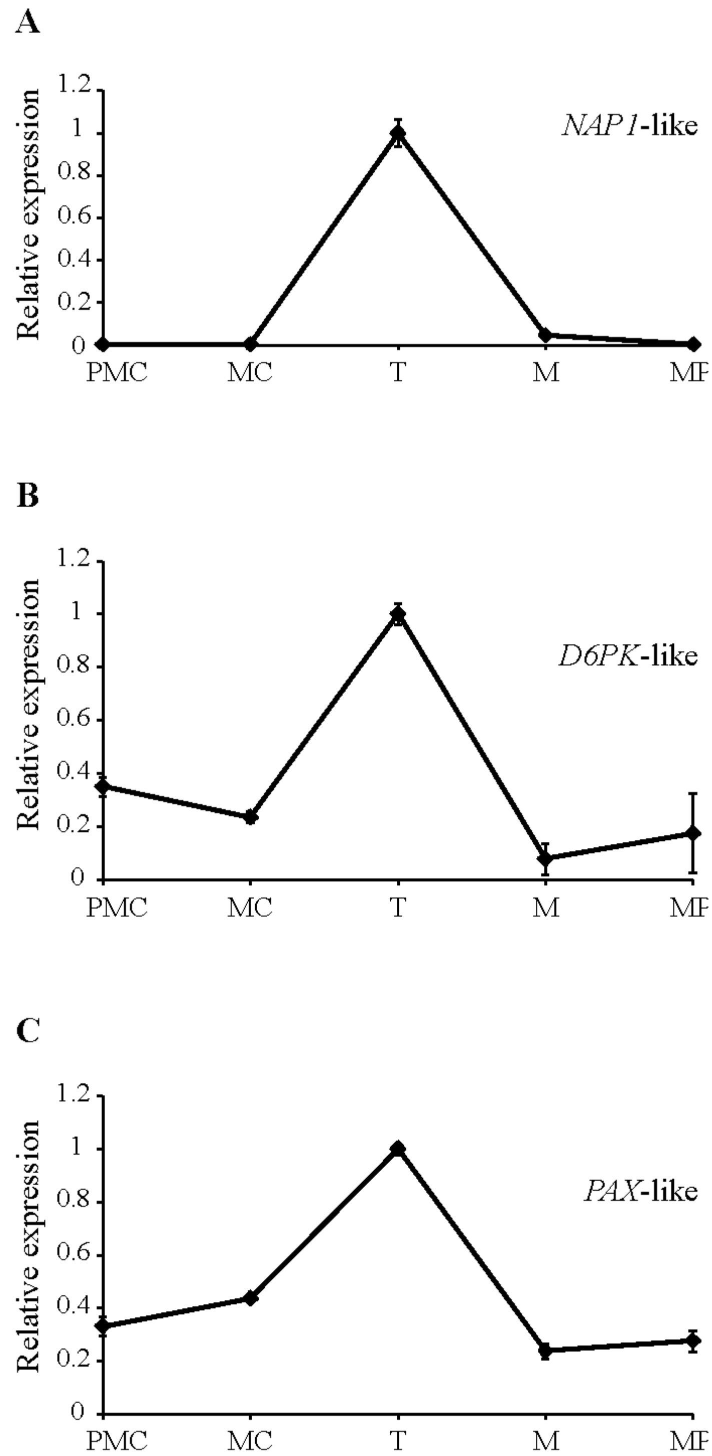


Supplementary Figure 3. The *EcUBCL1* sequence from *Eschscholzia californica* containing a fragment from the ubiquitin-like domain-containing CTD phosphatase and two dispersed fragments of *EcINP1*. The sequence shown corresponds to the reverse complement of the scaffold Eca_sc001433.1 (positions 421821..424882) obtained from the *Eschscholzia* Genome database. **(A)** Sequence map, created with SnapGene Viewer 5.2.1, showing the fragments with identity to the *EcINP1* gene (coding sequences, yellow; UTR's, grey) and the ubiquitin-like domain-containing CTD phosphatase region (red). Below the *EcINP1* coding regions, the domains that comprise them are specified. Primers used for VIGS experiment are shown above the sequence. **(B)** Alignment of *EcUBCL1* and *EcINP1* (reverse complement of the scaffold Eca_sc100701.1 positions 1579981..1582805 from the *Eschscholzia* Genome database). The shaded fragments in both sequences correspond to the regions that share significant identity (UTRs, grey; coding regions, yellow). Red shading in *EcUBCL1* indicates the sequence from a ubiquitin-like domain-containing CTD phosphatase. The *EcINP1* CDS is underlined and its domains are indicated; the sequence of the intron is indicated in lower case letters.



Supplementary Figure 4. Functional annotation of DEGs from VIGS-silenced plants of *Eschscholzia californica*. **(A)** Histogram showing the distribution by transcription factor families of the 30 transcription factors identified from Plant Transcription Factor DataBase. **(B)** Classification of each DEG into KEGG functional categories, using GhostKoala mapping tool.

Supplementary Figure 5. Maximum likelihood tree for the AGC1 group kinases, based on the protein sequences of the kinase domain, from *Eschscholzia californica* (Eca), *Papaver somniferum* (Psom), and *Arabidopsis thaliana*. The locus identifier is shown for each sequence. The sequences of the DEGs found in *E. californica* are in bold type. The sequences of the AGC3-group kinases from *A. thaliana* were used as an outgroup. The types of kinases indicated on the right for each clade are listed according to the names of the *A. thaliana* kinases.



Supplementary Figure 6. qRT-PCR-based expression analysis of **(A)** *NAP1*-like, Eca_sc004324.1_g2710.1; (**B**) *D6PK*-like, Eca_sc004639.1_g0030.1; **(C)** *PAX*-like, Eca_sc000153.1_g1520.1. The expression patterns were tested in anthers at different pollen development stages. Actin was used as a normalization control. PMC, pre-mother cell; MC, mother cell; T, tetrad; M, microspore; MP, mature pollen.

Supplementary Table 1. Primer sequences used in this study.

Primer	Sequence	Purpose
EcaINP1-F	ATGATCAAAGCTGCAGCTCGA	Intraspecific variability and phylogeny
EcaINP1-R	AATGCCTGATAATGGAATCTTGC	Intraspecific variability and phylogeny
EcaINP1-F2-EcoRI	GGAA_GAAT_TCCCTTCAAGATGATCAAAGCTGC	VIGS construct
EcaINP1-R2-BamHI	GGAA_GGATCC_GGAACCATTGATCTTAACCCACA	VIGS construct
EcaINP1-FwReal2	GCTGCAGCTCGATTTGGT	Gene expression qRT-PCR and semi-qRT-PCR
EcaINP1-RvReal4	AGAAACCAGAATTGAACGACGT	Gene expression qRT-PCR
EcaINP1-RvReal3	TGGACATGGGTAGAAAGTTGG	Gene expression semi-qRT-PCR
Actin2RTQFw	TTACAATGAGCTTCGTGTTGC	Gene expression qRT-PCR
Actin2RTQRv	CCCAGCACAATACCTGTAGTAC	Gene expression qRT-PCR
WAK2qPCRFw2	CCTTTCGTGGGATAACCGCT	Gene expression qRT-PCR
WAK2qPCRRv2	ACTCGACTTGACGTCCCTATG	Gene expression qRT-PCR
PAX2qPCRFw2	AGCGGCACTAGGTGCTATTT	Gene expression qRT-PCR
PAX2qPCRRv2	CTCTTTCTGTCTGAGCTCTGGT	Gene expression qRT-PCR
WAK3qPCR2Fw1	TCGAATGGGACCTTGCACC	Gene expression qRT-PCR
WAK3qPCRRv1	TCCTGCGGTTTCTGTAGCAAT	Gene expression qRT-PCR
NAC25qPCRFw2	GGACATGACGCACACTGCTT	Gene expression qRT-PCR
NAC25qPCRRv2	ACATCATCAGCGGGTTCTTG	Gene expression qRT-PCR
D6PKL1qPCRFw2	CTCATGTTTCATGCCGCGAT	Gene expression qRT-PCR
D6PKL1qPCRRv2	GCAATTAACTCAGGGCGTGC	Gene expression qRT-PCR
PAX_qPCR_Fw2	TCGAAGGTGTGAATTGGGCA	Gene expression qRT-PCR
PAX_qPCR_Rv2	CCCCAACCGGATCAACTGAA	Gene expression qRT-PCR
NAP1_qPCR_Fw2	TATGGGGTGGACCATACGGA	Gene expression qRT-PCR
NAP1_qPCR_Rv2	TTGTTGGAATGGGAACCAGTCT	Gene expression qRT-PCR
ANXUR1_qPCR_Fw2	TGTGATCACGGAGTTCAGCG	Gene expression qRT-PCR

ANXUR1_qPCR_Rv2	GTTCTTCACCTCCGGGGTTT	Gene expression qRT-PCR
Rho_qPCR_Fw2	GAAGAGTCGGTGGTCGCTTT	Gene expression qRT-PCR
Rho_qPCR_Rv2	AGTGTACCAGTAGAAGCAGCG	Gene expression qRT-PCR
NADPH_qPCR_Fw	ACTGCTCCACTCAAGAACGC	Gene expression qRT-PCR
NADPH_qPCR_Rv	ACCACCAGCGCTTACGATAG	Gene expression qRT-PCR
EcaUBCL1-XbaI-Fw3	GGAA_TCTAGA_AATCTCCTTGGATCATTTCTCAA	VIGS construct
EcaUBCL1-BamHI-Rv3	GGAA_GGATCC_AGAAGTACCCGTAAGCCG	VIGS construct

Supplementary Table 2. List of sequences used to test homology of the *Eschscholzia californica* EcINP1 protein by constructing the gene tree shown in Figure 1.

Species	**Locus name**	**Database**	**NCBI-ENA accession number (transcript/protein)**
Aquilegia coerulea	Aqcoe7G079000.1	Phytozome	
	Aqcoe7G079000.2	Phytozome	
Anemone pulsatilla	UPOG-2056161	1KP	
Arabidopsis thaliana	INP1 // AT4G22600	NCBI	NM_118386 NP_193991
Argemone mexicana	IRAF-2015101	1KP	
Brachipodium distachium	LOC100826863 // Bradi3g50820	NCBI	XM_024460675 XP_024316443
Eschscholzia californica	Eca_sc100701.1_g3140.1	Eschscholzia Genome Database	
	Eca TRINITY DN96950 c3 g1 i5	ENA	OU070349
Macleaya cordata	BVC80_1727g16	NCBI	MVGT01000886 OVA15152
Nandina domestica	YHFG-2062185/2005708/2005709	1KP	
Nelumbo nucifera	LOC104598573	NCBI	XM_010260712 XP_010259014

Nicotiana tomentosiformis	LOC104092836	NCBI	XM_009598514.3 XP_009596809
Oryza sativa	LOC4330222	NCBI	XM_026022761 XP_025878546
Papaver rhoeas	BEKN-2065739_GMAM-2014596	1KP	
Papaver somniferum	LOC113359130	NCBI	XM_026602809 XP_026458594
Papaver somniferum	LOC113284312	NCBI	XM_026533749 XP_026389534
Populus trichocarpa	LOC18093990 // POPTR_001G120700v3	NCBI	XM_024587580 XP_024443348
Solanum lycopersicum	LOC101247949	NCBI	XM_004245692.4 XP_004245740
Thalictrum thalictroides	KAF5188586	NCBI	JABWDY010026581 KAF5188586
Theobroma cacao	LOC18606310	NCBI	XM_007039852.2 XP_007039914
Vancouveria hexandra	vhe_VHALF1JP_Trinity_comp81110 c0_seq1	Phytometasyn	
Vitis vinifera	LOC100250621	NCBI	XM_010648242.2 XP_010646544
Zea mays	LOC100191973 // GRMZM2G112914	NCBI	XM_020553019 XP_020408608

Supplementary Table 3. Summary of sequencing and assembly for *EcINP1*-silenced and wild-type *Eschscholzia californica*.

