Vers une meilleure compréhension de l'identité et du développement du pétale de Petunia par séquençage ARN en cellule unique

Quentin Cavallini-Speisser

Université Claude Bernard Lyon 1

THESE de DOCTORAT DE
L'UNIVERSITE CLAUDE BERNARD LYON 1

Ecole Doctorale N°340
Biologie Moléculaire, Intégrative et Cellulaire

Discipline : Biologie, médecine et santé

Soutenue publiquement le 14/12/2023, par :
Quentin CAVALLINI-SPEISSER

Using single-cell RNA-Sequencing towards a better understanding of *Petunia* petal identity and development

-

Vers une meilleure compréhension de l'identité et du développement du pétale de *Petunia* par séquençage ARN en cellule unique

---

Devant le jury composé de :

BENSMIHEN, Sandra	Directrice de Recherche, CNRS, LIPME	Rapporteure
QUATTROCCHIO, Francesca	Professeure Associée, Université d'Amsterdam, Department of Plant Development and Genetics	Rapporteure
ARNAUD, Nicolas	Chargé de Recherche, INRAE, IJPB	Examinateur
COMTE, Gilles	Professeur des Universités, Université Lyon 1, LEM	Président
GANDRILLON, Olivier	Directeur de Recherche, CNRS, LBMC	Examinateur
MONNIAUX, Marie	Chargée de Recherche, CNRS, RDP	Directrice de thèse

ACKNOWLEDGMENTS / REMERCIEMENTS

First and foremost, I want to express my gratitude towards Marie for giving me the opportunity to work on this thrilling project under her direction. Thank you for the trust you placed in me, thank you for your guidance and thank you for your everyday positivism and cheerfulness that were very welcomed some days. I also want to take this opportunity to thank the other members of my jury, Nicolas Arnaud, Sandra Bensmihen, Gilles Comte, Olivier Gandrillon and Francesca Quattrocchio for accepting to evaluate my work. I am looking forward to meet you all and discuss it further. My warm thanks also go to the EvoDevo group at the RDP Lab. The relaxed and calm atmosphere enhanced by stimulating discussions when need be I found among you helped make these last three years both very enjoyable and enriching. Grateful thanks also go to our collaborators at the CRCL, especially Cyril and to the PSMN (Pôle Scientifique de Modélisation Numérique) of the ENS de Lyon for the computing resources. Thank you for your help and above all your calm and reassuring words when I was all stressed out during scRNA-Seq experiment days. Last but not least, I need to salute each and every member at the RDP Lab for contributing to the incredibly rich experience both human- and professionally-wise that is working at this place. Thank you so much. The RDP spirit is actually a thing! Keep it up!

Un grand merci à ma famille pour avoir toujours été là à mes côté. Lucille pour ces dernières années en collocation agrémentées de fous rires, de semaines sans presque se croiser et de discussions absurdes à essayer (en vain) de décortiquer physique quantique, géopolitique, philosophie et j'en passe. Isabelle et Bruno pour leurs visites régulières depuis leur Beauce sans bosse, maman ayant toujours de quoi renflouer les placards et papa toujours l'habitude de voir un problème de mathématiques partout. Les deux, toujours avec beaucoup de tendresse à partager. Nathanaël et Romain pour les brunchs du dimanche matin passés à discuter de tout et de rien. Romain encore pour les restaurants du vendredi midi. Mimi et Armin pour les séjours dans les Vosges et leurs petites attentions régulières. Gilou pour notre dernière ballade en Chartreuse dont le bol d'air était plus que bienvenu. Kiki, Babeth, Etienne, Camille, Victoria, Hugues, Valérie, Manon et Enzo pour les grandes tablées animées quand nous sommes, bien trop rarement, tous réunis. Je vous aime toutes et tous très fort.

Je remercie aussi Cédric, Étienne, Florent, Marine, Norman, Madjid, Marion, Remi, Marc, Anthony, les deux Mat', Bastien, PN, l'autre Florent, Morgan, Guillaume, et j'en oublie. Qui pour les soirées en ville. Qui pour un barbecue ou une raclette dans les Monts d'Or. Qui pour les soirées sushis/jeux-vidéos. Qui pour les nuits passées sur Discord et les week-ends à visiter la France. Pour tout ça et votre amitié sans faille, je vous embrasse toutes et tous.

RÉSUMÉ

Au sein du vivant, un haut niveau de contrôle et de synchronisation des mécanismes de division, croissance et différenciation cellulaire est nécessaire au bon développement des organes et à la mise en place de leurs fonctions. Identifier ces niveaux de contrôle et de synchronisation est l'une des questions centrales à la biologie du développement. Chez les angiospermes, les organes émergent à la périphérie d'organes génératifs, les méristèmes. Ces méristèmes abritent un groupe de cellules souches renouvelées constamment tant que le méristème est actif, maintenant ainsi leur fonction générative, et ces méristèmes sont organisés en couches cellulaires distinctes. Les cellules végétales étant incapables de mouvement au sein des tissus du fait de leur paroi rigide, cette organisation en couches cellulaires définies est propagée et maintenu dans les organes.

Chez *Arabidopsis thaliana*, modèle historique en biologie végétale, l'identité florale est mise en place, au sein du méristème d'inflorescence, par l'expression des gènes *LEAFY* (*LFY*) puis *APETALA1* (*AP1*), l'expression de ce dernier étant activée par le premier (Weigel et al., 1992; Bowman et al., 1993; Parcy et al., 1998). Une fois cette identité acquise, toujours sous le contrôle de *LFY*, l'expression d'autres gènes, encodant eux aussi des facteurs de transcription, va permettre le bon développement des différents organes floraux dans les bons verticilles. Ces gènes encodent pour la plupart des facteurs de transcription de la famille des MADS (d'après les quatre premiers membres identifiés, MCM1 chez *Saccharomyces cerevisiae*, AG chez *A. thaliana*, DEF chez *Antirrhinum majus* et SRF chez *Homo sapiens*) et ont été regroupés en trois familles, A, B et C, le chevauchement de leurs schémas d'expression régissant le type d'organe se développant dans un verticille donné (Schwarz-Sommer et al., 1990; Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994). La fonction A permet le développement des sépales dans le premier verticille ; l'expression simultanée de gènes A et B entraîne la formation des pétales dans le second ; B et C permettent le développement des étamines dans le troisième ; et la formation du pistil dans le quatrième verticille est sous le contrôle de la fonction C.

Chez *Petunia x hybrida*, le développement du pétale est sous le contrôle de trois gènes de classe B, *PhDEFICIENS* (*PhDEF*), *PhGLOBOSA1* (*PhGLO1*) et *PhGLOBOSA2* (*PhGLO2*), la formation d'hétérodimères PhDEF/PhGLO1 ou PhDEF/PhGLO2 étant absolument nécessaires au bon développement du pétale (Vandenbussche et al., 2004). *PhGLO1* et *PhGLO2* étant au moins partiellement fonctionnellement redondants, les fleurs de simples mutants *phglo1* ou *phglo2* possèdent bien des pétales dans le second verticille. Ce n'est pas le cas du simple mutant *phdef* chez qui on observe une conversion homéotique des pétales en sépales du fait de l'absence de protéine

PhDEF pouvant participer aux hétérodimères cités ci-dessus, dont au moins une version est nécessaire à la fonction B.

La fleur de *P. hybrida* sauvage présente normalement 5 pétales soudés formant un tube dans la partie inférieur de la fleur et s'ouvrant en de larges lobes pigmentés dans sa partie supérieure. En travaillant avec le simple mutant de transposition *phdef-151*, présentant une conversion homéotique des pétales en sépales, l'équipe où j'ai effectué ma thèse a observé deux types de phénotypes floraux issus d'événements de réversion, le transposon *dTph1* s'excisant du premier exon de *PhDEF* et restaurant l'expression d'une protéine fonctionnelle. Certaines fleurs, nommées *star*, présentent un tube comparable à celui de la fleur sauvage mais des lobes réduits et non pigmentés. D'autres au contraire, appelées *wico*, montrent une tube très court mais des lobes bien développés et pigmentés. Par hybridation *in-situ*, l'équipe a déterminé que ces deux phénotypes très différents sont la conséquence de la restauration de l'expression de *PhDEF* de manière couche cellulaire spécifique, seulement dans les couches internes chez *star*, seulement dans l'épiderme chez *wico* (Chopy et al., 2023). L'observation de phénotypes distincts dans ces deux situations suggère l'existence de réseaux de régulation couche cellulaire spécifiques impliquant *PhDEF* nécessaires au bon développement du pétale chez *P. hybrida*.

Afin de préciser ces réseaux de régulation couche cellulaire spécifiques, j'ai conçu puis appliqué un protocole de séquençage ARN en cellule unique (*scRNA-Seq*), depuis la production de protoplastes (des cellules végétale sans leur paroi) jusqu'à l'analyse bio-informatique des données de séquençage par un *pipeline* automatisé, sur des pétales de *P. hybrida*. Afin d'étudier plus précisément le développement du pétale, il était prévu d'appliquer ce protocole sur différents stades de développement de fleurs sauvages, *star*, *wico* et *phdef-151*. Cependant des difficultés techniques m'ont contraint à ne travailler que sur des pétales matures de fleurs sauvages, *star* et *wico*, ayant été incapable d'obtenir suffisamment de protoplastes pour les stades plus précoces ainsi que *phdef-151*.

Le *scRNA-Seq* est une technique novatrice développée depuis 2009 en biologie animale (Tang et al., 2009) et appliquée en biologie végétale depuis 2013 (Brennecke et al., 2013). Elle permet de récolter des données transcriptomiques au niveau de la cellule unique plutôt qu'au niveau de l'échantillon complet comme l'accomplit le séquençage ARN classique (*bulk RNA-Seq*), permettant d'exposer l'hétérogénéité des tissus biologiques et d'étudier spécifiquement certains types cellulaires parmi la diversité présente dans l'échantillon.

L'application du *scRNA-Seq* lors de ma thèse a permis la mise en évidence de différences transcriptomiques clefs entre les différents tissus composant le pétale de *P. hybrida*. De plus, l'utilisation de ces données couplées à d'autre techniques d'analyse comme le séquençage après

immunoprécipitation de chromatine (*ChIP-Seq*) et le *bulk RNA-Seq* sur nos génotypes d'intérêt a permis l'identification de potentiels partenaires et gènes cibles couche cellulaire spécifiques de *PhDEF*.

Ce travail ouvre la porte à une analyse fonctionnelle de ces gènes couche cellulaire spécifiques, et a fourni des données transcriptomiques au niveau de la cellule du pétale de *P. hybrida* sauvage et mutant, ainsi qu' un *pipeline* d'analyse *scRNA-Seq* documenté comme ressource pour la communauté de la biologie du développement végétal.

ABSTRACT

One of the central question of developmental biology is to understand the synchronization cues at work behind cell division, growth and differentiation, allowing robustness in organ shape, size and function. In flowering plants, organs emerge at the periphery of generative structures called meristems. Meristems shelter a pool of stem cells allowing their generative function and are organized in clonally-distinct cell layers. Since plant cells are unable to move within tissues because of their rigid cell-wall, this layered organization is carried over and maintained in organs. In *Petunia* flowers, petal identity and development is under the control of several B-class MADS-box genes, of which *PhDEF*. Total loss of *PhDEF* expression in developing flowers leads to the development of sepals in place of petals. But its layer-specific loss of function leads to distinct petal morphology defects whether the epidermis or the internal cell layers are knocked-out, hence suggesting cell-layer-specific *PhDEF*-mediated regulation networks driving correct petal development in *Petunia*.

During my PhD, I showed how single-cell RNA-Sequencing (scRNA-Seq), a novel technique allowing transcriptomic analysis at the single-cell level instead of the sample-level in bulk RNA-Seq, unveiled key transcriptomic heterogeneity between the different cell tissues of the petal of *Petunia x hybrida*. Moreover, coupling scRNA-Seq, bulk RNA-Seq and chromatin immunoprecipitation followed by sequencing (ChIP-Seq), I identified putative cell-layer-specific targets and partners of *PhDEF*.

This work opens the door to a functional study on the genes potentially involved in the development of the petal of *P. hybrida* in a cell-layer-specific manner alongside *PhDEF*, as well as providing single-cell level transcriptomic data of wild-type and mutant *P. hybrida* petal and a documented scRNA-Seq analysis pipeline that could be useful for the community of plant developmental biology.

TABLE OF CONTENTS

ABBREVIATIONS

LFY – LEAFY

AP1/3 – APETALA 1/3

MCM1 – MINICHROMOSOME MAINTENANCE 1

AG – AGAMOUS

DEF – DEFICIENS

SRF – SERUM RESPONSE FACTOR

P. hybrida – Petunia x hybrida

P. axillaris – Petunia axillaris

P. inflata – Petunia inflata

PhDEF – *P. hybrida* DEFICIENS

PhGLO1 – *P. hybrida* GLOBOSA 1

PhGLO2 – *P. hybrida* GLOBOSA 2

dTph1 – defective Transposable element *P. hybrida* 1

RNA – Ribonucleic Acid

ARN – Acide Ribonucleique

scRNA-Seq – single-cell RNA-Sequencing

ChIP-Seq – Chromatin Immunoprecipitation Sequencing

AN1/2/4/11 – ANTHOCYANIN 1/2/4/11

Ac/Ds – Activator Dissociator system

bp – base pair

DNA – Desoxyribonucleic Acid

ADN – Acide Desoxyribonucléique

Tst1 – Transposon *Solanum tuberosum* 1

Ac – Activator

rDt – retro Dotted Transposon

Act1 – Activator 1

Hi-C – High throughput Chromosome Conformation Capture

A. majus – Antirrhinum majus

A. thaliana – Arabidopsis thaliana

FBP11 – FLORAL BINDING PROTEIN 11

TF – Transcription Factor

N-ter – Amine-terminal

C-ter – Carboxyl-terminal

CarG box – CC-Arich-GG box (5′-CC(A/T)6GG-3′)

I domain – Intervening domain

K domain – Keratin-like coiled-coil domain

C domain – C-ter domain

SEP3 – SEPALLATA 3

STK – SEEDSTICK

GLO – GLOBOSA

ARF – Auxin Response Factor

FUL – FRUITFULL

PI – PISTILLATA

PhTM6 – P. hybrida TOMATO MADS 6

Y2H – yeast-two-hybrids assay

FM – Floral Meristem

UFO – UNUSUAL FLORAL ORGANS

bHLH – basic Helix Loop Helix

PDF2 – PROTODERMAL FACTOR 2

FDH – FIDDLEHEAD

AFI – ANTHIRRHINUM FIDDLEHEAD

ATML1 – MERISTEM LAYER 1

MAPK – Mitogen-Activated Protein Kinase

SEM – Scanning Electron Microscopy

NGS – Next Generation Sequencing

mRNA – messenger RNA

SAM – Shoot Apical Meristem

cDNA – coding DNA

wt – wild-type

sciRNA-Seq – single-cell combinatorial indexing RNA-Seq

snRNA-Seq – single-nuclei RNA-Seq

FACS – Fluorescence Activated Cell Sorting

INTACT – Isolation of Nuclei TAgged Cell Types

LCM – Laser Capture Microdissection

nt – nucleotide

GEM beads – Gel beads in EMulsion

UMI – Unique Molecular Identifier

RT – Reverse Transcription

PCR – Polymerase Chain Reaction

DGE – Differential Gene Expression

PCA – Principal Component Analysis

WGCNA

– Weighted Gene Correlation Network Analysis

PH4 – PH REGULATOR 4

DPL – DEEP PURPLE

GO – Gene Ontology

ROS – Reactive Oxygen Species

ABA – Abscisic Acid

ATP – Adenosine Triphosphate

tSNE

– t-Distributed Stochastic Neighbor Embedding

PC – Principal Component

Log2FC – Log 2 Fold Change

RNase – Ribonuclease

UMAP

– Uniform Manifold Approximation Projection

GL2 – GLABRA 2

SHN1 – SHINE 1

CYCA3 – CYCLIN A3

TPS1 – TERPENE SYNTHASE 1

RuBisCo

– Ribulose-1,5-bisphosphate carboxylase oxygenase

YAB1/2 – YABBY 1/2

IAA17 – AUXIN/INDOLE-3-ACETIC ACID 17

TSS – Transcription Start Site

TTS – Transcription Termination Site

F3H – Flavanone 3-Hydroxylase

PALa – phenylalanine ammonia-lyase a

HD-Zip – Homeodomain Leucine Zipper

TCP

– TEOSYNTE BRANCHED CYCLOIDE A PROLIFERATING CELL FACTOR

HDG – HOMEODOMAIN-GLABROUS

SPL

– SQUAMOSA PROMOTER BINDING PROTEIN-LIKE

GTF – Gene Transfer Format

WB – Western Blot

AB#1/#2 – Anti-PhDEF antibody #1/#2

HRP – Horse Radish Peroxidase

kDA – kilo Dalton

MW – Molecular Weight

TBS – Tris-HCl-Buffered Saline

PBS – Phosphate-buffered Saline

AF-568 – AlexFluor 568

ASE – Antibody Signal Enhancer

MTSB – Microtubule Stabilizing Buffer

mL – Milliliter

mm^2 – Square millimeter

H_2O – Water

$mOsm.kg^{-1}$ H_2O – Milliosmole per kilogram of water

M – Molar, $mol.L^{-1}$

BAM format – Binary Alignment Map format

SLURM

– Simple Linux Utility for Resource Management

HDF5 – Hierarchical Data Format 5

mM – Millimolar, $mmol.L^{-1}$

µL – Microliter

DAPI – 4',6-diamidino-2-phenylindole

FAA – Formalin-Acetic Acid-Alcohol

EDTA – Ethylenediaminetetraacetic Acid

CSV format – Comma Separated Values

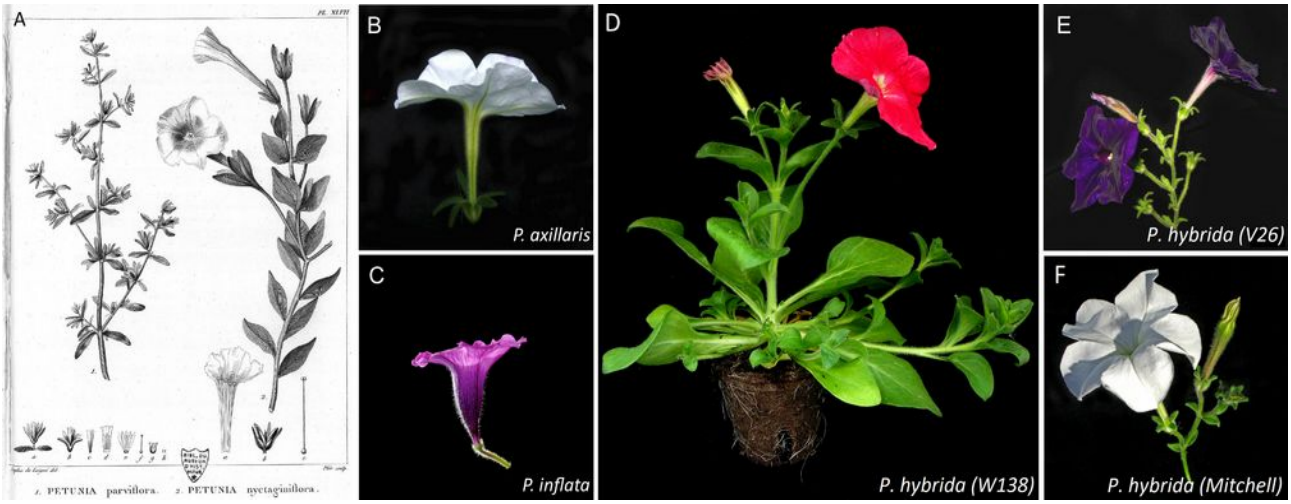


Figure 1.1: Anatomical drawing and photographs of *Petunia* plants, inflorescence and flowers

(A) Drawing of the 2 specimens used to establish the genus *Petunia* (Jussieu, 1803). **(B-F)** Photographs of **(B)** *P. axillaris*, **(C)** *P. inflata* flowers, **(D)** *P. hybrida* (W138 line) young whole plant, **(E)** *P. hybrida* (V26) and **(F)** *P. hybrida* (Mitchell) inflorescences. Adapted from (Vandenbussche et al., 2016).

INTRODUCTION

I – General introduction

One of the most potent question in developmental biology is to understand how cells and tissues synchronize their growth, division and differentiation, resulting in an organ with a specific and reproducible shape. In flowering plants, most organs result from generative structures, the meristems, that maintain a stem-cell pool while organizing differentiation patterns leading to the formation of organs at their periphery. A meristem is organized in distinct cell layers and this layered organization of the meristem is maintained throughout development by clonal divisions. Plant cells being unable to move within tissues, these clonally-distinct cell layers are carried over in organs. In mature organs, cell layers often have very different functions; in particular the epidermis is the site of interaction with the environment while the mesophyll tissue is often where photosynthesis takes place. Therefore, how these defined cell layers acquire their distinct features while synchronizing the development of organs is a key question in plant developmental biology, most of these mechanisms remaining unknown yet.

My work over the last three years within the EvoDevo group of the RDP laboratory aimed at clarifying such mechanisms by using the petal of *Petunia* as a model, using a set of cell layer-specific mutants in petal identity and various techniques, among which the novel approach that is single-cell RNA-Sequencing (scRNA-Seq).

II – *Petunia x hybrida*

II.1 – Brief origin, history, genetics and morphology

The *Petunia* genus was first established in 1803 by the botanist Antoine-Laurent de Jussieu based on 2 specimens collected in Uruguay by the naturalist Philibert Commerson (Jussieu, 1803) (Fig. 1.1, A). Nowadays classified as part of the Asterid clade in the order of Solanaceae, 14 wild species of *Petunia* are recognized as of 2009, all endemic of South America, in Brazil (13 known species), Argentina (5), Uruguay, Paraguay and Bolivia (2) (Stehmann et al., 2009). The garden *Petunia x hybrida* (*P. hybrida*) is an artificial hybrid of the white hawk moth-pollinated *Petunia axillaris* and the purple bee-pollinated *Petunia inflata* (Bombarely et al., 2016; Vandenbussche et al., 2016). It was most probably independently obtained multiple times from crosses of different accessions by European horticulturalists in the early 19th century (Sink, 1984; Gerats and Strommer, 2009) for aesthetics reasons, its flowers showing both the long tube and large corolla of *P. axillaris* flowers (Fig. 1.1, B) and for most varieties, the petal pigmentation of *P. inflata* (Fig. 1.1, C). (Fig. 1.1 D-F) show a few varieties of *P. hybrida*.

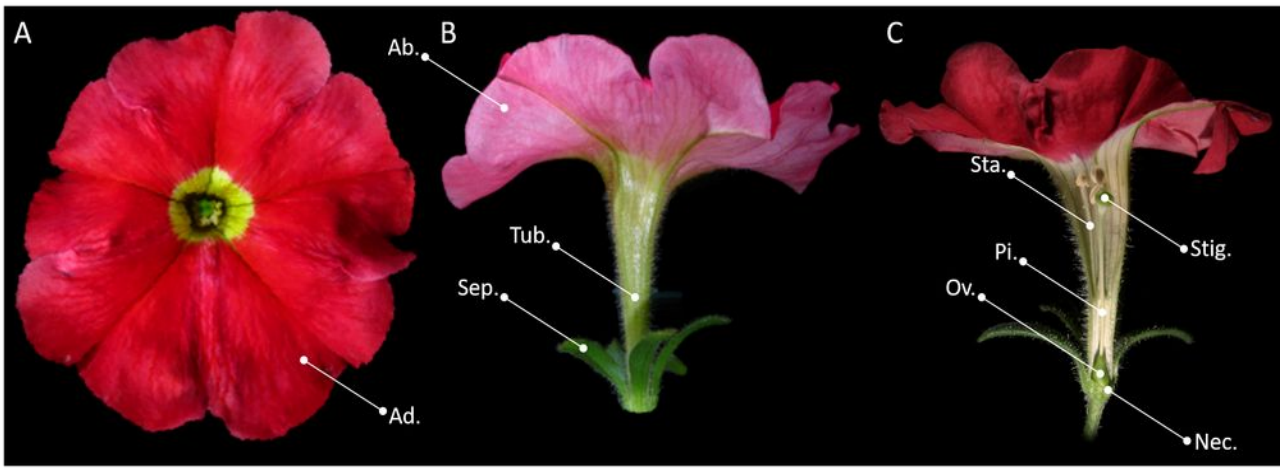


Figure 1.2: Photographs of *P. hybrida* (W138) flower

Top view **(A)**, side view **(B)**, side view after transversal dissection **(C)**. Ad.: adaxial side of the limb, Ab.: abaxial side of the limb, Sep.: sepal, Tub.: petal tube, Ov.: ovary, Sta.: stamen, Pi.: pistil, Stig.: pistil stigma, Nec.: nectary.

Figure 1.3: Photographs of *P. hybrida* (W138) flowers showing reverted red sectors

(A) Side view of an inflorescence, **(B)** top view of a flower.

Petunia species can be crossed together with little hassle and yield normal diploid progeny since species barriers are mainly pre-zygotic. Therefore *P. hybrida* has the same chromosome number (2n = 14) as its parental species (Vandenbussche et al., 2016).

P. hybrida grown in laboratory conditions (16h day at 22°C; 8h night at 18°C, 60% humidity) has a lifecyle of 4 months and can be maintained as long as two years with appropriate care if needed. Its cymose inflorescence (Kusters et al., 2015) regularly produces flowers and their associated bracts during the whole lifespan of the individual. The flowers are composed from outer to inner whorls of five sepals at the base of the flower (Fig. 1.2, Sep.), five fused petals forming the corolla composed of a tube in its lower half (Fig. 1.2, Tub.) and a large limb in its upper half (Fig. 1.2, Ad. And Ab.), five stamens of witch the lower half is fused with the tube and the upper half free (Fig. 1.2, Sta.) and two fused carpels into a single central pistil (Fig. 1.2, Pi.). The ovary is situated internally at the very base of the pistil (Fig. 1.2, Ov.), surrounded by nectaries (Fig. 1.2, Nec.). The sepals are highly chloroplastic and unpigmented otherwise, the petal tube slightly chloroplastic and unpigmented otherwise, the corolla non-chloroplastic and pigmented mainly by anthocyanins, strongly on the adaxial side of the petal, giving it a bright red color (Fig. 1.2, Ad.), and more faintly on the abaxial side of the petal giving it a pink hue (Fig. 1.2, Ab.). Stamens are unpigmented but carry bright yellow mature pollen grains while the pistil is unpigmented and mainly non-chloroplastic except for its stigma (Fig. 1.2, Stig.).

II.2 – The W138 line and the *dTph1* transposable element

P. hybrida as been used for decades in plant biology and particularly to study petal pigmentation. While working with in-bred lines of the red-flowered variety “Roter Vogel” (R27), (Bianchi et al., 1978) encountered white flowers progeny (W17 and W28) which when crossed with R27 gave rise to white flowers with reverted red sectors, W138 is the result of one of those crosses (Fig. 1.3). They showed the amount of red sector appearance is tied to temperature and more importantly several orders of magnitude too frequent to be caused by another mutation counteracting the effect of the one causing the flowers to be white. In this regard, they proposed that the regulation of the *ANTHOCYANIN1* (*AN1*) locus, at the time known to be situated on chromosome VI and to be involved in petal pigmentation (Smith et al., 1975), depends on a mutator-expressor system similar to the activator-dissociator (Ac/Ds) system described by (McClintock, 1950, 1965) in maize. The mutator being responsible for the activation of the *AN1* locus and the expressor responsible for its level of expression.

The possibility of the system involving transposable elements as described by (McClintock, 1950, 1965) was briefly discussed in (Bianchi et al., 1978) but the authors considered “the postulation of transpositions of the controlling element [as] unnecessary” in their system, although not proven wrong either. The idea that transposable elements could play a role in gene regulation was quite novel and even still debated at the time (Biémont, 2010) and the authors proposed a mechanism involving repeated sequences inside the mutator element undergoing deletions inhibiting the *AN1* locus, giving rise to white flowers, followed by DNA reparation mitigating these deletions over cell divisions, reverting anthocyanins biosynthesis back to normal in red sectors instead.

It was later shown using the line W138, that the observations mentioned above are the consequence of a non-autonomous bi-component transposable system involving an autonomous element situated on chromosome I and a non-autonomous element inside the *AN1* locus, nowadays known as a key regulator of anthocyanins biosynthesis (Spelt et al., 2000), situated on chromosome VI (Wijsman, 1986).

Study of the system continued and (Gerats et al., 1990) precised its molecular cues further showing that the non-autonomous element of the bi-component system is a small 284 bp. transposable element they named *dTph1* for *defective Transposable element petunia hybrida 1*, showing some homology with known transposable elements in other species like *Tst1* in potato (Köster-Töpfer et al., 1990) or *Ac* and *rDt* in maize (McClintock, 1987). To this day the second actor of the system, *Act1*, has not yet been characterized at the molecular level further than its location on chromosome I (Peterson, 2013). In summary, the line W138 has one copy of *dTph1* inserted into the *AN1* coding sequence which disrupts AN1 protein production since *dTph1* has STOP codons in all possible reading frames and in both orientations, resulting in overall white flowers. Excision of *dTph1*, either leaving a small in-frame footprint or no footprint at all, at different times of development results in restoration of AN1 function and pigmented sectors in the petal. The authors also determined that the line W138 has over 50 copies of *dTph1* in its genome although more recent studies use W138 individuals with over 200 copies (Peterson, 2013), making this peculiar *P. hybrida* line a fortunate tool, and the standard line in *Petunia* nowadays, for insertion mutants identification and study and therefore the line we use in the group.

II.3 – Available genome assemblies and annotations

As of August 2023 and to my knowledge, there is no published structural genome annotation available for *P. hybrida*. However, both parental genome assemblies and corresponding predicted structural annotations were published in 2016. It was determined that a majority of the protein-coding genes of *P. hybrida* probably originate from *P. axillaris*, few from *P. inflata*, respectively ~15,000 and ~700, ~1,500 genes seems to have a mixed parentage and ~2,000 potentially derive from an unknown ancestor among the ~20,000 genes studied (Bombarely et al., 2016). Therefore, and awaiting better alternative, most studies use *P. axillaris* as reference genome when working with *P. hybrida*.

To my knowledge there are three *P. axillaris* genome assemblies publicly available. The one published by (Bombarely et al., 2016) alongside two predicted structural annotations (a version 1 published in 2016 and an unpublished version 4 uploaded in 2020, both available in the Sol Genomics Network database (Fernandez-Pozo et al., 2015)). This draft genome assembly was later further scaffolded after Hi-C (adapted Chromosome Conformation Capture (3C) (Dekker et al., 2002) assay followed by High Throughput Sequencing) (Lieberman-Aiden et al., 2009) by DNA-Zoo (Dudchenko et al., 2017, 2018) into a newly available genome assembly and predicted structural annotation available since 2018. The most recent publicly available *P. axillaris* assembly and predicted structural annotation were uploaded to the Genome Evolution database (Lyons and Freeling, 2008) by Prof. Dr. Cris Kuhlemeier's group in 2022 and is the only genome assembly at chromosome level publicly available.

In the group we use a custom genome-annotation couple composed of the originally published predicted structural annotation (version 1) transferred onto the DNA-Zoo HiC genome assembly as described further in (Chopy et al., 2023).