	Wild 1	Wild 2	Wild 3	VIGS 1	VIGS 2	VIGS 3	Average
Sequenced reads	40072074	43677076	43973604	67984350	57866512	50223034	50632775
GC content	42	42.76	41.9	42.62	42.27	41.9	42.24
Q20(%)	98.13	97.87	98.04	97.83	97.68	97.74	97.88
Q30(%)	94.39	94.03	94.32	93.69	93.32	93.45	93.87
No HQ reads	24.482	61.764	50.048	113.308	83.572	62.008	65.864
HQ reads	40047592	43615312	43923556	67871042	57782940	50161026	50566911
Paired reads	20023796 (100%)	21807656 (100%)	21961778 (100%)	33935521 (100%)	28891470 (100%)	25080513 (100%)	25283456
Aligned paired reads	10344293 (51.66%)	11185146 (51.29%)	11740766 (53.46%)	18698472 (55.1%)	15913421 (55.08%)	13698976 (54.62%)	13596846 (53.53%)
Number of contigs >200 bases (merged assemble)	34729						
Average length of contigs	1006						

Chapter 2

Identification of candidate genes involved in the determinism of pollen grain aperture morphology by comparative transcriptome analysis in Papaveraceae

Ismael Mazuecos-Aguilera, Víctor N. Suárez-Santiago

Manuscript in preparation

Abstract

Background and Aims

Pollen aperture, areas of the wall where the exine is weakened or absent, are very varied in shape, number and position. In terms of shape, apertures can be elongated or round, both being present in all taxonomic groups of angiosperms and each being more adapted to certain environmental conditions. In the last decade some genes involved in the formation of pollen apertures have been discovered. However, only a protein family (ELMOD) has been found in *Arabidopsis* that acts upstream of the aperture formation pathway and in which the interaction of one of its members ELMOD_E with two other MCR/ELMOD_B and ELMOD_A can change the morphology of apertures from colpus (elongate) to pore (round). Here, we are prompted to identify candidate genes involved in the determinism of pollen aperture morphology in Papaveraceae (order Ranunculales) by comparative transcriptome analysis between two pairs of taxa with different aperture morphology.

Methods

We used two pairs of Papaveraceae species belonging to two different subfamilies and where within each pair one species has colpate pollen and the other porate (Fumarioideae: *Dactylicapnos torulosa*, 6-colpate and *Fumaria bracteosa*, pantoporate; Papaveroideae: *Eschscholzia californica* 5-7 colpate and *Roemeria refracta*, 6-porate). We sequenced and compared the transcriptome of the colpate and porate pollen when the pollen was in the tetrad development stage, time at which the process of formation of the apertures begins. We selected differentially expressed genes (DEGs) between the two groups and not differentially expressed within each group, which could potentially be involved in aperture shape determinism. The expression level of four genes was validated by qRT-PCR to confirm their differential expression.

Key Results

We found 531 DEGs between the colpate and porate pollen species groups that are not differentially expressed within each group. The expression levels obtained by qRT-PCR validate those obtained by transcriptome analysis. Among the DEGs we did not find any member of the ELMOD family, the only ones discovered so far that could be involved in determining the shape of pollen grain apertures. However, we did find genes related

to the formation of the callose wall or the organisation of the cytoskeleton, processes involved in the formation of the apertures.

Conclusions

Our results provide a list of genes that could determine the shape of the apertures in Papaveraceae. These results open up new avenues for functional studies of these genes to confirm their involvement in this process.

Keywords: pollen aperture, transcriptome analysis, genetic determinism, *Eschscholzia californica*, *Dactylicapnos torulosa*, *Roemeria refracta*, *Fumaria bracteosa*, RNA-seq, Papaveraceae, pollen

1. Introduction

Pollen apertures, sites that receive reduced amounts of exine (the name for the outermost wall of the pollen) or where it is completely absent, represent some of the most characteristic and well-defined elements of the pollen surface (Furness and Rudall, 2004; Zhou and Dobritsa, 2019). The apertural pattern of the pollen grain is defined by the shape, number, and position (polar or equatorial) of their apertures, and changes in this apertural pattern have been related to evolutionary advantages or variations in the efficiency of its functions. Thus, for example, the transition from pollen with one polar aperture –typical of basal angiosperms and monocots– to pollen with three apertures in equatorial positions, has been interpreted as a key innovation involved in the success and diversification of eudicots (Furness and Rudall, 2004).

As far as shape is concerned, Spermatophyte pollen show two predominant morphotypes of apertures: elongate, furrow-like apertures (colpus or sulcus), and round, pore-like apertures (porus or ulcus) (Walker and Doyle, 1975). The morphology of the apertures remains stable at the intraspecific level, although there are species with apertures that combine both types, with an elongated ectoaperture and a rounded endoaperture (colporus). Species with furrow-like apertures are more numerous than those with porate pollen (Furness and Rudall, 2004; PalDat, 2017). The significance of variation in aperture shape between taxa is poorly understood, although there seems to be a consensus that it is related to environmental xericity. One of the functions attributed to the apertures is to allow the rigid exine to adjust to changes in pollen volume due to dehydration/rehydration during pollination, the harmomegathy process

(Heslop-Harrison, 1979, Matamoro-Vidal *et al.*, 2016; Wodehouse, 1935). In response to the dehydration the membrane of the aperture sites folds inward, so that the edges of each aperture are touching each other, closing up the aperture site (Volkova *et al.*, 2013). Božič & Šiber (2020), in their study of the modelling of pollen infolding, showed that pollens with elongated apertures achieved successful closure with little stretching, furthermore their transition was graded allowing response to slight changes in humidity in the environment. In contrast, pore-like aperturates cannot adequately seal the aperture and suffer a large stretch at the margins of the aperture. In addition, if a porate pollen greatly reduces its volume, it adopts a mirror buckling geometry with a considerable amount of stretching energy concentrated at the edge. This geometry change requires a transition through a high energy state leading to bistability which would not allow response to small changes in humidity (Katifori *et al.*, 2010; Payne, 1972).

However, porate pollen is distributed throughout most of the angiosperm tree although usually restricted to a few species in a given group, and is seldom fixed at large taxonomic scales (Prieu *et al.*, 2017); and many of these species inhabit xeric environments. It has been observed that, unlike most pollen, many porate pollen grains disperse only partially dehydrated, which gives them a faster germination rate than dehydrated pollens (Franchi *et al.*, 2002). Prieu *et al.* (2017) on the basis of this evidence, explain why porate pollens have evolved many times during angiosperm history but have been less successful than furrow pollen. According to these authors, a faster germination could be a competitive advantage of porate pollen grains, which will be selected in the short term. However, their lower tolerance to dehydration would lead to their elimination in the long term. The independent evolution of porate pollen among the different groups of angiosperms suggests a common molecular mechanism for its emergence.

In the last decade some genes involved in the formation of pollen apertures have been discovered (*INAPERTURATE POLLEN 1*, *INP1*, Dobritsa and Coerper, 2012; *D6 PROTEIN KINASE-LIKE3*, *D6PKL3*, Lee *et al.*, 2018; *DEFECTIVE IN APERTURE FORMATION1*, *OsDAF1*, Zhang *et al.*, 2020; *INAPERTURATE POLLEN2*, *INP2*, Lee *et al.*, 2021), but only INP1 has been shown to affect apertural morphology by decreasing colpus length at lower doses of its protein (Dobritsa and Coerper, 2012). Only recently a protein family (ELMOD) has been found in *Arabidopsis* that acts upstream of the aperture formation pathway and in which the interaction of one of its

members, ELMOD_E, with two other, MACARON(MCR)/ELMOD_B and ELMOD_A, can change the morphology of apertures from colpus to pore (Zhou *et al.*, 2021). MCR and ELMOD_A are paralogs with redundancy in function (although MCR shows dominance) that participate in aperture domain specification by affecting the number of domains (dose dependent), but not the morphology of the apertures. *MCR* and *ELMOD_A* are both expressed at or near the tetrad stage of pollen development. The neomorphic activity of ELMOD_E occurs only when gene expression levels are elevated during the tetrad phase, as is the case when its expression is induced with the MCR promoter. Experiments by Zhou et al. (2021) show that high levels of MCR counteract the neomorphic activity of ELMOD_E, suggesting that both proteins compete for the same interactors. In *Arabidopsis*, a colpate species, *ELMOD_E* expression during tetrads is low, so it does not influence the aperture pattern. In contrast to INP1, which has been shown to be conserved in several species of key clades among angiosperms (Li *et al.*, 2018; Mazuecos-Aguilera *et al.*, 2021), how conserved the function of the ELMOD protein family is in other species is unknown, as is the role of ELMOD_E in the development of round apertures in porate species and/or if others molecular players are involved in determining the shape of the pollen apertures.

In the present study, we performed a comparison of the gene expression profiles among species with colpate and porate pollen grains during the tetrad stage to identify potential genes involved in the determinism of the aperture shape. The selected species belong to the family Papaveraceae, an interesting family because of its phylogenetic position at the base of the large clade of the Eudicots (order Ranunculales), and which is also an euripaline family (Pérez-Gutiérrez *et al.*, 2015) and therefore allows comparative studies between phylogenetically closely related taxa with different aperture pollen types (reducing the differences in expression due to the phylogenetic relationships). We used two pairs of species, each pair belonging to different subfamilies of Papaveraceae and where within each pair, one species is colpate and the other porate (Papaveroideae: *Eschscholzia californica* Cham., 5-7 colpate and *Roemeria refracta* DC., 6-porate; Fumarioideae: *Dactylicapnos torulosa* Hook.f. & T. Thomson, 6-colpate and *Fumaria bracteosa* Pomel,, pantoporate). The *de novo* transcriptomes assembled and annotated were used to find differentially expressed genes (DEGs) common between the colpate and poate species of each pair.

2. Material and methods

2.1. Plant material

Plants of *Dactylicapnos torulosa*, *Eschscholzia californica, Fumaria bracteosa and Roemeria refracta*, were sown in pots (9x9x9 cm3) with universal substrate and vermiculite mixed in a 3:1 ratio and kept in a greenhouse at a temperature range between 26 ºC to 14 ºC under a light/dark cycle of 16/8 h. Each pot was fertilized once at the beginning of the experiment and watered every day. When pollen was at the tetrad stage, confirmed by optical microscopy (Olympus-CX31, Japan), buds were collected from three plants per species (three independent biological replicates) and frozen in liquid nitrogen and stored at -80 °C until use.

2.2. RNA extraction, library construction, sequencing and read filtering

Total RNA was extracted from anthers with pollen in the tetrad development stage using the NucleoSpin® RNA Plant kit (Macherey-Nagel GmbH and Co., Germany), following the manufacturer's instructions. Quantity and quality was ensured using a Nanovue Plus Spectrometer (Biochrom, Cambridge, UK) and by agarose gel electrophoresis. Library preparation and sequencing were performed at Macrogen Inc. (South Korea). RNA libraries were prepared with an Illumina TruSeq Stranded Total RNA sample Preparation kit with Ribo-Zero Plant and sequenced on an Illumina Hiseq 2500 platform with paired-end reads of 150 bases. Generation of raw data was performed using Illumina package bcl2fastq.

The quality control of raw single reads (in FASTQ format) was evaluated using FastQC v0.11.83 (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). TrimGalore v0.6.4 (https://github.com/FelixKrueger/TrimGalore; parameters: --paired --phred33 -e 0.1 -q 20), was used for removing adaptors and low-quality sequences from the data set. Unpaired reads were also discarded for the remainder of the assembly pipeline. After trimming, FastQC was used again to examine the characteristics of the libraries and to verify trimming efficiency.

2.3. *De novo* transcriptome assembly and transcript reconstruction

The high-quality reads from *D. torulosa*, *F. bracteosa*, and *R. refracta* were *de novo* assembled by Trinity v1.8 (Grabherr *et al.*, 2011; parameters: --seqType fq --JM 10G \ -

-left reads.ALL.left.fq –right reads.ALL.right.fq \--SS_lib_type RF --CPU 6). For *E. californica*, as its genome is available (ftp://ftp.kazusa.or.jp/pub/*Eschscholzia*/ECA_r1.0.cds.fa.gz), it was used as reference.

All reads across three biological replicates from each species were combined to generate a single reference Trinity assembly per species. Basic statistical information over *de novo* assemblies was obtained running the Trinity package utility script TrinityStats.pl.

Then, a transcript reconstruction was carried out to filter the best-generated transcripts for *de novo* assembled species. For it, TransDecoder v5.5.0 (Haas and Papanicolaou, 2016; https://transdecoder.github.io/) was used to select the single best open reading frame (ORF) per transcript longer than 100 amino acids through the command TransDecoder.LongOrfs, and to predict the coding sequences with the command TransDecoder.Predict.

To obtain a set of non-redundant transcripts, we then clustered highly similar coding sequences using CD-HIT v4.8.1 (Fu *et al*., 2012) with an amino-acid sequence identity threshold of 0.99.

2.4. Transcript quantification

RNA-seq reads of different samples were aligned to their relevant Trinity assembly references or genome reference in case of *E californica*, using the software HISAT2 v2.1.0 (https://ccb.jhu.edu/software/hisat2/; Pertea *et al*., 2016). The aligned reads were assembled and quantified using the software StringTie v2.0 (https://ccb.jhu.edu/software/stringtie/), with the merge option to merging the assemblies of three biological samples together (Pertea *et al*., 2016). Assembly information was obtained through GffCompare v0.10.1 program (https://ccb.jhu.edu/software/stringtie/gffcompare.shtml; Pertea *et al*., 2016).

2.5. Transcriptome annotation and functional classification

The transcripts were annotated using the Trinotate annotation pipeline following the method outlined at (http://trinotate.github.io/). Initially, they were searched against SwissProt database (Bairoch and Boeckmann, 1992; https://data.broadinstitute.org/Trinity/Trinotate_v3_RESOURCES/uniprot_sprot.pep.gz) using BLASTX allowing one hit and with output in tabular format. The expected

protein translations were obtained using TransDecoder and then searched against SwissProt using BLASTP. The same BLAST parameters were used as for the BLASTX searches. The BLAST searches were loaded into the Trinotate SQLite database v3.0.2 (http://trinotate.github.io/) and an annotation report was generated. An e-value of 1e-5 was used as the threshold for the BLAST results during the report generation.

To assign a function to each transcript, annotated transcripts were further functionally classified with Gene Ontology (GO; Ashburner *et al*., 2000) and Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa and Goto, 2000) databases using Blast2GO v5.2.5 application (Götz *et al*., 2008) and GhostKoala mapping tool (Kanehisa *et al*., 2016), respectively. To predict putative plant transcription factor (TF) among transcripts, coding sequences were aligned to TF domains from Plant Transcription Factor Database (PlantTFDB - Plant Transcription Factor Database @ CBI, PKU; Zhang *et al.* 2011).

2.6. Analysis of DEGs

We compared gene expression profiles of two colpate pollen species (*D. torulosa* and *E. californica*) with two porate pollen species (*F. bracteosa* and *R. refracta*). Differential gene expression analyses were performed with the DESeq2 R package (v1.24.0) using the coverage produced by StringTie and merging transcript of different species by annotation obtained with SwissProt database using BLASTX. DEGs were filtered considering p-value and p-adj < 0.05 and $\log_2$-fold change >2. DEGs were classified functionally using the GhostKoala mapping tool as described above. Hierarchical clustering of log-transformed expression data, principal components analysis and heatmap was carried out using the DESeq2 R package (v1.24.0).

2.7. Validation of RNA-Seq analysis by qRT-PCR

To verify the differences observed in gene expression, qRT-PCR were performed for 4 selected DEGs. RNA was extracted as described above, and 1 µg of RNA was reverse-transcribed to cDNA using SuperScript III Reverse Transcriptase (Invitrogen, Karlsruhe, Germany) and oligo(dT)18 (ThermoFisher Scientific, USA). cDNA was diluted to 50 ng/µL and was used to perform qRT-PCR. *ACTIN* served as the reference gene for calculating the relative expression intensities in qRT-PCR analyses, using the $2^{\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). qRT-PCR was carried out using the

FastGene IC Green 2x qPCR mix (NIPPON Genetics, Tokyo, Japan) according to the manufacturer's instructions, and the qTower 2.2 real-time PCR thermocycler (Analytik Jena, Germany). Gene-specific primers were designed using the software Primer3 (Untergasser *et al*., 2012; Table Supplementary1). All experiments were repeated with three biological and three technical replicates.

3. Results

3.1. Transcriptome sequencing and assembly

Sequencing of the Ribo-Zero RNA-Seq libraries, from buds with pollen in tetrad development stage, of the four species yielded 278 million raw paired-end reads (Supplementary Table 2). Around 98% and 94% of reads presented a phred quality score >20 and >30 (Q20 and Q30) respectively, indicating a high quality of sequences. After filtering out low quality reads, 133,927,570, 145,092,452, 138,590,018, 138,351,862 and clean reads were obtained for *D. torulosa*, *E. californica*, *F. bracteosa*, and *R. refracta* respectively (Supplementary Table 2).

De novo assembly of reads led to the construction of 294,904, 235,229 and 412,427 contigs with average lengths of 1,177, 1,321 and 782 bases from *D. torulosa*, *F. bracteosa*, and *R. refracta*, respectively (Supplementary Table 2). When single best ORF per transcript longer than 100 amino acids was selected, 127,049, 115,642 and 147,231 transcripts were retained in *D. torulosa*, *F. bracteosa*, and *R. refracta*, respectively. After removing redundant transcripts, 56,689, 52,752 and 92,051 transcripts were kept for further analysis in *D. torulosa*, *F. bracteosa*, and *R. refracta*, respectively (Supplementary Table 2).

Eschscholzia californica reads assembly produced 37,400 transcripts after the reads alignment with his genome reference (Supplementary Table 2).

3.2. Transcriptome functional annotation and classification

BLAST searches in the SwissProt database with the assembled sequences resulted in 209,199 (70.94 %), 32,996 (88.22 %), 169,720 (72.15 %), and 340,571 (82.57 %) annotated transcripts in *D. torulosa*, *E. californica*, *F. bracteosa*, and *R. refracta* respectively (Supplementary Table 2; S3; S4; S5; S6). Using GhostKoala, 20677 (36,5%; *D. torulosa*), 11474 (33,6%; *E. californica*), 20979 (39,8%; *F. bracteosa*) and

28195 (30,6%; *R. refracta*) entries were classified in different functional categories; being genetic information processing and protein families involved in genetic information processing the most common categories (Supplementary Table 7). Through PlantTFDB 1,401, 1,559, 1,614 and 2,181 transcripts were identified as transcription factors in *D. torulosa*, *E. californica*, *F. bracteosa* and *R. refracta* respectively, with bHLH and B3 being the most abundant families (Supplementary Table 8).

3.3. Differential gene expression analysis

To find genes potentially involved in the shape of pollen apertures, gene expression profiles at the pollen tetrad stage of colpate vs. porate species were compared. The filtering of the genes identified 531 differentially expressed genes (DEGs), of which 231 were up-regulated and 300 were down-regulated in colpate species in comparison to porate species (Supplementary Table 9).

The hierarchical clustering, heatmap and principal component analyses showed the DEGs between colpate and porate species samples (Figure 1), irrespective of the Papaveraceae subfamily they belong to. To determine the functions performed by the different DEGs, they were functionally annotated. Through GhostKoala, 171 (32.2%) DEGs were classified, being "Genetic Information Processing" the most frequent category (Figure 2). Through PlantTFDB 12 DEGs were identified as transcription factors, repeating WRKY and ERF transcription factor twice each. (Supplementary Table 9). Using Blast2GO 420 DEGs were blasted, of which 236 were mapped and 221 were annotated; the 221 annotated DEGs were classified in three categories: biological processes, cellular components, and molecular functions (Supplementary Table 10). It is noteworthy to highlight the large number of predicted cellular components identified as integral components of membrane, others have been identified as components of the cytoskeleton. Many DEGs are involved in processes of transmembrane transport, regulation of transcription carbohydrate metabolism and organization, transport or depolymerization of components of the cytoskeleton. Important signal transduction elements were categorized such as putative serine/threonine protein kinases and transcription factors and other DEGs have DNA-binding and RNA-binding activity. Among the DEGs is *INP1*, the first aperture factor discovered, and also genes involved in the deposition or degradation of callose, which are related to the aperture formation

process. However, genes of the ELMOD family, the only ones described so far which could be involved in aperture shape determinism, were not found among the DEGs.

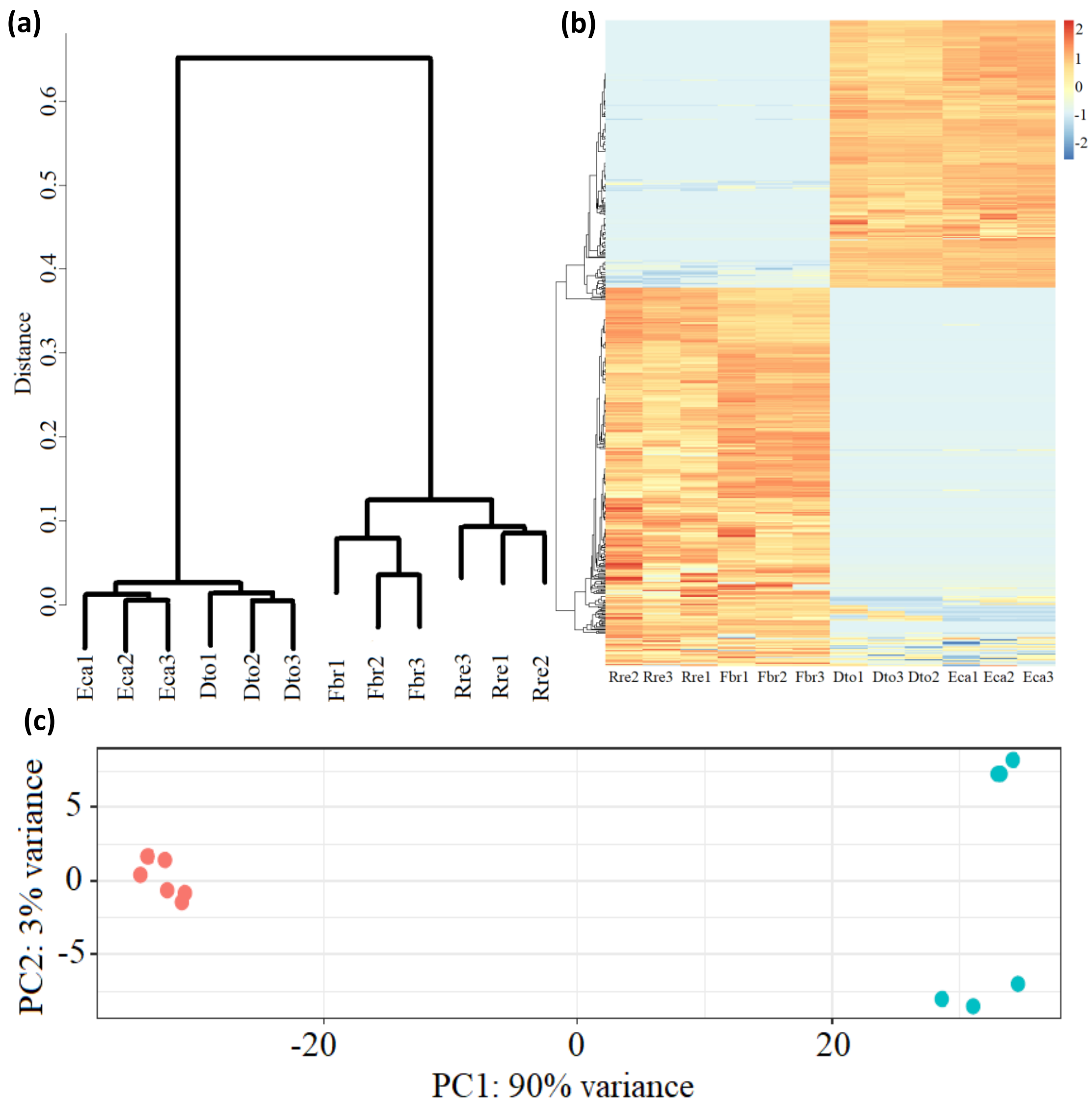


Figure 1. Analysis of differentially expressed genes (DEGs) between colpate pollen and porate pollen species **(a)** Hierarchical clustering shows dissimilarity among the transcriptome samples; distance is calculated by Pearson correlation coefficient. **(b)** Heatmap of transcriptomes of colpate and porate species. Heatmap scale bars indicate $\log_2$fold changes. **(c)** Principal component analysis of the transcriptome samples. Red points represent colpate species; blue points represent porate species. Eca, *Eschscholzia californica*; Dto, *Dactylicapnos torulosa*; Fbr, *Fumaria bracteosa*; Rre, *Roemeria refracta*.