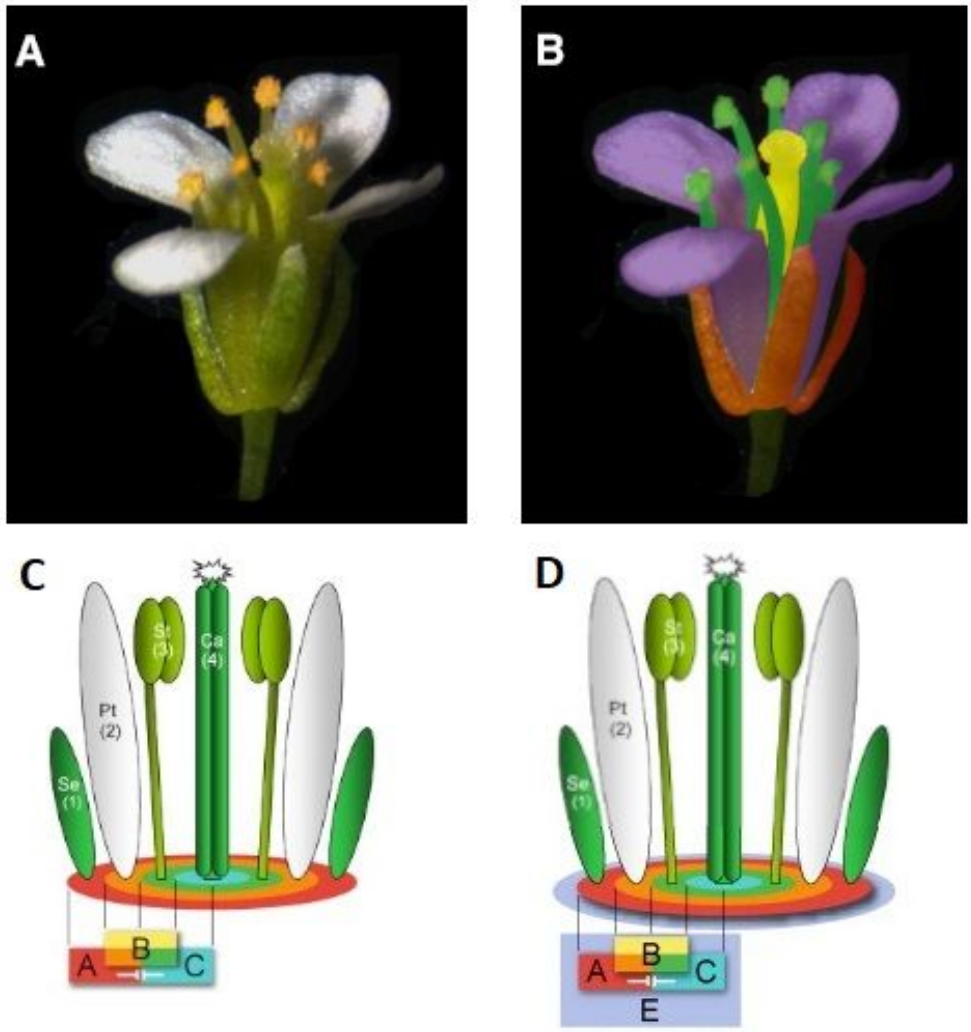


Figure 1.4: The historical ABC model

(A) Side view of an *Arabidopsis thaliana* flower. **(B)** Side view of an *Arabidopsis thaliana* flower with false colors to put sepals (red), petals (purple), stamens (green) and carpels (yellow) in evidence. **(C)** Diagram of the expression patterns of the A-, B- and C-function genes in the flower of *A. thaliana* as historically proposed in the ABC model. **(D)** Diagram of the expression patterns of the A-, B-, C- and E-function genes in the flower of *A. thaliana*. Adapted from (Causier et al., 2010; Irish, 2017).

III – Flower development and *P. hybrida* petal development

III.1 – Historical ABC(D)E flower identity and development model

In the early 90s, the study of several homeotic mutants, displaying well-formed organs at the wrong location, in *Antirrhinum majus* (*A. majus*) and *Arabidopsis thaliana* (*A. thaliana*) flowers led to the proposal of the "ABC flower development model" (Schwarz-Sommer et al., 1990; Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994). Typical flowers of most angiosperms have four main organs arranged concentrically in four whorls. From exterior to interior, sepals (Fig. 1.4, A-B, red) and petals (Fig. 1.4, A-B, purple) surround a bisexual axis composed of male reproductive stamens (Fig. 1.4, A-B, green) and female reproductive carpels (Fig. 1.4, A-B, yellow). The aforementioned ABC model described three classes of homeotic genes, genes supervising the development of organs, of functions A, B and C. Each gene is supposed to be expressed in two consecutive whorls, A-class genes in sepals and petals, B-class genes in petals and stamens, C-class genes in stamens and carpels. The overlapping expression patterns define which organ is formed at a given location, A-function alone in the first whorl controlling sepals formation, combined expression of A- and B-class genes in the second whorl giving rise to petals, concomitant B- and C-function in the third whorl leading to the emergence of stamens and in the fourth whorl the C-function enabling carpels development (Fig. 1.4, C). A- and C-function were predicted to mutually repressed themselves, meaning a loss-of-function mutant for either one of the two functions will lead to the invasion of its expression zone by the antagonist function, hence the observation of homeotic conversions of organs into others in A-, B- and C-function mutants.

This historical ABC model was broadly adopted by the scientific community after several independent ectopic expression studies corroborated its functioning although several other contemporary experiments didn't fully verified the model, already at that time suggesting a more complex system (Causier et al., 2010).

And complexity indeed arised. The addition of a D-function to the model was proposed by (Colombo et al., 1995) after study of the gene *FBP11* in *P. hybrida*, the D-function being responsible, in cooperation with the C-function, for ovule development. But more importantly regarding the development of the main floral organs that are sepals, petals, stamens and carpels, another group of genes affecting B- and C-function were described first in tomato (Pnueli et al., 1994, 5) and *Petunia* (Angenent et al., 1994), *Antirrhinum* (Davies et al., 1996), and few years after *Arabidopsis* (Pelaz et al., 2000). Nowadays known as E-class genes, their expression is required for the expression of B- and C-function genes and are considered to control the establishment of the floral context needed for floral organ identity genes to function (Causier et al., 2010) (Fig. 1.4, D).

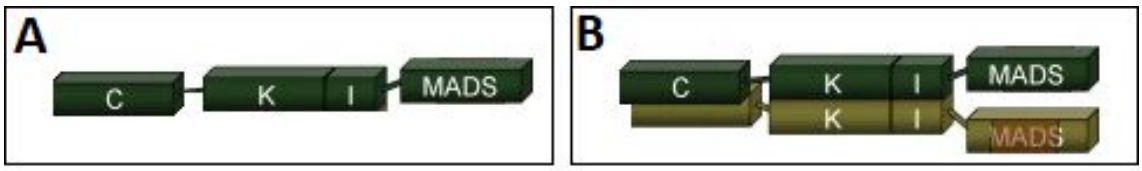


Figure 1.5: The structure of plant MADS-box proteins

(A) Schematic view of the structure of a plant MADS-box protein. **(B)** Schematic view of the structure of a plant MADS-box protein heterodimer. Adapted from (Causier et al., 2010).

Finally, it was proposed that the A-function is not widely shared and the general role of A-function genes is mainly to repress B- and C-functions to prevent their expression in the outer floral whorls, but they have no clear role in defining sepal and petal identity *per se* (Causier et al., 2010; Morel et al., 2017; Monniaux and Vandenbussche, 2018).

More and more additions and corrections to the (A)BC(D)E were made since it was proposed almost 30 years ago now, as well as several species-specific discoveries showing its limits outside of the historically described models, but it remains central in most floral-development-oriented studies as a strong foundation for new inquiries and discoveries.

III.2 – The MADS-box proteins and quartet model

As stated earlier, the interactions of A-, B-, C- and E-function within the flower lead to specific organs developing at specific locations. This combinatorial model suggested some form of interactions between these genes and/or the proteins they encode. Cloning the C-function gene *AGAMOUS* (*AG*) in *Arabidopsis* and the B-class gene *DEFICIENS* (*DEF*) in *Antirrhinum* showed homologies between their respective predicted proteins AG (Yanofsky et al., 1990) and DEF (Sommer et al., 1990) with known transcription factors (TFs) human SRF (Norman et al., 1988) and yeast MCM1 (Passmore et al., 1989). These four proteins founded the MADS-box TF family (MCM1, AG, DEF, SRF), a family now rich of hundreds of predicted or described proteins with its distinctive MADS DNA-binding motif being conserved throughout evolution from yeasts, to mammals, to plants, to nematodes and to insects (Messenguy and Dubois, 2003).

The MADS-box TFs of class II, to which the ABC proteins belong, are composed of four domains, from N-terminal (N-ter) to C-terminal (C-ter), MADS, I, K and C (Fig. 1.5, A). The highly conserved MADS domain encodes DNA-binding, recognizing consensus sequences named CArG boxes (for ‘CC-Arich-GG’, 5′-CC(A/T)6GG-3′) (Melzer and Theissen, 2009), nuclear localization and protein dimerization functions. Both I and K domains are moderately conserved. The Intervening (I) domain contributes to heterodimerization specificity and plays an indirect role in DNA-binding by stabilizing the MADS domain (Lai et al., 2021) while the Keratin-like (K) coiled-coil domain is necessary for dimerization and tetramerization (Puranik et al., 2014) (Fig. 1.5, B). Finally the highly variable Carboxyl-terminal (C) domain promotes transcriptional activity, high order MADS complexes formation and plays a role in functional specificity (Jack, 2004).

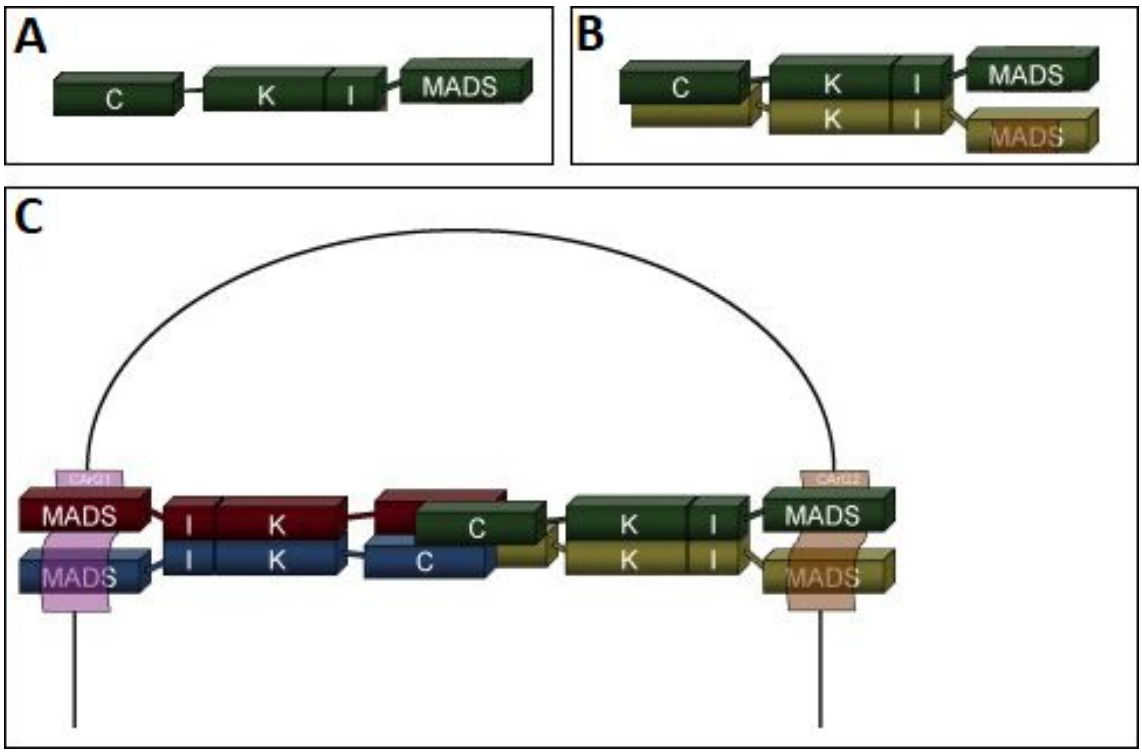


Figure 1.6: The structure of plant MADS-box proteins

(A) Schematic view of the structure of a plant MADS-box protein. **(B)** Schematic view of the structure of a plant MADS-box protein heterodimer. **(C)** Schematic view of the structure of a plant MADS-box protein tetramer binding DNA on two CarG consensus sequences. Adapted from (Causier et al., 2010).

The combinatorial nature of the (A)BC(D)E model, alongside the organization of MADS-box proteins with multiple protein-protein interaction domains led to the hypothesis that A-, B-, C- and E-class proteins interact together to play their biological roles. This was formulated as the "quartet" molecular model proposing that MADS-box proteins combine themselves in tetramers to bind pairs of CArG consensus sequences in close proximity of floral organ identity effector genes, controlling their expression to establish sepals, petals, stamens and carpels (Theissen and Saedler, 2001; Theißen, 2001) (Fig. 1.6, C).

This model is now broadly consensual among the scientific community although strong *in vivo* evidence was long to finally emerge. Following earlier work by (Melzer and Theißen, 2009) showing *in vitro* evidence of the quartet model, (Pajoro et al., 2014) showed *in vivo* genome-wide evidence that MADS-box proteins co-localize on the same regulatory sequences, supporting the existence of the quartet model even more. Since its proposal this model has been characterized at several precise levels. (Melzer et al., 2009) demonstrated that the E-function protein SEPALLATA3 (SEP3) is able to loop DNA *in vitro*. The looping of DNA by MADS-box proteins was further characterized by (Mendes et al., 2013), showing that STK-SEP3 (SEEDSTICK-SEPALLATA3, D- and E-function proteins) heterodimers induce short-range DNA looping in its target genes promoter by binding two neighbor CArG sequences, supporting the idea that a precise distance between CArG boxes is necessary for MADS-box protein action. Moreover, there is some evidence of protein-protein interactions specificity within the MADS-box proteins. For instance, the B-class proteins DEFICIENS (DEF) and GLOBOSA (GLO) are strict (or obligate) heterodimers and do not form dimers with other MADS-box proteins that we know of (Melzer et al., 2014). MADS-box proteins DNA-binding specificity is another striking aspect of the complexity of the quartet model. Each one of them having specific preferential binding CarG sequences variants leading to target-gene regulation specificity as a consequence of the different heterodimers at play (Smaczniak et al., 2017).

MADS-box proteins seem to participate in other protein-protein interactions. Multiple other protein interactors have been identified by *in vitro* assays (yeast two-hybrid or co-immunoprecipitation assays), although no strong *in vivo* evidence has yet been established in this regard to my knowledge. For instance, (Smaczniak et al., 2012) found that MADS-box proteins also interact with chromatin remodelers and Auxin Response Factors (ARFs) alongside other TFs, the latter also supported by literature (Bemer et al., 2017). Recently, the protein interaction network of FRUITFULL (FUL), a MADS-box protein involved in floral transition and pistil development, has been characterized in different tissues and at different developmental stages.

This revealed a highly dynamic network of protein-protein interactions across tissues, which changes FUL DNA-binding specificity, resulting in the regulation of different target genes (van Mourik et al., 2023).

The quartet molecular model added an extra level of complexity to the MADS-box TFs function. The findings made in its light obviously show that an even more intricate mechanism is at play, involving multiple levels of interactions involving multiple molecular families. However, this high complexity is not really a surprise, indeed, the amount and depth of biological changes at play under the influence of MADS-box TFs during flower development and that are needed for organ identity establishment and maintenance could only call for such complexity, complexity that we for sure do not fully understand as of yet, if we ever will.

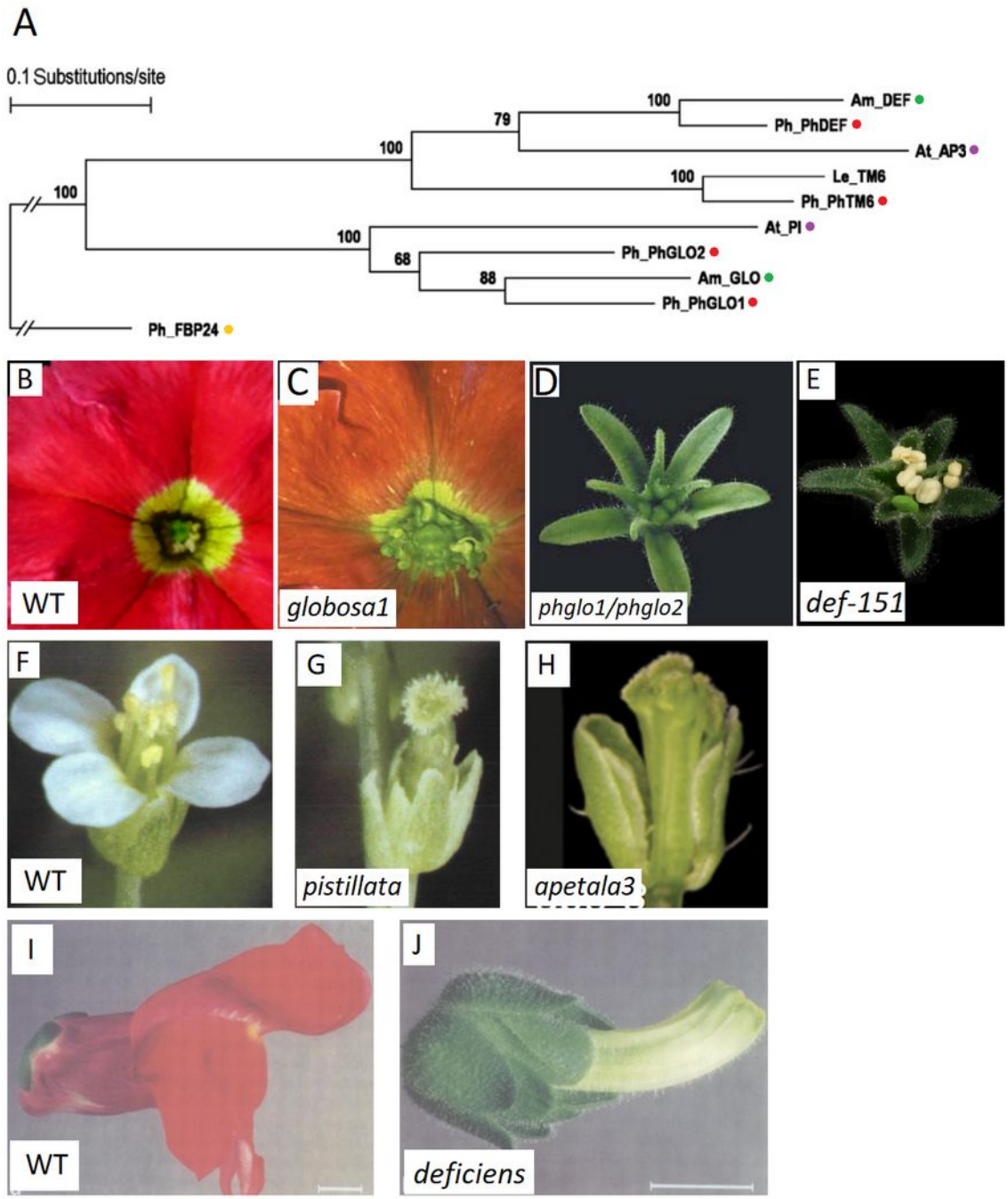


Figure 1.7: Neighbor-joining tree of B-class MADS-box genes across different species and a set of the corresponding mutants in *P. hybrida*

(A) neighbor-joining tree of B-class MADS-box genes from *Petunia hybrida* (Ph, red), *Arabidopsis thaliana* (At, purple), *Antirrhinum majus* (Am, green), and a tomato (Le) *TM6* lineage gene. Tree was rooted with *FBP24*, a *P. hybrida* member of the Bsister (Bs) MADS-box subfamily. Adapted from (Gerats and Strommer, 2009). **(B-E)** top view of *P. hybrida* wild-type, *phglo1* and *phdef-151* mutants flowers. **(F-H)** side view of *A. thaliana* wild-type, *pistillata* and *apetala3* mutants flowers. **(I-J)** side view of *A. majus* wild-type and *deficiens* mutant flowers. Adapted from (Sommer et al., 1990; Bowman et al., 1991; Vandenbussche et al., 2004; Wuest et al., 2012).

III.3 – *P. hybrida* B-class genes and petal development

Differing with the situation described in *Arabidopsis* and *Antirrhinum* in which petal identity is specified, in partnership with A- and E-class genes, by B-function heterodimers formed respectively by APETALA3 (AP3) and PISTILLATA (PI) or by DEF and GLO, *Petunia* petal identity results of a more intricate set of heterodimers of B-function proteins (Vandenbussche et al., 2004). It has been showed that several duplication events occurred in the *AP3/DEF* and *PI/GLO* lineages (Fig. 1.7, A) leading to functional redundancy and diversification inside both lineages, a major duplication event inside the *AP3/DEF* lineage coinciding with the higher eudicots radiation (Kramer et al., 1998).

In *Petunia* the aforementioned duplication events lead to the identification of 2 expressed genes inside each clade, *PhGLO1* (*P. hybrida GLOBOSA1*) and *PhGLO2* (*P. hybrida GLOBOSA2*) in the *GLO*/*PI* clade, *PhDEF* (*P. hybrida DEFICIENS*) and *PhTM6* (*P. hybrida TOMATO MADS BOX GENE6*) in the *DEF*/*AP3* clade. Interestingly, single-mutants for those genes show very different phenotypes (Vandenbussche et al., 2004).

On one hand, *phglo1* or *phglo2* single-mutants mostly show a wild-type-like phenotype (Fig. 1.7, B-C), in contradiction to what has been described in *Arabidopsis* for *AP3* and *PI* single-mutants (Fig. 1.7, F-H) (Bowman et al., 1989) or in *Antirrhinum* for *DEF* or *GLO* single-mutants (Fig. 1.7, I, J) (Sommer et al., 1990), all of these mutants showing homeotic conversions of petals into sepals and stamens into carpels. However, the double *phglo1*;*phglo2* mutant indeed shows a full homeotic conversion of petals into sepals and stamens into carpels (Fig. 1.7, D). This demonstrates the functional redundancy of *PhGLO1* and *PhGLO2*, one rescuing the missing function of the other in both petals and stamens.

On the other hand, *phdef* mutant show an homeotic conversion of petals into sepals while stamens are unaffected (Fig. 1.7, E), again in contradiction to what has been described in *Arabidopsis* or *Antirrhinum*, and in opposition with what is observed within the *GLO*/*PI* clade. This observation is coherent with the idea of a subfunctionalization of *PhTM6* leading to the absence of *PhTM6* activity inside petals, making it unable to rescue B-function in this single organ if *PhDEF* is knocked-out while still rescuing B-function in stamens, hence their normal development in *PhDEF* single-mutants.

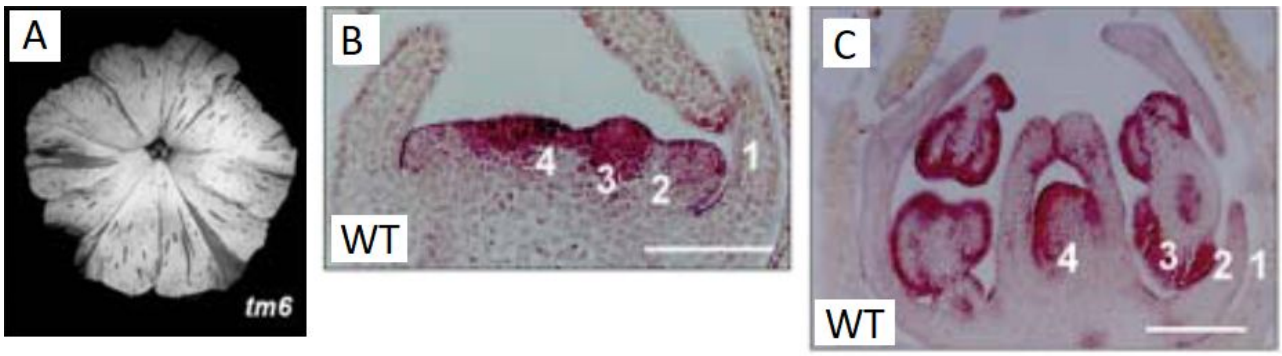


Figure 1.8: *phtm6* mutant flower and *in-situ* hybridization against *PhTM6*

(A) top view of a *phtm6* mutant flower. **(B-C)** *in-situ* hybridization against *PhTM6* mRNA in wild-type background at petal initiation and later in flower development. 1: sepal; 2: petal; 3 stamen; 4:carpel. Adapted from (Vandenbussche et al., 2004; Rijpkema et al., 2006).

The *phtm6* single-mutant has no visible petal phenotype (Fig. 1.8, A) (Rijpkema et al., 2006). *PhTM6* expression pattern showed that it behaves more like a C-class gene, although its sequence being closer to one of B-class genes (Fig. 1.8, B, C) (Vandenbussche et al., 2004).

These findings, backed with reverse transcription quantitative polymerase chain reaction (RT-qPCR), yeast-two-hybrids (Y2H) and *in*-situ localization assays led to the proposal of *Petunia* petal development being largely under the control of redundant PhDEF/PhGLO1 and PhDEF/PhGLO2 heterodimers while PhTM6 has likely no role in it. Hence, in *Petunia,* a single-mutant for *PhDEF* is sufficient to observe an homeotic conversion of petals into sepals without affecting stamens.

IV – From cell identity to organ identity and morphology

IV.1 – From organ identity to organ morphology

As previously described, the overlapping patterns of expression of several MADS-box genes in the flower leads to the formation of sepals, petals, stamens and carpels at precise locations from the outside to the inside. How does the organ identity, which is established early by the expression of these genes, determine final organ morphology and cell identities, is a central question that has been driving research since the ABC model proposal. In the following few paragraphs, I will try to illustrate the different levels of hierarchical regulations at play linking organ-identity, cell-identity and organ-morphology in *A. thaliana* flower development (but likely applicable to all flowering plants).

Considering flower development, the upper level of regulation is considered to be held by *LEAFY* (*LFY*), encoding a “pioneer” transcription factor (TF). LFY protein promotes floral meristem (FM) identity by activating the expression of the A-class and FM identity gene *APETALA1* (*AP1*) (Weigel et al., 1992; Bowman et al., 1993; Parcy et al., 1998). LFY is also responsible for the activation of all the other MADS-box genes involved in flower development, but interestingly in interactions with other players and not alone. LFY and UNUSUAL FLORAL ORGANS (UFO) partnership promotes the expression of the B-class genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) (Weigel et al., 1992; Honma and Goto, 2000). Finally, the C-class gene *AGAMOUS* (*AG*) expression is under the control of LFY and the homeobox protein WUSCHEL (WUS) (Lohmann et al., 2001). Once activated by LFY and its partners, the aforementioned MADS-box TFs will in turn activate hundreds if not thousands of downstream genes. These structural and effector genes all together will be responsible for the various major cellular shifts leading to cell differentiation and organ development.

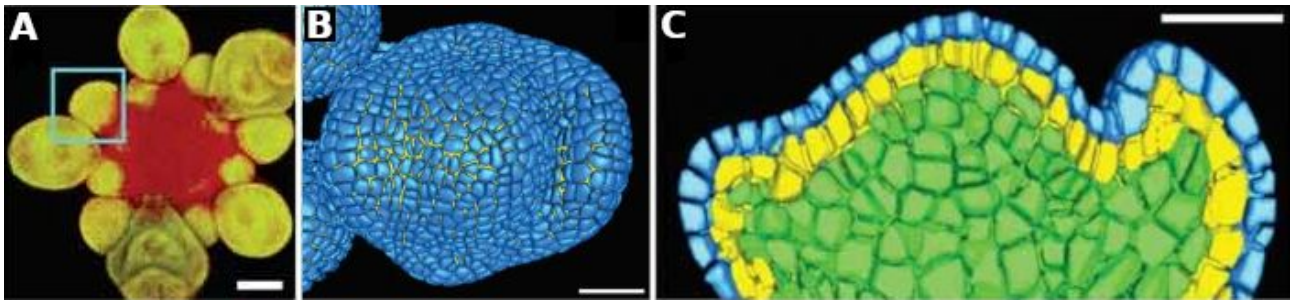


Figure 1.9: Informatics reconstruction of the clonal cell-layer organization of the flower meristem

(A) Early *Arabidopsis thaliana* inflorescence expressing a flower-specific GFP marker (pLEAFY::ER-GFP), the flower of interest is comprised within the blue square. **(B)** 3D segmentation resulting of signal analysis of the flower of interest. **(C)** Virtual section of the flower of interest with color coded cell layers; blue: L1; yellow: L2; green: L3 and deriving inner layers. Scale bars: 50 µm. Adapted from (Fernandez et al., 2010).

Once activated by LFY, *AP3* and *PI* are also capable of self-activation through a positive feedback loop driven by their obligate proteins heterodimerization (Lenser et al., 2009), a loop that is actively maintained until petals are mature, demonstrating their role not only in petal identity establishment through downstream genes regulation but also in petal identity maintenance.

An example of such downstream genes of the petal identity genes is *MIXTA* in *Antirrhinum majus*, encoding a R2R3-MYB TF that has been shown to drive petal epidermal cell formation (Martin et al., 2002). Displaying a very peculiar phenotype, usually conical (papillate), frequently pigmented and producing volatile scent molecules, these cells express a large set of very specific genes giving them their striking characteristics. In *Petunia*, anthocyanins biosynthesis-related genes are quite well described and participate in these characteristics by giving *Petunia* flowers their wide set of colors. Similarly to the establishment of cell-identity by MADS-box genes, anthocyanins biosynthesis in *Petunia* petal is under the control of protein complexes (MBW) formed by a MYB protein (among others ANTHOCYANIN2 (AN2)), the bHLH protein AN1 and the WD40 protein AN11 (Koes et al., 2005). Our group recently proposed that *AN2*, and possibly *AN1*, is activated by *PhDEF*, driving petal limb pigmentation, again showing the link between MADS-box organ-identity-establishing genes and cell-identity (Chopy et al., 2023).

Parallel to organ- and cell-identity, layer identity is of major importance in plants. Since plant cells are fixed into position by their cell wall and cannot migrate within tissues, organs are organized in clonal layers (Fig.1.9). The epidermis derives from the L1 which divides itself periclinally (Fig. 1.9, C, in blue). The mesophyll derives from the L2 and L3 layers, the L2 dividing periclinally and forming subepidermis tissues (Fig. 1.9, C, in yellow) while the L3 divides periclinally and anticlinally to form inner tissues (Fig. 1.9, C, in green). Cell layer identity is established very early at the embryo stage and maintained throughout development by genes such as *ARABIDOPSIS THALIANA MERISTEM LAYER* 1 (*ATML1*), *PROTODERMAL FACTOR2* (*PDF2*) in *Arabidopsis*, and later-acting genes such as *FIDDLEHEAD (FDH) or ANTIRRHINUM FIDDLEHEAD* (*AFI*) in *Antirrhinum* maintain this epidermal identity in mature organs (Yephremov et al., 1999; Efremova et al., 2004). Most of their underlying molecular regulation networks are still uncharacterized, although it was recently showed that *ATML1* expression in the epidermis is driven by mechanical stress and implies MAPK signaling and proteasome activity (Iida and Takada, 2021; Iida et al., 2023). Layer identity is also of crucial importance regarding organ morphology as demonstrated in *Antirrhinum (Perbal et al., 1996)* and *Petunia* (Chopy et al., 2023) flower chimeras.

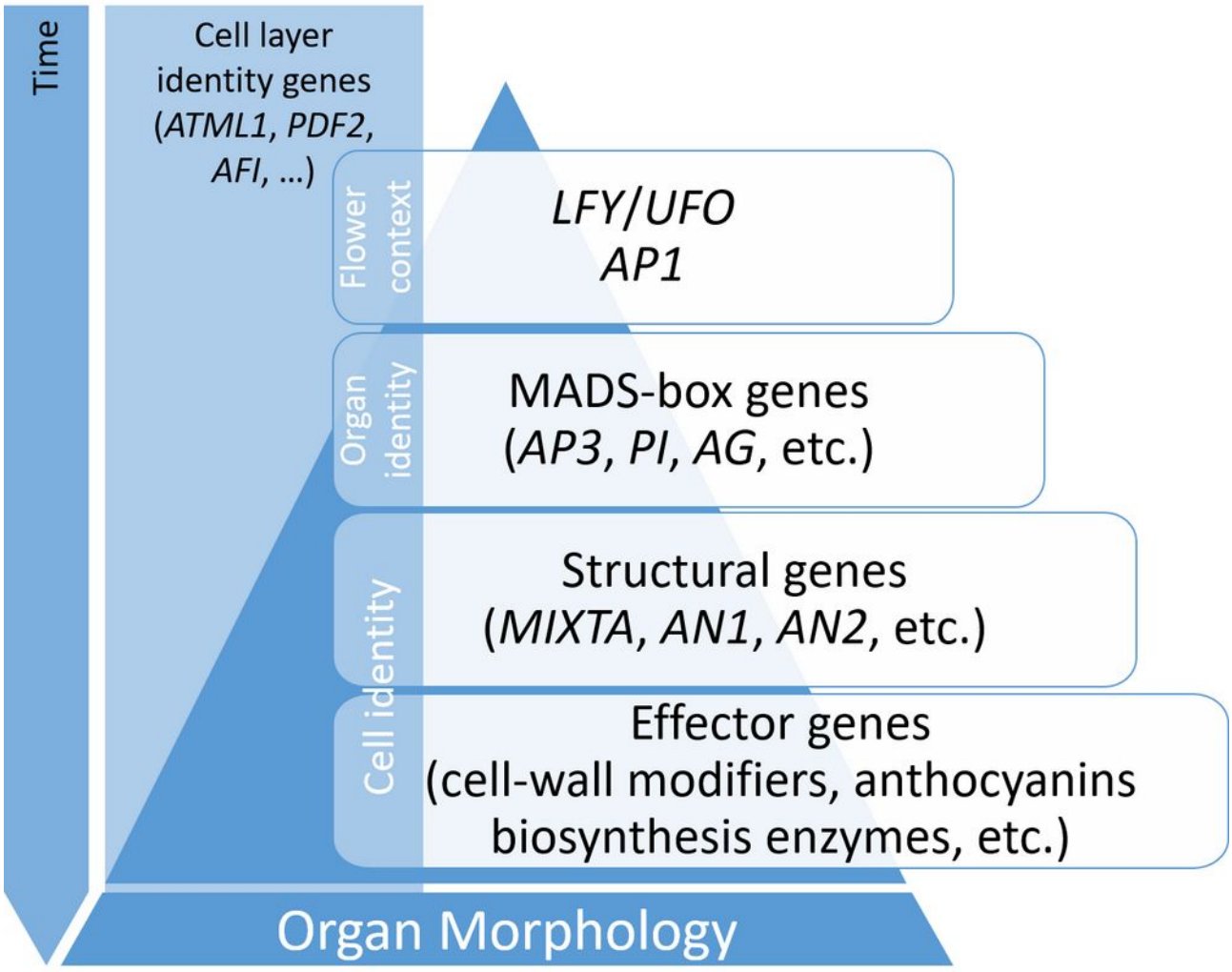


Figure 1.10: Temporal and gene expression regulation hierarchy of the main gene families involved in floral development

Schematic view of the intricate superposition of temporal, positional and gene regulation networks cues leading to the establishment of the flowering context, organ identity and cell identity towards organ morphology robustness during flower development.