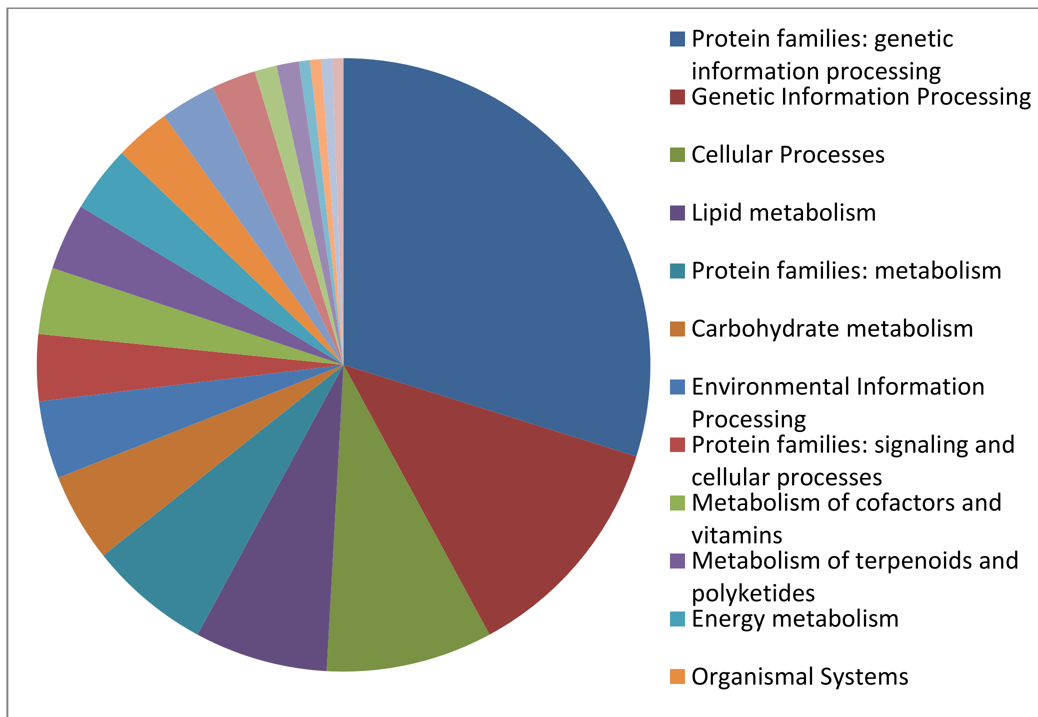


Figure 2. Functional classification of differentially expressed genes (DEGs) between colpate and porate species. Classification of each DEG into KEGG functional categories, using GhostKoala mapping tool.

3.4. Validation of RNA-Seq analysis

To validate the results obtained by RNA-Seq, for four DEGs their differential expression was confirmed with qRT-PCR; 3 of these putatively involved in genetic determinism of the apertural system. The qRT-PCR results were consistent with that of RNA-Seq analysis. Transcription factor *DYSFUNCTIONAL TAPETUM1* (*DYT1*) was more expressed in porate pollen (*R. refraca* and *F. bracteosa*), *MADS-BOX TRANSCRIPTION FACTOR 16* (*MAD16*) and *INAPERTURATE POLLEN 1* (*INP1*) were more expressed in colpate pollen (*D. torulosa* and *E. californica*), and the transcription factor *ABORTED MICROSPORES* (*AMS*) was more expressed in *Eschscholzia californica* than in other species (Figure 3). These results support those obtained by RNA-Seq analysis.

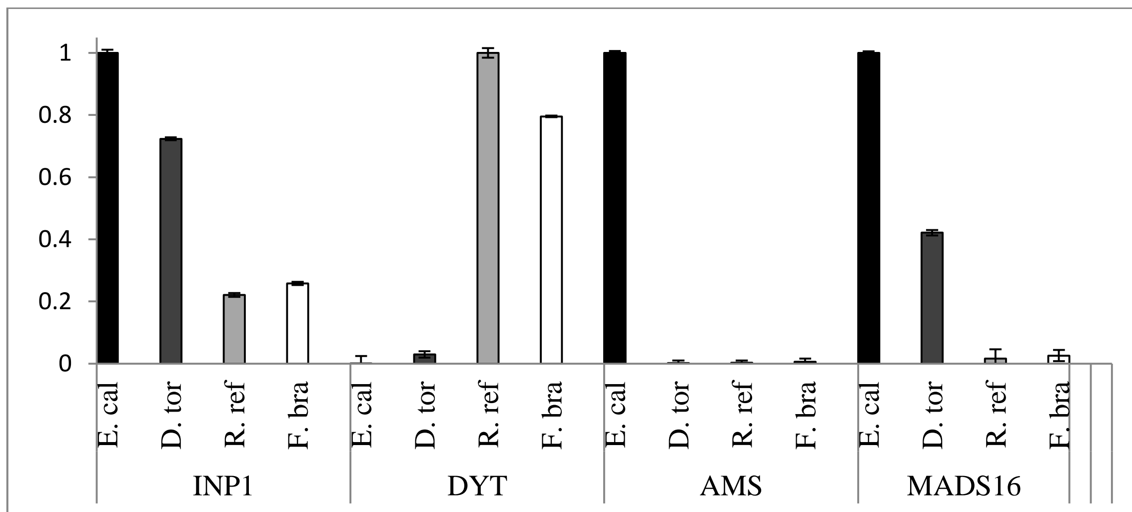


Figure 3. Results of qRT-PCR analysis to confirm the differential expression of four transcripts between the differents studied species. INP1, *INAPERTURATE POLLEN 1*; DYT, *DYSFUNCTIONAL TAPETUM1*; AMS, *ABORTED MICROSPORE*; MADS16, *MADS-BOX TRANSCRIPTION FACTOR 16*. E. cal, *Eschscholzia californica;* D. tor, *Dactylicapnos torulosa;* R. ref, *Roemeria refracta;* F. bra, *Fumaria bracteosa.*

4. Discussion

Pollen apertures play an important role in the reproductive process of spermatophytes, favouring the viability and germination of the male gametophyte, and evolutionary changes in the aperture system have come to be interpreted as key innovations that have led to the success of certain groups such as the Eudicots (Furness and Rudall, 2004). The shape of the aperture is often related to the tolerance to dehydration of the male gametophyte . Furrow-like apertures tend to close and isolate the gametophyte from the dry external environment, even gradually with varying degrees of dryness; whereas pore-like apertures cannot adequately seal the aperture (Božič and Šiber, 2020; Katifori *et al*., 2010; Payne, 1972). However, the distribution of porate pollen throughout most of the angiosperm groups, including species from xeric environments suggests some kind of selection of this apertural type, which could be related to a higher germination speed of the porate pollen because it is dispersed only partially dehydrated, and could lead to a higher reproductive success in the short term (Franchi *et al*., 2002; Prieu *et al*. (2017). Whether this selection involves the selection of a molecular mechanism common to angiosperms for the emergence of pore phenotypes is unknown. The molecular basis of pollen aperture shape remains largely unknown. In this study, we took advantage of the natural variation of apertural morphology in Papaveraceae to adopt a comparative perspective, between colpate and porate species, to find candidate

molecular players involved in apertural shape determinism and their conservation among taxonomic groups.

Recently, Zhou *et al*. (2021) discovered some members of the ELMOD protein family involved in the control of early steps in aperture domain formation in *Arabidopsis*. While the ELMOD_A and ELMOD_B/MCR paralogs specify position and number of aperture domains with redundancy in function, ELMOD_E can interacts with MCR and ELMOD_A activities changing the aperture morphology from colpus to pore; although it does not naturally occurs in *Arabidopsis* (Zhou *et al*., 2021). *ELMOD_A* and *MCR* are both expressed at or near the tetrad stage of pollen development. Experiments by Zhou *et al.* (2021) showed that the neomorphic activity of ELMOD_E occurs only when gene expression levels were elevated during the tetrad phase; however, during this stage *ELMOD_E* expression is low and therefore it does not influence the colpate apertures of *Arabidopsis*. Thus, differential expression of these *ELMOD* genes could be a way to regulate the occurrence of porate pollen. *ELMOD_A* and *MCR* are paralogs that diverged in the common ancestor of the Brassicaceae (Zhou *et al*., 2021); but many species from the main angiosperm groups included in the "A/B clade" of the phylogeny in Zhou *et al*. (2021) have at least two paralogs of this lineage. These authors suggested that these independent duplications could respond to a strong selective pressure to maintain more than one A/B gene and their redundancy. *ELMOD_E* corresponds to a basal ELMOD lineage. All three *ELMOD* genes (2 A/B genes and *ELMOD_E*-like) were found in our four transcriptomes, but they were not differentially expressed between morphs at the tetrad stage we investigated. The most relevant of this result is that it does not support the hypothesis that the interaction among ELMOD_E with the proteins of the A/B lineage regulates the change from colpus to pore in Papaveraceae (a basal eudicot), and then the conserved function of ELMOD_E. The co-expression of A/B lineage genes during the tetrad stage is in agreement with what has been observed in *Arabidopsis* and their possible redundant role in determining aperture domains, although further studies are needed to ascertain whether their function is conserved in other plant groups.

Although ELMOD_E can determine aperture morphology in *Arabidopsis* (Zhou *et al*., 2021), our results suggest that this is not the case in basal eudicots. Our transcriptome analysis comparing species with colpate and porate pollen presents a

basis for identifying DEGs that may represent some candidates involved in the aperture morphology.

Among the 531 DEGs we found one of the genes involved in aperture system determinism, *INP1*, which has been seen to affect aperture morphology to some extent, and it does so only when its expression levels are low. In these cases, furrow-like apertures are significantly reduced in length, something that has been found in *Arabidopsis* and *Eschscholzia californica* (Dobritsa and Coerper 2012; Mazuecos-Aguilera *et al*., 2021). *INP1* is one of the DEGs we have identified among colpate and porate species, being downregulated in the porate species. The direct involvement of *INP1* in the change from colpus to pore is unlikely, as different complementation experiments performed by transforming *Arabidopsis* with *INP1* from other species with different apertures resulted in the typical *Arabidopsis* apertures (Li *et al*., 2018; Lee *et al*., 2021). It is possible that the differential expression of *INP1* is related to the smaller surface area of the pore-like apertures. In the tetrad stage, INP1 accumulates in the plasma membrane of pollen at the sites where the apertures will appear. Then INP1 prevents the deposition of sporopollenin in these areas and the apertures are formed. Thus, the porate pollen will need less protein to cover the pre-established aperture domains than the colpate pollen.

However, the recently described *INAPERTURATE POLLEN 2* (*INP2*), a species-specific partner for *INP1* in *Arabidopsis*, is upregulated in porate species, but is also differentially expressed between the two porate species. Therefore, it does not show a similar expression pattern to *INP1* as would be expected because these two proteins interact with each other for aperture formation and have co-evolved (Lee *et al.,* 2021). Thus, perhaps the function of *INP2* is not conserved in angiosperms as a whole or perhaps in this species another species-specific partner of INP1 fulfills the function of INP2.