IV.2 – Petal Cellular Identities

To sum up, from the establishment of the identity of cell layers and organs arise cell-identity and organ morphology (Fig. 1.10). As a way to better appreciate the state of knowledge about the characteristics and roles of the different cell types and tissues within the petals of angiosperm in my early thesis, Marie and I wrote the review featured below. It gives an overview of the two main layers composing a mature petal, the upper (adaxial) and lower (abaxial) epidermises and the mesophyll lying between. Both cell-type composition and tissue function were explored. It also glances over the petal organ identity establishment which I extensively covered in the previous part of this Introduction. Finally, it explores in more details the interplay between cell and organ identity by discussing petal conical cells and the layers organizing the petal. The Scanning Electron Microscopy (SEM) pictures featured in the review were taken by another member of the EvoDevo group at the RDP, Patrice Morel, I thank him for giving me the opportunity to use them as illustration.

frontiers in Plant Science

MINI REVIEW
published: 27 October 2021
doi: 10.3389/fpls.2021.745507

Petal Cellular Identities

*Quentin Cavallini-Speisser, Patrice Morel and Marie Monniaux**

Laboratoire de Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRAE, Lyon, France

Petals are typified by their conical epidermal cells that play a predominant role for the attraction and interaction with pollinators. However, cell identities in the petal can be very diverse, with different cell types in subdomains of the petal, in different cell layers, and depending on their adaxial-abaxial or proximo-distal position in the petal. In this mini-review, we give an overview of the main cell types that can be found in the petal and describe some of their functions. We review what is known about the genetic basis for the establishment of these cellular identities and their possible relation with petal identity and polarity specifiers expressed earlier during petal development, in an attempt to bridge the gap between organ identity and cell identity in the petal.

Keywords: petal, cell type, conical cell, mesophyll, epidermis, cell identity, petal polarities

OPEN ACCESS

Edited by:
Deshu Lin,
Fujian Agriculture and Forestry University, China

Reviewed by:
Tengbo Huang,
Shenzhen University, China
Amy Litt,
University of California, Riverside, United States

****Correspondence:***
Marie Monniaux
marie.monniaux@ens-lyon.fr

Specialty section:
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

Received: *22 July 2021*
Accepted: *04 October 2021*
Published: *27 October 2021*

Citation:
Cavallini-Speisser Q, Morel P and Monniaux M (2021) Petal Cellular Identities. Front. Plant Sci. 12:745507. doi: 10.3389/fpls.2021.745507

INTRODUCTION

Diversity in petal shape, size, color, and number is a key contributor to the dazzling variety of floral forms observed in the wild. The petal is often described as a very simple laminar structure, reminiscent of a leaf in its shape. The *Arabidopsis* petal could not be much simpler: a flat organ with a basal greenish claw and a distal white blade and only few different cell types (Irish, 2008). This simplicity makes it an excellent model to study plant organogenesis and cell type differentiation processes (Irish, 2008; Szécsi et al., 2014; Huang and Irish, 2016). However, *Arabidopsis* is only one among more than 350,000 flowering plant species (The Plant List, 2013), whose petal structures can be much more complex (Endress, 2001; Moyroud and Glover, 2017). Petals can display complex elaborations, such as lobes, fringes, nectary spurs, or hair pads (Endress and Matthews, 2006). In most asterid species, petals are fused together; therefore, the proximal (tube) and distal (limbs) parts of the fused petals can appear very different (Endress, 2001). Moreover, within a single flower, all petals are not the same, particularly in bilaterally symmetric flowers: Legume flowers develop distinct dorsal, lateral, and ventral petals (Ojeda et al., 2009). Petals also display an abaxial-adaxial polarity, the adaxial side of the petal being the upper/inner one (closest to the main stem), while the abaxial side is the lower/outer one. Finally, petal cells also have a layer identity, since petals generally derive from 2 (sometimes 3) layers from the shoot apical meristem that generates all aerial organs (Satina and Blakeslee, 1941; Jenik and Irish, 2000). Mature petals are thus typically composed of an adaxial epidermal layer (L1-derived), a few layers of mesophyll cells (L2-derived), and an abaxial epidermal layer (L1-derived).

In this mini-review, we will give an overview of the diversity of cell types that can be encountered on this apparently simple structure that is the petal. We will first focus on the two petal epidermises in which we find conical cells, together with many other cell types. We will next explore cell types and functions in the petal mesophyll, containing the petal vasculature surrounded by parenchyma cells. Finally, we will review the molecular mechanisms

involved in cell differentiation in the petal epidermis and their potential link with petal identity and polarity specifiers.

THE PETAL EPIDERMIS: CONICAL CELLS, STRIATIONS, TRICHOMES, AND STOMATA

Petal epidermal cells display striking differentiation features. The typical petal epidermal cell is conical (also called papillate), and this particular cell shape, readily observable by light microscopy or scanning electron microscopy, is often used as a marker for petal cell identity; indeed, it is found in 75–80% of angiosperm petals (Kay et al., 1981). Conical cells are generally found on the adaxial (upper) surface of the petal limb, and their shape and size can be extremely different among angiosperm species (Kay et al., 1981; Whitney et al., 2011a). They have been shown to increase petal color intensity and cause its sparkling appearance, increase pollinator's grip on the flower, affect overall petal shape, and decrease its wettability (Gorton and Vogelmann, 1996; Baumann et al., 2007; Whitney et al., 2009a, 2011a,b). They are also in most cases where pigments are produced (Kay et al., 1981) and frequently where scent is released (Baudino et al., 2007). All of the aforementioned traits potentially improve attraction and interaction with pollinators and therefore likely lead to a higher pollination success (Whitney et al., 2011a). Conical cells can thus be viewed as a key cellular innovation of flowering plants.

Other cell types are frequently found in the petal, and their distribution depends on their position in the petal. To explore this distribution along the petal proximo-distal axis, we chose the example of the petunia petal (*Petunia x hybrida*, **Figure 1A**). Petunia petals are fused, like petals from the vast majority of asterid flowers (Endress, 2001), and are organized in a tube and limbs (**Figure 1A**). In the limbs, cells are conical and smooth, and their density increases toward the center of the flower, which might influence petal color intensity and levels of emission of volatiles (Skaliter et al., 2021). At the most distal part of the tube, cells appear elongated and covered with striations (**Figure 1A**, tube 1). Striations are regular folds of the waxy cuticle of the outer epidermal cell wall and are frequently observed on petal epidermal cells (Antoniou Kourounioti et al., 2013). When regularly spaced and parallel oriented, these striations can cause light diffraction and iridescence of the petal, a visible cue for pollinators (Whitney et al., 2009b). Around the middle of the petunia petal tube, epidermal cells appear elongated with a small central papilla and still slightly striated (**Figure 1A**, tube 2). These striations progressively disappear as we progress toward the proximal part of the tube, and the central papilla becomes more and more pronounced (**Figure 1A**, tube 3). The function of this central papilla on tube cells is unknown.

Cell identity usually appears quite different on the two sides of the petal: Abaxial cells are flatter (lenticular) than adaxial conical cells, but they often contain pigments, and they can be a site of scent production (Kay et al., 1981; Baudino et al., 2007). Additionally, petal epidermal cells are often interspersed with trichomes, either glandular (for instance producing scent, nectar or defense compounds) or non-glandular ones, with various structures, shapes, and sizes. For instance in cotton flowers, both sides of the petals are covered in long non-glandular trichomes entangled together, resulting in the anchoring of adjacent petals together and their correct unfolding (Tan et al., 2016). In snapdragon flowers (*Antirrhinum majus*), glandular trichomes form very locally inside the corolla tube where they produce scent to attract pollinators and trap the pollen that they carry (Kolosova et al., 2001; Perez-Rodriguez et al., 2005). Finally, stomata are sometimes found on the petal epidermis, although their density is much more reduced than in leaves (Roddy et al., 2016; Zhang et al., 2018). They participate in gas exchange for photosynthesis in the petal (Zhang et al., 2018), and they might also be involved in maintenance of correct turgor pressure of the petal to avoid precocious wilting and have been proposed to play a role in flower opening in tulip (Azad et al., 2007).

This description of petal epidermal cell types is not exhaustive, and cell types in this tissue can be manifold. In elaborate petals, this diversity can be quite extreme. As an example, the *Nigella arvensis* flower forms highly elaborate petals of a complex shape with bifurcations and lobes, eyebrow-like stripes, long hairs, short trichomes, nectaries, and pseudo-nectaries (**Figure 1B**; Yao et al., 2019). Ten different subdomains can be defined in these petals, each displaying a distinct epidermal cell identity, among which conical cells, pavement cells, secretory cells, or polygonal cells, to cite just a few (Yao et al., 2019). One might argue that these petals are extremely derived and thus a particular case, but there is also strong variation in epidermal cell types on the petals of legume flowers, which are simple petals with a classical appearance (Dong et al., 2005; Ojeda et al., 2009).

THE PETAL MESOPHYLL: LIFE AND DEATH OF THE PETAL

In between the two epidermises stands the petal mesophyll, a spongy tissue whose thickness greatly varies between species: a single-cell layer in poppies (van der Kooi and Stavenga, 2019) but several dozens in the giant *Rafflesia* flower (Nikolov et al., 2013; Mursidawati et al., 2020). The petal mesophyll comprises the vascular bundles of the petal, surrounded by parenchyma cells that are roundish cells without any striking visual features.

One obvious role of the mesophyll is for petal nutrition. Vascular bundles embedded within the parenchyma supply the water and metabolites necessary for petal function. Additionally, in some species like petunia, mesophyll parenchyma cells contain chloroplasts, even in the mature petal (Weiss et al., 1988; Vainstein and Sharon, 1993). Coupled to the presence of stomata on the petal epidermis and lacunae in the mesophyll favoring gas exchange, conditions are gathered for active photosynthesis to take place in petunia petals, although it is not as intense nor as efficient as in leaves (Weiss et al., 1988, 1990).

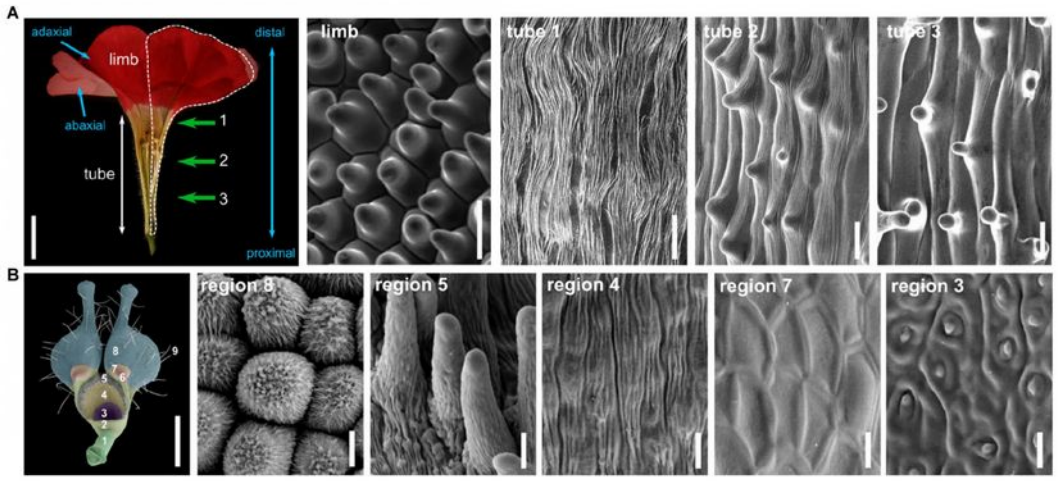


FIGURE 1 | Cellular identities in the petal epidermis. **(A)** Half-flower from *Petunia x hybrida*, cut open longitudinally to display the tube and limb regions (scale bar = 1 cm). The contour of one petal is shown with a white dotted line. Proximo/distal and adaxial/abaxial polarities of the petal are indicated with blue arrows. Scanning electron micrographs (scale bars = 20 µm) of the adaxial surface of petals in the limbs, and at three different regions from the tube, indicated by green arrows and numbers in the flower picture. These pictures were obtained with a HIROX SH-1500 bench top environmental scanning electron microscope equipped with a cooling stage (−10°C, 5 kV). **(B)** Petal from *Nigella arvensis* viewed from its adaxial side (scale bar = 1 mm), with nine regions with different cellular identities as identified in Yao et al. (2019). A 10th region is only visible on the abaxial side of the petal. Scanning electron micrographs of cells from five of these regions, giving an overview of cellular diversity in this organ (scale bars = 10 µm). Region 8: conical cells; region 5: short trichomes; region 4: oblong cells; region 7: polygonal cells with smooth surfaces; region 3: secretory cells. Pictures are reproduced from Yao et al. (2019) with permission from the authors.

This photosynthetic activity does not provide enough energy for the organ to be self-sustainable but, in particular, anthocyanin production appears to strongly depend on it (Weiss and Halevy, 1991).

The mesophyll is also involved in petal growth: In tulips, the mesophyll is considered to be the main driver of late petal growth by cell expansion (van Doorn and Van Meeteren, 2003), and in petunia, we recently showed that the mesophyll is the main driver for the growth of the petal tube (mainly by cell expansion), similarly to what had been previously observed in snapdragon flowers (Perbal et al., 1996; Efremova et al., 2001; Vincent et al., 2003; Chopy et al., 2021). In tulips and crocus flowers, temperature variation between lighted (outer) and shaded (inner) parts of the petal causes differential expansion of the parenchyma cell layers, resulting in flower opening (Wood, 1953). Similarly, in rose flowers, endoreduplication of parenchyma cells specifically on the adaxial side of the petal base, under the control of ethylene signaling, results in asymmetric growth of the petal mesophyll and flower opening (Cheng et al., 2021). Interestingly, only parenchyma cells toward the adaxial side of the petal respond to ethylene (Cheng et al., 2021), suggesting prior differentiation of mesophyll cells along the adaxial-abaxial axis.

The mesophyll also participates in petal pigmentation and therefore possibly in pollinator attraction. For instance in wallflowers petals (*Erysimum*), the epidermis is pigmented but the parenchyma cells also contain many chromoplasts and large pigmented cytoplasmic vesicles (Weston and Pyke, 1999). In the blue-flowered members of the *Boraginaceae* and *Liliaceae* families, the parenchyma cells contain anthocyanins and are the main contributor to petal pigmentation (Kay et al., 1981). The mesophyll can also influence the appearance of petals by reflecting or diffusing light. For example, buttercup petals (yellow-colored *Ranunculus*) have a reflective starch-containing parenchyma cell layer just underneath their epidermis, participating to the glossy and reflective petal surface (Parkin, 1928, 1931; Vignolini et al., 2012; van der Kooi et al., 2017). By a similar mechanism, the mesophyll of poppies and kingcup (*Caltha palustris*) petals contains large air cavities, creating a difference in refractive indices of the petal tissues and therefore strong light reflection and scattering, participating to the shiny appearance of the petals (Whatley, 1984; van der Kooi and Stavenga, 2019).

Finally, mesophyll cells are often the first site of petal senescence (van Doorn and Woltering, 2008). In petunia and lilies, this process begins in the petal parenchyma as early as 2 days after pollination, as evidenced by signs of autophagy (granules formation, loss of membrane integrity or expression of programmed cell death markers; Shibuya et al., 2013; Mochizuki-Kawai et al., 2015). This suggests that resource relocation after pollination, from the petal to the ovary, first relies on mesophyll degradation. In Iris flowers, mesophyll cell death begins at the apical part of the petal and progresses toward the base (van Doorn et al., 2003), suggesting that the mesophyll is not entirely homogeneous in this respect and that the process is influenced by petal polarity.

In summary, mesophyll cells play various specific roles over the course of petal development. Although parenchyma cells display only subtle differentiation features and therefore might not be classified into different cell types within this tissue, there can be a zonation of their activity and function along the different petal axes.

FROM ORGAN IDENTITY TO CELL IDENTITY

As proposed in the ABCE model of floral organ identity, petal identity is specified in a region of the floral meristem by expression of B-class genes in a floral context, defined by A- and E-class genes, most of them being MADS-box genes (Schwarz-Sommer et al., 1990; Coen and Meyerowitz, 1991; Pelaz et al., 2000; Causier et al., 2010; Thomson and Wellmer, 2019). This is generally well conserved within all angiosperms (Soltis et al., 2007; Irish, 2009). The question then arises as to how expression of a small number of MADS-box genes results in the specification of the different petal cell types that we have described in the previous paragraphs. To our knowledge, in the petal mesophyll, nothing specific is known about the molecular players downstream of MADS-box genes that could define cell identity. In contrast, the acquisition of cell identity in the petal epidermis has been well characterized at the molecular level, specifically for conical cells and trichomes. Interestingly, and although these two cell types can appear quite different, it might be relatively simple to switch from one to the other.

Major molecular players in conical cell formation are MIXTA and MIXTA-like proteins, belonging to the large group of R2R3-MYB (MYB proteins with two repeats of the MYB DNA-binding domain) transcription factors (TFs). *MIXTA* was first identified in snapdragon petals; it is sufficient to drive both conical cell and trichome differentiation when overexpressed in tobacco leaves, but since its endogenous expression pattern is only late during petal development, it only directs conical cell differentiation *in vivo* (Glover et al., 1998; Martin et al., 2002). Indeed, another *MIXTA*-like gene, *AmMYBML1*, is expressed early in the ventral petal, and because of this early expression, it directs both conical cell and trichome differentiation (Perez-Rodriguez et al., 2005). This suggests that conical cell and trichome specification processes are closely related to the molecular level, and that shifts in the spatio-temporal pattern of *MIXTA*-like genes expression are sufficient to drive conical cell and/or trichome specification, and therefore the patterning of these cell types at the petal scale.

More generally, the identities of various plant epidermal cell types are determined by MBW protein complexes, composed of one MYB TF, one bHLH TF, and one WD40 repeat protein (Ramsay and Glover, 2005; Robinson and Roeder, 2015). In the different species (mainly *Arabidopsis*, snapdragon, maize, and petunia) and tissues (root, leaf, seed, and flower) where these complexes have been studied, they can trigger the formation of different cell types (trichomes, stomata, pavement cells, or root hairs), the production of pigmentation (in the whole plant, the seed coat or the petal), or of other epidermal features (seed mucilage). The WD40 protein appears to have a general scaffolding role, and there has been only one WD40 protein identified per species, while there are few bHLH proteins and many different MYB proteins, thereby resulting in a combination of specific MBW complexes (Ramsay and Glover, 2005). In petals, the specific role of these complexes has been elucidated in particular when exploring the petal pigmentation patterns in different petunia lines. Production of anthocyanins in petunia petals is controlled by MBW complexes composed of the WD40 protein ANTHOCYANIN11 (AN11), the bHLH protein AN1, and an R2R3-MYB protein that can be any among AN2, AN4, DEEP PURPLE (DPL), or PURPLE HAZE (PHZ), which will, respectively, result in pigmentation in the limbs (AN2), in the tube and anthers (AN4), in the petal veins (DPL), or during blushing of the petal under high light (PHZ; Quattrocchio et al., 1993, 1999; de Vetten et al., 1997; Spelt et al., 2000; Albert et al., 2011). These complexes regulate the expression of several structural genes in the anthocyanin pathway (Quattrocchio et al., 1993; Huits et al., 1994). The diversity and specificity of action of each MYB protein grant high modularity to the petal pigmentation system and the potential to evolve subtle changes in pigmentation patterns while avoiding to loose anthocyanin production entirely (Ramsay and Glover, 2005). Cell identity and pigmentation are thus specified by similar protein complexes in the petal epidermis.

Additionally, the petal appears to be pre-patterned to specify particular cell fates when the right regulators are expressed at the right time and place (**Figure 2**). For instance, as seen previously, *MIXTA-like* genes do not direct the development of the same cell fates when expressed at a different time and place. What could this petal pre-patterning be? Briefly, markers of layer identity, such as the *HD-ZIP class IV* genes *MERISTEM L1 LAYER* (*ATML1*) or *PROTODERMAL FACTOR2* (*PDF2*) in *Arabidopsis* (Lu et al., 1996; Abe et al., 2003), specify epidermal identity from the embryonic stage onwards. Later, as floral organs initiate, their adaxial/abaxial polarity is established by genes, such as the *KANADI* and *YABBY* genes (abaxial side) and *HD-ZIP class III* genes (adaxial side; Siegfried et al., 1999; Kerstetter et al., 2001; Emery et al., 2003; Manuela and Xu, 2020), and their proximo-distal polarity is established by genes, such as *BLADE ON PETIOLE1* (*BOP1*) and *BOP2*, *TCP* genes or *JAGGED* (Hepworth et al., 2005; Norberg et al., 2005; Sauret-Gueto et al., 2013; Huang and Irish, 2015). More or less simultaneously, the B-class MADS-box genes specify petal identity, in a floral context specified by A- and E-class genes. Their initial expression appears quite homogeneous in all layers of the petal primordia (Urbanus et al., 2009; Prunet et al., 2017), but these genes are expressed throughout organ development and their expression pattern can be quite dynamic (Dornelas et al., 2011; Wuest et al., 2012). For instance in *Arabidopsis*, the E-class SEP3 protein is mostly expressed in the epidermis of the developing petal and more strongly on its adaxial side; similarly, the A-class AP1 protein accumulates more at the tip of developing sepals than at their base

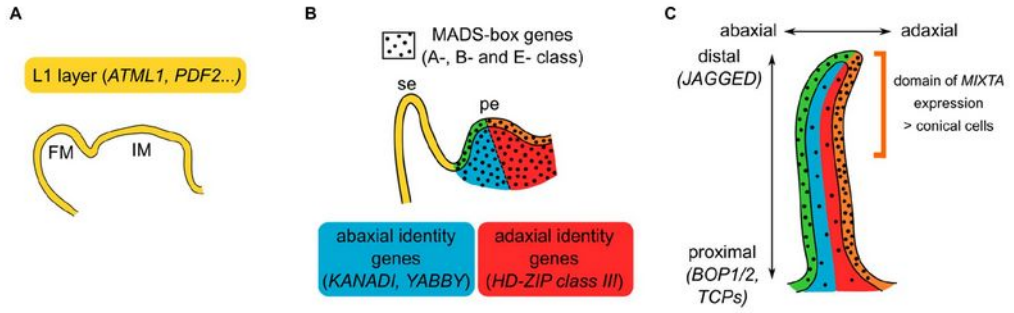


FIGURE 2 | Model for the combinatorial specification of petal identity and polarities. **(A)** In the inflorescence and flower meristems (IM and FM, respectively), layer identity is already specified. In particular, the meristematic L1 layer and the epidermis that derives from it express a specific set of genes. **(B)** Initiating sepal (se) and petal (pe). Petal identity is defined by the expression of B-class MADS-box genes, in a floral context defined by A- and E-class genes, in all layers of the petal primordia (black dots). Abaxial (blue) and adaxial (red) sides of the petal are specified by a set of genes. Intersection of epidermal identity with abaxial or adaxial identity results in petal subdomains represented in green and orange, respectively. **(C)** Developing petal. Layer identity and abaxial/adaxial polarity are still maintained, and proximo-distal polarity establishes as the petal grows. MADS-box genes expression is not necessarily homogeneous in the developing petal: here is depicted the case of *Arabidopsis* SEP3 protein, enriched in the petal epidermis and particularly on its adaxial side. The combination of these different positional signals can result in the specification of distinct cell fates: In snapdragon, *MIXTA* is only expressed in the epidermal, adaxial, and distal part of the petal, driving there the formation of conical cells.

(Urbanus et al., 2009; Dornelas et al., 2011). Interestingly, mutations in epidermal specifier genes from the *PDF2* family result in alterations of petal identity with reduced expression of the B-class gene *APETALA3* (*AP3*), suggesting that *AP3* might be a particularly prominent target of these epidermal specifiers (Kamata et al., 2013a,b). MADS-box gene expression and/or protein localization might thus depend on layer identity, abaxial/adaxial, and proximal/distal polarity specifiers, through molecular mechanisms unknown so far. *Vice versa*, members from the *HD-ZIP class IV*, *KANADI, YABBY, HD-ZIP class III*, or *TCP* gene families, as well as *BOP1*, are found within the direct regulatory targets of B-class proteins in *Arabidopsis* (Wuest et al., 2012), suggesting a feedback loop between petal identity and positional signals within the petal.

How could these different positional signals relate to the different cell identities observed in the petal? Quite similarly to the combinatorial ABCE model proposed for floral organ identity, we propose that the combination of positional signals in the petal specifies the patterning of different cell types at the petal scale (**Figure 2**). The example of *MIXTA-like* genes, the main specifiers of conical cell fate, can illustrate this idea: In snapdragon, *MIXTA* is specifically expressed in the adaxial epidermis of the petal, particularly at the distal part where conical cells develop (Glover et al., 1998). This specific expression pattern can be interpreted as the result of the presence of petal and epidermal markers, together with distal and adaxial polarity specifiers. Indeed, pieces of genetic or molecular evidence support a link between *MIXTA-like* genes expression or function and positional signals: *MIXTA-like* gene expression is genetically downstream of petal identity, proximo/distal, and adaxial/abaxial specifiers (Eshed et al., 2001; Perez-Rodriguez et al., 2005; van Es et al., 2018), and MIXTA-like proteins can directly interact with HD-Zip class IV and TCP proteins (Yan et al., 2018; Camoirano et al., 2021). Therefore, one can imagine that petal positional signals activate *MIXTA-like* genes expression in the right time and place, driving cell differentiation toward the conical cell fate, later reinforced by the direct interaction of MIXTA-like proteins with proteins specifying position in the petal.

Downstream this layer of regulatory genes, effector genes act to modify the cytoskeleton arrangement and the cell wall, to give the petal cells their final shape and function, participating to their identity. Most of the knowledge on this topic comes from *Arabidopsis* conical cells, in which it was found that a circumferential arrangement of cortical microtubules, controlled by proteins such as KATANIN1, SPIKE1, or ROPs, supports cellulose deposition and cone formation (Ren et al., 2016, 2017). Other players, such as RHAMNOSE BIOSYNTHESIS 1, control cell wall composition in conical cells and thus correct cell and petal shape (Saffer et al., 2017), while striations on the surface of petal epidermal cells depend on enzymes from the cutin synthesis pathway (Li-Beisson et al., 2009). The direct link between those various effector genes and the upstream regulatory genes is not established yet, but a glimpse of the whole regulatory network is beginning to emerge (Irish, 2008; Huang and Irish, 2016). Additional molecular evidence is needed to understand how cell types are specified in the petal and surely, the processes of interest here are complex, continuous, and overlapping with each other, with extensive cross-talk involved throughout petal development.

CONCLUSION AND FUTURE DIRECTIONS

Although the petal is a simple laminar structure, it contains several different cell types whose identity is specified by a wide range of signals. How these signals are integrated at the

molecular level and result in a specific gene expression profile and cellular function is mostly unknown. Today, the petal should not be viewed as an organ with a single identity, but rather as a population of cells in a petal specification context, each with a slightly different combination of lineage and positional signals (Xu et al., 2021). Single-cell technologies (transcriptome, proteome, interactome, chromatin accessibility, metabolome…) will surely lead to breakthroughs in the understanding of cell type specification in the petal and the molecular basis for its variation between species.

AUTHOR CONTRIBUTIONS

PM performed the electron micrographs. QC-S and MM wrote the article. All authors contributed to the article and approved the submitted version.

FUNDING

This work is supported by grants to QC-S and MM from the Agence Nationale de la Recherche (grant ANR-19-CE13-0019, FLOWER LAYER).

ACKNOWLEDGMENTS

We thank the PLATIM platform (SFR BioSciences Lyon, UAR3444/CNRS, US8/Inserm, ENS de Lyon, UCBL) for electron microscopy technical support, Mathilde Chopy for kindly providing us with the petunia limb electron micrograph, and Clément Verez for his help in electron microscopy. We also thank Hongzi Kong and colleagues for agreeing that we reproduce their pictures from *Nigella arvensis* petals. We also thank Michiel Vandenbussche for critical reading of the manuscript.

REFERENCES

Abe, M., Katsumata, H., Komeda, Y., and Takahashi, T. (2003). Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in *Arabidopsis*. *Development* 130, 635–643. doi: 10.1242/dev.00292

Albert, N. W., Lewis, D. H., Zhang, H., Schwinn, K. E., Jameson, P. E., and Davies, K. M. (2011). Members of an R2R3-MYB transcription factor family in petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant J.* 65, 771–784. doi: 10.1111/j.1365-313X.2010.04465.x

Antoniou Kourounioti, R. L., Band, L. R., Fozard, J. A., Hampstead, A., Lovrics, A., Moyroud, E., et al. (2013). Buckling as an origin of ordered cuticular patterns in flower petals. *J. R. Soc. Interface* 10:20120847. doi: 10.1098/rsif.2012.0847

Azad, A. K., Sawa, Y., Ishikawa, T., and Shibata, H. (2007). Temperature-dependent stomatal movement in tulip petals controls water transpiration during flower opening and closing. *Ann. Appl. Biol.* 150, 81–87. doi: 10.1111/j.1744-7348.2006.00111.x

Baudino, S., Caissard, J.-C., Bergougnoux, V., Jullien, F., Magnard, J.-L., Scalliet, G., et al. (2007). Production and emission of volatile compounds by petal cells. *Plant Signal. Behav.* 2, 525–526. doi: 10.4161/psb.2.6.4659

Baumann, K., Perez-Rodriguez, M., Bradley, D., Venail, J., Bailey, P., Jin, H., et al. (2007). Control of cell and petal morphogenesis by R2R3 MYB transcription factors. *Development* 134, 1691–1701. doi: 10.1242/dev.02836

Camoirano, A., Alem, A. L., Gonzalez, D. H., and Viola, I. L. (2021). *Arabidopsis thaliana* TCP15 interacts with the MIXTA-like transcription factor MYB106/NOECK. *Plant Signal. Behav.* 16:1938432. doi: 10.1080/15592324.2021.1938432

Causier, B., Schwarz-Sommer, Z., and Davies, B. (2010). Floral organ identity: 20 years of ABCs. *Semin. Cell Dev. Biol.* 21, 73–79. doi: 10.1016/j.semcdb.2009.10.005

Cheng, C., Yu, Q., Wang, Y., Wang, H., Dong, Y., Ji, Y., et al. (2021). Ethylene-regulated asymmetric growth of the petal base promotes flower opening in rose (*Rosa hybrida*). *Plant Cell* 33, 1229–1251. doi: 10.1093/plcell/koab031

Chopy, M., Cavallini-Speisser, Q., Chambrier, P., Morel, P., Just, J., Hugouvieux, V., et al. (2021). Cell layer-specific expression of the B-class MADS-box gene PhDEF drives petal tube or limb development in petunia flowers. bioRxiv [Preprint]. doi: 10.1101/2021.04.03.438311

Coen, E. S., and Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353, 31–37. doi: 10.1038/353031a0

de Vetten, N., Quattrocchio, F., Mol, J., and Koes, R. (1997). The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Genes Dev.* 11, 1422–1434. doi: 10.1101/gad.11.11.1422

Dong, Z., Zhao, Z., Liu, C., Luo, J., Yang, J., Huang, W., et al. (2005). Floral patterning in *Lotus japonicus*. *Plant Physiol.* 137, 1272–1282. doi: 10.1104/pp.104.054288

Dornelas, M. C., Patreze, C. M., Angenent, G. C., and Immink, R. G. H. (2011). MADS: the missing link between identity and growth? *Trends Plant Sci.* 16, 89–97. doi: 10.1016/j.tplants.2010.11.003

Efremova, N., Perbal, M. C., Yephremov, A., Hofmann, W. A., Saedler, H., and Schwarz-Sommer, Z. (2001). Epidermal control of floral organ identity by class B homeotic genes in *Antirrhinum* and *Arabidopsis*. *Development* 128, 2661–2671. doi: 10.1242/dev.128.14.2661

Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., et al. (2003). Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. *Curr. Biol.* 13, 1768–1774. doi: 10.1016/j.cub.2003.09.035

Endress, P. K. (2001). Origins of flower morphology. *J. Exp. Zool.* 291, 105–115. doi: 10.1002/jez.1063

Endress, P. K., and Matthews, M. L. (2006). Elaborate petals and staminodes in eudicots: diversity, function, and evolution. *Org. Divers. Evol.* 6, 257–293. doi: 10.1016/j.ode.2005.09.005

Eshed, Y., Baum, S. F., Perea, J. V., and Bowman, J. L. (2001). Establishment of polarity in lateral organs of plants. *Curr. Biol.* 11, 1251–1260. doi: 10.1016/S0960-9822(01)00392-X

Glover, B. J., Perez-Rodriguez, M., and Martin, C. (1998). Development of several epidermal cell types can be specified by the same MYB-related plant transcription factor. *Development* 125, 3497–3508. doi: 10.1242/dev.125.17.3497

Gorton, H. L., and Vogelmann, T. C. (1996). Effects of epidermal cell shape and pigmentation on optical properties of *Antirrhinum* petals at visible and ultraviolet wavelengths. *Plant Physiol.* 112, 879–888. doi: 10.1104/pp.112.3.879

Hepworth, S. R., Zhang, Y., McKim, S., Li, X., and Haughn, G. W. (2005). BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in *Arabidopsis*. *Plant Cell* 17, 1434–1448. doi: 10.1105/tpc.104.030536

Huang, T., and Irish, V. F. (2015). Temporal control of plant organ growth by TCP transcription factors. *Curr. Biol.* 25, 1765–1770. doi: 10.1016/j.cub.2015.05.024

Huang, T., and Irish, V. F. (2016). Gene networks controlling petal organogenesis. *J. Exp. Bot.* 67, 61–68. doi: 10.1093/jxb/erv444

Huits, H. S., Gerats, A. G., Kreike, M. M., Mol, J. N., and Koes, R. E. (1994). Genetic control of dihydroflavonol 4-reductase gene expression in *Petunia hybrida*. *Plant J.* 6, 295–310. doi: 10.1046/j.1365-313X.1994.06030295.x

Irish, V. F. (2008). The *Arabidopsis* petal: a model for plant organogenesis. *Trends Plant Sci.* 13, 430–436. doi: 10.1016/j.tplants.2008.05.006

Irish, V. F. (2009). Evolution of petal identity. *J. Exp. Bot.* 60, 2517–2527. doi: 10.1093/jxb/erp159

Jenik, P. D., and Irish, V. F. (2000). Regulation of cell proliferation patterns by homeotic genes during *Arabidopsis* floral development. *Development* 127, 1267–1276. doi: 10.1242/dev.127.6.1267

Kamata, N., Okada, H., Komeda, Y., and Takahashi, T. (2013a). Mutations in epidermis-specific HD-ZIP IV genes affect floral organ identity in *Arabidopsis thaliana*. *Plant J.* 75, 430–440. doi: 10.1111/tpj.12211

Kamata, N., Sugihara, A., Komeda, Y., and Takahashi, T. (2013b). Allele-specific effects of PDF2 on floral morphology in *Arabidopsis thaliana*. *Plant Signal. Behav.* 8:e27417. doi: 10.4161/psb.27417

Kay, Q. O. N., Daoud, H. S., and Stirton, C. H. (1981). Pigment distribution, light reflection and cell structure in petals. *Bot. J. Linn. Soc.* 83, 57–83. doi: 10.1111/j.1095-8339.1981.tb00129.x

Kerstetter, R. A., Bollman, K., Taylor, R. A., Bomblies, K., and Poethig, R. S. (2001). KANADI regulates organ polarity in *Arabidopsis*. *Nature* 411, 706–709. doi: 10.1038/35079629

Kolosova, N., Sherman, D., Karlson, D., and Dudareva, N. (2001). Cellular and subcellular localization of *S*-adenosyl-l-methionine:benzoic acid carboxyl methyltransferase, the enzyme responsible for biosynthesis of the volatile ester methylbenzoate in snapdragon flowers. *Plant Physiol.* 126, 956–964. doi: 10.1104/pp.126.3.956

Li-Beisson, Y., Pollard, M., Sauveplane, V., Pinot, F., Ohlrogge, J., and Beisson, F. (2009). Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *Proc. Natl. Acad. Sci. U. S. A.* 106, 22008–22013. doi: 10.1073/pnas.0909090106

Lu, P., Porat, R., Nadeau, J. A., and O'Neill, S. D. (1996). Identification of a meristem L1 layer-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox genes. *Plant Cell* 8, 2155–2168. doi: 10.1105/tpc.8.12.2155

Manuela, D., and Xu, M. (2020). Patterning a leaf by establishing polarities. *Front. Plant Sci.* 11:568730. doi: 10.3389/fpls.2020.568730

Martin, C., Bhatt, K., Baumann, K., Jin, H., Zachgo, S., Roberts, K., et al. (2002). The mechanics of cell fate determination in petals. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 357, 809–813. doi: 10.1098/rstb.2002.1089

Mochizuki-Kawai, H., Niki, T., Shibuya, K., and Ichimura, K. (2015). Programmed cell death progresses differentially in epidermal and mesophyll cells of lily petals. *PLoS One* 10:e0143502. doi: 10.1371/journal.pone.0143502

Moyroud, E., and Glover, B. J. (2017). The evolution of diverse floral morphologies. *Curr. Biol.* 27, R941–R951. doi: 10.1016/j.cub.2017.06.053

Mursidawati, S., Wicaksono, A., and Teixeira da Silva, J. A. (2020). *Rafflesia patma* Blume flower organs: histology of the epidermis and vascular structures, and a search for stomata. *Planta* 251:112. doi: 10.1007/s00425-020-03402-5

Nikolov, L. A., Endress, P. K., Sugumaran, M., Sasirat, S., Vessabutr, S., Kramer, E. M., et al. (2013). Developmental origins of the world's largest flowers, Rafflesiaceae. *Proc. Natl. Acad. Sci. U. S. A.* 110, 18578–18583. doi: 10.1073/pnas.1310356110

Norberg, M., Holmlund, M., and Nilsson, O. (2005). The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. *Development* 132, 2203–2213. doi: 10.1242/dev.01815

Ojeda, I., Francisco-Ortega, J., and Cronk, Q. C. B. (2009). Evolution of petal epidermal micromorphology in Leguminosae and its use as a marker of petal identity. *Ann. Bot.* 104, 1099–1110. doi: 10.1093/aob/mcp211

Parkin, J. (1928). The glossy petal of ranunculus. *Ann. Bot.* os-42, 739–755. doi: 10.1093/oxfordjournals.aob.a090138

Parkin, J. (1931). The structure of the starch layer in the glossy petal of ranunculus. *Ann. Bot.* os-45, 201–205. doi: 10.1093/oxfordjournals.aob.a090266

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., and Yanofsky, M. F. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 405, 200–203. doi: 10.1038/35012103

Perbal, M. C., Haughn, G., Saedler, H., and Schwarz-Sommer, Z. (1996). Non-cell-autonomous function of the *Antirrhinum* floral homeotic proteins DEFICIENS and GLOBOSA is exerted by their polar cell-to-cell trafficking. *Development* 122, 3433–3441. doi: 10.1242/dev.122.11.3433

Perez-Rodriguez, M., Jaffe, F. W., Butelli, E., Glover, B. J., and Martin, C. (2005). Development of three different cell types is associated with the activity of a specific MYB transcription factor in the ventral petal of *Antirrhinum majus* flowers. *Development* 132, 359–370. doi: 10.1242/dev.01584

Prunet, N., Yang, W., Das, P., Meyerowitz, E. M., and Jack, T. P. (2017). SUPERMAN prevents class B gene expression and promotes stem cell termination in the fourth whorl of *Arabidopsis thaliana* flowers. *Proc. Natl. Acad. Sci. U. S. A.* 114, 7166–7171. doi: 10.1073/pnas.1705977114

Quattrocchio, F., Wing, J. F., Leppen, H. T. C., Mol, J. N. M., and Koes, R. E. (1993). Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. *Plant Cell* 5, 1497–1512. doi: 10.2307/3869734

Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N., Mol, J., et al. (1999). Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *Plant Cell* 11, 1433–1444. doi: 10.1105/tpc.11.8.1433

Ramsay, N. A., and Glover, B. J. (2005). MYB–bHLH–WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* 10, 63–70. doi: 10.1016/j.tplants.2004.12.011

Ren, H., Dang, X., Cai, X., Yu, P., Li, Y., Zhang, S., et al. (2017). Spatio-temporal orientation of microtubules controls conical cell shape in *Arabidopsis thaliana* petals. *PLoS Genet.* 13:e1006851. doi: 10.1371/journal.pgen.1006851

Ren, H., Dang, X., Yang, Y., Huang, D., Liu, M., Gao, X., et al. (2016). SPIKE1 activates ROP GTPase to modulate petal growth and shape. *Plant Physiol.* 172, 358–371. doi: 10.1104/pp.16.00788

Robinson, D. O., and Roeder, A. H. K. (2015). Themes and variations in cell type patterning in the plant epidermis. *Curr. Opin. Genet. Dev.* 32, 55–65. doi: 10.1016/j.gde.2015.01.008

Roddy, A. B., Brodersen, C. R., and Dawson, T. E. (2016). Hydraulic conductance and the maintenance of water balance in flowers: hydraulic structure-function of flowers. *Plant Cell Environ.* 39, 2123–2132. doi: 10.1111/pce.12761

Saffer, A. M., Carpita, N. C., and Irish, V. F. (2017). Rhamnose-containing cell wall polymers suppress helical plant growth independently of microtubule orientation. *Curr. Biol.* 27, 2248.e4–2259.e4. doi: 10.1016/j.cub.2017.06.032

Satina, S., and Blakeslee, A. F. (1941). Periclinal chimeras in datura stramonium in relation to development of leaf and flower. *Am. J. Bot.* 28, 862–871. doi: 10.1002/j.1537-2197.1941.tb11017.x

Sauret-Gueto, S., Schiessl, K., Bangham, A., Sablowski, R., and Coen, E. (2013). JAGGED controls *Arabidopsis* petal growth and shape by interacting with a divergent polarity field. *PLoS Biol.* 11:e1001550. doi: 10.1371/journal.pbio.1001550

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., and Sommer, H. (1990). Genetic control of flower development by homeotic genes in *Antirrhinum majus*. *Science* 250, 931–936. doi: 10.1126/science.250.4983.931

Shibuya, K., Niki, T., and Ichimura, K. (2013). Pollination induces autophagy in petunia petals via ethylene. *J. Exp. Bot.* 64, 1111–1120. doi: 10.1093/jxb/ers395

Siegfried, K. R., Eshed, Y., Baum, S. F., Otsuga, D., Drews, G. N., and Bowman, J. L. (1999). Members of the YABBY gene family specify abaxial cell fate in *Arabidopsis*. *Development* 126, 4117–4128. doi: 10.1242/dev.126.18.4117

Skaliter, O., Kitsberg, Y., Sharon, E., Shklarman, E., Shor, E., Masci, T., et al. (2021). Spatial patterning of scent in *Petunia corolla* is discriminated by bees and involves the ABCG1 transporter. *Plant J.* 106, 1746–1758. doi: 10.1111/tpj.15269

Soltis, D. E., Chanderbali, A. S., Kim, S., Buzgo, M., and Soltis, P. S. (2007). The ABC model and its applicability to basal angiosperms. *Ann. Bot.* 100, 155–163. doi: 10.1093/aob/mcm117

Spelt, C., Quattrocchio, F., Mol, J. N. M., and Koes, R. (2000). anthocyanin1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* 12, 1619–1631. doi: 10.1105/tpc.12.9.1619

Szécsi, J., Wippermann, B., and Bendahmane, M. (2014). Genetic and phenotypic analyses of petal development in *Arabidopsis*. *Methods Mol. Biol.* 1110, 191–202. doi: 10.1007/978-1-4614-9408-9_8

Tan, J., Walford, S.-A., Dennis, E. S., and Llewellyn, D. (2016). Trichomes control flower bud shape by linking together young petals. *Nat. Plants* 2:16093. doi: 10.1038/nplants.2016.93

The Plant List (2013). Version 1.1. Published on the Internet. Available at: http://www.theplantlist.org/ (Accessed October 22, 2021).

Thomson, B., and Wellmer, F. (2019). Molecular regulation of flower development. *Curr. Top. Dev. Biol.* 131, 185–210. doi: 10.1016/bs.ctdb.2018.11.007

Urbanus, S. L., de Folter, S., Shchennikova, A. V., Kaufmann, K., Immink, R. G., and Angenent, G. C. (2009). In planta localisation patterns of MADS domain proteins during floral development in *Arabidopsis thaliana*. *BMC Plant Biol.* 9:5. doi: 10.1186/1471-2229-9-5

Vainstein, A., and Sharon, R. (1993). Biogenesis of petunia and carnation corolla chloroplasts: changes in the abundance of nuclear and plastid-encoded photosynthesis-specific gene products during flower development. *Physiol. Plant.* 89, 192–198. doi: 10.1111/j.1399-3054.1993.tb01805.x

van der Kooi, C. J., Elzenga, J. T. M., Dijksterhuis, J., and Stavenga, D. G. (2017). Functional optics of glossy buttercup flowers. *J. R. Soc. Interface* 14:20160933. doi: 10.1098/rsif.2016.0933

van der Kooi, C. J., and Stavenga, D. G. (2019). Vividly coloured poppy flowers due to dense pigmentation and strong scattering in thin petals. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 205, 363–372. doi: 10.1007/s00359-018-01313-1

van Doorn, W. G., Balk, P. A., van Houwelingen, A. M., Hoeberichts, F. A., Hall, R. D., Vorst, O., et al. (2003). Gene expression during anthesis and senescence in iris flowers. *Plant Mol. Biol.* 53, 845–863. doi: 10.1023/B:PLAN.0000023670.61059.1d

van Doorn, W. G., and Van Meeteren, U. (2003). Flower opening and closure: a review. *J. Exp. Bot.* 54, 1801–1812. doi: 10.1093/jxb/erg213

van Doorn, W. G., and Woltering, E. J. (2008). Physiology and molecular biology of petal senescence. *J. Exp. Bot.* 59, 453–480. doi: 10.1093/jxb/erm356

van Es, S. W., Silveira, S. R., Rocha, D. I., Bimbo, A., Martinelli, A. P., Dornelas, M. C., et al. (2018). Novel functions of the *Arabidopsis* transcription factor TCP5 in petal development and ethylene biosynthesis. *Plant J.* 94, 867–879. doi: 10.1111/tpj.13904

Vignolini, S., Thomas, M. M., Kolle, M., Wenzel, T., Rowland, A., Rudall, P. J., et al. (2012). Directional scattering from the glossy flower of *ranunculus*: how the buttercup lights up your chin. *J. R. Soc. Interface* 9, 1295–1301. doi: 10.1098/rsif.2011.0759

Vincent, C. A., Carpenter, R., and Coen, E. S. (2003). Interactions between gene activity and cell layers during floral development. *Plant J.* 33, 765–774. doi: 10.1046/j.1365-313X.2003.01666.x

Weiss, D., and Halevy, A. H. (1991). The role of light reactions in the regulation of anthocyanin synthesis in *Petunia corollas*. *Physiol. Plant.* 81, 127–133. doi: 10.1111/j.1399-3054.1991.tb01724.x

Weiss, D., Schönfeld, M., and Halevy, A. H. (1988). Photosynthetic activities in the *Petunia corolla*. *Plant Physiol.* 87, 666–670. doi: 10.1104/pp.87.3.666

Weiss, D., Shomer-Ilan, A., Vainstein, A., and Halevy, A. H. (1990). Photosynthetic carbon fixation in the corollas of *Petunia hybrida*. *Physiol. Plant.* 78, 345–350. doi: 10.1111/j.1399-3054.1990.tb09046.x

Weston, E. L., and Pyke, K. A. (1999). Developmental ultrastructure of cells and plastids in the petals of wallflower (*Erysimum cheiri*). *Ann. Bot.* 84, 763–769. doi: 10.1006/anbo.1999.0981

Whatley, J. M. (1984). The ultrastructure of plastids in the petals of *Caltha palustris* L. *New Phytol.* 97, 227–231. doi: 10.1111/j.1469-8137.1984.tb04126.x

Whitney, H. M., Bennett, K. M. V., Dorling, M., Sandbach, L., Prince, D., Chittka, L., et al. (2011a). Why do so many petals have conical epidermal cells? *Ann. Bot.* 108, 609–616. doi: 10.1093/aob/mcr065

Whitney, H. M., Chittka, L., Bruce, T. J. A., and Glover, B. J. (2009a). Conical epidermal cells allow bees to grip flowers and increase foraging efficiency. *Curr. Biol.* 19, 948–953. doi: 10.1016/j.cub.2009.04.051

Whitney, H. M., Kolle, M., Andrew, P., Chittka, L., Steiner, U., and Glover, B. J. (2009b). Floral iridescence, produced by diffractive optics, acts as a cue for animal pollinators. *Science* 323, 130–133. doi: 10.1126/science.1166256

Whitney, H. M., Poetes, R., Steiner, U., Chittka, L., and Glover, B. J. (2011b). Determining the contribution of epidermal cell shape to petal wettability using isogenic *Antirrhinum* lines. *PLoS One* 6:e17576. doi: 10.1371/journal.pone.0017576

Wood, W. M. L. (1953). Thermonasty in tulip and crocus flowers. *J. Exp. Bot.* 4, 65–77. doi: 10.1093/jxb/4.1.65

Wuest, S. E., O'Maoileidigh, D. S., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., et al. (2012). Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proc. Natl. Acad. Sci. U. S. A.* 109, 13452–13457. doi: 10.1073/pnas.1207075109

Xu, X., Smaczniak, C., Muino, J. M., and Kaufmann, K. (2021). Cell identity specification in plants: lessons from flower development. *J. Exp. Bot.* 72, 4202–4217. doi: 10.1093/jxb/erab110

Yan, T., Li, L., Xie, L., Chen, M., Shen, Q., Pan, Q., et al. (2018). A novel HD-ZIP IV/MIXTA complex promotes glandular trichome initiation and cuticle development in *Artemisia annua*. *New Phytol.* 218, 567–578. doi: 10.1111/nph.15005

Yao, X., Zhang, W., Duan, X., Yuan, Y., Zhang, R., Shan, H., et al. (2019). The making of elaborate petals in *Nigella* through sdevelopmental repatterning. *New Phytol.* 223, 385–396. doi: 10.1111/nph.15799

Zhang, F.-P., Murphy, M. R. C., Cardoso, A. A., Jordan, G. J., and Brodribb, T. J. (2018). Similar geometric rules govern the distribution of veins and stomata in petals, sepals and leaves. *New Phytol.* 219, 1224–1234. doi: 10.1111/nph.15210

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

IV.3 – Cell identities: a simple concept hiding a complex truth

Since their first description in a cork sample by Robert Hooke in 1665, cells have been historically defined by their morphology, localization, ontogeny and function into cell types. Nowadays, three main identification keys are commonly used for such classification: their phenotype, function and, with the emergence of Next-Generation-Sequencing (NGS) technologies, transcriptome (Xu et al., 2021). The first part of this chapter aimed to describe the multiple levels of dynamics at play to establish cell identity leading to organ morphology in plants. The second part presented current knowledge about petal cell identities. This last part intends to show known sources of heterogeneity in cell types as an attempt to temper the idea that all cells belonging to a given cell type are absolutely identical to one another.

A first level of heterogeneity is a consequence of the stochastic nature of most biological processes. Gene expression has been shown to be stochastic at multiple levels (Lipniacki et al., 2006). Gene expression noise can therefore be divided in two groups (Elowitz et al., 2002). Intrinsic noise due to inherent variations in transcription and translation rates caused by the multiple layers of biological molecules interactions, protein-protein, protein-DNA, protein-RNA, among others. Extrinsic noise caused by cellular and environmental states, or the influence of cell-autonomous and non cell-autonomous cues introducing bias towards one outcome or another into the aforementioned intrinsic stochasticity.

A second level of heterogeneity is caused by the interactions between multiple gradients of signaling molecules. Within a tissue, each cell is constantly adapting its transcriptomic state in response to hundreds of molecular signals diffusing or being actively transported from neighboring cells or even from further away in a systemic manner. Each cell being at a given position at a given time, perceived signal composition and strength is specific and thus each cell has its own transcriptional state at a given time. A classical example of such mechanism would be the auxin concentration-dependent regulation of auxin response genes by auxin response factors (ARFs) proteins, mechanism involved in most aspects of plant development (Li et al., 2016).

Mechanical signaling also causes heterogeneity. For instance, as said earlier, it has been shown that *ATML1* expression in leaf epidermis is driven by mechanical stress. Its effect can even be so drastic that the removal of mechanical stress is sufficient to drive differentiation of outermost mesophyll cells into epidermal cells (Iida et al., 2023).

Speaking of differentiation, several scRNA-Seq datasets show a large variety of concomitant stages of differentiation within a given cell type, blurring the outline of this concept even more (Shahan et al., 2020; Zheng et al., 2021).

Finally, another factor of heterogeneity is the cell cycle. Indeed, in most somatic tissues cycling is asynchronous and most cells are not even actively cycling (Dolezel et al., 1999). Cell cycle drives major changes to allow DNA replication and cell division, resulting in very different transcriptional states between two cells, one dividing and the other quiescent, even if they belong to the same cell type. This heterogeneity has once again been observed in several scRNA-Seq datasets leading to the development of tools to mitigate their effect on analysis (Buettner et al., 2015; Barron and Li, 2016; Zhang et al., 2021a).

Recently the major technical advance that represents single-cell RNA sequencing (scRNA-Seq), about which I devote a whole other chapter in this Introduction, reminded everyone this heterogeneity by unveiling it for study. I believe the concept of cell type still is a valid approach to study biological systems. Indeed, grouping cells into bigger functional groups seems pertinent when studying multi-cellular organisms. One must however keep in mind the true cell-level heterogeneity that exists behind the monolithic idea of a cell type.

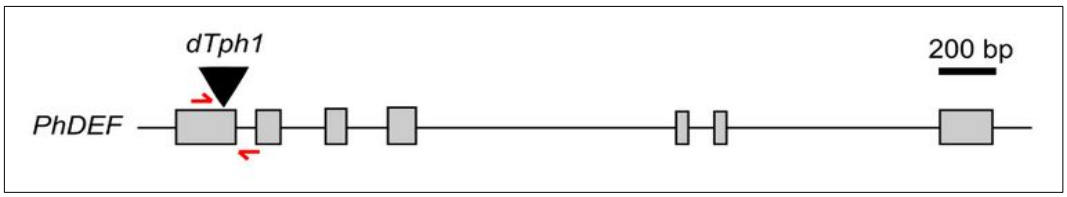


Figure 1.11: *PhDEF* structure and site of insertion of the *dTph1* transposon in *phdef-151* mutant

Exons are in gray, introns and regulatory regions a plain black line, the black inverted triangle the site of insertion of *dTph1* and the red arrows the position of sequencing and genotyping PCR primers.

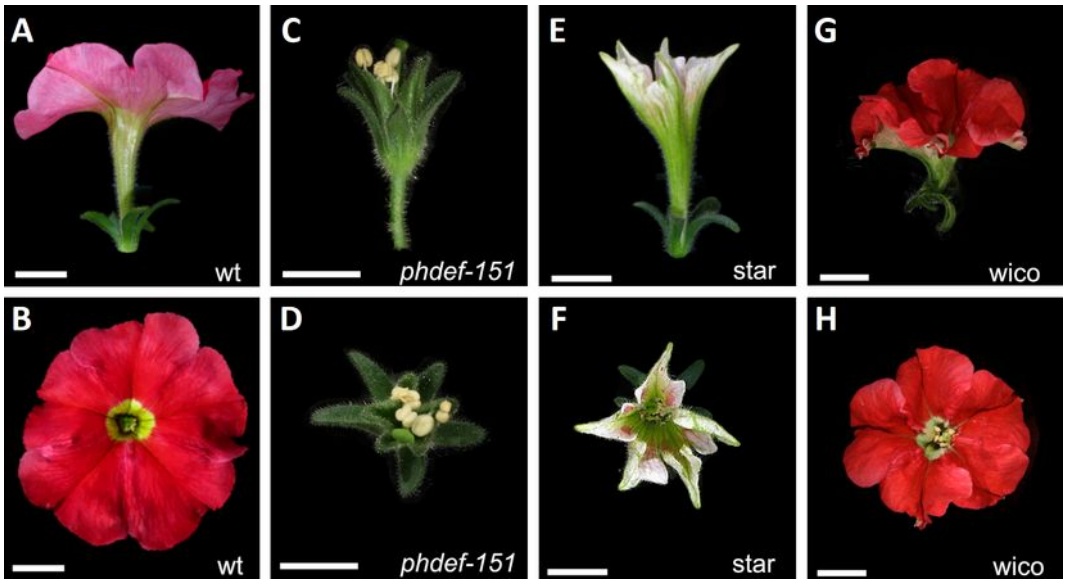


Figure 1.12: Morphological traits of *star* and *wico*

(A-B) wild-type *Petunia hybrida* W138 flower, side and top views. **(C-D)** *phdef-151* single-mutant flower, side and top views. **(E-F)** *star* flower, side and top views. **(G-H)** *wico* flower, side and top views. Scale bars: 1 cm. Adapted from (Chopy et al., 2023).

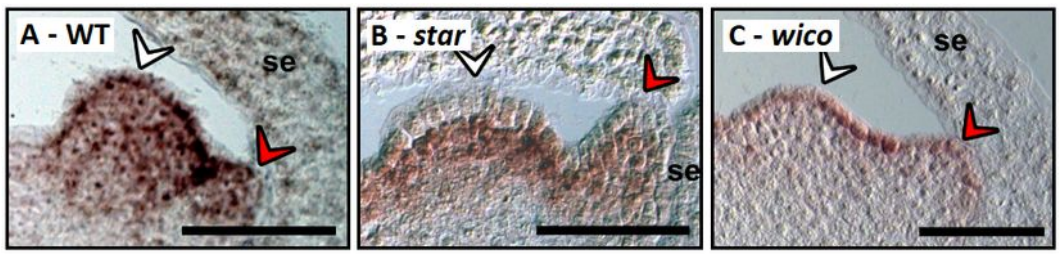


Figure 1.13: *in-situ* hybridization against *PhDEF* mRNA

(A) wild-type floral meristem. **(B)** *star* floral meristem. **(C)** *wico* floral meristem. White arrowheads: stamen primordium. Red arrowheads: petal primordium. Se: sepal.

V – *star* and *wico* mutants in a nutshell

This part gives a brief overview of the characteristics of the set of mutants I have worked with these last years to enlighten the last parts of this Introduction, and in order to bring the reader to the thesis questions and objectives. Full fledged study of these mutants is featured later-on in part I of the Results of this manuscript in (Chopy et al., 2023).

V.1 – Origin and morphology

While working with the B-class gene transposon-insertion loss of function single-mutant *phdef-151* in the *P. hybrida* W138 line background (Fig. 1.11), the group observed peculiar flowers that, instead of showing a full homeotic conversion of petals into sepals as expected (Fig. 1.11, C-D), showed 2 different intermediate petal phenotypes more related to wild-type phenotype.

One was named *star*, after its shape and the other *wico*, for “wide-corolla”. Compared to a wild-type flower (Fig. 1.12, A-B), the *star* flowers present a wild-type-like tube but their limb is poorly developed and mostly unpigmented (Fig. 1.12, E-F). On the other hand, the petals of *wico* show a very reduced tube while the limb remains largely wild-type-like, although smaller in surface and showing subtle ripples (Fig. 1.12, G-H).

V.2 – *star* and *wico* are periclinal chimeras

Further study of the genetic cues at play demonstrated that *star* and *wico* are chimeras for the expression of the B-class gene *PhDEF*. Both showed a reversion, caused by *dTph1* transposon excision from one *PhDEF* allele, of the wild-type gene sequence or in-frame variations, restoring in both case what appears to be a functional PhDEF protein. However, these reversions proved to be cell-layer-specific; therefore star and wico flowers have cell layers of a different genotype (homozygous or heterozygous mutant for *phdef-151*) and are so-called periclinal chimeras (Frank and Chitwood, 2016). In wild-type petal, *PhDEF* is expressed in all cell-layers (Fig. 1.13, A). In *star* petals, the epidermis remains knocked-out for the expression of *PhDEF* while it is restored in the mesophyll (Fig. 1.13, B). In *wico* petals the opposite was observed, *PhDEF* expression is restored in the epidermis while it is still knocked-out in inner layers (Fig. 1.13, C).

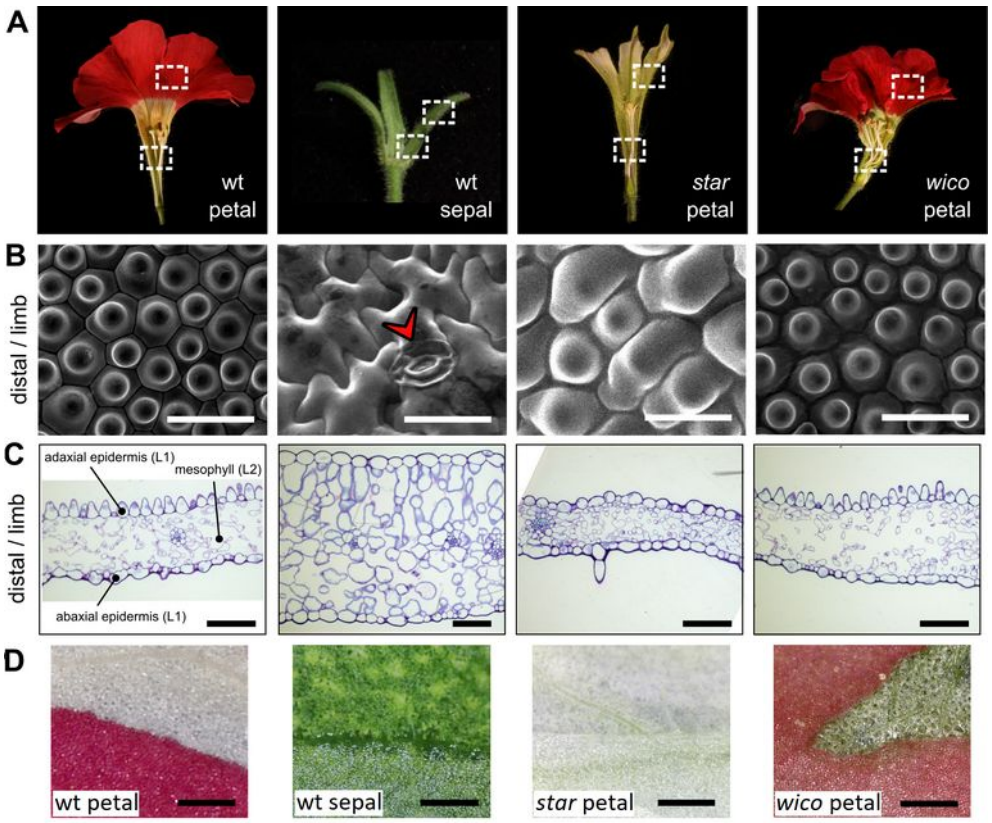


Figure 1.14: Histological traits of *star* and *wico*

(A) wild-type petal and sepal and *star* and *wico* petal longitudinally sectioned, the upper dotted square locates the distal (limb) zone featured in the observations underneath. **(B)** representative scanning electron micrographs of wild-type petal and sepal limb and *star* and *wico* petal limb epidermis, scale bars: 30 µm. **(C)** representative cross sections of wild-type petal and sepal limb and *star* and *wico* petal limb, scale bars: 100 µm. **(D)** area of wild-type petal and sepal limb and *star* and *wico* petal limb epidermis after a portion of the adaxial epidermis was manually peeled off. Adapted from (Chopy et al., 2023).