The callose wall seems to play an essential role in the development of apertures (Wang and Dobritsa, 2018). Apertures develop where membrane ridges maintain close contact with the callose wall, avoiding primexine deposition and thus preventing the exine development; INP1 appears to be involved in maintaining these ridges near the callose wall (Dobritsa *et al.* 2018; Wang and Dobritsa, 2018). In addition, it has been shown that the position of the apertures is related to where additional depositions of

callose are formed after cytokinesis, and that these depositions have the same shape as the future apertures, suggesting they must determine the shape of the aperture (Albert *et al*., 2010; Prieu *et al*., 2017b). So far, no molecular factor involved in these deposits is known. Interestingly we found five DEGs annotated as genes that affect in some way the deposition or degradation of callose. One of them, upregulated in porate pollen, is a homologue of *DYSFUNCTIONAL TAPETUM1* (*DYT1*), which encodes a bHLH transcription factor. *DYT1* is strongly expressed in the tapetum, where it is important for the expression of about 1,000 anther genes, including genes involved in callose synthesis and degradation, and pollen wall development (Feng *et al*.; 2012; Zhou *et al*., 2017). *DYT1* expression declines rapidly at the end of meiosis of meiocytes. The *dyt1* mutant exhibits abnormal anther morphology and meiocytes do not have a thick callose wall although they are able to complete meiosis (Zhang *et al*., 2006). According to Feng *et al*. (2012), it is possible that DYT1 regulates genes that are essential for active metabolism in the tapetal cells and for export of materials to meiocytes, such as those for callose wall formation. More research is needed to see if it affects the additional deposition of callose that determines the position and perhaps the shape of the apertures. Interestingly, in the study by Feng *et al*. (2012) on the regulation exerted by DYT1 on *Arabidopsis* anther genes, one of the genes downregulated in *dyt1* knockout mutants, and mutants of its paralogue *AMS* (*ABORTED MICROSPORES*), was *ELMOD_E* (see Supplementary Tables 1, 11 in Feng *et al.*, 2012); directly linking *DYT1* and *AMS* to the only known gene affecting aperture morphology in *Arabidopsis*. It has been shown that tapetum transcription factors may affect the development of the pollen wall. Lou *et al.* (2014) showed a direct regulation by *AMS* of tapetum genes for the sexine and nexine formation, showing that a transcriptional cascade in the tapetum specifies the development of pollen wall. Thus, *ELMOD_E* could be one of the targets of this transcriptional cascade. One of the genes regulated by *AMS* and which specifies formation of the nexine layer, *TRANSPOSABLE ELEMENT SILENCING VIA AT-HOOK* (*TEK*), has recently been shown to be directly involved in callose synthesis by negatively regulating *CALLOSE SYNTHASE5* (*CalS5*) after the tetrad stage (Xiong *et al*., 2020), when the callose wall is believed to begin to be degraded by tapetally secreted callase activity (Stieglitz, 1977).

On the other hand, among our DEGs we found two β-1,3-glucanases upregulated in porate pollen, which are involved in callose degradation. One of them, annotated as

Glucan endo-1,3-beta-glucosidase 10/AtBG-pap and encoded by the gen *PUTATIVE PLASMODESMAL ASSOCIATED PROTEIN* (*PPAP*) (At5g42100; Levy *et al*., 2007), is a plasmodesmal-associated membrane protein involved in plasmodesmal callose degradation. This β-1,3-glucanase could be recruited, in addition to plasmodesmata, to other specific membrane domains such as additional callose deposits where it would act by degrading callose in porate pollen and thus contribute to the determination of the shape of the apertures. However, this needs further studies to be confirmed. The other β-1,3-glucanase was annotated as *β-1,3-GLUCANASE 5*/*BG5*, whose protein is localised extracellularly (Delp and Palva, 1999).

Two genes more involved in the callose deposition were annotated as the two homologues of *Oryza sativa BIFUNCTIONAL NUCLEASE IN BASAL DEFENSE RESPONSE 1* (*OsjBBD1*) and *BIFUNCTIONAL NUCLEASE IN BASAL DEFENSE RESPONSE 2* (O*sjBBD2*), nucleases with both RNase and DNase activities. BBD1 functions in cell wall reinforcement through the abscisic acid-derived callose deposition that is induced following infection by a necrotrophic pathogen, probably through activation of *PDF1.2*, *ABA1*, and *AtSAC1* gene expression (You *et al*., 2010); function of BBD2 is not clear but seems to be redundant to BBD1 (Huque *et al*., 2020). In *Arabidopsis*, each AtBBD protein forms a homodimer but they do not form a heterodimer (Huque *et al*., 2020). Our results showed that *BBD1*-like was upregulated in porate pollen, while *BBD2*-like was upregulated in colpate pollen, suggesting differential dominance according to the type of apertures. The role of this gene in pollen is unknown.

Thus, any of these genes that directly or indirectly affect the deposition and degradation of callose could act by determining the shape of additional deposition of callose and thereby affect the shape of the apertures.

Among the Go categories we found that many DEGs are integral components of the membrane. Seven of them, 2 upregulated in colpate pollen and 5 upregulated in porate pollen, code for transmembrane transporters. Two of the five upregulated in porate pollen were annotated as belonging to the plant drug/metabolite exporter (DME) family, homologues of the *Arabidopsis* genes *WAT1(WALLS ARE THIN1)-related protein7* (*WTR7*) and *WAT1-related protein8* (*WTR8*). WAT1 is required for secondary cell-wall deposition (Ranocha *et al*., 2010), and this and other proteins of this family are

auxin-induced proteins involved in different development processes (Busov *et al*., 2004; Pomares-Viciana *et al*., 2019; Ranocha *et al*., 2010). In *Arabidopsis*, this protein family involves at least 38 members, most of unknown function; according to Busov *et al.* (2004), this multigene structure suggests functional divergence of its members. Because the callose wall (surrounding the microspore mother cell (MMC) and tetrads) is quite impermeable to many primexin components and other components involved in exine patterning (Wang and Dobritsa, 2018), most exine precursors must be synthesized by the MMCs and/or microspores and transported across the membrane to occupy their final positions. Thus, differentially expressed transmembrane transporters suggest the need to transport different factors, or in different quantities, depending on the aperture morphology.

Other DEGs were classified as involved in the organisation of actin filaments, the cytoskeleton or microtubules, others as having an actin-binding function, and others as components of microtubules or the cytoskeleton. Several studies have described a relationship between the organisation of cytoskeletal elements and the formation of apertures. In *Nicotiana*, the distribution of microtubules in post-meiotic cytokinesis is related to the number of apertures (Ressayre *et al*., 2002). In *Vigna*, there is a spatial correlation between microtubules and exine, with cytoplasmic patches of microtubules being observed where apertures will form during the tetrad stage (Muñoz *et al*., 1995); and in *Lilium henryi*, the microtubular organizing centers (MTOCs) participate in locating the endoplasmic reticulum in the area where the aperture will appear, which acts as a barrier for the deposition of sporopollenin (Sheldon and Dickinson 1986). In addition, a certain organization of the microtubules can produce undulations in the plasma membrane, which is associated with the construction process of the primexine framework and the apertures (Ariizumi and Toriyama 2011).

One of our DEGs related to the cytoskeleton, upregulated in porate pollen, was annotated as a homologue of *NETWORKED 4B* (*NET4B*) of *Arabidopsis*, a member of Networked (NET) superfamily. NET proteins possess an actin-binding region (NAB domain) and are membrane-associated. These proteins specifically link actin filaments to cell membranes to specify different membrane domains (Deeks *et al*., 2012). In *Arabidopsis*, NET superfamily is composed of 4 subfamilies (NET1-4). *NET1A* is a gene highly expressed in plasmodesmata, where the actin cytoskeleton, plasma membrane, and endoplasmic reticulum are brought together. Plasmodesmata are small

and defined areas in walls, such as apertures. Genes from *NET2* subclade are expressed preferentially in pollen, which could indicate interactions with pollen-specific ligands (Deeks *et al*., 2012). NET4A protein localises to highly constricted regions of the vacuolar membrane and contributes to vacuolar morphology (Kaiser *et al*., 2019).

Other DEG involved in cytoskeleton organization was annotated as an homologue of *VILLIN5* (*VLN5*), which is preferentially expressed in the *Arabidopsis* pollen, and which was upregulated in porate pollen. VLN5 is an actin binding protein and plays a key role in the formation of higher-order structures from actin filaments and in the regulation of actin dynamics in eukaryotic cells. Villin family members from plants have been shown to sever, cap, and bundle actin filaments (Zhang *et al*., 2010). Loss of VLN5 function retarded pollen tube growth and actin filaments were more sensitive to depolymerization in *Arabidopsis* pollen grains and tubes (Zhang *et al*., 2010). In addition, VLN5 functions in concert with oscillatory calcium gradients in pollen. In relation to this, we find different DEGs related to calcium, such as calcium transporters, calcium-binding proteins, or calcium channels.

Other DEGs are involved in remodelling of wall. One of them was annotated as *α-L-Arabinofuranosidase* (*ASD2*), a bifunctional a-L-arabinofuranosidase/b-D-xylosidase. Mutants of *ASD2* present differences in cell wall composition and structure (Chávez Montes *et al*., 2008).

Also, among the DEGs we found homologues of two *WALL-ASSOCIATED RECEPTOR KINASES LIKE* (*WAKL*), *WAKL22* and *WAKL6*, encoding cell wall-associated receptor-like kinase (RLKs). WAK have the potential to directly signal cellular events through their cytoplasmic kinase domain, in addition, WAK protein may have an effect on other cell wall proteins (He *et al*., 1999; Kohorn, 2001). WAKL22 is involved in response to pathogen infection (Diener and Ausubel 2005), however WAKL6 function has not been described yet. On the other hand, D6PKL3, whose involvement in aperture formation has been confirmed, requires its kinase activity for its correct functioning (Zhou and Dobritsa, 2019). Therefore, these WAKLs, like *D6PKL3*, could act by phosphorylating other aperture factors in the aperture domains. Also many other RLKs have been found among the DEGs.

In addition to DYT, other DEGs were annotated as transcription factors, such as the one annotated as *OsMADS16*, which specifies the identity of stamens and lodicles in

rice. It will therefore act upstream of the genes discussed above and regulate many of them (Xiao *et al*., 2003). We also found interesting proteins due to their location close to the wall, such as transmembrane proteins. These proteins, by their place of action, could be involved in the formation of distinct aperture domains. Functional analysis will be necessary to confirm whether some of the proposed DEGs are involved in determination of aperture shape.

References

Albert B, Nadot S, Dreyer L, Ressayre A. 2010. The influence of tetrad shape and intersporal callose wall formation on pollen aperture pattern ontogeny in two eudicot species. *Annals of Botany* **106**: 557-564.

Ariizumi T, Toriyama K. 2011. Genetic regulation of sporopollenin synthesis and pollen exine development. *Annual review of plant biology* **62**: 437-460.

Ashburner M, Ball CA, Blake JA, *et al*. 2000. Gene ontology: tool for the unification of biology. *Nature genetics* **25**: 25-29.

Bairoch A, Boeckmann B. 1992. The SWISS-PROT protein sequence data bank. *Nucleic acids research* **20**: 2019.

Blackmore S, Wortley AH, Skvarla JJ, Rowley JR. 2007. Pollen wall development in flowering plants. *New Phytologist* **174**: 483-498.

Božič A, Šiber A. 2020. Mechanical design of apertures and the infolding of pollen grain. *Proceedings of the National Academy of Sciences of the United States of America* **117**: 26600-26607.

Busov VB, Johannes E, Whetten RW, Sederoff RR, Spiker SL, Lanz-Garcia C, Goldfarb B. 2004. An auxin-inducible gene from loblolly pine (Pinus taeda L.) is differentially expressed in mature and juvenile-phase shoots and encodes a putative transmembrane protein. *Planta* **218**: 916-927.

Chávez Montes RA, Ranocha P, Martinez Y. 2008. Cell wall modifications in Arabidopsis plants with altered α-l-arabinofuranosidase activity. *Plant Physiology* **147**: 63-77.

Deeks MJ, Calcutt JR, Ingle EK. 2012. A superfamily of actin-binding proteins at the actin-membrane nexus of higher plants. *Current Biology* **22**: 1595-1600.

Delp G, Palva ET. 1999. A novel flower-specific Arabidopsis gene related to both pathogen-induced and developmentally regulated plant β-1, 3-glucanase genes. *Plant molecular biology* **39**: 565-575.

Diener AC, Ausubel FM. 2005. RESISTANCE TO FUSARIUM OXYSPORUM 1, a dominant Arabidopsis disease-resistance gene, is not race specific. *Genetic* **171**: 305-321.

Dobritsa AA, Coerper D. 2012. The novel plant protein INAPERTURATE POLLEN1 marks distinct cellular domains and controls formation of apertures in the Arabidopsis pollen exine. *The Plant Cell* **24**: 4452-4464.

Edlund AF, Swanson R, Preuss D. 2004. Pollen and stigma structure and function: the role of diversity in pollination. *The Plant Cell* **16(suppl 1)**: S84-S97.

Feng B, Lu D, Ma X, Peng Y, Sun Y, Ning G, Ma H. 2012. Regulation of the Arabidopsis anther transcriptome by DYT1 for pollen development. *The plant journal* **72:** 612-624.

Franchi GG, Nepi M, Dafni A, Pacini E. 2002. Partially hydrated pollen: taxonomic distribution, ecological and evolutionary significance. *Plant Systematics and Evolution* **234**: 211-227.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**: 3150-3152.

Furness CA, Rudall PJ. 2004. Pollen aperture evolution–a crucial factor for eudicot success?. *Trends in plant science* **9**: 154-158.

Götz S, García-Gómez JM, Terol J, *et al*. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research* **36**: 3420-3435.

Grabherr MG, Haas BJ, Yassour M, *et al*. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature biotechnology* **29**: 644.

Haas B, Papanicolaou A. 2016. TransDecoder (find coding regions within transcripts).

He ZH, Cheeseman I, He D, Kohorn BD. 1999. A cluster of five cell

wall-associated receptor kinase genes, Wak1–5, are expressed in specific organs

of Arabidopsis. *Plant Mol. Biol. 39*: 1189–1196.

Heslop-Harrison J. 1979. An interpretation of the hydrodynamics of pollen. *American Journal of Botany* **66**: 737-743.

Huque AKM, So WM, You MK, Shin JS. 2020. Phylogenetic Analysis and In Vitro Bifunctional Nuclease Assay of Arabidopsis BBD1 and BBD2. *Molecules* **25**: 2169.

Kaiser S, Eisa A, Kleine-Vehn J, Scheuring D. 2019. NET4 modulates the compactness of vacuoles in Arabidopsis thaliana. *International journal of molecular sciences* **20**: 4752.

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**: 27-30

Kanehisa M, Sato Y, Morishima K, *et al*. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology* **428**: 726-731.

Katifori E, Alben S, Cerda E, Nelson DR, Dumais J. 2010. Foldable structures and the natural design of pollen grains. *Proceedings of the National Academy of Sciences* **107**: 7635-7639.

Kohorn BD. 2001. WAKs; cell wall associated kinases. *Current Opinion in Cell Biology* **13**: 529-533.

Lee BH, Weber ZT, Zourelidou M, *et al.* 2018. Arabidopsis protein kinase D6PKL3 is involved in the formation of distinct plasma membrane aperture domains on the pollen surface. *The Plant Cell* **30**: 2038-2056.

Lee BH, Wang R, Moberg IM, *et al.* 2021. A species-specific functional module controls formation of pollen apertures. *Nature Plants* **7**: 966-978.

Levy A, Erlanger M, Rosenthal M, Epel BL. 2007. A plasmodesmata-associated β-1, 3-glucanase in Arabidopsis. *The Plant Journal* **49**: 669-682.

Li P, Ben-Menni Schuler S, Reeder SH, Wang R, Suárez Santiago VN, Dobritsa AA. 2018. INP1 involvement in pollen aperture formation is evolutionarily conserved and may require species-specific partners. *Journal of Experimental Botany* **69**: 983-996.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. *Methods* **25**: 402-408.

Lou Y, Xu XF, Zhu J, Gu JN, Blackmore S, Yang ZN. 2014. The tapetal AHL family protein TEK determines nexine formation in the pollen wall. *Nature communications* **5**: 1-9.

Matamoro-Vidal A, Raquin C, Brisset F, et al. 2016. Links between morphology and function of the pollen wall: an experimental approach. Botanical journal of the Linnean Society **180**: 478-490.

Mazuecos-Aguilera I, Romero-García AT, Klodová B, *et al*. 2021. The role of INAPERTURATE POLLEN1 as a pollen aperture factor is conserved in the basal eudicot Eschscholzia californica (Papaveraceae). *Frontiers in Plant Science* **1405**.

Muñoz CA, Webster BD, Jernstedt JA, *et al*. 1995. Spatial congruence between exine pattern, microtubules and endomembranes in Vigna pollen. *Sexual Plant Reproduction* **8**: 147-151.

Nadot S, Forchioni A, Penet L, Sannier J, Ressayre A. 2006. Links between early pollen development and aperture pattern in monocots. *Protoplasma* **228**: 55.

PalDat 2017. PalDat – A Palynological Database, www.paldat.org (accessed 2021)

Payne WW. 1972. Observations of harmomegathy in pollen of Anthophyta. *Grana* **12**: 93-98.

Pérez-Gutiérrez MA, Romero-García AT, Fernández MC, *et al*., 2015. Evolutionary history of fumitories (subfamily Fumarioideae, Papaveraceae): An old story shaped by the main geological and climatic events in the Northern Hemisphere. *Molecular phylogenetics and evolution* **88**: 75-92.

Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature protocols.* **11**: 1650.

Pomares-Viciana T, Del Río-Celestino M, Román B, Die J, Pico B, Gómez P. 2019. First RNA-seq approach to study fruit set and parthenocarpy in zucchini (Cucurbita pepo L.). *BMC plant biology* **19**: 1-20.

Prieu C, Sauquet H, Gouyon PH, Albert B. 2017. More than sixty origins of pantoporate pollen in angiosperms. *American journal of botany* **104**: 1837-1845.

Prieu C, Toghranegar Z, Matamoro-Vidal A, Nadot S, Albert B. 2017b. Additional callose deposits are located at the future apertural regions in sulcate, ulcerate, porate,

colporate, colpate and syncolpate pollen grains. *Botanical Journal of the Linnean Society* **183**: 271-279.

Ranocha P, Denancé N, Vanholme R, *et al.* 2010. Walls are thin 1 (WAT1), an Arabidopsis homolog of Medicago truncatula NODULIN21, is a tonoplast-localized protein required for secondary wall formation in fibers. *The Plant Journal* **63**: 469-483.

Ressayre A, Raquin C, Mignot A, Godelle B, Gouyon PH. 2002. Correlated variation in microtubule distribution, callose deposition during male post-meiotic cytokinesis, and pollen aperture number across Nicotiana species (Solanaceae). *American Journal of Botany* **89**: 393-400.

Sheldon JM, Dickinson HG. 1986. Pollen wall formation in Lilium: the effect of chaotropic agents, and the organisation of the microtubular cytoskeleton during pattern development. *Planta* **168**: 11-23.

Stieglitz H. 1977. Role of β-1, 3-glucanase in postmeiotic microspore release. *Developmental biology* **57**: 87-97.

Untergasser A, Cutcutache I, Koressaar T, *et al.* 2012. Primer3 – new capabilities and interfaces. *Nucleic Acids Research* **40**: e115-e115.

Volkova OA, Severova EE, Polevova SV. 2013. Structural basis of harmomegathy: evidence from Boraginaceae pollen. *Plant Systematics and Evolution* **299**: 1769–1779.

Wodehouse R. 1935. *Pollen grains, their structure, identification and significance in science and medicine*. New York: McGraw-Hill.

Walker JW, Doyle JA. 1975. The bases of angiosperm phylogeny: palynology. *Annals of the Missouri Botanical Garden* 664-723.

Wang R, Dobritsa AA. 2018. Exine and aperture patterns on the pollen surface: their formation and roles in plant reproduction. *Annual Plant Reviews Online* 589-628.

Wellmer F, Alves-Ferreira M, Dubois A, Riechmann JL, Meyerowitz EM. 2006. Genome-wide analysis of gene expression during early Arabidopsis flower development. *PLoS Genet* **2**: e117.

Xiao H, Wang Y, Liu D, *et al.* 2003. Functional analysis of the rice AP3 homologue OsMADS16 by RNA interference. *Plant molecular biology* **52**: 957-966.

Xiong SX, Zeng QY, Hou JQ, *et al.* 2020. The temporal regulation of TEK contributes to pollen wall exine patterning. *PLoS genetics* **16**: e1008807.

You MK, Shin HY, Kim YJ, *et al.* 2010. Novel bifunctional nucleases, OmBBD and AtBBD1, are involved in abscisic acid-mediated callose deposition in Arabidopsis. *Plant physiology* **152**: 1015-1029.

Zhang H, Qu X, Bao C, *et al.* 2010. Arabidopsis VILLIN5, an actin filament bundling and severing protein, is necessary for normal pollen tube growth. *The Plant Cell* **22**: 2749-2767.

Zhang H, Jin J, Tang L, *et al*. 2011. PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. *Nucleic Acids Res* **39**: D1114–D1117.

Zhang W, Sun Y, Timofejeva L, *et al*. 2006. Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor. *Development* **133**: 3085-3095.

Zhang X, Zhao G, Tan Q, *et al*. 2020. Rice pollen aperture formation is regulated by the interplay between OsINP1 and OsDAF1. *Nature Plants* **6:** 394-403.

Zhou S, Zhang H, Li R, Hong Q, Li Y, Xia Q, Zhang W. 2017. Function identification of the nucleotides in key cis-element of Dysfunctional Tapetum1 (DYT1) promoter. *Frontiers in plant science* **8**: 153.

Zhou Y, Dobritsa AA. 2019. Formation of aperture sites on the pollen surface as a model for development of distinct cellular domains. *Plant Science* **288**: 110222.

Zhou Y, Amom P, Reeder SH, Lee BH, Helton A, Dobritsa AA. 2021. Members of the ELMOD protein family specify formation of distinct aperture domains on the Arabidopsis pollen surface. *Elife* **10**: e71061.

Supporting Information

Supplementary Table 1. Primer sequences used in this study.

Primer	Sequence
ActinRrefqPCRFw1	GTTGCACCACCAGAGAGGAA
ActinRrefqPCRRv1	TGACTCGTCGTACTCCCCTT
ActinDtorqPCRFw1	AGCTCGCATATGTGGCTCTT
ActinDtorqPCRRv1	ACCATCAGGCAGCTCGTAAC
ActinFbraqPCRFw2	GCCATCCTTCGTTTGGACCT
ActinFbraqPCRRv1	ACAATTTCCCGCTCAGCAGT
ActinEcaqPCRFw2	TTACAATGAGCTTCGTGTTGC
ActinEcaqPCRRv2	CCCAGCACAATACCTGTAGTAC
EcaINP1Fw	ATGATCAAAGCTGCAGCTCGA
EcaINP1Rv	AATGCCTGATAATGGAATCTTGC
INP1FbraqPCRFw2	TCGAACAGTGCCAGATTCCC
INP1FbraqPCRRv2	CAGTCGCAGAAAGCCCCATA
INP1DtorqPCRFw3	CCTACTCTCGTTACCCGTGC
INP1DtorqPCRRv3	GCCTTTGATACCTCCTGCCT
INP1RrefqPCRFw3	TCCCAACTTGGCACAACTCA
INP1RrefqPCRRv3	TCATCGTCGTCGTAATTGAGTGA
DYT_FbraFw2	TCAGTTTCAGAGGAGCAACCC
DYT_FbraRv2	AGCAAGAAGTTCCTCACCTGT
DYT_RrefFw1	GCCACTCTTGACGACGCTAT
DYT_RrefRv1	CTTCAACGGCTCGAGGAACT
DYT_DtorFw1	GGAACCGACCAGCTGAATGAT
DYT_DtorRv1	ATCCAAGAACAAGCCAGCACA
MAD16_DtorFw2	TGCTATCGTAAGCCTGGTGG
MAD16_DtorRv2	TGCGTGCGAGGTTTTTACAC
MAD16_FbraFw2	TGCACTAGCAAACGAAGGGG
MAD16_FbraRv2	AAGATTGGGCTGACTTGGCT

MAD16_RrefFw2	CATGTTCTGCACTCGCACAC
MAD16_RrefRv2	AGCCGCATAAACTGCACGTA
MAD16_EcaFw2	GAAATGGCGAATGGTGGTGTT
MAD16_EcaRv2	AAGCAAGGCGTAGATCGTGA
AMS_DtorFw1	AAGCTGCCAGTACTCACGTC
AMS_DtorRv1	GGGTCGGCCATACGAAAGAA
AMS_EcaFw2	TGCGCGAACGTGATTCTTTC
AMS_EcaRv2	CTCCAGAAATTCCCCTGTCCC
AMS_RrefFw1	TGCTGCAAAACATGTACGCC
AMS_RrefRv1	AGCATTACGCCATCCATCCA
AMS_FbraFw1	GCCATGGGTTATCTCGGACT
AMS_FbraRv1	ACATCACTTGTTTGAGGAGGT