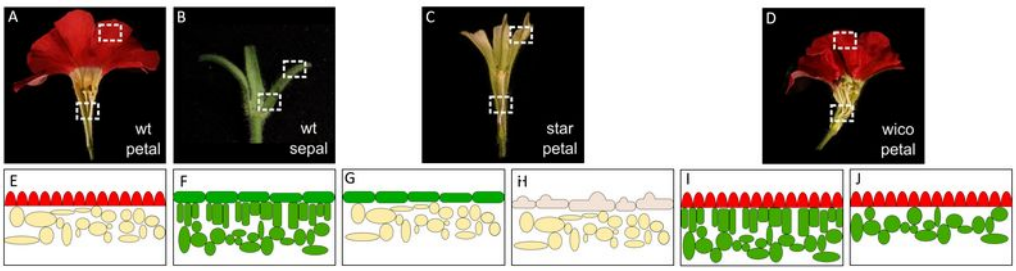


Figure 1.15: Schematic representations of wild-type petal and sepal and *star* and *wico* petal histology

(A-D) wild-type petal and sepal and star and wico petal longitudinally sectioned. **(E, F, H, J)** schematic representations of the observed histology of wild-type petal and sepal and star and wico petal limb. **(G, I)** schematic representations of awaited *star* and *wico* petal limb histology. Adapted from (Chopy et al., 2023).

V.3 – *PhDEF* expression and cell identity

Further histological analysis showed ambiguous cell morphology and tissue organization in both *star* and *wico*. Wild-type petal limb in *Petunia* presents typical pigmented conical epidermal cells and a very lacunous mesophyll (Fig. 1.14, A-D, first column). Wild-type sepal shows large, puzzle-piece shaped and flat chloroplastic epidermal cells delimiting a chloroplastic mesophyll composed of two distinct layers, an upper layer of loose palisadic cells while the underlying layer is even more lacunous, although not as strongly as in wild-type petal limb (Fig. 1.14, A-D, second column).

Knowing the chimeral nature of *star* and *wico*, one could imagine that *star* petal limb would present lacunous wild-type-petal-like mesophyll, since *PhDEF* expression is restored in these layers, and sepal-like flat epidermal cells where *PhDEF* expression is still knocked-out, as sketched in (Fig. 1.15, G). However although the mesophyll structure is close to one of a wild-type petal limb (Fig. 1.14, C, third column, Fig. 1.15, E), the epidermis shows large domed cells very different-looking to those of a wild-type sepal epidermis (Fig. 1.14, B, third column, Fig. 1.14, H).

Following same principles, one could imagine *wico* petal limb presenting wild-type-petal-like conical epidermal cells since *PhDEF* expression is restored in the epidermis, alongside a chloroplastic wild-type-sepal-like two-layer palisadic and lacunous mesophyll as proposed in (Fig. 1.15, I). But again, although presenting typical conical epidermal cells very comparable to wild-type petal limb epidermis (Fig. 1.14, B, fourth column, Fig. 1.15, E), the palisadic cell-layer awaited in wild-type-sepal-like mesophyll is missing (Fig. 1.14, C, fourth column, Fig. 1.15, J).

Similar histological discrepancies between awaited and observed tissue structures were found in the petal tube section of *star* and *wico* petal. Complementary descriptions in this regard are fully available in (Chopy et al., 2023).

These observations advocate for acquisition of petal identity at cellular and cell-layer levels under *PhDEF* control. Moreover, the presence of mixed cellular identities (domed *star* petal limb epidermal cells) and mixed cell-layer identities (absence of palisadic mesophyll in *wico* petal limb) in the cell layers still knocked-out for the expression of *PhDEF* lead to hypothesize that non-autonomous signaling involving *PhDEF* plays a role in *Petunia* petal formation, although the nature of this mechanism is unknown. Finally, the characteristics of *star* and *wico* petal support the idea of a cell-layer-specific role of *PhDEF* since its expression in all petal layers is required for good petal development and its cell-layer-specific expression leads to very distinct phenotypes. During my PhD, I used scRNA-Seq as a tool to unravel *PhDEF* cell-layer-specific regulation networks involved in *Petunia* petal development and suggested by *star* and *wico* phenotypes.

VI – scRNA-Seq: a novel and powerful tool to unravel *PhDEF* cell-layer-specific role in petal identity

VI.1 – What is scRNA-Seq?

VI.1.a – RNA-Seq and scRNA-Seq

RNA sequencing, abbreviated RNA-Seq, allows the identification and quantification of RNA molecules within a sample. Performing such an assay on a messenger RNA (mRNA) - enriched sample captures a snapshot of the transcripts that it contains, *i.e.* its transcriptome, at the time of sampling, allowing to appreciate the expression levels of thousands of genes and explore their regulation networks.

High-throughput next-generation sequencing (NGS) emerged in the mid 2000s. Its first display recognized as such by the scientific community yielded approximately 250,000 reads 100 nucleotide-long (Margulies et al., 2005). Nowadays, techniques yield either billions reads hundreds nucleotide-long or millions reads hundred-thousands nucleotide-long. In the past two decades, RNA-Seq, especially on mRNAs, has proven to be a central tool for biological research by helping deciphering gene regulation networks at play in biological systems.

However, this technique by design considers all mRNA molecules to be of the same origin, the sample. This aspect hides the true heterogeneity, cell-type-wise, and cell-state-wise, of the complex structures that are biological tissues, by averaging out transcriptional states of millions of cells into one reading. Tissue cell-heterogeneity is of key importance in multiple systems, for instance cancers are known since several decades now to be extremely heterogeneous, leading to a variety of responses to treatments for a given cancer type (Marusyk et al., 2012; Gerdes et al., 2014). In plants, cellular-heterogeneity in mechanical cues is known to influence cell growth and division in the shoot apical meristem (SAM) of *Arabidopsis thaliana (Long et al., 2020)*, while in roots there is evidence of cell-type-specific protein interactions (Long et al., 2017).

In order to address the loss of tissue-complexity information a novel technique arose in the last decade, single-cell RNA sequencing (scRNA-Seq). Various methods have been developed and all have the same goal, retrieve cell-level instead of tissue-level transcriptome, exposing the true heterogeneous nature of biological tissues which is not allowed by bulk RNA-Seq.

VI.1.b – Brief history overview and main scRNA-Seq techniques

Early single-cell experiment aimed to study precious, rare and/or small cell sub-populations (Svensson et al., 2018a). The first published single-cell transcriptome analysis was described using single mouse cells by (Tang et al., 2009). Cells were manually isolated by pipetting before ongoing

mRNA extraction, coding DNA (cDNA) library preparation and sequencing. Enough data was gathered in this world premiere to show differentially expressed genes between mutants and wild-type oocytes and splice-variants within blastomers. When (Guo et al., 2010) identified three distinct cell-types within 500 isolated mouse blastocysts in parallel, it became obvious to the scientific community that collecting and analyzing larger number of single-cells would yield new insights regarding cell-heterogeneity within tissues (Svensson et al., 2018a). This approach brought (Islam et al., 2011) to develop semi-automated cell isolation in 96-well plates followed by well-specific cDNA barcoding before multiplexing and sequencing, finally giving rise to the identification of all known mouse cortex cell-types using scRNA-Seq (Zeisel et al., 2015). In parallel, other isolation techniques were researched, for instance (Macosko et al., 2015) described a droplet-based microfluidic cell isolation method, Drop-Seq. A microfluidic chip mixes in very precise quantities an aqueous cell suspension, an aqueous reaction mix alongside DNA barcoded beads and an inert carrier oil into an emulsion of aqueous droplets containing a single cell and a single barcoded bead dispersed within the oil. Each droplet becomes an independent bioreactor were cells are lysed, their mRNAs captured by the beads and reverse-transcribed into barcoded cDNAs. After droplets disruption and cleanup, library amplification and sequencing can be performed as in bulk RNA-Seq. In the following years, another method was described by (Cao et al., 2017), sciRNA-Seq (for single-cell combinatorial indexing RNA-Seq). This approach is very different to the previous one since it never isolate cells. Instead, cells are fixed and permeabilized but they keep most of their structural integrity. Cells are then distributed into a 384-well plate. First strand reverse-transcription adds a first well-specific barcode and all wells are pooled together before once again being distributed into a 384-well plate. A second barcode is then added by PCR, these last two steps being repeatable. This method results in statistically uniquely barcoded cDNAs for each cell, the probability of a same cell visiting the same sequence of wells and getting the same barcodes decreasing exceedingly quickly with the number of time the last two steps are repeated. This last approach has proven to work in single-nuclei RNA-Seq (snRNA-Seq) using isolated nuclei instead of isolated cells (Rosenberg et al., 2018).

Over the years a still growing number of other techniques, most of them variations of micro-well plate, droplet or fluorescence-activated cell sorting (FACS) sorting, were developed. Among others, one can cite STRT-Seq (Natarajan, 2019), Smart-Seq (Wang et al., 2021b; Hagemann-Jensen et al., 2022), DNBelabC4 (Liu et al., 2019; Wang et al., 2021a), Drop-Seq (Macosko et al., 2015) and its commercial variation, 10x Genomics Chromium (10x Genomics, Pleasanton, CA, USA).

Although most of the scRNA-Seq techniques were originally described using animal models, scRNA-Seq usage is exploding in plant science also, although with some delay due to necessary technical adjustments. Early trials to gather better resolved transcriptome among different cell-types of a given plant tissue relied on FACS, isolation of nuclei tagged in specific cell types (INTACT) (Deal and Henikoff, 2011) or laser capture micro-dissection (LCM) (Asano et al., 2002) but these methods were still merely averaging the expression profiles of several cells within each cell-type and not strictly speaking at the single-cell resolution (Shaw et al., 2021). The first two proper scRNA-Seq studies in plant-science were published by (Brennecke et al., 2013) and (Efroni et al., 2015). In both studies fluorescent-tagged *A. thaliana* root-tip protoplasts, *i.e.* plant cells without their cell-wall, were manually isolated in micro-tubes before RNA-Seq. Possibilities opened by these two pioneer studies regarding transcriptome study in plant biology are vast, but two main problems remained considering the technique in use. First, manually isolating cells does not allow high-throughput studies of thousands of cells. Second, using fluorescent-tagged protoplasts limits the application to models were such techniques are developed and work. However, in 2019, several studies demonstrated that using droplet-based approaches yielded good results on *A. thaliana* root-tip protoplasts, allowing the study of thousands of individual cells in a single run (Denyer et al., 2019a; Jean-Baptiste et al., 2019; Liu et al., 2019; Ryu et al., 2019; Shulse et al., 2019). Since then, this approach has been widely used by plant biology research groups on an ever-growing variety of models such as maize shoot apical meristem (Satterlee et al., 2020), *Arabidopsis* lateral root (Gala et al., 2021), *Arabidopsis* and maize leaves (Bezrutczyk et al., 2021; Kim et al., 2021), *Nicotiana* flower petal (Kang et al., 2022) or woody angiosperms stem (Tung et al., 2023).

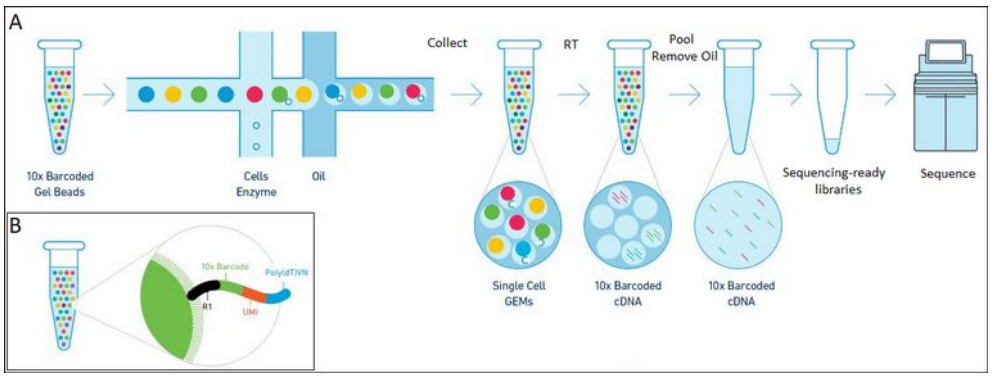


Figure 1.16: 3' Chromium Next GEM Single Cell Gene Expression Assay workflow overview

Adapted from 10x Genomics.

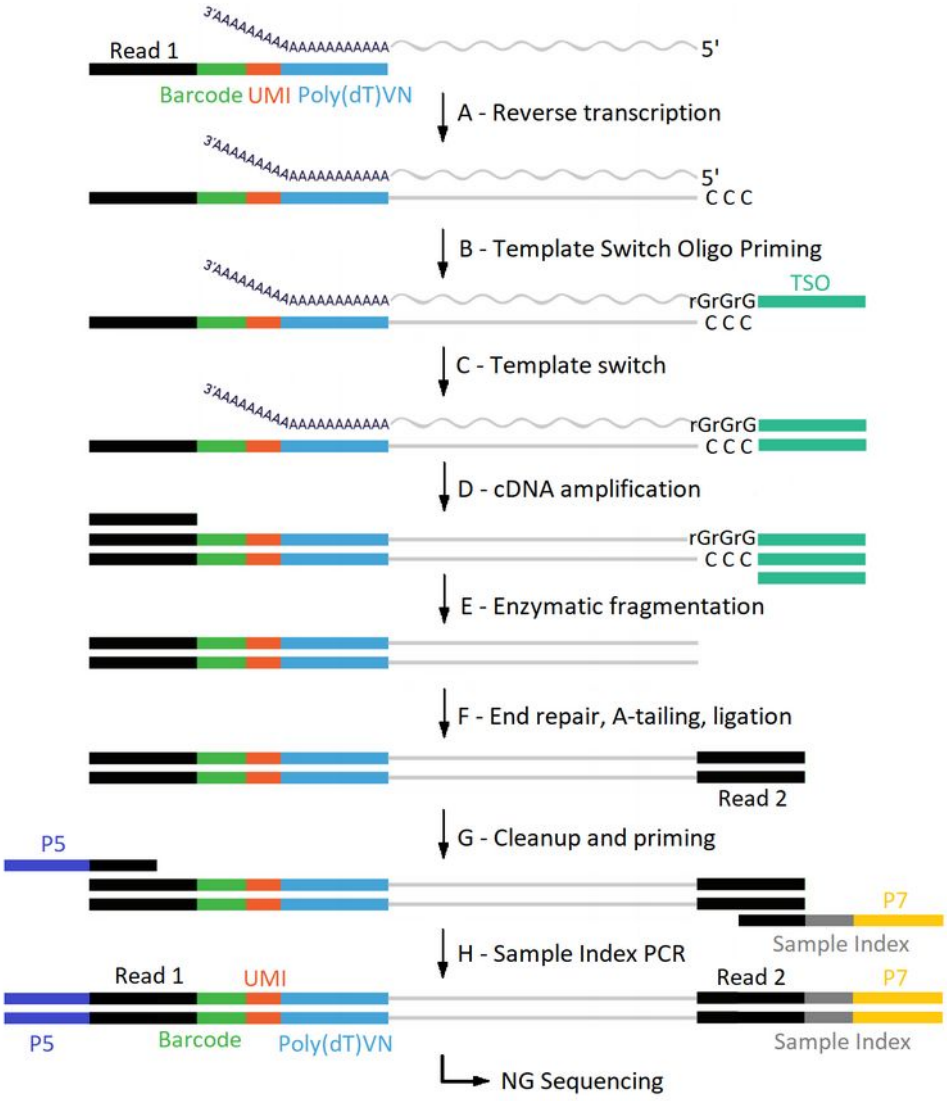


Fig. 1.17: 3' Chromium sequencing library preparation

Adapted from 10x Genomics.

VI.1.c – 10x Genomics Chromium scRNA-Seq in more details

During my thesis I used a droplet-based approach, derivative of Drop-Seq (Macosko et al., 2015), developed by 10x Genomics, 3' Chromium Next GEM Single Cell Gene Expression Assay, of which here is a brief overview. As described earlier, this technique relies on isolating aqueous droplets each containing a single cell alongside a barcoded gel bead within a lipidic matrix by using microfluidics (Fig. 1.16, A). Generating the emulsion is quite straightforward, once a cell suspension of the necessary quality and cell-concentration is obtained in earlier steps of the workflow. The cell suspension, the reaction mix with barcoded gel beads and the carrier oil are loaded in a custom microfluidic chip which is then loaded into the Chromium controller.

This bench-top apparatus precisely pressurizes the chip, mixing the previously loaded fractions together into an emulsion in about 15 minutes. The emulsion can then be processed for library preparation, the cDNA reverse transcription and barcoding taking place inside the individual droplets, later on disrupted by heat before library amplification followed by sequencing (Fig. 1, A).

The barcoded gel beads are covered by single-strand DNA adapters allowing capture and barcoding of mRNA molecules. The barcodes are composed of 4 main parts, a 22-nt partial Illumina TruSeq Read 1 sequence (Fig. 1.16, B, R1), a 16-nt barcode (Fig. 1.16, B, 10x Barcode) identifying the bead, a 12-nt unique molecular identifier (Fig. 1.16, B, UMI) allowing for bias corrections during analysis and a 30-nt poly-dT capture tail (Fig. 1.16, B, Poly(dT)VN). All adapters have the same barcode for a given bead while UMIs are random.

Once encapsulation is complete, the cells are lysed and their mRNAs captured by the beads through the poly-dT tails of the adapters. After reverse-transcription of each mRNA into its complementary cDNA (Fig. 1.17, A), a template switch oligonucleotide (TSO) is hybridized to the cytosines (Fig. 1.17, B, “CCC”) added by the reverse transcriptase (RT) upon reaching the end of the mRNA (Fig. 1.17, B), allowing the RT to further add a known sequence to the previously synthesized cDNA 3’ end (Fig. 1.17, C). The mRNA is discarded and the cDNA amplified (Fig. 1.17, D). TSO sequences are removed by enzymatic digestion (Fig. 1.17, E) and a 22-nt partial Illumina TruSeq Read 2 sequence ligated to former TSO ends (Fig. 1.17, F). A sample index and P5 and P7 sequencing primers are finally added by PCR (Fig. 1.17, G-H) and the library can be sequenced with Illumina.

VI.1.d – scRNA-Seq classical data analysis pipeline

The shear amount of data generated by high-throughput NGS calls for data handling and analysis through bioinformatics. scRNA-Seq is of no exception in this regard since a single sample can yield billions of reads needing analysis. Although quite a recent method, the ever-growing popularity among the scientific community of scRNA-Seq has for consequence that multitude of analysis tools exist, which is a relief for non-bioinformaticians such as me, but can also lead in some pitfalls since the choice is so wide it can be overwhelming.

The standard workflow of a scRNA-Seq dataset analysis is the following. The first step is to preprocess the reads obtained after sequencing, ensuring good quality, perform trimming if needed and normalizing the read counts to account for sequencing depth variations, batch effects, etc. At this step, amplification bias can also be corrected using the barcode-UMI-read sequence, the UMI being used to distinguish sequenced reads that originate from unique mRNA molecules vs PCR duplicates.

Once this first step performed, the reads need to be given their cell-barcode identity, counted and mapped onto a reference transcriptome. These secondary results are usually stored inside a matrix linking read counts to their cell-barcode and gene identity.

This matrix needs to be further normalized and its dimensions reduced to greatly decrease computing time and power needed to analyze the dataset. The main method in use is the principal component analysis (PCA) (Pearson, 1901; Jolliffe and Cadima, 2016, 2016) which helps determine an optimal reduced number of dimensions retaining the essence of the original data.

The fourth main step consists of clustering cells together by transcriptome resemblance. The basic idea is to compare gene expression levels of all cells to one another and establish a graph of differences and similarities to be able to regroup similar cells together for further analysis.

Finally the most interesting step can be performed, differential gene expression analysis (DGE). Very similar to what is done on bulk RNA-Seq data comparing different samples using R and DESeq2 (Love et al., 2014), the idea is to compare cell clusters together to detect differentially expressed genes between them. This step should first help assigning a cell-type identity to the detected clusters and help refine the clustering by regrouping or splitting clusters with known marker genes if need be. Once the clustering is refined, DGE data can be used further to study a biological mechanism in a cell-type-specific manner and even at the single-cell level, reaching the intended usage of scRNA-Seq.

VI.1.e – Technical drawbacks of scRNA-Seq

Since an already substantial and still growing variety of techniques exists, there is no full set of drawbacks that will match each and every technique. However there are a few difficulties that are shared by all current scRNA-Seq methods.

First of all, due to the reduced amount of starting material, the smallest unit of observation window being a single cell in scRNA-Seq instead of several millions in bulk RNA-Seq, the capture efficiency is low and therefore dropouts events, not detecting a given transcript, are highly frequent. Technical limits of even the most sensitive techniques do not fully overcome this issue as of mid 2023, therefore the number of detected expressed genes in scRNA-Seq is usually lower when compared to bulk-RNA-Seq (Haque et al., 2017).

Second, the known transcriptional stochasticity in gene expression (Lipniacki et al., 2006) as well as the blunt heterogeneity of cells composing a sample, cause scRNA-Seq data to be much more variable and noisy than bulk RNA-Seq. For instance, the cell cycle state is known to influence greatly the transcriptome of a cell, independent of its identity.

VII – PhD biological questions and objectives

As presented during this Introduction, *star* and *wico* flowers demonstrate the existence of a cell-layer-specific role for the B-class gene and petal identity master regulator *PhDEF* in *Petunia*. Indeed, the poorly developed and unpigmented petal limb in *star* shows that *PhDEF* epidermis-expression specifically plays a role in petal limb development and pigmentation. Conversely, the nearly absent petal tube in *wico* suggests that *PhDEF* inner-expression specifically drives petal tube growth. On the other hand, mixed cellular and cell-layer identities present in both *star* and *wico* suggest that non-cell-autonomous effects, whether of molecular or mechanical nature, are also at play during petal development in *Petunia*.

The broad biological question that can circumvent my PhD work is: what is the cell-layer-specific role of *PhDEF* during *Petunia* petal development? More specifically, I asked the following:

1. What are the genes regulated by PhDEF in the two cell layers of the petal? Is there a large difference in the number and function of these genes?

2. Can we identify key cell layer-specific genes involved in tube versus limb development?

3. Can we identify key cell layer-specific genes involved the establishment of specific petal cell types?

In order to answer these questions various technical objectives were reached. The main one was to put together a full scRNA-Seq pipeline, from cell isolation to semi-automated bioinformatics analysis. The original goal was to perform scRNA-Seq on wild-type, *def-151* single mutant, *star* and *wico* petals at four developmental stages. However, technical challenges in protoplast isolation have changed these original plans and scRNA-Seq could only be applied on mature petals of wild-type, *star* and *wico* flowers. This major renouncement led to most of the developmental aspect of the question of *PhDEF* role in *Petunia* petal development (question 2) to be put aside since dynamics are lost. However, the scRNA-Seq data I gathered still revealed novel insights into the processes behind cell layer identities in the petal.

Characterization of the protein PhDEF localization in the petal is key for a broader understanding of its function. Therefore, two additional objectives were to develop fluorescence-tagged PhDEF reporter lines in order to assess its localization *in vivo* in wild-type plants, and to perform immunolocalization assays of PhDEF protein in wild-type, *def-151* single mutant, *star* and *wico* petal cross sections to confirm its localization in wild-type and assess putative protein movement between the different petal cell layers. Unfortunately, these approaches were unsuccessful.

Finally, in order to identify the direct targets for PhDEF in a cell-layer-specific manner, chromatin-immunoprecipitation and sequencing (ChIP-Seq) was also performed on wild-type, *def-151* single mutant, *star* and *wico* petals. Since this experiment took place quite late during my project and the analysis was completed only a few weeks ago, only preliminary results will be featured in this manuscript, but they show to be promising.

RESULTS

I – Cell-layer-specific expression of the B-class MADS-box gene *PhDEF* drives petal tube or limb development in petunia flowers

I.1 – Contribution to Chopy *et al*., 2023

The group recently proposed that the B-class gene *PhDEF* cell-layer-specific expression is responsible for petal tube or limb development in *Petunia*. The full soon-to-be-published work is featured afterwards, it has just been accepted for publication in The Plant Cell under (Chopy et al., 2023). This work was mainly accomplished by Mathilde Chopy, former PhD student in the group under the direction of Michiel Vandenbussche, and constitute the early milestones of the project I took over and worked on these last three years.

I contributed to this work by bringing additional insight regarding *Petunia* wild-type, *phdef-151* single-mutant, *star* and *wico* petal histology by providing and characterizing resin-embedded petal cross sections. This work highlighted the presence of mixed cell and cell-layer identities in cell layers still knocked-out for the expression of *PhDEF* and support the hypothesis that non-cell-autonomous effects involving *PhDEF* are necessary for good petal development in *Petunia*. I also performed bioinformatics analysis on bulk RNA-Seq data of wild-type, *phdef-151* single-mutant, *star* and *wico* samples previously gathered by Mathilde. Preliminary differential gene expression analysis data exploration is featured in (Chopy et al., 2023). Further investigations of this dataset are described below but were eventually not included in the manuscript, for reasons explained later.

Cell layer-specific expression of the homeotic MADS-box transcription factor PhDEF contributes to modular petunia petal morphogenesis

Chopy M.[1,2], Cavallini-Speisser Q.[1], Chambrier P.[1], Morel P.[1], Just J.[1], Hugouvieux V.[3], Rodrigues Bento S.[1], Zubieta C.[3], Vandenbussche M.[1,*] and Monniaux M.[1,*]

[1] Laboratoire de Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRAE, 69007 Lyon, France.
[2] Current address: Institute of Plant Sciences, University of Bern, Altenbergrain 21, Bern, CH-3013, Switzerland.
[3] Laboratoire de Physiologie Cellulaire et Végétale, Université Grenoble-Alpes, CNRS, CEA, INRAE, IRIG-DBSCI, 38000 Grenoble, France.
* Authors for correspondence: michiel.vandenbussche@ens-lyon.fr, marie.monniaux@ens-lyon.fr

Short title: Role of cell layers in petunia petal morphogenesis

Material distribution: The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell/pages/General-Instructions) are: Marie Monniaux (marie.monniaux@ens-lyon.fr) and Michiel Vandenbussche (michiel.vandenbussche@ens-lyon.fr).

ABSTRACT

Floral homeotic MADS-box transcription factors ensure the correct morphogenesis of floral organs, which are organized in different cell layers deriving from the meristematic L1, L2 and L3 layers. How cells from these distinct layers acquire their respective identity and coordinate their growth to ensure normal floral organ morphogenesis is unresolved. Here, we study petunia petals that form a limb and tube through congenital fusion, a complex morphology that coevolved with pollinators. We have identified petunia mutants expressing the B-class MADS-box gene *PhDEF* in the epidermis or in the mesophyll of the petal only, called wico and star respectively. Strikingly, wico flowers form a strongly reduced tube while their limbs are almost normal, while star flowers form a normal tube but very reduced and unpigmented limbs, showing that petunia petal morphogenesis is highly modular. Comparative transcriptome analysis of star, wico and wild-type petals revealed a strong down-regulation of the anthocyanin production pathway in star petals including its major regulator *ANTHOCYANIN2 (AN2).* We found that PhDEF directly binds to *AN2* regulatory sequence *in vitro* by gel shift assay, and *in vivo* by chromatin immunoprecipitation, suggesting that PhDEF directly activates the petal pigmentation pathway by activating *AN2*. Altogether, we show that cell-layer specific homeotic activity in petunia petals differently impacts tube and limb development, revealing the relative importance of the different cell layers in the modular architecture of petunia petals.

INTRODUCTION

All plant aerial organs derive from clonally-distinct layers, named L1, L2 and L3 in the shoot apical meristem (SAM) (Satina et al., 1940). Within the L1 and L2 layers, cells divide anticlinally, thereby maintaining a clear layered structure in all aerial organs produced by the SAM (Meyerowitz, 1997; Stewart and Burk, 1970; Scheres, 2001). Already at the embryonic stage, meristematic cell layers express different genes and have distinct identities (Abe et al., 1999; Lu et al., 1996), that are maintained in the adult SAM (Yadav et al., 2014). During flower development, floral organ identity will be appended on top of layer identity by the combinatorial expression of homeotic floral genes, most of which are MADS-box genes (Coen and Meyerowitz, 1991; Schwarz-Sommer et al., 1990). How these master floral regulators specify all floral organ features, such as organ size, shape, pigmentation, and cellular properties, while maintaining layer-specific identities, is unknown.

Petals are often the most conspicuous organs of the flower, and they display a tremendous diversity in size, shape and pigmentation across flowering plants (Moyroud and Glover, 2017). Floral organ identity is specified by a combination of A-, B- and C-class identity genes as proposed by the classical ABC model established on *Arabidopsis thaliana* (Arabidopsis) and *Antirrhinum majus* (snapdragon), and B-class genes are particularly important for petal identity (Coen and Meyerowitz, 1991; Schwarz-Sommer et al., 1990; Morel et al., 2017). B-class proteins, belonging to MADS-box transcription factors, are grouped in the DEF/AP3 and the GLO/PI subfamilies, named after the snapdragon/Arabidopsis B-class proteins DEFICIENS/APETALA3 and GLOBOSA/PISTILLATA (Purugganan et al., 1995; Theißen et al., 1996). These proteins act as obligate heterodimers consisting of one DEF/AP3 and one GLO/PI protein, and this complex activates its own expression for maintenance of high expression levels all along petal and stamen development (Tröbner et al., 1992). In petunia, gene duplication has generated four B-class genes, namely *PhDEF* and *PhTM6* belonging to the *DEF/AP3* subfamily, and *PhGLO1* and *PhGLO2* belonging to the *GLO/PI* subfamily (Vandenbussche et al., 2004; Rijpkema et al., 2006; van der Krol et al., 1993; Angenent et al., 1992). Mutating the two members of each subfamily (*phdef phtm6* or *phglo1 phglo2* double mutants) results in a classical B-function mutant phenotype with homeotic transformation of petals into sepals and stamens into carpels (Vandenbussche et al., 2004; Rijpkema et al., 2006). Additionally, gene copies within the *DEF/AP3* subfamily have diverged in function: while *PhDEF* exhibits a classical B-class expression pattern largely restricted to developing petals and stamens, *PhTM6* is atypically expressed in stamens and carpels, and its

upregulation depends on the petunia C-function genes (Rijpkema et al., 2006; Heijmans et al., 2012a). As a consequence, the single *phdef* mutant displays a homeotic conversion of petals into sepals, while the stamens are normal due to functional redundancy with *PhTM6* (Rijpkema et al., 2006). The petunia *phdef* mutant is therefore an interesting model to study the mechanism of petal identity specification alone since it displays a single-whorl complete homeotic transformation, which is quite rare for floral homeotic mutants that generally show defects in two adjacent whorls.

Flowers from the *Petunia* genus develop five petals, that arise as individual primordia and fuse congenitally (Vandenbussche et al., 2009). Mature petals are fully fused and the corolla is organized in two distinct domains: the tube and the limb. Variation in the relative size of the tube and the limb are observed among wild species of *Petunia*, where flowers with a long tube grant nectar access to long-tongued hawkmoths or hummingbirds, while wide and short tubes are easily accessible to bees (Galliot et al., 2006). The short- and long-tube species cluster separately on a phylogeny of wild *Petunia* species, and the short-tube phenotype is likely the ancestral one (Reck-Kortmann et al., 2014). Pollinator preference assays and field observations have confirmed that tube length and limb size are discriminated by pollinators and thereby might play a role in reproductive isolation, together with multiple other traits of the pollination syndromes such as limb pigmentation or volatile emission (Venail et al., 2010; Hoballah et al., 2007; Galliot et al., 2006). Tube and limb therefore appear to act as different functional modules in the petunia flower.

Although the petunia petal tube and limb seem to play important ecological roles, the mechanisms driving their development are mostly unknown. Tube and limb develop as relatively independent entities in flowers from the Solanaceae family, to which petunia belongs: for instance, tube length and limb width are uncorrelated traits in intra-specific crosses performed in *Nicotiana* or *Jaltomata* (Bissell and Diggle, 2008; Kostyun et al., 2019). Moreover, tube and limb identities can be acquired independently: this is strikingly observed in the petunia *blind* mutant, a partial A-class mutant that forms an almost wild-type tube topped by functional anthers, due to ectopic C-class activity in the second whorl (Cartolano et al., 2007). Apart from the petal identity genes, the molecular players involved in petunia tube or limb morphogenesis are mostly unknown. General growth factors affect petal development as a whole (both tube and limb) together with other vegetative or reproductive traits (Vandenbussche et al., 2009; Terry et al., 2019; Brandoli et al., 2020), but very few genes have been found to specifically affect growth of one subdomain of the petal (Zenoni et al., 2004). Therefore, the mechanisms of petunia tube and limb morphogenesis remain to be fully explored. In contrast, the genetic and molecular bases of petunia petal

pigmentation are extremely well characterized, thanks to the plethora of mutants that have been isolated over decades of breeding and research (Bombarely et al., 2016; Tornielli et al., 2009). Petunia limb pigmentation is mainly due to the accumulation of anthocyanins in the vacuole of adaxial epidermal cells. Briefly, the earliest steps of anthocyanin production are ensured by a MBW regulatory complex composed of an R2R3-MYB transcription factor (either ANTHOCYANIN2 (AN2), AN4, DEEP PURPLE (DPL) or PURPLE HAZE), a bHLH transcription factor (AN1 or JAF13), and a WD-40 repeat protein (AN11), which drives the expression of anthocyanin biosynthesis enzymes and proteins involved in vacuolar acidification of epidermal cells (Albert et al., 2011; de Vetten et al., 1997; Spelt et al., 2000; Quattrocchio et al., 1998, 1999, 1993). How this pathway is activated, after regulators such as PhDEF have specified petal identity, has not been elucidated so far.

In this work, we present petunia flowers with strongly affected tube or limb development, that we respectively named wico and star, and that spontaneously arose from plants mutant for *PhDEF*. We provide genetic and molecular evidence that both of these flower types are periclinal chimeras, resulting from the layer-specific excision of the transposon inserted into the *PhDEF* gene, restoring *PhDEF* activity either in the epidermis or in the mesophyll of the petal. The star and wico phenotypes indicate that in the petunia petal, the epidermis mainly drives limb morphogenesis while the mesophyll mainly drives tube morphogenesis. This is seemingly different from previous studies in snapdragon flowers, another species with fused petals, where *def* periclinal chimeras indicated that epidermal *DEF* expression was making a major contribution to overall petal morphology (Perbal et al., 1996; Vincent et al., 2003; Efremova et al., 2001). We characterized in detail the star and wico petal phenotypes at the tissue and cellular scale, and found evidence for non-cell-autonomous effects affecting cell identity between layers. We sequenced the total petal transcriptome from wild-type (wt), wico and star flowers at three developmental stages, and we found that a large proportion of the genes involved in anthocyanin production were downregulated in star petal samples, as could be expected from their white petals. We further showed, by gel shift assay and chromatin immunoprecipitation, that PhDEF binds to the terminator region of *AN2*, thereby likely directly activating its expression and triggering the first steps of limb pigmentation. Our results and our unique flower material promise to improve our understanding of tube and limb morphogenesis in petunia, and address the broader question of how organ identity and cell layer identity superimpose during organ development.

Results

Spontaneous appearance of two phenotypically distinct classes of partial revertants from the *phdef-151* locus

Previously described null alleles for the *PhDEF* gene (also named *GP* or *pMADS1*) were obtained by either ethyl methanesulfonate (EMS) mutagenesis (de Vlaming et al., 1984; Rijpkema et al., 2006) or by γ-radiation (van der Krol et al., 1993). Because neither of these alleles were straightforward to genotype in a heterozygous state, we screened our sequence-indexed *dTph1* transposon mutant population in the W138 genetic background (Vandenbussche et al., 2008) for new insertions into *PhDEF*. We identified a new mutant allele named *phdef-151*, referring to the *dTph1* insertion position 151 bp downstream of the ATG in the first exon of the *PhDEF* gene, predicted to fully disrupt the MADS-domain in the protein sequence by premature termination of the first exon due to multiple STOP codons in the different reading frames of *dTph1*. As observed for previously identified *phdef* null alleles, *phdef-151* flowers display a complete homeotic conversion of petals into sepals, while heterozygous or homozygous wild-type siblings display red-coloured wt petals. *phdef-151* is thus very likely a null mutant allele.

While growing homozygous *phdef-151* individuals during several seasons, we repeatedly observed the spontaneous appearance of inflorescence side branches that developed flowers with a partial restoration of petal development (Figure 1, Supplemental Figure 1), suggesting excision of the *dTph1* transposon from the *phdef-151* allele specifically in these sidebranches. Remarkably, these partially revertant flowers could be classified as belonging to either one of two contrasting phenotypic classes, that we named star and wico, and that could even occur simultaneously in different branches on the same plant (Fig. 1A). For both phenotypic classes, we obtained more than 15 independent reversion events. The star flowers (Fig. 1D-F), named in reference to their star-shaped petals, grow an elongated tube similar to wt flowers, but their limbs are underdeveloped: they appear to mainly grow around the mid-vein with strongly reduced lateral expansion, hence losing the typical round shape of wt limb. Moreover, they have almost white petals, suggesting strongly reduced accumulation of anthocyanins. We quantified these changes in flower morphology (Fig. 1K-N) and found that total limb area was reduced almost 5-fold in star flowers (Fig. 1M). In contrast, total tube length was only slightly reduced in star as compared to wt (Fig. 1L), and this was mainly due to a reduction in length of domain D1, corresponding to the part of the tube fused with stamens (as defined in (Stuurman et al., 2004), Fig. 1K), while length of the rest of the tube (domain D2) remained unchanged (Fig. 1L, Supplemental Figure 2). As a result, the ratio between

limb area and tube length, which we use as a simple measure for overall corolla morphology, is reduced about 4-fold in star flowers as compared to wt (Fig. 1N). In addition, we occasionally observed fully pigmented secondary revertant sectors of various sizes in the star genetic background, in some cases leading to the development of a single wt-like petal in a star flower background (Fig. 1J). These revertant sectors, observed multiple times, always exhibited simltaneous restauration of pigmentation and normal petal limb growth patterns, demonstrating that the strongly reduced pigmentation in star petals was due to impaired PhDEF function, and not to an additional mutation in the pigmentation pathway.

The wico flowers, named after their wide corolla, grow round-shaped and pigmented limb while their tube remains underdeveloped (Fig. 1G-I). Limb pigmentation ranged from pink to bright red, and green sepaloid tissue was observed around the mid-veins, commonly well visible in all wico flowers on the abaxial side of the petals (see for instance Supplemental Figure 1E). Total tube length was reduced about 3-fold in wico flowers, with domain D1 being absent since stamens were totally unfused to the tube (Supplemental Figure 2), while domain D2 was significantly reduced in size (Fig. 1L). Limb area was also about 2-fold reduced in wico as compared to wt flowers (Fig. 1M), but the ratio between limb area and tube length was higher than in wt flowers (Fig. 1N), indicating the larger contribution of limb tissue to total corolla morphology in wico flowers. In summary, the star flowers form an almost normal tube but small, misshaped and unpigmented limb, while the wico flowers form almost normally shaped and pigmented limb but a tube strongly reduced in length. These contrasting phenotypes suggest that tube and limb development can be uncoupled in petunia flowers, at least to some degree.

The star and wico flowers result from excision of the *dTph1* transposon from the *phdef-151* locus

Reversion of a mutant phenotype towards a partial or a complete wt phenotype is classically observed in unstable transposon insertion mutant alleles. In the petunia W138 line from which *phdef-151* originates, the *dTph1* transposon is actively transposing (Gerats et al., 1990). We assumed therefore that the star and wico flowers were caused by the excision of *dTph1* from the *PhDEF* locus. *dTph1* transposition is generally accompanied by an 8-bp duplication of the target site upon insertion, and excision can have various outcomes depending on the length and nature of the remaining footprint (van Houwelingen et al., 1999). Hence, we first hypothesized that the

distinct star and wico phenotypes were caused by different types of alterations of the *PhDEF* coding sequence after the excision of *dTph1*.

To test this hypothesis, we characterized the *phdef-151* locus from in total 14 star and 14 wico independent reversion events (Figure 2). For this, we extracted genomic DNA from sepals or petals of star and wico flowers, and we amplified the part of the *PhDEF* locus containing the *dTph1* transposon with primers flanking the insertion site (Fig. 2A). All samples produced a mixture of *PhDEF* fragments, some containing the *dTph1* transposon and some where *dTph1* had been excised (Fig. 2B). We specifically sequenced the small fragments resulting from *dTph1* excision in star and wico petal samples, including *phdef-151* second whorl organs as a control (Fig. 2C). In *phdef-151*, the *dTph1*-excised alleles were always out-of-frame, with either 7 or 8 additional nucleotides as compared to the wt sequence. Due to a reading frame shift, both of these alleles are expected to produce an early truncated protein likely not functional (Fig. 2C), in line with the normal *phdef* mutant phenotype observed in these plants. In contrast, in both star and wico flowers we could find either wt sequences (found 1 time and 3 times independently in star and wico flowers respectively) or in-frame footprint alleles consisting of various additions of 6 nucleotides (alleles further named *PhDEF+6*, found 13 times and 11 times independently in star and wico flowers respectively, Fig. 2C). These last insertions are predicted to result in proteins with 2 additional amino-acids inserted towards the end of the DNA-binding MADS domain (Fig. 2C). Together these results demonstrate that wico and star revertant flowers depend on the presence of an in-frame *def-151* derived excision allele that partially restores petal development. However, and in contrast to our initial expectations, there was no correlation between the sequence of the locus after excision and the phenotype of the flower, and both star and wico flowers could be found with a wt *PhDEF* excision allele or with an identical *PhDEF+6* allele (e.g. the 6-bp GTCTGG footprint allele was frequently found both in wico and star flowers). This indicates that the phenotypic difference between the star and wico flowers cannot be explained by a differently modified *PhDEF* sequence after *dTph1* excision. Secondly, since the *phdef* mutation is fully recessive (Vandenbussche et al., 2004), the presence of one transposon mutant allele combined with the wt revertant sequence, normally should lead to wt flowers. Together this implied that another molecular mechanism was causing the difference between wico and star flowers.

The wico flowers are L1 periclinal chimeras

Excision of *dTph1* from a gene can occur at different times during plant development: if happening at the zygotic stage, then the whole plant will have a *dTph1*-excised allele. If excision occurs later, this will result in a genetic mosaic (chimera) with a subset of cells carrying the *dTph1* insertion at the homozygous state and others having a *dTph1*-excised allele. This typically leads to branches or flowers with a wt phenotype on a mutant mother plant (supposing a recessive mutation). Furthermore, since all plant organs are organized in clonally-independent cell layers, excision can happen in one cell layer only, thereby creating a periclinal chimera, *i.e.* a branch or flower where cell layers have different genotypes (Frank and Chitwood, 2016; De Keukeleire et al., 2001).

Analyzing the progeny of wico flowers suggested that they were periclinal chimeras, since the wico phenotype turned out not to be heritable (in consequence, they had to be maintained by cuttings of revertant branches). Instead, we found that the progeny of the wico flowers displayed a *phdef* mutant phenotype at a proportion close to 100%, undistinguishable from the parental *phdef-151* allele (Table 1). This suggested that the gametes generated by the wico flowers exclusively carried the mutant *phdef-151* allele, hence resulting in homozygous *phdef-151* mutants in the progeny. Gametes are exclusively derived from the L2 layer in flowering plants (Tilney-Bassett, 1986), therefore indicating that L2-derived germ cells were homozygous mutant for *phdef-151* in wico flowers, which should result in a *phdef* phenotype if the epidermal tissue had the same genotype. This discrepancy suggested that the L1 layer of wico flowers was probably carrying a functional *PhDEF* allele.

To test this hypothesis, we localized the *PhDEF* transcript in wico flowers by *in situ* hybridization (Figure 3, Supplemental Figure 3). In wt flowers, the *PhDEF* transcript was first detected in the stamen initiation domain, then shortly after in incipient stamen and petal primordia (Fig. 3A, B). At all stages observed, *PhDEF* expression appeared quite homogeneous in all cell layers of the organs, with a stronger expression in the distal part of the petal at later stages of development (Fig. 3C, Supplemental Figure 3). In contrast, in wico flowers *PhDEF* expression was restricted to the L1 and epidermis, all throughout petal development (Fig. 3G-I, Supplemental Figure 3). Therefore, we conclude that wico flowers are the result of an early *dTph1* excision event in one cell from the L1 meristematic layer, resulting in a chimeric flower expressing *PhDEF* only in the epidermis (L1-derived cells) of petals. Wico flowers are therefore L1-periclinal chimeras.

The star flowers are L2 periclinal chimeras

Similarly, we analyzed the progeny of the star flowers, and the star phenotype also turned out not to be heritable, and hence maintained by cuttings of revertant branches. The progeny of the star flowers with a *PhDEF+6* allele yielded three different phenotypic classes (in a proportion close to 1:1:2; Table 1): plants displaying a *phdef* phenotype, plants having wt flowers, and plants carrying flowers with a wild-type architecture but with altered pigmentation, further referred to as « pink wt » (Supplemental Figure 4). We genotyped the *PhDEF* locus in plants descendant from one star parent and carrying flowers with a wt architecture (Supplemental Table 2). We found that all plants with a pink wt phenotype were heterozygous with an out-of-frame *phdef* allele and an in-frame *PhDEF+6* allele, while fully red wt flowers had in-frame *PhDEF+6* alleles at the homozygous state. This indicates that the PhDEF protein with 2 additional amino acids is not 100% fully functional, as it leads to a reduction in limb pigmentation when combined with an out-of-frame allele. The fact that it can ensure normal petal development when at the homozygous state indicates that this is dosage dependent. In summary, the segregation ratio shows that the star gametes carried either the *phdef-151* allele or an in-frame *PhDEF* allele at a 1:1 ratio, and hence that the germ cells generating these gametes were heterozygous for these two alleles. This suggested that in star flowers, the L2 layer was carrying a functional *PhDEF* allele while the L1 layer was homozygous mutant for *phdef-151*.

In support of this, in star flowers *PhDEF* expression was absent from the L1 and epidermis (Fig. 3D-F, Supplemental Figure 3). At the petal margins, underlying layers were also devoid of *PhDEF* expression (Fig. 3F), which likely corresponds to the restricted petal area where cells of L1 origin divide periclinally and invade the mesophyll (Satina and Blakeslee, 1941a). Therefore, we conclude that star flowers are the result of an early *dTph1* excision event in one cell from the L2 meristematic layer, resulting in a chimeric flower expressing *PhDEF* only in the mesophyll (L2-derived cells) of petals. Star flowers are therefore L2-periclinal chimeras. Considering the star and wico phenotypes, we can conclude that the petal epidermis is the main driver for limb morphogenesis (growth, shape and pigmentation), while the mesophyll mainly drives tube morphogenesis (growth and shape).

Non-cell-autonomous effects of layer-specific *PhDEF* expression on cell identity

Having determined the genetic basis of the star and wico phenotypes, we next wondered how layer-specific *PhDEF* expression affects the determination of cell identity, in the layer where *PhDEF* is expressed (cell-autonomous effect) but also in the layer devoid of *PhDEF* expression (non-cell-

283 autonomous effect). For this, we observed petal adaxial epidermal cells by scanning electron
284 microscopy, and mesophyll cells on petal cross-sections, in wt petals and sepals, and in star and
285 wico petals (Figure 4).
286 On the adaxial side of the wt petal, cells from the limb are round and adopt the classical
287 conical shape found in many angiosperm petal limb, while cells from the tube are elongated with a
288 central cone (Fig. 4B) (Cavallini-Speisser et al., 2021). In contrast, the adaxial epidermis of wt
289 sepals (indistinguishable from *phdef-151* second whorl organs) displays typical leaf-like features
290 (Morel et al., 2019), with puzzle-shaped cells interspersed with stomata and trichomes (Fig. 4B).
291 Epidermal cell identity can thus be clearly determined on the basis of cell shape. In wico petals,
292 epidermal limb cells are conical, similar to wt cells from the same area, although marginally bigger
293 (Fig. 4B, D). In contrast, cells from the tube, albeit displaying a similar shape than wt cells, are
294 strongly reduced in length (Fig. 4B, E). This suggests that, in addition to the absence of the D1
295 region of the tube (Fig. 1K, Supplemental Figure 2), a defect in cell elongation in the D2 region is
296 (at least partly) responsible for overall tube length reduction in wico petals. In star petal tubes,
297 epidermal cells have a similar appearance as in a wt petal tube but are slightly less elongated (Fig.
298 4B, E). In contrast, epidermal cells from the star limb are very different to both wt petal conical
299 cells and wt sepal puzzle cells: they are slightly bulging cells, more or less roundish, and about 3-
300 times larger than wt conical cells (Fig. 4D). Together with the observation that the star limb area is
301 around 5-times smaller compared to wt limb (Fig. 1M), this indicates that normal activity of PhDEF
302 in the epidermis is required not only for the typical conical cell shape but also for a high cell
303 division rate during petal development. As mentioned earlier, we occasionally observed pigmented
304 revertant sectors on star flowers, resulting from an additional independent *dTph1* excision in the
305 epidermis, generating wt sectors on a star flower (Fig. 1I). These sectors allow the immediate
306 comparison between star and wt epidermal cells on a single sample, confirming the difference in
307 conical cell size, shape and colour (Supplemental Figure 5). Moreover, the star limb adaxial
308 epidermis occasionally forms trichomes (Supplemental Figure 5), a feature that is normally not
309 observed in the wt limb adaxial epidermis. Altogether, these observations suggest that epidermal
310 cells from star limb have an intermediate identity between petal and sepal cells. Since star petals do
311 not express *PhDEF* in their epidermis, these observations show that non-cell-autonomous effects
312 are at stake to specify cell identity.
313 Mesophyll cell identity was investigated by analyzing petal cross-sections stained with
314 toluidine blue (Fig. 4C). In the wt petal tube, mesophyll cells are big and round and the tissue is

315 loosely arranged. In the limb, mesophyll cells appear smaller and with a more elongated shape,
316 while still very loosely arranged. Sepal mesophyll cells are bigger than petal mesophyll cells, and
317 they display the typical leaf mesophyll organization with an upper palisade layer (elongated and
318 parallel cells) and a lower spongy layer (dispersed cells). Hence mesophyll cell size, shape and
319 tissue-level organization are characteristic features allowing to distinguish between sepal and petal
320 mesophyll tissue. In star petals, we observed no visible difference in mesophyll cell features with wt
321 petals, suggesting that petal cell identity is normally specified in the star petal mesophyll. In wico
322 petals, mesophyll cells also appeared similar to wt: they were round and big in the tube, and slightly
323 more elongated in the limb. Their organization was clearly distinct from the one found in sepals
324 since no palisade layer was observed. However, peeling the epidermis from wico limb revealed that
325 the underlying mesophyll was chloroplastic, similar to a sepal mesophyll and in striking contrast
326 with the white mesophyll of wt petal limb (Fig. 4F). Thus, the *phdef* mutant mesophyll in wico
327 flowers has an intermediate identity between sepal and petal, suggesting again the existence of non-
328 autonomous effects influencing cell identity across layers. The interpretation of these effects is
329 summarized in Supplemental Figure 6.
330 We wondered if the non-cell-autonomous effects that we observed between layers in the star
331 petals were also happening within a single layer. The revertant sectors observed on star flowers
332 showed a very abrupt transition between pigmented and non pigmented epidermal cells, together
333 with a quite sharp transition in conical cell shape and size (Supplemental Figure 5). In particular, we
334 found a clear file of pigmented cells on a star petal and the scanning electron micrograph revealed
335 that these cells were also conical, in stark contrast with the surrounding flat elongated cells of the
336 star petal mid-vein (Supplemental Figure 5). Therefore, we conclude that within the epidermal
337 layer, cell shape and pigmentation are defined cell-autonomously, suggesting that different
338 processes are at stake for cell-cell communication across and within layers.

339 **Transcriptome sequencing of star and wico petals**
340 To better understand the molecular basis for the star and wico phenotypes, we performed RNA-Seq
341 on total petal tissue at three developmental stages, including wt and *phdef-151* samples (Figure 5).
342 We chose an early stage (stage 4 as defined in (Reale et al., 2002)), an intermediate stage (stage 8)
343 when tube length is at half its final size, and a late stage (stage 12) before limb is fully expanded
344 (Fig. 5A). For *phdef-151* we only sequenced second-whorl sepal tissue at stage 12 (before anthesis).
345 Principal component analysis showed that developmental stage is the first contributor to variation in

346 gene expression, while genotype corresponds to the second axis of variation (Fig. 5B). All
347 genotypes clustered separately except wico and wt samples which were highly similar at the two
348 later stages. We analyzed one-to-one differential gene expression between mutant and wt samples
349 with DESeq2 (Love et al., 2014) and we found on average 5,818 deregulated genes in *phdef-151*, as
350 compared to 1,854 and 1,115 deregulated genes in star and wico respectively, when averaging for
351 all stages (Fig. 5C, Supplemental Dataset 1). There were generally more upregulated genes than
352 downregulated ones in mutant or chimeric genotypes, and the number of deregulated genes
353 increased with ageing of the petal in both star and wico (Fig. 5C). At stage 12, a large proportion of
354 genes (58-61%) deregulated in wico or star petals were also deregulated in *phdef-151* (Fig. 5D), as
355 expected since wico and star flowers are mutant for *PhDEF* in one cell layer. Genes uniquely
356 deregulated in star or wico flowers represented 36% of deregulated genes for each, and only 16-
357 29% of deregulated genes were jointly deregulated in star and wico flowers, consistent with the
358 very different phenotypes of these flowers. These proportions indicate that the star and wico
359 phenotypes are mostly subtended by the deregulation of sets of genes also deregulated in *phdef-151*,
360 together with the deregulation of a unique set of genes for each genotype.
361 In star and wico petals, we found that *PhDEF* was down-regulated about two-fold at all
362 stages (Supplemental Figure 7), as expected since *PhDEF* is expressed in one cell layer only. In
363 contrast, *PhTM6* was not deregulated in star and wico nor in *phdef-151* (Supplemental Figure 7), as
364 expected since this atypical B-class gene is mostly expressed in stamens and carpels and its
365 upregulation depends on the C-function genes (Rijpkema et al., 2006; Heijmans et al., 2012b).
366 Unexpectedly, we observed that the B-class genes *PhGLO1* and *PhGLO2* were not down-regulated
367 in wico petals, and only modestly in star petals, although their expression was almost null in the
368 *phdef-151* mutant (Supplemental Figure 7). The fact that *PhGLO1/PhGLO2* expression does not
369 strictly mirror the expression of *PhDEF* in star and wico petals, which is what we would have
370 expected since the B-class heterodimers are known to activate their own expression, suggests that
371 *PhGLO1/PhGLO2* expression is not entirely dependent on the B-class heterodimeric complexes, in
372 particular in the epidermal layer of the petal.

373 **PhDEF directly binds *in vivo* to the terminator region of *AN2*, encoding a major regulator of**
374 **petal pigmentation**
375 The transcriptomes of star and wico petals constitute a promising dataset to identify genes involved
376 in the establishment of petal epidermis and mesophyll identities, and in tube and limb development.

To evaluate their potential to decipher the gene regulatory networks underlying petal development, we decided to focus our attention on genes involved in petal pigmentation. Indeed, the players and regulatory pathways involved in anthocyanin biosynthesis in the petal epidermis have been extremely well described but their relationship with the specifiers of petal identity, to whom PhDEF belongs, is so far unknown. The absence of pigmentation in star petals, the restoration of pigmentation in L1-revertant sectors and the phenotype of the pink wt flowers prompted us to investigate the direct link between *PhDEF* expression and petal pigmentation. For this, we examined the 451 genes down-regulated in both *phdef-151* and star samples (at any stage) but not deregulated in wico samples (Supplemental Dataset 2), and we found 23 anthocyanin-related genes in this gene set (Supplemental Figure 7), out of a total of 42 in the whole genome, which constitutes an exceptionally high enrichment for this gene function ($p < 0.001$, Fisher's exact test). We payed a particular attention at the genes possibly involved in the first steps of anthocyanin production, ie encoding proteins involved in the MBW complexes activating anthocyanin biosynthesis (AN1, AN2, AN4, AN11, JAF13, DPL and PURPLE HAZE). We found that *AN1*, *AN2*, *DPL* and *JAF13* were downregulated both in *phdef-151* and star samples (Supplemental Figure 7, Supplemental Dataset 2). DPL is involved in the limb venation pattern (Albert et al., 2011; Zhang et al., 2021) and JAF13 has only a moderate contribution to limb pigmentation (Bombarely et al., 2016), therefore we decided to focus our attention on the two major activators of anthocyanin biosynthesis AN1 and AN2 (Figure 6). Indeed, the *an1* mutant has fully white petals and the *an2* mutant has strongly reduced limb pigmentation (Quattrocchio et al., 1999; Spelt et al., 2000). Furthermore, *AN2* was shown to act as an upstream activator of *AN1* since overexpressing *AN2* in petunia leaves is sufficient to activate *AN1* expression, and for anthocyanins to accumulate (Quattrocchio et al., 1998; Spelt et al., 2000). We observed that both genes were already expressed at stage 4 of wt petal development, before any pigmentation is visible, and their expression levels strongly increased from stage 4 to stage 12, while both being strongly downregulated in star petals and *phdef-151* second whorl organs, but not in wico flowers (Fig. 6A, B). *AN2* was expressed at higher levels than *AN1* at all stages, consistent with its most upstream role in the anthocyanin pigmentation pathway

We aimed to test if PhDEF could be a direct activator of *AN1* or *AN2* expression. For this, we first attempted to predict PhDEF binding on the genomic sequences of *AN1* and *AN2*. We used the high-quality transcription factor (TF) binding profile database Jaspar (Fornes et al., 2020; Sandelin et al., 2004), using position weight matrices for each TF to compute relative binding scores that reflect *in vitro* binding preferences (Stormo, 2013). The exact DNA-binding specificity

of PhDEF has not been characterized, but the one of its Arabidopsis homologs AP3 and PI has (Riechmann et al., 1996b). However, since PhDEF DNA-binding specificity might be slightly different to those of AP3 and PI, we decided to predict binding for all MADS-box TFs available in Jaspar 2020, accounting for 23 binding profiles including those of AP3 and PI (Fornes et al., 2020). We hypothesized that sequences predicted to be bound by several MADS-box TFs might be high-confidence CArG boxes (the binding site for MADS-box proteins). As a validation of this strategy, we analyzed the genomic sequence of *PhDEF* and found a high-confidence CArG box in the *PhDEF* promoter (visible by the presence of good predicted binding sites for several MADS-box proteins and therefore appearing as a clear black line in Fig. 6C). This CArG box is extremely conserved between distantly-related flowering plants (Rijpkema et al., 2006) and it was shown to be important for *AP3* petal-specific expression and for its auto-activation in Arabidopsis (Hill et al., 1998; Wuest et al., 2012), and for DEF function and binding to its own promoter in Anthirrhinum (Schwarz-Sommer et al., 1992). We next applied this predictive approach to the genomic sequences of *AN1* and *AN2*. For *AN1*, we predicted a high-confidence CarG box (*AN1-bs1*) with a very high score for several MADS-box proteins and for AP3 and PI in particular, in the terminator region (Fig. 6D). For *AN2*, we also predicted one promising candidate binding site (*AN2-bs3*), again in the terminator region of the gene (Fig. 6E), although its binding score was more modest in comparison to *AN1-bs1*.

To determine if PhDEF could indeed bind to *AN1-bs1* and *AN2-bs3* and potentially regulate *AN1* and *AN2* expression, we performed gel shift assays using *in vitro* translated PhDEF and/or PhGLO1 proteins (Fig. 6F). We found that, when incubating a 60-bp fragment containing *AN1-bs*1 in its center with either PhDEF or PhGLO1, no shift in migration was visible, indicating that neither protein could bind to this site alone. However, when incubating *AN1-bs1* with both PhDEF and PhGLO1 proteins, we observed a clear shift in migration, consistent with the obligate heterodimerization of these proteins necessary for DNA binding (Riechmann et al., 1996a). Similarly, a 60-bp fragment containing *AN2-bs3* in its center, and incubated with PhDEF and PhGLO1 proteins, resulted in a clear shift in migration. In contrast, a control 60-bp fragment named *AN1-bs2*, located in the *AN1* terminator region but predicted to have a very low binding score (relative score under 0.8 both for AP3 and PI), was not bound by the PhDEF + PhGLO1 protein complex, showing that our assay was specific. Therefore PhDEF, when dimerized with PhGLO1, is able to bind to sites in putative regulatory regions in *AN1* and *AN2*, suggesting that it might directly regulate the expression of these two genes.

Next, we tested if PhDEF could bind *in vivo* to genomic regions containing *AN1-bs1* and *AN2-bs3* by chromatin immunoprecipitation (ChIP). We produced recombinant PhDEF protein devoid of its highly conserved MADS domain, to avoid cross-reactivity with other MADS-box proteins, and generated a polyclonal antibody against this truncated PhDEF protein. We performed the ChIP assay on second whorl organs (petal or sepal) from wt, *phdef-151* or *phglo1 phglo2* plants at an intermediate stage of development (stage 8). In wt petal samples, we found a significant binding enrichment for some of the genomic fragments (GF) that we tested, and in particular *PhDEF*^{GF1} (Fig. 6G), containing the high-confidence CArG box previously described (Fig. 6C), which is expected since PhDEF activates its own expression. We also observed a significant binding enrichment in *AN2*^{GF3} (Fig. 6G), containing the previously identified *AN2-bs3* binding site (Fig. 6E). In contrast, no strong enrichment was detected in any of the *AN1* genomic fragments, even the one containing the *AN1-bs1* strong *in vitro* binding site for PhDEF (AN1^{GF3}). Our ChIP assay was specific, since no enrichment was detected for the *phdef-151* mutant, nor for the *phglo1 phglo2* mutant (Fig.6G). The *phglo1 phglo2* samples constitute an indirect control for PhDEF binding, since the PhDEF protein partners PhGLO1/PhGLO2 are absent, thereby indirectly preventing PhDEF binding on DNA. The fact that we do not detect any binding enrichment in these plants shows that our ChIP assay is robust. Therefore, we conclude that PhDEF binds to the terminator region of *AN2 in planta*.

Discussion

In this work, we identified periclinal chimeras expressing the B-class MADS-box gene *PhDEF* in different cell layers of the flower. This layer-specific expression resulted in the correct development of sub-domains of the petal only, showing that epidermal *PhDEF* expression mainly drives limb morphogenesis while its expression in the mesophyll is more important for tube morphogenesis. This indicates that cell layer-specific actions of PhDEF are different and contribute in a complementary fashion to overall petal development.

Contribution of cell layers to mature petunia petals

The SAM of all flowering plants is organized in three independent layers. Generally, it is assumed that L1-derived cells form the epidermis, L2-derived cells produce the mesophyll and sub-epidermal tissue, and L3-derived cells generate the ground tissues (inner mesophyll, vasculature, pith of the stem). However, there is variation to this general pattern between organs; for instance Arabidopsis sepals, stamens and carpels derive from these three layers, while petals derive from the L1 and L2 layers only (Jenik and Irish, 2000). Moreover, the contribution of cell layers can vary between the same organ in different species: for instance Datura petals are derived from all three layers, in contrast to petals from Arabidopsis (Satina and Blakeslee, 1941b). Finally, even in one organ from a single species, cell layer contribution is not always homogeneous in different parts of the organ: in Datura petals, the L3 only participates to the vasculature at the base of the organ but does not contribute to the distal part of the petal, and the L1 invades the mesophyll at the petal edges (Satina and Blakeslee, 1941b).

In fact, the contribution of cell layers to mature organ organization can only be strictly assessed by clonal analysis, where one follows cell lineage using trackable cell-autonomous markers. In petunia, no clonal analysis has been performed so far, hence one can only assume which cell layers participate to petal development based on clonal analyses performed in closely-related species. In Datura, member of the Solanaceae family like petunia, periclinal chimeras induced by colchicine treatment and refined histological observations have provided a detailed clonal analysis for cell layers in floral organs (Satina and Blakeslee, 1941b). The first visible event of petal initiation is a periclinal cell division from the L2 layer, and further growth of the petal depends primarily on cell divisions from the L2, both anticlinal and periclinal. The L3 layer only contributes to the vascular tissue at the very base of the petal. L1-derived cells form the epidermis by anticlinal divisions, except at the petal edges where periclinal divisions are observed, leading to L1-derived

cells invading the mesophyll. Hence, the Datura petal is formed by all 3 layers with a major contribution of the L1 and L2 layers, and a relative enrichment in L1-derived cells (by thinning of the mesophyll) as we progress from the base towards the tip of the petal. In this work, we hypothesized that the petunia petal is formed similarly. Consistently, we only obtained two phenotypic classes of periclinal chimeras, star and wico, suggesting that L3-specific *PhDEF* expression probably only leads to a *phdef* mutant phenotype.