Supplementary Table 2. Summary of sequencing, assembly and annotation of *Dactylicapnos torulosa, Fumaria bracteosa, Roemeria refracta* and *Eschscholzia californica.*

		Dactylicapnos torulosa			*Fumaria bracteosa*			*Roemeria refracta*			*Eschscholzia californica*		
	Sample	1	2	3	1	2	3	1	2	3	1	2	3
Raw data sequencing	Read bases (bp)	6723757898	6425362872	7085366960	6968752378	7452993070	6517429384	7147818144	6295312612	7462078740	8042227988	6888987232	6995468506
	Reads	44528198	42552072	46922960	46150678	49357570	43161784	47336544	41690812	49417740	53259788	45622432	46327606
	Total reads (paired reads)	556328184 (278164092)											
	Q20(%)	97.889	97.955	97.889	98.082	98.094	97.924	98.063	98.058	97.836	97.954	97.905	98.003
	Q30(%)	94.016	94.083	94.042	94.404	94.344	94.040	94.307	94.334	94.102	94.095	93.975	94.156
	No HQ reads	24652	23548	27460	25714	26586	27714	26396	28512	38326	48188	35444	33742
	HQ reads	133927570			138590018			138351862			145092452		
	Total 'genes'	181486			138846			277766			34156		
	Total transcripts	294904			235229			412427			37400		
	Percent GC	38.85			38.9			39.05			38.93		
Stats ALL transcript contigs	Contig N50	2253			2487			1331					
	Median contig length	598			697			422					
	Average contig length	1177			1321			782					
	Total assembled bases	311793519			273138652			322858588					
Stats ONLY LONGEST ISOFORM	Contig N50	1255			1480			677					
	Median contig length	416			439			333					
	Average contiglength	775.52			850.2			548.1					
	Total assembled bases	140745253			118055325			152243515					
	Transdecoder filter transcripts	127049			115642			147231					
	Non-redundant transcript	56689			52752			92051					
	Swissprot annot blastx	209199 (70.94 %)			169720 (72.15 %)			340571 (82.57 %)			32996 (88.22 %)		
	Swissprot annot blastp	122989 (41.7 %)			112115 (47.66 %)			142477 (34.54 %)					

Supplementary Table 3. Trinotate annotation report for *Dactylicapnos torulosa*. Annotation through blastx and blastp for predict transcript against Swissprot database.

Available at FigShare repository.

Supplementary Table 4. Trinotate annotation report for *Fumaria bracteosa*. Annotation through blastx and blastp for predict transcript against Swissprot database.

Available at FigShare repository.

Supplementary Table 5. Trinotate annotation report for *Eschscholzia californica*. Annotation through blastx and blastp for predict transcript against Swissprot database.

Available at FigShare repository.

Supplementary Table 6. *Eschscholzia californica* annotation through blastx against Swissprot database for transcripome transcript aligned to reference. Qseqid, query or source (e.g., gene) sequence id; seqid, subject or target (e.g., reference genome) sequence id; pident, percentage of identical matches; length, alignment length (sequence overlap); mismatch, number of mismatches; gapopen, number of gap openings; qstart, start of alignment in query; qend, end of alignment in query; sstart, start of alignment in subject; send, end of alignment in subject; evalue, expect value; bitscore, bit score.

Available at FigShare repository.

Supplementary Table 7. Functional classification for transcripts of *Fumaria bracteosa, Roemeria refracta, Dactylicapnos torulosa* and *Eschscholzia californica* transcriptome assembly. Classification into KEGG functional categories, using GhostKoala mapping tool.

Category \ Specie	*Fumaria bracteosa*	*Roemeria refracta*	*Dactylicapnos torulosa*	*Eschscholzia californica*
Protein families: genetic information processing	4,780	6,276	4,751	2,264
Genetic Information Processing	3,668	6,425	3,758	2,270
Carbohydrate metabolism	1,489	1,706	1,528	948
Protein families: signaling and cellular processes	1,467	1,727	1,451	682
Protein families: metabolism	1,466	1,736	1,510	682
Environmental Information Processing	1,213	1,948	1,146	754
Cellular Processes	965	1,062	940	518
Lipid metabolism	768	878	689	429

Unclassified: metabolism	753	866	779	425
Organismal Systems	706	852	644	451
Amino acid metabolism	640	957	663	399
Metabolism of cofactors and vitamins	522	593	467	229
Energy metabolism	427	439	383	334
Glycan biosynthesis and metabolism	426	360	390	162
Human diseases	408	784	434	250
Metabolism of terpenoids and polyketides	355	320	312	139
Nucleotide metabolism	273	339	286	158
Byosinthesis of other secondary metabolites	243	390	188	140
Unclassified	144	167	96	46
Unclassified: signaling and cellular processes	134	137	123	93
Metabolism of other aminoacids	85	188	84	70
Unclassified: genetic information processing	35	38	50	25
Xenobiotics degradation and metabolism	9	8	5	6
Total	20,979 (39.8%)	28,195 (30.6%)	20,677 (36.5%)	11,474 (33.6%)
Entradas	52752	92051	56689	34156

Supplementary Table 8. Transcripts annotated as transcription factor through PlantTFDB. Available at FigShare repository.

Supplementary Table 9. Genes differentially expressed between colpate and porate species. Annotation through BLASTX searching against the SwissProt Database.

Available at FigShare repository.

Supplementary Table 10. Functional classification of differentially expressed genes (DEGs) between colpate and porate species using Blast2GO software. Sheet 1, Blast2GO output file with annotations and functional classification for each DEG. Sheet2, summary of the number of DEGs for each functional annotation within the three different functional categories, note that each DEG can have several annotation possibilities.

Available at FigShare repository.

General discussion

General discussion

In this thesis we have confirmed that the protein INAPERTURATE POLLEN 1 (INP1) retains its function in *Eschscholzia californica*, being indispensable for the formation of pollen grain apertures (Mazuecos-Aguilera *et al*., 2021). INP1 function had already been confirmed in *Arabidopsis thaliana*, member of eudicots, in which the gene *INP1* was first described, and in maize (*Zea mays*) and rice (*Oryza sativa*), members of monocots (Dobritsa and Coerper, 2012; Li *et al*., 2018; Zhang *et al*., 2020). While *Arabidopsis* (Brassicaceae, Rosales) has the typical tricolpate apertural system, and rice and maize the typical ulcerate pollen of the gramineae, *Eschscholzia californica*, a member of the family Papaveraceae of the most basal order of eudicots (Ranunculales), has a very different apertural system with 5-7 colpate pollen (Mazuecos-Aguilera *et al*., 2021). The confirmation of INP1 function in a member of Papaveraceae and with an apertural system different from the previously described species, allowed us to confirm that INP1 conserves its function in the transition from monocots to eudicots, and probably in the whole of the angiosperms. On the other hand, the expression pattern of *EcINP1* in *E. californica* is similar to that of *Arabidopsis* and rice. The maximum expression of *EcINP1* at the tetrad stage might indicate that, as in *Arabidopsis*, at this stage of microsporogenesis EcINP1 is anchored in the aperture domains between the plasma membrane and the callose wall surrounding the microspores, preventing primexine deposition (Dobritsa *et al*., 2018).

We found that apertures are indispensable for pollen grain germination in *Eschscholzia californica*, as in *Arabidopsis* and unlike grasses (Dobritsa and Coerper, 2012; Li *et al*., 2018; Zhang *et al*., 2020). This could be due to differences in wall morphology such as thickness and tectum sculpture or due to differences in the fluid exchange process with the stigma (Edlund *et al*., 2016; Mazuecos-Aguilera *et al*., 2021). Thus the pollen tube of these species may choose the shortest route into the stigma in pollen germination (Zhou and Dobritsa, 2019).

Through generation of transient virus-induced gene silencing (VIGS) mutants, we corroborated that EcINP1 is not involved in determining the shape, position or number of apertures. Like *Arabidopsis*, decreasing the EcINP1 protein dose results in shorter and shallower apertures, but does not change the aperture pattern (Dobritsa and Coerper, 2012).

Among the 971 differentially expressed genes (DEGs) between *E. californica inp1* mutant and wild-type *E. californica* we found the *NUCLEOSOME ASSEMBLY PROTEIN 1* (*NAP1*) gene. *NAP1* in Vitis vinifera subsp. sylvestris, together with a deleted homologue of *INP1*, could be involved in pollen sterility in female flowers possibly due to the lack of apertures (Badouin *et al*., 2020). We also found two DEGs belonging to AGC1 kinases: one related to *D6 PROTEIN KINASE-LIKE 3* (*D6PKL3*) from *Arabidopsis* and the other to *PROTEIN KINASE ASSOCIATED WITH BRX* (*PAX)*. D6PKL3 act upstream of INP1 and although less essential than INP1, is required for aperture formation, and its protein dosage determines the number of apertures generated (Lee *et al*., 2018). D6PKL3 marks the membrane domains where INP1 must act to prevent sporopollenin deposition. On the other hand, D6PKL3 needs INP1 to remain anchored to the membrane (Lee *et al*., 2018; Zhou and Dobritsa, 2019). PAX is another AGC1 kinase involved in membrane domain formation (Lee *et al*., 2018; Zourelidou et al., 2009). These three genes also show a similar temporal expression in *E. californica* as *INP1*, with a maximum expression at the tetrad stage, making their involvement in the formation of apertures and their interaction with INP1 very probable.

By comparative transcriptome analysis of species with colpate pollen versus species with porate pollen, we found 531 DEGs that could be involved in aperture shape determinism. Among the DEGs we find the INP1 gene, although this has already been discarded as a determinant of the shape of the apertures (Coerper and Dobritsa, 2012; Mazuecos-Aguilera *et al*., 2021); so it possibly appears due to the difference in surface area occupied by the colpate and porate apertures, since it is attracted to the membrane domains where the apertures will appear. In contrast, its co-evolved partner in *Arabidopsis, INAPERTURATE POLLEN 2* (*INP2*) is not differentially expressed so its function may not be conserved in Papaveraceae. This, together with the fact that INP1 degradation is delayed in *E. californica* compared to *Arabidopsis*, could indicate that INP1 interacts with different protein partners in Papaveraceae.

On the other hand, we can emphasise other DEGs related to the deposition or degradation of callose, since the callose wall has been implicated in the formation of apertures and is necessary for the correct functioning of INP1 (Dobritsa *et al.* 2018; Wang and Dobritsa, 2018). One of them is the transcription factor *DYSFUNCTIONAL TAPETUM1* (*DYT1*), which, among others, regulates genes involved in the formation of the callose wall (Feng *et al*.; 2012; Zhang *et al*., 2006). We also found other DEGS

related to the organisation of cytoskeletal elements, which are also related to the formation of apertures (Ressayre *et al*., 2002; Muñoz *et al*., 1995; Sheldon and Dickinson 1986).

On the other hand, the *ELMOD_E* gene, which has been proposed as a possible molecular player involved in determining the shape of the apertures (Zhou *et al*., 2021), was not found among the DEGs. When *ELMOD_E* is overexpressed in *Arabidopsis* during the tetrad stage a change in the shape of the apertures from colpo to pore is produced (Zhou *et al*., 2021). However, the expression of colpate species is not higher than that of porate species in Papaveraceae. Furthermore *ELMOD_E* is also not more expressed than the other members of the *ELMOD_A* and *ELMOD_B*/*MACARON*(MCR) gene family, which control the number of apertures in *Arabidopsis*. Therefore, at least in Papaveraceae, *ELMOD_E* does not regulate the change from colpus to pore.

Regarding future prospects, with the functional analysis of the DEGs obtained from the two comparative analyses, we could confirm the involvement of some of them in the formation processes of the different apertural systems existing in the basal eudicot family Papaveraceae. Due to its phylogenetic position, the study of the genes involved in aperture formation in this family is fundamental to understanding this process in the angiosperms as a whole, providing different information to that obtained from the study of these genes in model plants. Recently a protocol for obtaining stable mutants in *E. californica* has been described (Lotz et al., 2021). This, together with the relative ease of sequencing and de novo transcriptome assembly, allows us to investigate aperture factors in Papaveraceae, for which fewer resources are available than in model species, as mutant libraries.

On the other hand, the transcriptomes generated, and especially the de novo transcriptome assembly of *Dactylicapnos torulosa*, *Fumaria bracteosa* and *Roemeria refracta*, whose genomes have not been sequenced, are an important new scientific resource which could be used for various developmental studies and in an evolutionary context.

Despite advances in the understanding of the molecular mechanisms of aperture formation, a lot of questions remain. It is necessary to know how INP1 is targeted and transported to the aperture domains and whether INP1 interacts directly with callose. Furthermore, it should be investigated whether other proteins, such as those

found among the DEGs, interact with the callose to form the apertures. On the other hand, it would be interesting to investigate whether the homologues of *INP2*, *D6PKL3*, *MCR* and *ELMOD_A* conserve their function in Papaveraceae and the rest of angiosperms. In addition, many other aperture factors remain to be discovered, including those responsible for determining the enormous variety of aperture patterns in angiosperms.

References

Badouin H, Velt A, Gindraud F, *et al*. 2020. The wild grape genome sequence provides insights into the transition from dioecy to hermaphroditism during grape domestication. *Genome Biology* **21**: 1-24.

Dobritsa AA, Coerper D. 2012. The novel plant protein INAPERTURATE POLLEN1 marks distinct cellular domains and controls formation of apertures in the Arabidopsis pollen exine. *The Plant Cell* **24**: 4452-4464.

Dobritsa AA, Kirkpatrick AB, Reeder SH, Li P, Owen HA. 2018. Pollen aperture factor INP1 acts late in aperture formation by excluding specific membrane domains from exine deposition. *Plant Physiology* **176**: 326-339.

Edlund AF, Zheng Q, Lowe N, *et al*. 2016. Pollen from Arabidopsis thaliana and other Brassicaceae are functionally omniaperturate. *American Journal of Botany* **103**: 1006-1019.

Feng B, Lu D, Ma X, Peng Y, Sun Y, Ning G, Ma H. 2012. Regulation of the Arabidopsis anther transcriptome by DYT1 for pollen development. *The plant journal* **72:** 612-624.

Furness CA, Rudall PJ. 2004. Pollen aperture evolution–a crucial factor for eudicot success?. *Trends in plant science* **9**: 154-158.

Lee BH, Weber ZT, Zourelidou M, *et al.* 2018. Arabidopsis protein kinase D6PKL3 is involved in the formation of distinct plasma membrane aperture domains on the pollen surface. *The Plant Cell* **30**: 2038-2056.

Li P, Ben-Menni Schuler S, Reeder SH, Wang R, Suárez Santiago VN, Dobritsa AA. 2018. INP1 involvement in pollen aperture formation is evolutionarily conserved and may require species-specific partners. *Journal of Experimental Botany* **69**: 983-996.

Lotz D, Imani J, Ehlers K, Becker A. 2021. Towards A Genetic Model Organism: An Efficient Method For Stable Genetic Transformation of Eschscholzia Californica (Ranunculales). Preprint.

Mazuecos-Aguilera I, Romero-García AT, Klodová B, *et al*. 2021. The role of INAPERTURATE POLLEN1 as a pollen aperture factor is conserved in the basal eudicot Eschscholzia californica (Papaveraceae). *Frontiers in Plant Science* **1405**.

Muñoz CA, Webster BD, Jernstedt JA, *et al*. 1995. Spatial congruence between exine pattern, microtubules and endomembranes in Vigna pollen. *Sexual Plant Reproduction* **8**: 147-151.

Ressayre A, Raquin C, Mignot A, Godelle B, Gouyon PH. 2002. Correlated variation in microtubule distribution, callose deposition during male post-meiotic cytokinesis, and pollen aperture number across Nicotiana species (Solanaceae). *American Journal of Botany* **89**: 393-400.

Sheldon JM, Dickinson HG. 1986. Pollen wall formation in Lilium: the effect of chaotropic agents, and the organisation of the microtubular cytoskeleton during pattern development. *Planta* **168**: 11-23.

Wang R, Dobritsa AA. 2018. Exine and aperture patterns on the pollen surface: their formation and roles in plant reproduction. *Annual Plant Reviews Online* 589-628.

Wege S, Scholz A, Gleissberg S, Becker A. 2007. Highly efficient virus-induced gene silencing (VIGS) in California poppy (Eschscholzia californica): an evaluation of VIGS as a strategy to obtain functional data from non-model plants. *Annals of Botany* **100**: 641-649.

Zhang W, Sun Y, Timofejeva L, *et al*. 2006. Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor. *Development* **133**: 3085-3095.

Zhang X, Zhao G, Tan Q. 2020. Rice pollen aperture formation is regulated by the interplay between OsINP1 and OsDAF1. *Nature Plants* **6:** 394-403.

Zhou Y, Dobritsa AA. 2019. Formation of aperture sites on the pollen surface as a model for development of distinct cellular domains. *Plant Science* **288**: 110222.

Zhou Y, Amom P, Reeder SH, Lee BH, Helton A, Dobritsa AA. 2021. Members of the ELMOD protein family specify formation of distinct aperture domains on the Arabidopsis pollen surface. *Elife* **10**: e71061.

Zourelidou M, Müller I, Willige BC, *et al.* 2009. The polarly localized D6 PROTEIN KINASE is required for efficient auxin transport in *Arabidopsis thaliana*. *Development* **136**: 627-636.

Conclusions | Conclusiones

Conclusions

1. The homolog gene of *INAPERTURATE POLLEN 1* (*INP1*) in *Eschscholzia californica*, *EcINP1*, is involved in the formation of pollen grain apertures and is essential for this process. Due to the basal phylogenetic position of *Eschscholzia californica* within eudicots, we can confirm that its role is conserved in the transition between monocots and core eudicots, in which its function has already been confirmed.

2. Peak expression of *EcINP1* occurs in anthers at the tetrad stage of pollen development, when the first events of pollen wall formation occur; similar to the model species *Arabidopsis thaliana* and *Oryza sativa*. The concentration of EcINP1 protein remains high until the free microspore stage, as in rice and unlike in *Arabidopsis thaliana*, in which it is rapidly degraded after the tetrad stage. This delay in EcINP1 degradation suggests a possible functional diversification of the protein, as has been demonstrated in rice.

3. The apertures are dispensable for pollen germination of *Eschscholzia californica*, like *Arabidopsis thaliana* and other Brassicaceae and unlike grasses. Several authors consider that the independence of pollen tube exit from the presence of apertures may be due to differences in exine morphology and/or pollen and stigma physiology.

4. The tetrad-stage pollen transcriptome of silenced EcINP1 plants changes significantly from that of wild-type plants, allowing us to identify possible candidate genes with which EcINP1 may interact during the aperture formation process. Among the 971 differentially expressed genes (DEGs), the homologues in *E. californica* of NUCLEOSOME ASSEMBLY PROTEIN 1 (NAP1), D6 PROTEIN KINASE LIKE 3 (D6PKL3) and PROTEIN KINASE ASSOCIATED WITH BRX (PAX) stand out for their involvement in pollen development, aperture formation and cell signalling. These genes showed similar expression patterns to EcINP1, suggesting their involvement in the formation of the apertures and/or possible interaction of their products with the EcINP1 protein.

5. Comparative transcriptome analysis of two pairs of Papaveraceae species belonging to two different subfamilies, each pair comprising a colpate and a porate species, allowed the identification of genes potentially involved in the determination of aperture shape. The 531 DEGs include the INP1 gene and other genes implicated in processes possibly involved in aperture formation, such as the synthesis or degradation of callose or the organisation of cytoskeletal elements.

6. The Papaveraceae homologue of the ELMOD_E gene, the only one described so far that could be involved in aperture shape determinism in *Arabidopsis*, is not differentially expressed between colpate and porate species during the tetrad stage. At this stage, ELMOD_E-like is also not differentially expressed compared to the other two members of the A/B lineage gene family, which in *Arabidopsis* are involved in determining the number and position of apertures, ELMOD_A and ELMOD_B/MCR. All this evidence does not support the hypothesis described in *Arabidopsis* suggesting that the interaction between ELMOD_E with A/B lineage proteins regulates colpo-to-pore shift in Papaveraceae, and thus the conserved function of ELMOD_E.

Conclusiones

1. El gen homólogo de *INAPERTURATE POLLEN 1* (*INP1*) en *Eschscholzia californica, EcINP1*, está implicado en la formación de las aperturas del grano polínico, siendo imprescindible para que este proceso ocurra. Debido a la posición filogenética de *Eschscholzia californica* en la base de las Eudicotiledóneas podemos confirmar que su papel está conservado en la transición entre monocotiledóneas y eudicotiledóneas núcleo, en las que su función ya había sido confirmada.

2. La expresión máxima de *EcINP1* ocurre en anteras en el estadio de desarrollo del polen de tétradas, momento en el que ocurren los primeros eventos de la formación de la pared del polen; al igual que ocurre en las especies modelo *Arabidopsis thaliana* y *Oryza sativa*. La concentración de la proteína EcINP1 permanece elevada hasta el estadio de microspora libre, al igual que en arroz y a diferencia de *Arabidopsis thaliana*, en la que se degrada rápidamente tras el estadio de tétrada. Este retraso en la degradación de EcINP1 sugiere una posible diversificación funcional de la proteína, como se ha comprobado en arroz.

3. Las aperturas son prescindibles para la germinación del polen de *Eschscholzia californica,* al igual que *Arabidopsis thaliana* y otras Brassicaceae y a diferencia de gramíneas. Diversos autores consideran que la independencia de la salida de los tubos polínicos respecto a la presencia de aperturas puede deberse a la existencia de diferencias en la morfología de la exina y/o en la fisiología del polen y del estigma.

4. El transcriptoma polínico en estadío de tétradas de las plantas con el gen *EcINP1* silenciado cambia significativamente respecto al de las plantas salvajes, permitiéndonos identificar posibles genes candidatos con los que EcINP1 puede interaccionar durante el proceso de formación de las aperturas. Entre los 971 genes diferencialmente expresados (DEGs) destacan, por su implicación en el desarrollo del polen, formación de las aperturas y señalización celular, los homólogos en *E. californica* de *NUCLEOSOME ASSEMBLY PROTEIN 1* (*NAP1*), *D6 PROTEIN KINASE LIKE 3* (*D6PKL3*) y *PROTEIN KINASE ASSOCIATED WITH BRX* (*PAX*). Estos genes mostraron patrones de expresión coincidentes con los de *EcINP1*, que apoyan su implicación en la formación de las aperturas y/o posible interacción de sus productos con la proteína EcINP1.

5. El análisis comparativo de transcriptomas de dos pares de especies de Papaveraceae pertenecientes a dos subfamilias diferentes y donde cada par lo compone una especie colpada y otra porada, permitió identificar genes potencialmente implicados en la determinación de la forma de las aperturas. Entre los 531 DEGs aparece el gen *INP1* y otros genes implicados en procesos que posiblemente intervienen en la formación de las aperturas, como los relacionados con la síntesis o degradación de la calosa o con la organización de elementos del citoesqueleto.

6. El homólogo en Papaveraceae del gen *ELMOD_E*, el único descrito hasta el momento que podría estar implicado en el determinismo de la forma de las aperturas en *Arabidopsis*, no se expresa diferencialmente entre especies colpadas y poradas durante el estadío de tétradas. En este estadío, *ELMOD_E*-like tampoco se expresa de forma diferencial respecto a los otros dos miembros de la familia génica del linaje A/B y que en *Arabidopsis* están implicados en el determinismo del número y posición de las aperturas, *ELMOD_A* y *ELMOD_B*/*MCR*. Todo estas evidencias no apoyan la hipótesis descrita en *Arabidopsis* de que la interacción entre ELMOD_E con las proteínas del linaje A/B regula el cambio de colpo a poro en Papaveraceae, y por tanto la función conservada de ELMOD_E.