The contribution of L1- and L2-derived tissues is heterogeneous in the petunia petal. Indeed, cross-sections in the middle of the petal tube indicate that the mesophyll is thick, with several layers of cells (Fig. 4C). The mesophyll tissue is quite dense in this part of the tube, with lacunae between cells being relatively small. In contrast in the limb, mesophyll cells are very small and interspersed with large lacunae. There is also a general thinning of the mesophyll as we progress from the base of the petal towards its edges, whereas the epidermis always appears as a single layer of tightly connected cells. Therefore, it is rather logical that in the petal limb, whose mesophyll is extremely reduced, morphogenesis is driven by the epidermal layer. However, one could not have easily guessed that tube morphogenesis would be mostly driven by the petal mesophyll.

Different cell layers drive tube and limb morphogenesis

The star and wico phenotypes revealed that in petunia petals, the epidermis is the main driver for limb morphogenesis while the mesophyll is the main driver for tube morphogenesis. The epidermis has been proposed to be the layer in control of organ morphogenesis, since it is a layer under tension that restricts growth of the underlying inner tissues that tend to expand (Kutschera and Niklas, 2007). In particular, epidermal expression of the brassinosteroid receptor BRI1 is sufficient to restore normal leaf morphogenesis in a *bri1* mutant (Savaldi-Goldstein et al., 2007). Similarly, the expression of the auxin transporter PIN1 in the L1 of the SAM is sufficient to restore normal phyllotaxis in a *pin1* mutant (Kierzkowski et al., 2013). However, pieces of evidence suggest that organ inner layers can have an active role in morphogenesis: for instance, mesophyll-specific expression of *ANGUSTIFOLIA* (*AN*) is sufficient to restore normal leaf width in the Arabidopsis *an* mutant (Bai et al., 2010); leaf shape is controlled by the L2- and L3-derived tissues in tobacco (McHale and Marcotrigiano, 1998); and the leaf mesophyll is the main player for leaf flatness in Arabidopsis (Zhao et al., 2020). Moreover, expressing *BRI1* in the root phloem also restaures *bri1* plant dwarfism (Graeff et al., 2020). The contribution of cell layers to organ morphogenesis is thus a complex process that varies between organs, species and the genetic systems investigated.

Our work has confirmed that the petunia petal has a modular structure, since tube and limb can develop relatively independently from each other in the star and wico flowers. This modularity is consistent with previous observations in the literature (described in the Introduction), and in line with the different ecological roles of the tube and the limb for the interaction with pollinators. Our results highlight that a homeotic factor, PhDEF, can participate to the establishment of this modular structure. Indeed, although PhDEF is normally present in all cell layers of the wild-type petal, its action in the different cell layers is mainly responsible for tube or limb development. This provides a possible mechanism, at the tissue level, for the establishment of the modular structure of petunia petals by homeotic genes. It also participates to the understanding of how homeotic genes can specify at the same time the overall identity of an organ and the coordinated development of its different functional modules.

One may wonder if our findings apply outside of petunia flowers. In snapdragon and Arabidopsis flowers, periclinal chimeras for orthologs of *PhDEF* (*DEF* and *AP3* respectively) or *PhGLO1/PhGLO2* (*GLO* and *PI* respectively) have been previously obtained (Perbal et al., 1996; Vincent et al., 2003; Efremova et al., 2001; Bouhidel and Irish, 1996; Jenik and Irish, 2001; Urbanus et al., 2010b). In snapdragon, expression of *DEF* only in the L1 layer largely restores petal development, particularly in the limb, in contrast to the L2/L3 specific *DEF* or *GLO* expression which causes reduced limb growth (Perbal et al., 1996; Vincent et al., 2003; Efremova et al., 2001). Petals are fused into a tube in snapdragon flowers, but the tube is much more reduced than in petunia, hence conclusions on tube length restoration in the chimeras were not drawn by the authors. However, in light of our results, it is clear that snapdragon chimeras expressing *DEF* or *GLO* in the L2/L3 layers restore tube development to a higher degree than limb development, similar to what we observed. In Arabidopsis that has simple and unfused petals, petal shape and size were never fully restored when *AP3* was expressed in one cell layer only (Jenik and Irish, 2001; Urbanus et al., 2010b); in contrast epidermal expression of *PI* was sufficient to restore normal petal development (Bouhidel and Irish, 1996). Therefore, it seems that the contribution of different cell layers to petal development varies across species and depending on the petal identity gene under investigation.

Autonomous and non-autonomous effects of *PhDEF* expression on petal traits

Our study revealed that petal traits were affected differently by layer-specific *PhDEF* expression (Fig. S6). For instance, epidermal pigmentation is a clearly autonomous trait, since star petals are

not pigmented except when wt revertant sectors arise. On the contrary, epidermal cell shape appears to behave as a partially autonomous trait since star epidermal cells have an intermediate phenotype between wt petal conical cells and sepal epidermal cells. Finally, organ size and shape are specified non-autonomously in sub-domains of the petal: *PhDEF* expression in the L1 or L2 is sufficient to specify correct shape of the limb or correct size and shape of the tube respectively, suggesting that in these petal domains, layer-specific *PhDEF* expression is sufficient to signal cells from the other layer to grow normally. The mechanisms for this inter-layer communication remain unknown. Our attempts to detect the PhDEF protein in petal tissue by immuno-histochemistry have been unsuccessful, therefore we do not know if the PhDEF protein itself might be moving between layers, which would be the simplest mechanistic explanation for the non-autonomous traits that we observe. Indeed, in Antirrhinum petals expressing *DEF* in the L2/L3 layers, the DEF protein was found in the epidermis and it is likely why petals from these chimeras are pigmented (Perbal et al., 1996), hence suggesting that no such movement occurs in the star petals that are mostly white. In contrast, Arabidopsis AP3 and PI GFP-fusion proteins are unable to move between cell layers, although they can move within the epidermal layer (Urbanus et al., 2010a, 2010b). In any case, even if the PhDEF protein would move between layers in our chimeric flowers, it is likely to be in small amounts only, otherwise both flower types would have a wt phenotype. Therefore, it is unlikely to be the reason for tube and limb correct development in the star and wico flowers. Alternatively, the non-autonomous effects that we observed might be triggered by mechanical signals transmitted between layers. For instance, in star flowers normal growth of the mesophyll could merely drag along epidermal cells, since cells are connected by their cell walls, which could be sufficient to trigger their expansion and division. Other features, like conical cell shape, might be directly influenced by mechanical signals. Indeed, conical cells are shaped by a circumferential microtubule arrangement controlled by the microtubule-severing protein KATANIN, and altering this arrangement affects conical cell shape (Ren et al., 2017). Microtubule arrangement responds to mechanical signals (Hamant et al., 2008), which are likely to be transmitted between layers. Therefore, it is possible that the formation of bulging cells in the star epidermis is merely triggered by mechanical signals from the growing underlying layer, independent of any petal identity specifier, as was recently evidenced from the observation of conical-like bulges on the hypocotyl of the tubulin kinase mutant *nek6* (Takatani et al., 2020). The molecular or physical nature of the signals involved in communication between layers remains to be explored in full depth.

Towards the gene regulatory networks of petal development

Our star and wico material granted the opportunity to explore the gene regulatory networks driving petal development in petunia, more specifically by decoupling tube vs. limb development on one hand, and epidermis vs. mesophyll development on the other hand. However, these effects are confounded in our dataset, since both epidermis and limb development are affected in star flowers, whereas both mesophyll and tube development are affected in wico flowers. Further analyses, like for instance sequencing the transcriptome from star and wico limb and tube tissues separately, would help uncouple these effects, but it is not obvious to clearly separate these different domains during early stages of development, which are crucial stages for petal morphogenesis. Spatial transcriptomics techniques would be ideal to precisely dissect transcriptional changes between layers and domains of the petal at young developmental stages. Still, we exploited our transcriptomic dataset by focusing our analysis on anthocyanin-related genes, because the molecular link between the early establishment of petal identity by homeotic transcription factors, such as PhDEF, and the late establishment of petal maturation traits, such as anthocyanin accumulation, was unknown. For this, we examined the presence of anthocyanin-related genes among genes downregulated both in star and *phdef-151* samples, but not deregulated in wico samples. We found a very strong enrichment of anthocyanin-related genes in this dataset, suggesting that the initial triggering event for most of the anthocyanin production pathway was missing in star flowers.

Finally, we investigated the direct link between PhDEF and petal pigmentation and found that, *in vitro*, the PhDEF + PhGLO1 protein complex directly binds to good predicted binding sites in the regulatory regions of *AN1* and *AN2*. We confirmed that PhDEF binds to the genomic region of *AN2 in plant*a by ChIP, but binding to *AN1* was not observed, confirming that *in vitro* binding does not necessarily imply *in vivo* binding, the last being strongly influenced by the local chromatin landscape. The binding site of PhDEF on *AN2* lies in the terminator region of the gene (and the next gene on the chromosome is more than 100 kb away), which although unusual, is not incompatible with an activating role in transcription, through DNA looping to the promoter (Jash et al., 2012) or by promoting transcription termination and reinitiation (Wang et al., 2000). Together with the fact that *AN2* expression is strongly down-regulated in the *phdef-151* transcriptome, our data indicates that PhDEF directly activates *AN2* expression in the petal. Ectopic expression of *AN2* in petunia leaves is sufficient to trigger anthocyanin accumulation in this tissue, by inducing *AN1* expression among others (Spelt et al., 2000; Quattrocchio et al., 1998). Therefore, the fact that PhDEF activates *AN2* expression should be sufficient to launch the whole pigmentation pathway in the wt petal limb.

615 A direct link between petal identity and pigmentation has never been evidenced before, although
616 genetic evidence in orchid flowers strongly implied that different B-class proteins heteromeric
617 complexes are responsible for specific pigmentation spots in the different petal types, but physical
618 binding of these B-class protein complexes on pigmentation genes was not tested (Hsu et al., 2021).
619 The direct target genes of B-class proteins have been identified by ChIP-Seq and transcriptomic
620 analyses in Arabidopsis (Wuest et al., 2012), but this species has unpigmented petals, thereby
621 preventing to draw any possible link between petal identity and pigmentation. Therefore, to our
622 knowledge, our results show for the first time the direct activation of a petal pigmentation regulator
623 by a petal homeotic gene, which contributes to fill the « missing link » between the identity of a
624 floral organ and its final appearance (Dornelas et al., 2011).

METHODS

Plant material, growth conditions and plant phenotyping

The *phdef-151* plants were obtained from the *Petunia x hybrida* W138 line and were grown in a culture room in long day conditions (16h light 22°C; 8h night 18°C; 75-WValoya NS12 LED bars; light intensity: 130 µE). The wico and star flowers were repeatedly obtained from several different *phdef-151* individuals and were maintained by cuttings. Plant and flower pictures were obtained with a CANON EOS 450D camera equipped with objectives SIGMA 18-50mm or SIGMA 50mm. To measure tube length, the flower was cut longitudinally and photographed from the side. To measure limb area, the limbs were flattened as much as possible on a glass slide covered with transparent tape and photographed from the top. The photographs were used to measure D1 and D2 lengths and limb area with ImageJ.

Genotyping

Extraction of genomic DNA from young leaf tissue was performed according to Edwards et al., 1991. The region spanning the *dTph1* insertion site in *PhDEF* was amplified using primers MLY0935/MLY0936 (Table S1). PCR products were separated on a 2% agarose gel, fragments of interest were purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel), and sequenced with Eurofins SupremeRun reactions.

In situ RNA Hybridization

Floral buds from wt, 2 wico and 1 star lines were fixated overnight in FAA (3.7% formaldehyde, 5% acetic acid, 50% ethanol), cleared in Histo-clear and embedded in paraffin to perform 8 µm sections. *PhDEF* cDNA was amplified from wt petunia inflorescence cDNAs with primers MLY1738/MLY1739 (Supplementary Table 1), generating a 507 bp fragment excluding the part encoding the highly conserved DNA-binding domain. The digoxigenin-labeled RNA probe was synthesized from the PCR fragment by *in vitro* transcription, using T7 RNA polymerase (Boehringer Mannheim). RNA transcripts were hydrolyzed partially for 42 min by incubation at 60°C in 0.1 M Na_2CO_3/$NaHCO_3$ buffer, pH 10.2. Later steps were performed as described by (Cañas et al., 1994). For imaging, slides were mounted in Entellan (Sigma) and imaged with a Zeiss Axio Imager M2 light microscope equipped with a Zeiss Axio Cam HRc camera.

Petal cross-sections

Small pieces (around 5 mm^2) of tissue were harvested from the proximal and distal parts of wt mature sepals, and from the tube and limbs of wt, star and wico mature petals. Samples were fixated overnight in FAA (3.7% formaldehyde, 5% acetic acid, 50% ethanol) and dehydrated in an ethanol

series. Preinfiltration was performed in a 1:1 mixture of ethanol:Technovit 7100 (Electron Microscopy Sciences) for 4 h under light agitation, then overnight in a 1:3 ethanol:Technovit 7100 mixture. Infiltration was performed in the infiltration solution for 1.5 h under vacuum, then for one night followed by one additional week. Samples were disposed in the moulds with the polymerization solution for 2 h at room temperature, then mounted with the Technovit 3040 resin to relieve the blocks from the moulds. Blocks were sectioned with a microtome to generate 3-7 µm-thick sections. Slides were incubated for 10 minutes in a 0.1% toluidine blue solution and imaged with a Zeiss Axio Imager M2 light microscope equipped with a Zeiss Axio Cam HRc camera.

Scanning Electron Microscopy (SEM)

Scanning electron micrographs were obtained with a HIROX SH-1500 bench top environmental scanning electron microscope equipped with a cooling stage. Samples were collected and quickly imaged to limit dehydration, at -5°C and 5 kV settings. For cell area and length measurements, pictures were taken from 3 petal tubes and 3 petal limbs from different wt, star and wico flowers. For each sample, 3 pictures were taken and 5 cells (for the tube) or 10 cells (for the limb) were measured for each picture. Measures were performed with ImageJ by manually drawing the outline or length of the cells.

RNA-Seq

Petal tissue was collected at 1 pm from several plants stemming from a single star line, a single wico line, and several individual wt plants (progeny of a single star flower) and *phdef-151* plants (progeny of the same star flower). Tube length was macroscopically measured to compare stages, the corolla was cut open and stamens were removed as much as possible from the corolla by pulling on the filaments fused to the tube. One biological replicate contains total petal tissue from 2 flowers. Tissue was grounded in liquid nitrogen and RNA was extracted with the Spectrum Plant Total RNA Kit (Sigma) including on-column DNase digestion (Sigma). RNA integrity and quantity were determined by a Bioanalyzer RNA 6000 Nano assay (Agilent). Libraries were prepared with poly-A enrichment and single-end 75-bp sequencing was performed on a NextSeq 500 platform (Illumina). 16 to 23 million reads were recovered per library. Reads were checked for quality with FastQC v0.11.4 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), adaptors and low-quality ends were trimmed with Cutadapt v 1.16 (Martin, 2011) and custom Perl scripts. The reference genome sequence used for transcriptome analysis is the *Petunia axillaris* v1.6.2 HiC genome published in (Bombarely et al., 2016) and further scaffolded by HiC by DNAzoo (Dudchenko et al., 2017, 2018); gene annotations were transferred from the published assembly to

689 the HiC-scaffolded version using Blat (Kent, 2002), Exonerate (Slater and Birney, 2005) and
690 custom Perl scripts. In the rare cases when gene annotations from the published genome mapped to
691 several regions in the HiC-scaffolded genome, these different putative genes were identified by a
692 letter added at the end of the gene identifier (for instance Peaxi162Scf00179g00121a). The
693 complete set of reads was mapped on the reference genome sequence using HISAT2 v2.2.1 (Kim et
694 al., 2015) to identify splicing sites, before performing mapping sample per sample. Reads per gene
695 were counted using FeatureCounts v1.5.1 (Liao et al., 2014). DESeq2 version 3.12 (Love et al.,
696 2014) was used with R version 4.0.3 to perform the Principal Component Analysis and the
697 differential gene expression analysis. Genes having less than 10 reads in the sum of all samples
698 were considered as non-expressed and discarded. Genes were considered to be deregulated if
699 log2FoldChange > 1 or < -1, and p-adjusted value < 0.01. Venn diagrams were built with
700 InteractiVenn (Heberle et al., 2015). Due to the automatic gene name annotation pipeline used in
701 (Bombarely et al., 2016) based on homology with tomato proteins, many of the previously
702 characterized petunia genes have not been annotated according to their first described name, making
703 interpretation of some of the RNA-Seq results less straightforward. We have manually added
704 annotations of 42 genes from the anthocyanin biosynthesis pathway based on the Supplementary
705 Note 7 from (Bombarely et al., 2016), and 31 type-II MIKC-C MADS-box genes based on previous
706 studies from the literature ; these annotations can be found in the Supplemental Datatset 1 of this
707 manuscript. We noticed that the gene annotations from three major pigmentation genes, *DFR*
708 (Peaxi162Scf00366g00630), *CHSa* (Peaxi162Scf00047g01225) and *PH1*
709 (Peaxi162Scf00569g00024) were lost during the gene annotation transfer procedure, because they
710 lie in regions of the genome that are still badly resolved. Therefore, we manually searched the
711 position of these transcripts in the HISAT2 output and we were able to map part of the *DFR* and
712 *CHSa* genes to two small scaffolds, while *PH1* position was not found. We added the transcript
713 positions of *DFR* and *CHSa* in the gtf/gff files before running FeatureCounts. The read counts for
714 *DFR* and *CHSa* reported in Supplemental Figure 7 are therefore an under-estimation of their actual
715 expression levels, since we miss part of the genes.

716 **Prediction of MADS-box TF binding sites**

717 Genomic sequences from *AN1*, *AN2* and *PhDEF* from the *Petunia x hybrid*a R27 line, starting 3 kb
718 upstream the START codon and ending 1 kb downstream the STOP codon, were scanned with all
719 MADS-box TF matrices included in the Jaspar 2020 database (http://jaspar.genereg.net), only
720 removing matrices from AGL42 and AGL55 which are much shorter than the other matrices and

721 therefore yield much higher scores. Relative scores above 0.86 were plotted against their genomic
722 position.
723 **Electrophoretic Gel Shift Assays (EMSAs)**
724 CDS sequences from *PhDEF* and *PhGLO1* were amplified from *Petunia x hybrida* R27
725 inflorescence cDNAs with primers MLY2382/MLY2383 and MLY2384/2385 respectively
726 (Supplementary Table 1) and cloned into the *in vitro* translation vector pSPUTK (Stratagene) by
727 NcoI/XbaI restriction. From these vectors, the PhDEF and PhGLO1 proteins were produced with
728 the TnT SP6 High-Yield Wheat Germ Protein Expression System (Promega) according to the
729 manufacturer's instructions. The terminator regions from *AN1* (0.8 kb) and *AN2* (1kb) were
730 amplified from *Petunia x hybrida* R27 genomic DNA with primers from Table S1 and cloned into
731 pCR-BluntII-TOPO (ThermoFisher). Binding sites were amplified from these plasmids with
732 primers listed in Supplementary Table 1, with the forward primer labelled with Cy5 in 5'. The
733 labelled DNA was purified and incubated with the TnT *in vitro* translation mixture as described in
734 (Silva et al., 2015) before loading on a native acrylamide gel.
735 **PhDEF protein and antibody production**
736 The *PhDEF* truncated cDNA (without the sequence coding for the MADS domain) was chemically-
737 synthesized with optimization for expression in *E. coli* and cloned into a pT7 expression vector by
738 Proteogenix (www.proteogenix.science). The expected PhDEF protein starts at aminoacid 60
739 (PSITT...) and ends at the last aminoacid of the sequence (...FALLE), and a 6xHis tag was added at
740 the N-terminal part of the protein. The 6xHis-PhDEF protein was purified by affinity column with a
741 Nickel resin under denaturing conditions (8M urea) by Proteogenix. The purified protein was
742 injected in two rabbits for immunization by Proteogenix, to generate PhDEF-directed polyclonal
743 antibodies, that were purified by affinity against the antigen. Both lots of purified antibodies were
744 validated by Western-Blot in petal or sepal tissues from wt, *phdef-151* and *phtm6* samples.
745 **Chromatin Immunoprecipitation (ChIP)**
746 Samples (wt: full corolla from 2 flowers; *phdef-151*: second whorl sepals from 3 flowers; *phglo1*
747 *phglo2*: second whorl sepals from 3/4 flowers) at stage 8 were collected and ground in liquid
748 nitrogen. Ground tissue was resuspended into 10 mL fixation buffer (10 mM Hepes pH7.6, 0.5 M
749 sucrose, 5 mM KCl, 5 mM $MgCl_2$, 5 mM EDTA pH8, Complete Protease Inhibitor Cocktail
750 (Merck), 14 mM 2-mercaptoethanol) and a double cross-linking was performed at room
751 temperature (1 hour with disuccinimidyl glutarate at 2.5 mM with gentle shaking, and 5 minutes
752 with formaldehyde 1%). Cross-linking was stopped by adding glycine at 200 mM and samples were

put directly on ice. Cells were lysed with a 40 mL-Dounce tissue grinder (Duran Wheaton Kimble), Triton X-100 was added at 0.6% and the lysate was filtered subsequently through 100 µm and 40 µm nylon meshes to recover nuclei. Nuclei were pelleted for 10 minutes at 3,000 g at 4°C, and the pellet was resuspended in 300 µL of cold nuclear isolation buffer (i.e. fixation buffer without 2-mercaptoethanol), carefully deposited on 600 µL of a 15% Percoll solution (15 % Percoll, 10 mM Hepes pH8, 0.5 M sucrose, 5 mM KCl, 5 mM $MgCl_2$, 5 mM EDTA pH8) and centrifuged for 5 minutes at 2,000 g at 4°C. The pellet was resuspended into 900 µL of cold nuclear lysis buffer (50 mM Tris-HCl pH7.5, 0.1% SDS, 10 mM EDTA pH8) to lyse the nuclei, and chromatin was sonicated twice for 15 minutes with a Covaris S220 sonicator (peak power 105, Duty factor 5, Cycles/Burst 200 for 900s). For each sample, 25 µL of magnetic protein-A Dynabeads and 25 µL of magnetic protein-G Dynabeads (Invitrogen) were washed twice with 100 µL of cold ChIP dilution buffer (15 mM Tris-HCl pH7.5, 150 mM NaCl, 1% Triton X-100, 1 mM EDTA pH8) using a magnetic rack (MagRack 6, Cytiva). Beads were mixed with 2.5 µg of anti-PhDEF antibody and 1.8 mL of cold ChIP dilution buffer, and incubated for 2 hours at 4°C on a rotating wheel. Sonicated chromatin was centrifuged for 5 minutes at 15,000 g at 15°C, and 25 µL of supernatant (for wt samples) or 50 µL of supernatant (for *phdef-151* and *phglo1 phglo2* samples) was added to the mix of beads and antibody, and incubated overnight at 4°C on a rotating wheel. Beads were washed twice (one quick wash and one long wash with 15 minutes incubation on a rotating wheel) with each of the following buffers: low salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA pH8, 20 mM Tris-HCl pH8, 150 mM NaCl), high salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA pH8, 20 mM Tris-HCl pH8, 500 mM NaCl), LiCl wash buffer (0.25 M LiCl, 1% NP40/Igepal,1% deoxycholate, 1 mM EDTA pH8, 20 mM Tris-HCl pH8) and TE buffer. Elution was performed twice with 250 µL of elution buffer (0.1 M $NaHCO_3$, 1% SDS) at 65°C. IP and input samples were decrosslinked overnight at 65°C by adding NaCl at 200 mM, then incubating for 2 h at 42°C with 20 µg proteinase K in 10 mM EDTA pH8 and 40 mM Tris-HCl pH6.5. DNA was purified with phenol:chloroform:isoamyl alcohol (25:24:1) followed by chloroform:isoamyl alcohol (24:1), precipitated with ethanol at -20°C and the pellet was washed with ethanol 70 %. The dry pellet was recovered in 50 µL TE and 1 µL was used for each qPCR reaction, which were performed in technical triplicates for each biological replicate (3 for wt and *phdef-151*, 2 for *phglo1 phglo2* and the control without antibody). The qPCR reaction was performed with 1X FastStart Universal SYBR Green (Merck) and 0.3 µM primer mix (Supplementary Table 1), for 40 cycles (15 seconds at 95°C, 1 minute at 60°C) in a QuantStudio 6 Flex instrument (ThermoFisher). Percentage

of input was calculated as 100* e^(CtIN – log2(DF) – CtIP), with e the efficiency of the primer pair, CtIN the average Ct value for the Input sample, DF the dilution factor and CtIP the average Ct value for the IP sample), as described in (Solomon et al., 2021).

Accession numbers

Sequence data from this article can be found in the EMBL/GenBank data libraries under accession numbers OQ418981 (AN1), OQ418982 (AN2 and OQ418983 (PhDEF).

Supplemental Data

Supplemental Figure 1. Additional pictures of star and wico flowers.

Supplemental Figure 2. Stamens are unfused to the tube in wico flowers.

Supplemental Figure 3. Additional pictures of *PhDEF* transcript *in situ* hybridization in wt, star and wico flowers.

Supplemental Figure 4. Wt and pink wt flowers observed in the progeny of a star parent.

Supplemental Figure 5. Epidermal revertant sectors on star petals.

Supplemental Figure 6. Autonomous and non-autonomous effects in star and wico petals.

Supplemental Figure 7. Expression of B-class genes and a subset of pigmentation genes in wt, star, wico and *phdef-151* samples.

Supplemental Table 1. List of primers used in this study.

Supplemental Table 2. Genotyping results of the progeny of a star flower.

Supplemental Dataset 1. Differential gene expression calculated by DESeq2.

Supplemental Dataset 2. List of the 451 genes downregulated in star and *phdef-151* samples, and not deregulated in wico samples.

ACKNOWLEDGMENTS

We thank Patrice Bolland, Justin Berger and Alexis Lacroix for plant care assistance, the PLATIM platform (SFR BioSciences Lyon, UAR3444/CNRS, US8/Inserm, ENS de Lyon, UCBL) for electron microscopy technical support, Benjamin Gillet and Sandrine Hugues from the sequencing platform of the Institut de Génomique Fonctionnelle de Lyon for library preparation and sequencing of the transcriptomes of this study, Rémy Belois for assistance for *in situ* hybridization experiments and Daniel Bouyer for assistance for chromatin immunoprecipitation experiments. This work was supported by a PhD fellowship to M.C. from the French Ministry of Higher Education and Research, by a grant to Q.C.S and M.M. from the Agence Nationale de la Recherche (grant ANR-

19-CE13-0019, FLOWER LAYER), by a grant to M.M. from IDEXLYON (Université de Lyon, grant ELAN-ERC), and by a grant to V.H. and C.Z. from the Agence Nationale de la Recherche (grant ANR-16-CE92-0023, FLOPINET).

AUTHOR CONTRIBUTIONS

M.M. and M.V conceived and designed the experiments. M.C., Q.C.S, P.M., P.C., V.H. and S.R.B. performed the experiments. M.C., Q.C.S., J.J., M.V. and M.M. analyzed the data. M.C., C.Z., M.V. and M.M. wrote the article.

		Phenotype of the progeny (% of the total)		
		phdef	wt	pink wt
Parent flower	wico-1	15 (94%)		1 (6%) *
	wico-2	14 (88%)	1 (6%) *	1 (6%) *
	wico-3	16 (100%)		
	wico-4	15 (94%)		1 (6%) *
	wico-5	16 (100%)		
	wico-6	12 (100%)		
	wico-7	12 (100%)		
	star-1	11 (46%)	4 (17%)	9 (38%)
	star-2	4 (25%)	4 (25%)	8 (50%)
	star-3	7 (29%)	5 (21%)	12 (50%)
	star-4	3 (19%)	3 (19%)	10 (63%)

Table 1. Progeny of the star and wico flowers after selfing.
7 wico flowers and 4 star flowers have been selfed and their progeny has been phenotyped and classified into *phdef*, wt or pink wt phenotype. Summing the star progeny for the 4 parents gives 25 *phdef*, 16 wt and 39 pink wt plants, which is not significantly different to a 1:1:2 ratio (chi-square test, p = 0.22). * For wico, we found 4 plants with wt or pink wt flowers in the progeny, and all of them were linked to the presence of a *de novo* transposon excision from the *PhDEF* locus, restoring either a *PhDEF+6* (in the case of pink wt progeny) or a wild-type *PhDEF* (in the case of the wt progeny) allele.

REFERENCES

Abe, M., Takahashi, T., and Komeda, Y. (1999). Cloning and characterization of an L1 layer-specific gene in Arabidopsis thaliana. Plant Cell Physiol **40**: 571–580.

Albert, N.W., Lewis, D.H., Zhang, H., Schwinn, K.E., Jameson, P.E., and Davies, K.M. (2011). Members of an R2R3-MYB transcription factor family in Petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. Plant J **65**: 771–784.

Angenent, G.C., Busscher, M., Franken, J., Mol, J.N., and van Tunen, A.J. (1992). Differential expression of two MADS box genes in wild-type and mutant petunia flowers. Plant Cell **4**: 983–993.

Bai, Y., Falk, S., Schnittger, A., Jakoby, M.J., and Hülskamp, M. (2010). Tissue layer specific regulation of leaf length and width in Arabidopsis as revealed by the cell autonomous action of ANGUSTIFOLIA. Plant J **61**: 191–199.

Bissell, E.K. and Diggle, P.K. (2008). Floral Morphology in Nicotiana: Architectural and Temporal Effects on Phenotypic Integration. International Journal of Plant Sciences **169**: 225–240.

Bombarely, A. et al. (2016). Insight into the evolution of the Solanaceae from the parental genomes of Petunia hybrida. Nat Plants **2**: 16074.

Bouhidel, K. and Irish, V.F. (1996). Cellular Interactions Mediated by the HomeoticPISTILLATAGene Determine Cell Fate in theArabidopsisFlower. Developmental Biology **174**: 22–31.

Brandoli, C., Petri, C., Egea-Cortines, M., and Weiss, J. (2020). The clock gene Gigantea 1 from Petunia hybrida coordinates vegetative growth and inflorescence architecture. Sci Rep **10**: 275.

Cañas, L.A., Busscher, M., Angenent, G.C., Beltrán, J.-P., and Tunen, A.J.V. (1994). Nuclear localization of the petunia MADS box protein FBP1. The Plant Journal **6**: 597–604.

Cartolano, M., Castillo, R., Efremova, N., Kuckenberg, M., Zethof, J., Gerats, T., Schwarz-Sommer, Z., and Vandenbussche, M. (2007). A conserved microRNA module exerts homeotic control over Petunia hybrida and Antirrhinum majus floral organ identity. Nat Genet **39**: 901–905.

Cavallini-Speisser, Q., Morel, P., and Monniaux, M. (2021). Petal Cellular Identities. Front Plant Sci **12**: 745507.

Coen, E.S. and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. Nature **353**: 31–37.

De Keukeleire, P., Maes, T., Sauer, M., Zethof, J., Van Montagu, M., and Gerats, T. (2001). Analysis by Transposon Display of the behavior of the dTph1 element family during ontogeny and inbreeding of Petunia hybrida. Mol Genet Genomics **265**: 72–81.

Dornelas, M.C., Patreze, C.M., Angenent, G.C., and Immink, R.G.H. (2011). MADS: the missing link between identity and growth? Trends Plant Sci **16**: 89–97.

Dudchenko, O. et al. (2018). The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. bioRxiv: 254797.

Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., and Aiden, E.L. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science **356**: 92–95.

Efremova, N., Perbal, M.-C., Yephremov, A., Hofmann, W.A., Saedler, H., and Schwarz-Sommer, Z. (2001). Epidermal control of floral organ identity by class B homeotic genes in Antirrhinum and Arabidopsis. Development **128**: 2661–2671.

Fornes, O. et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Res **48**: D87–D92.

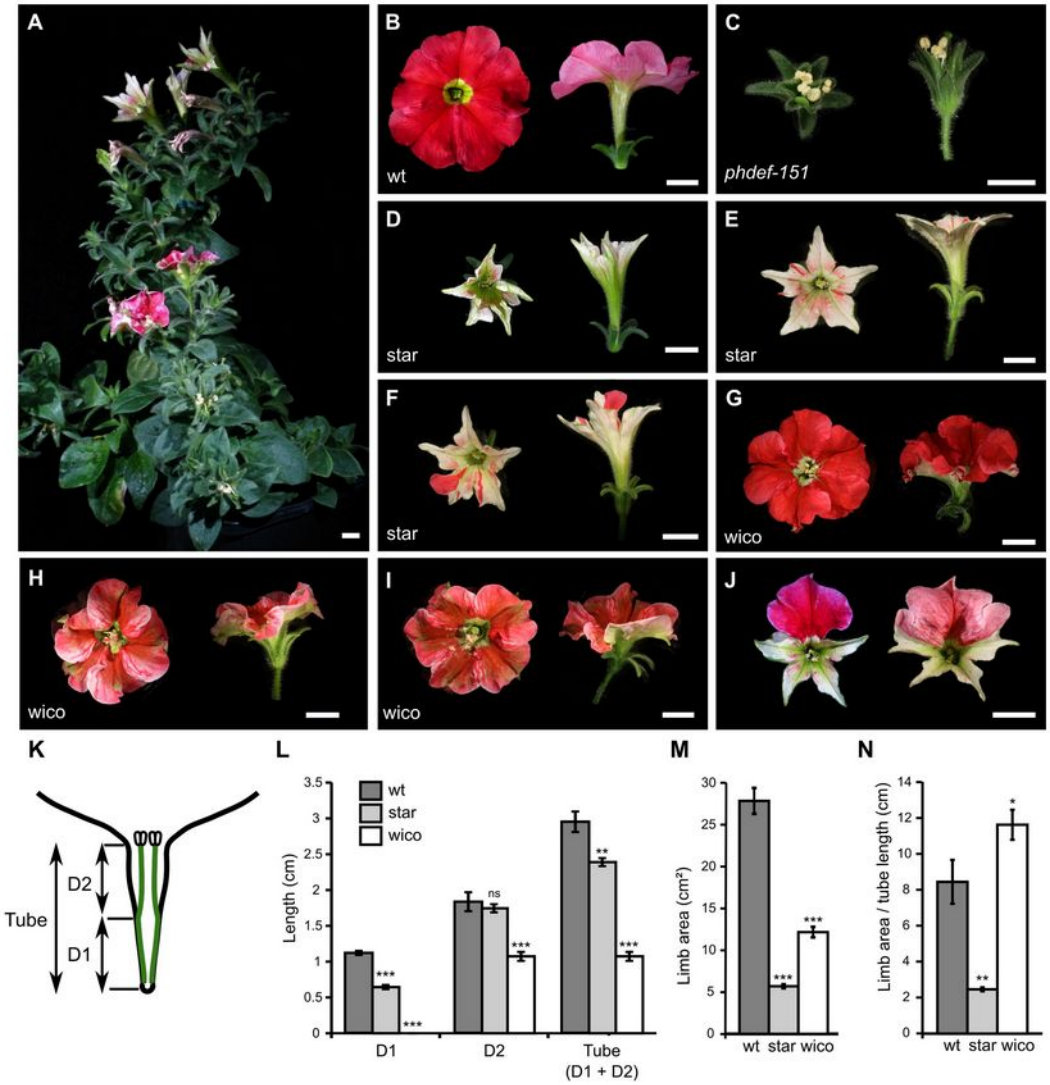
Frank, M.H. and Chitwood, D.H. (2016). Plant chimeras: The good, the bad, and the “Bizzaria.” Dev. Biol. **419**: 41–53.
Galliot, C., Stuurman, J., and Kuhlemeier, C. (2006). The genetic dissection of floral pollination syndromes. Curr Opin Plant Biol **9**: 78–82.
Gerats, A.G., Huits, H., Vrijlandt, E., Marana, C., Souer, E., and Beld, M. (1990). Molecular characterization of a nonautonomous transposable element (dTph1) of petunia. Plant Cell **2**: 1121–1128.
Graeff, M., Rana, S., Marhava, P., Moret, B., and Hardtke, C.S. (2020). Local and Systemic Effects of Brassinosteroid Perception in Developing Phloem. Curr Biol **30**: 1626-1638.e3.
Hamant, O., Heisler, M.G., Jönsson, H., Krupinski, P., Uyttewaal, M., Bokov, P., Corson, F., Sahlin, P., Boudaoud, A., Meyerowitz, E.M., Couder, Y., and Traas, J. (2008). Developmental patterning by mechanical signals in Arabidopsis. Science **322**: 1650–1655.
Heberle, H., Meirelles, G.V., da Silva, F.R., Telles, G.P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics **16**: 169.
Heijmans, K., Ament, K., Rijpkema, A.S., Zethof, J., Wolters-Arts, M., Gerats, T., and Vandenbussche, M. (2012a). Redefining C and D in the petunia ABC. Plant Cell **24**: 2305–2317.
Heijmans, K., Ament, K., Rijpkema, A.S., Zethof, J., Wolters-Arts, M., Gerats, T., and Vandenbussche, M. (2012b). Redefining C and D in the petunia ABC. Plant Cell **24**: 2305–2317.
Hill, T.A., Day, C.D., Zondlo, S.C., Thackeray, A.G., and Irish, V.F. (1998). Discrete spatial and temporal cis-acting elements regulate transcription of the Arabidopsis floral homeotic gene APETALA3. Development **125**: 1711–1721.
Hoballah, M.E., Gubitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., Dell’Olivo, A., Arnold, M., and Kuhlemeier, C. (2007). Single gene-mediated shift in pollinator attraction in Petunia. Plant Cell **19**: 779–790.
van Houwelingen, A., Souer, E., Mol, J., and Koes, R. (1999). Epigenetic interactions among three dTph1 transposons in two homologous chromosomes activate a new excision-repair mechanism in petunia. Plant Cell **11**: 1319–1336.
Hsu, H.-F., Chen, W.-H., Shen, Y.-H., Hsu, W.-H., Mao, W.-T., and Yang, C.-H. (2021). Multifunctional evolution of B and AGL6 MADS box genes in orchids. Nat Commun **12**: 902.
Jash, A., Yun, K., Sahoo, A., So, J.-S., and Im, S.-H. (2012). Looping mediated interaction between the promoter and 3’ UTR regulates type II collagen expression in chondrocytes. PLoS One **7**: e40828.
Jenik, P.D. and Irish, V.F. (2000). Regulation of cell proliferation patterns by homeotic genes during Arabidopsis floral development. Development **127**: 1267–1276.
Jenik, P.D. and Irish, V.F. (2001). The Arabidopsis floral homeotic gene APETALA3 differentially regulates intercellular signaling required for petal and stamen development. Development **128**: 13–23.
Kent, W.J. (2002). BLAT—The BLAST-Like Alignment Tool. Genome Res. **12**: 656–664.
Kierzkowski, D., Lenhard, M., Smith, R., and Kuhlemeier, C. (2013). Interaction between meristem tissue layers controls phyllotaxis. Dev Cell **26**: 616–628.
Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods **12**: 357–360.
Kostyun, J.L., Gibson, M.J.S., King, C.M., and Moyle, L.C. (2019). A simple genetic architecture and low constraint allow rapid floral evolution in a diverse and recently radiating plant genus. New Phytol **223**: 1009–1022.
van der Krol, A.R., Brunelle, A., Tsuchimoto, S., and Chua, N.H. (1993). Functional analysis of petunia floral homeotic MADS box gene pMADS1. Genes Dev **7**: 1214–1228.

Kutschera, U. and Niklas, K.J. (2007). The epidermal-growth-control theory of stem elongation: an old and a new perspective. J. Plant Physiol. **164**: 1395–1409.
Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics **30**: 923–930.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**: 550.
Lu, P., Porat, R., Nadeau, J.A., and O'Neill, S.D. (1996). Identification of a meristem L1 layer-specific gene in Arabidopsis that is expressed during embryonic pattern formation and defines a new class of homeobox genes. Plant Cell **8**: 2155–2168.
Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **17**: 10–12.
McHale, N.A. and Marcotrigiano, M. (1998). LAM1 is required for dorsoventrality and lateral growth of the leaf blade in Nicotiana. Development **125**: 4235–4243.
Meyerowitz, E.M. (1997). Genetic control of cell division patterns in developing plants. Cell **88**: 299–308.
Morel, P., Chambrier, P., Boltz, V., Chamot, S., Rozier, F., Rodrigues Bento, S., Trehin, C., Monniaux, M., Zethof, J., and Vandenbussche, M. (2019). Divergent Functional Diversification Patterns in the SEP/AGL6/AP1 MADS-Box Transcription Factor Superclade. Plant Cell **31**: 3033–3056.
Morel, P., Heijmans, K., Rozier, F., Zethof, J., Chamot, S., Bento, S.R., Vialette-Guiraud, A., Chambrier, P., Trehin, C., and Vandenbussche, M. (2017). Divergence of the Floral A-Function between an Asterid and a Rosid Species. Plant Cell **29**: 1605–1621.
Moyroud, E. and Glover, B.J. (2017). The Evolution of Diverse Floral Morphologies. Curr. Biol. **27**: R941–R951.
Perbal, M.C., Haughn, G., Saedler, H., and Schwarz-Sommer, Z. (1996). Non-cell-autonomous function of the Antirrhinum floral homeotic proteins DEFICIENS and GLOBOSA is exerted by their polar cell-to-cell trafficking. Development **122**: 3433–3441.
Purugganan, M.D., Rounsley, S.D., Schmidt, R.J., and Yanofsky, M.F. (1995). Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. Genetics **140**: 345–356.
Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N., Mol, J., and Koes, R. (1999). Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. Plant Cell **11**: 1433–1444.
Quattrocchio, F., Wing, J.F., Leppen, H.T.C., Mol, J.N.M., and Koes, R.E. (1993). Regulatory Genes Controlling Anthocyanin Pigmentation Are Functionally Conserved among Plant Species and Have Distinct Sets of Target Genes. Plant Cell **5**: 1497–1512.
Quattrocchio, F., Wing, J.F., van der Woude, K., Mol, J.N., and Koes, R. (1998). Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J **13**: 475–488.
Reale, L., Porceddu, A., Lanfaloni, L., Moretti, C., Zenoni, S., Pezzotti, M., Romano, B., and Ferranti, F. (2002). Patterns of cell division and expansion in developing petals of Petunia hybrida. Sex Plant Reprod **15**: 123–132.
Reck-Kortmann, M., Silva-Arias, G.A., Segatto, A.L.A., Mader, G., Bonatto, S.L., and de Freitas, L.B. (2014). Multilocus phylogeny reconstruction: new insights into the evolutionary history of the genus Petunia. Mol Phylogenet Evol **81**: 19–28.
Ren, H., Dang, X., Cai, X., Yu, P., Li, Y., Zhang, S., Liu, M., Chen, B., and Lin, D. (2017). Spatio-temporal orientation of microtubules controls conical cell shape in Arabidopsis thaliana petals. PLOS Genetics **13**: e1006851.

Riechmann, J.L., Krizek, B.A., and Meyerowitz, E.M. (1996a). Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. Proc Natl Acad Sci U S A **93**: 4793–4798.
Riechmann, J.L., Wang, M., and Meyerowitz, E.M. (1996b). DNA-binding properties of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucleic Acids Res **24**: 3134–3141.
Rijpkema, A.S., Royaert, S., Zethof, J., Weerden, G. van der, Gerats, T., and Vandenbussche, M. (2006). Analysis of the Petunia TM6 MADS Box Gene Reveals Functional Divergence within the DEF/AP3 Lineage. The Plant Cell **18**: 1819–1832.
Sandelin, A., Alkema, W., Engström, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res **32**: D91-94.
Satina, S. and Blakeslee, A.F. (1941a). Periclinal Chimeras in Datura Stramonium in Relation to Development of Leaf and Flower. American Journal of Botany **28**: 862–871.
Satina, S. and Blakeslee, A.F. (1941b). Periclinal Chimeras in Datura Stramonium in Relation to Development of Leaf and Flower. American Journal of Botany **28**: 862–871.
Satina, S., Blakeslee, A.F., and Avery, A.G. (1940). Demonstration of the Three Germ Layers in the Shoot Apex of Datura by Means of Induced Polyploidy in Periclinal Chimeras. American Journal of Botany **27**: 895–905.
Savaldi-Goldstein, S., Peto, C., and Chory, J. (2007). The epidermis both drives and restricts plant shoot growth. Nature **446**: 199–202.
Scheres, B. (2001). Plant cell identity. The role of position and lineage. Plant Physiol **125**: 112–114.
Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lönnig, W.E., Saedler, H., and Sommer, H. (1992). Characterization of the Antirrhinum floral homeotic MADS-box gene deficiens: evidence for DNA binding and autoregulation of its persistent expression throughout flower development. EMBO J. **11**: 251–263.
Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., and Sommer, H. (1990). Genetic Control of Flower Development by Homeotic Genes in Antirrhinum majus. Science **250**: 931–936.
Silva, C.S., Puranik, S., Round, A., Brennich, M., Jourdain, A., Parcy, F., Hugouvieux, V., and Zubieta, C. (2015). Evolution of the Plant Reproduction Master Regulators LFY and the MADS Transcription Factors: The Role of Protein Structure in the Evolutionary Development of the Flower. Front Plant Sci **6**: 1193.
Slater, G.S.C. and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics **6**: 31.
Solomon, E.R., Caldwell, K.K., and Allan, A.M. (2021). A novel method for the normalization of ChIP-qPCR data. MethodsX **8**: 101504.
Spelt, C., Quattrocchio, F., Mol, J.N.M., and Koes, R. (2000). anthocyanin1 of Petunia Encodes a Basic Helix-Loop-Helix Protein That Directly Activates Transcription of Structural Anthocyanin Genes. The Plant Cell **12**: 1619–1631.
Stewart, R.N. and Burk, L.G. (1970). Independence of Tissues Derived from Apical Layers in Ontogeny of the Tobacco Leaf and Ovary. American Journal of Botany **57**: 1010–1016.
Stormo, G.D. (2013). Modeling the specificity of protein-DNA interactions. Quant Biol **1**: 115–130.
Stuurman, J., Hoballah, M.E., Broger, L., Moore, J., Basten, C., and Kuhlemeier, C. (2004). Dissection of floral pollination syndromes in Petunia. Genetics **168**: 1585–1599.
Takatani, S., Verger, S., Okamoto, T., Takahashi, T., Hamant, O., and Motose, H. (2020). Microtubule Response to Tensile Stress Is Curbed by NEK6 to Buffer Growth Variation in the Arabidopsis Hypocotyl. Curr Biol **30**: 1491-1503.e2.

Terry, M.I., Pérez-Sanz, F., Díaz-Galián, M.V., Pérez de Los Cobos, F., Navarro, P.J., Egea-Cortines, M., and Weiss, J. (2019). The Petunia CHANEL Gene is a ZEITLUPE Ortholog Coordinating Growth and Scent Profiles. Cells **8**.
Theißen, G., Kim, J.T., and Saedler, H. (1996). Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. J Mol Evol **43**: 484–516.
Tilney-Bassett, R.A.E. (1986). Plant chimeras (E. Arnold: London; Baltimore, Md., U.S.A).
Tornielli, G., Koes, R., and Quattrocchio, F. (2009). The Genetics of Flower Color. In Petunia: Evolutionary, Developmental and Physiological Genetics, T. Gerats and J. Strommer, eds (Springer: New York, NY), pp. 269–299.
Tröbner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P., Lönnig, W.E., Saedler, H., Sommer, H., and Schwarz-Sommer, Z. (1992). GLOBOSA: a homeotic gene which interacts with DEFICIENS in the control of Antirrhinum floral organogenesis. EMBO J **11**: 4693–4704.
Urbanus, S.L., Dinh, Q.D.P., Angenent, G.C., and Immink, R.G.H. (2010a). Investigation of MADS domain transcription factor dynamics in the floral meristem. Plant Signal Behav **5**: 1260–1262.
Urbanus, S.L., Martinelli, A.P., Dinh, Q.D.P., Aizza, L.C.B., Dornelas, M.C., Angenent, G.C., and Immink, R.G.H. (2010b). Intercellular transport of epidermis-expressed MADS domain transcription factors and their effect on plant morphology and floral transition. Plant J **63**: 60–72.
Vandenbussche, M., Horstman, A., Zethof, J., Koes, R., Rijpkema, A.S., and Gerats, T. (2009). Differential recruitment of WOX transcription factors for lateral development and organ fusion in Petunia and Arabidopsis. Plant Cell **21**: 2269–2283.
Vandenbussche, M., Janssen, A., Zethof, J., van Orsouw, N., Peters, J., van Eijk, M.J.T., Rijpkema, A.S., Schneiders, H., Santhanam, P., de Been, M., van Tunen, A., and Gerats, T. (2008). Generation of a 3D indexed Petunia insertion database for reverse genetics. Plant J **54**: 1105–1114.
Vandenbussche, M., Zethof, J., Royaert, S., Weterings, K., and Gerats, T. (2004). The duplicated B-class heterodimer model: whorl-specific effects and complex genetic interactions in Petunia hybrida flower development. Plant Cell **16**: 741–754.
Venail, J., Dell'olivo, A., and Kuhlemeier, C. (2010). Speciation genes in the genus Petunia. Philos Trans R Soc Lond B Biol Sci **365**: 461–468.
de Vetten, N., Quattrocchio, F., Mol, J., and Koes, R. (1997). The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev **11**: 1422–1434.
Vincent, C.A., Carpenter, R., and Coen, E.S. (2003). Interactions between gene activity and cell layers during floral development. The Plant Journal **33**: 765–774.
de Vlaming, P., Gerats, A.G.M., Wiering, H., Wijsman, H.J.W., Cornu, A., Farcy, E., and Maizonnier, D. (1984). Petunia hybrida: A short description of the action of 91 genes, their origin and their map location. Plant Mol Biol Rep **2**: 21–42.
Wang, Z., Bai, L., Hsieh, Y.-J., and Roeder, R.G. (2000). Nuclear factor 1 (NF1) affects accurate termination and multiple-round transcription by human RNA polymerase III. The EMBO Journal **19**: 6823–6832.
Wuest, S.E., O'Maoileidigh, D.S., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., Lohan, A.J., Loftus, B., Graciet, E., and Wellmer, F. (2012). Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. Proc. Natl. Acad. Sci. U.S.A. **109**: 13452–13457.
Yadav, R.K., Tavakkoli, M., Xie, M., Girke, T., and Reddy, G.V. (2014). A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche. Development **141**: 2735–2744.

Zenoni, S., Reale, L., Tornielli, G.B., Lanfaloni, L., Porceddu, A., Ferrarini, A., Moretti, C., Zamboni, A., Speghini, A., Ferranti, F., and Pezzotti, M. (2004). Downregulation of the Petunia hybrida alpha-expansin gene PhEXP1 reduces the amount of crystalline cellulose in cell walls and leads to phenotypic changes in petal limbs. Plant Cell **16**: 295–308.
Zhang, B., Xu, X., Huang, R., Yang, S., Li, M., and Guo, Y. (2021). CRISPR/Cas9-mediated targeted mutation reveals a role for AN4 rather than DPL in regulating venation formation in the corolla tube of Petunia hybrida. Hortic Res **8**: 116.
Zhao, F. et al. (2020). Microtubule-Mediated Wall Anisotropy Contributes to Leaf Blade Flattening. Curr Biol **30**: 3972-3985.e6.



831 **Figure 1.** Macroscopic description of the star and wico flowers.

832 **(A)** *phdef-151* mutant plant harboring one branch with wico revertant flowers and one branch with

833 star revertant flowers. Scale bar: 1 cm. **(B-I)** Representative wt (B), *phdef-151* (C), star (D-F) and

834 wico (G-I) flowers from a top (left) and side (right) view. The star and wico flowers come from

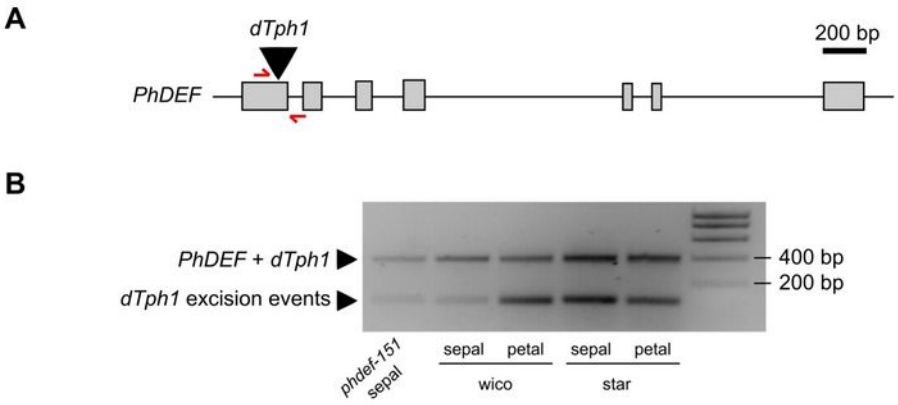
835 independent reversion events (from different *phdef-151* plants or from different branches of a single

836 *phdef-151* plant). Scale bar: 1 cm. **(J)** Two star flowers with additional L1-revertant sectors in one

837 petal (left) or one petal and two half petals (right). **(K)** Schematic cross-section of a wt flower,

838 showing stamens (in green) partially fused to the petal tube. The region of the tube fused to stamens

839 is named D1, and the region of the tube where stamens are free is named D2, as defined in
840 (Stuurman et al., 2004). **(L)** Average length of regions D1, D2 and total tube length in wt, star and
841 wico flowers. **(M)** Average limb area in wt, star and wico flowers. **(N)** Average ratio between limb
842 area and tube length in wt, star and wico flowers. n = 7 wt flowers, n = 12 star flowers from 4
843 different branches, n = 18 wico flowers from 5 different branches. Student's t test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$). Error bars represent ± s.e.m.



C

Phenotype	Allele	Nucleotidic sequence	Protein Sequence	n
wt	*PhDEF*	CCAGTA------CTGG--------CAAGCTTCAT	DAKVSIIMISS--TG--KLHEFIS	
phdef	*phdef(+8)*	CCAGTA------CTGGCAGTCTGGCAAGCTTCAT	DAKVSIIMISS--TGSLASFMNSL(+22aa)*	1
	phdef(+7)	CCAGTA------CTGGC-GTCTGGCAAGCTTCAT	DAKVSIIMISS--TGVWQAS*	1
star	*PhDEF*	CCAGTA------CTGG--------CAAGCTTCAT	DAKVSIIMISS--TG--KLHEFIS	1
	phdef (+6a)	CCAGTA------CTGG--GTCTGGCAAGCTTCAT	DAKVSIIMISS--TGSGKLHEFIS	7
	phdef (+6b)	CCAGTA------CTGGCA-T-TGGCAAGCTTCAT	DAKVSIIMISS--TGIGKLHEFIS	6
wico	*PhDEF*	CCAGTA------CTGG--------CAAGCTTCAT	DAKVSIIMISS--TG--KLHEFIS	3
	phdef (+6a)	CCAGTA------CTGG--GTCTGGCAAGCTTCAT	DAKVSIIMISS--TGSGKLHEFIS	7
	phdef (+6c)	CCAGTA------CTGGCA--CTGGCAAGCTTCAT	DAKVSIIMISS--TGTGKLHEFIS	3
	phdef (+6d)	CCAGTAGCCAGTCTGG--------CAAGCTTCAT	DAKVSIIMISSSQSG--KLHEFIS	1

845 **Figure 2.** Sequencing the *PhDEF* excision alleles in star and wico flowers.
846 **(A)** *PhDEF* gene model indicating the position of the *dTph1* insertion in the first exon (black
847 triangle) and the primers used for subsequent amplification and sequencing (in red). **(B)** Amplicons
848 generated with primers spanning the *dTph1* insertion site, on genomic DNA from *phdef-151* second
849 whorl organs and star and wico sepals and petals. The large fragment still contains the *dTph1*
850 transposon inserted (expected size: 407 bp), while small fragments result from different events of
851 *dTph1* excision (expected size: 115 bp) and were subsequently sequenced. **(C)** The small *PhDEF*
852 fragments from (B) were sequenced in the second whorl organs of flowers with a *phdef* (n = 2), star
853 (n = 14) and wico (n = 14) phenotype. The nucleotidic sequence and predicted protein sequence are
854 indicated, with STOP codons represented by a star. Additional nucleotides or amino-acids as
855 compared to the wt sequences are indicated in red. n = number of independent reversion events
856 where the same excision footprint was found.

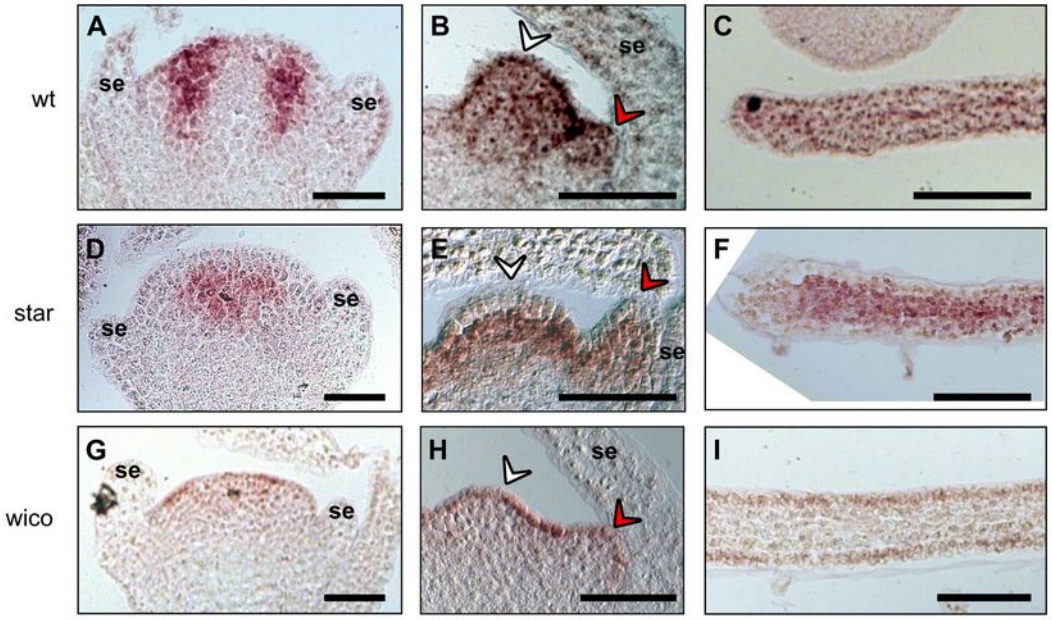
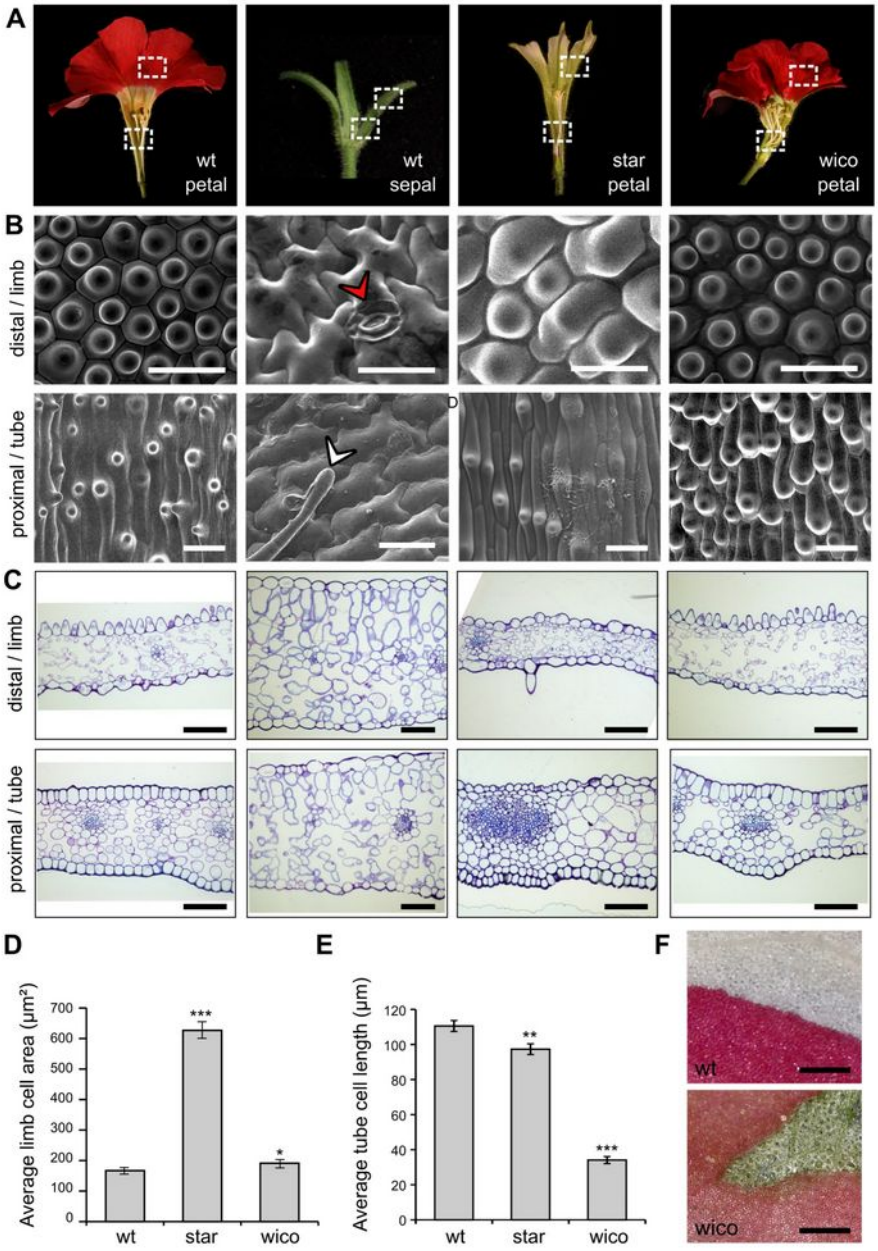
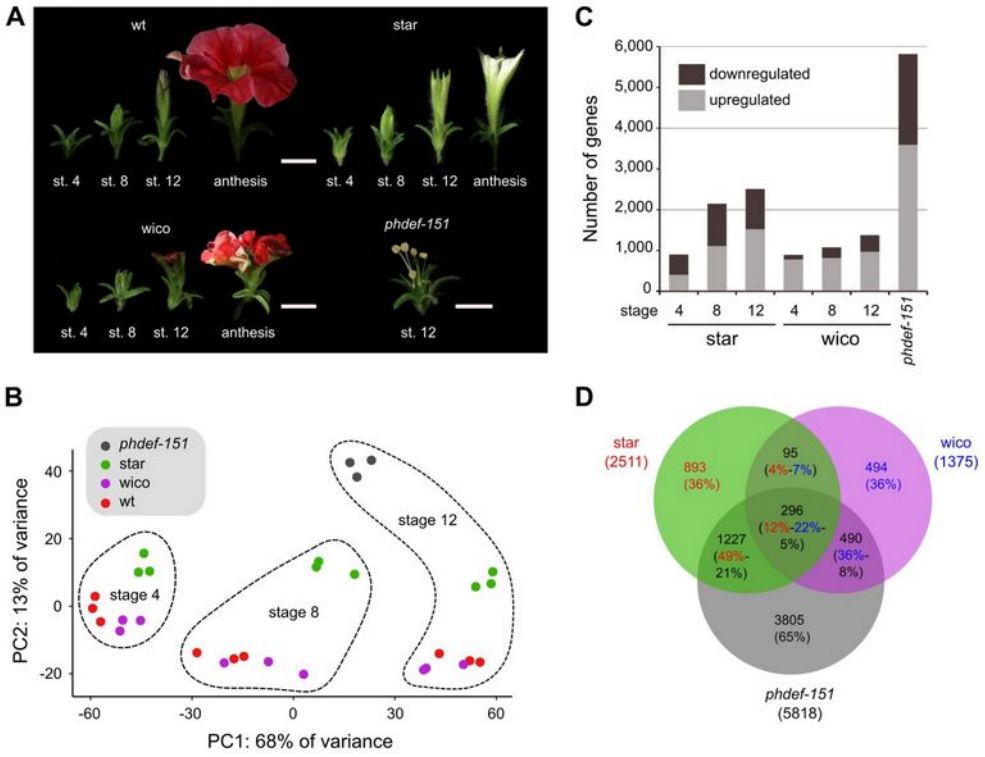


Figure 3. Localization of the *PhDEF* transcript in wt, star and wico flowers by *in situ* hybridization. Longitudinal sections of wt (A, B, C), star (D, E, F) and wico (G, H, I) flowers or young petals hybridized with a DIG-labelled *PhDEF* antisense probe. At the earliest stage chosen (A, D, G), sepals are initiating and *PhDEF* is expressed in the future petal / stamen initiation domain. Note that if the section was not performed at the center of the flower, the *PhDEF* signal might artificially appear to be in the middle of the flower (as in D) whereas it is actually on its flanks. At the middle stage chosen (B, E, H), stamens (white arrowhead) and petals (red arrowhead) are initiating, and *PhDEF* is expressed in both primordia. *PhDEF* expression is also detected at the tip of young petal limb (C, F, I). se: sepals. Scale bar: 50 µm.



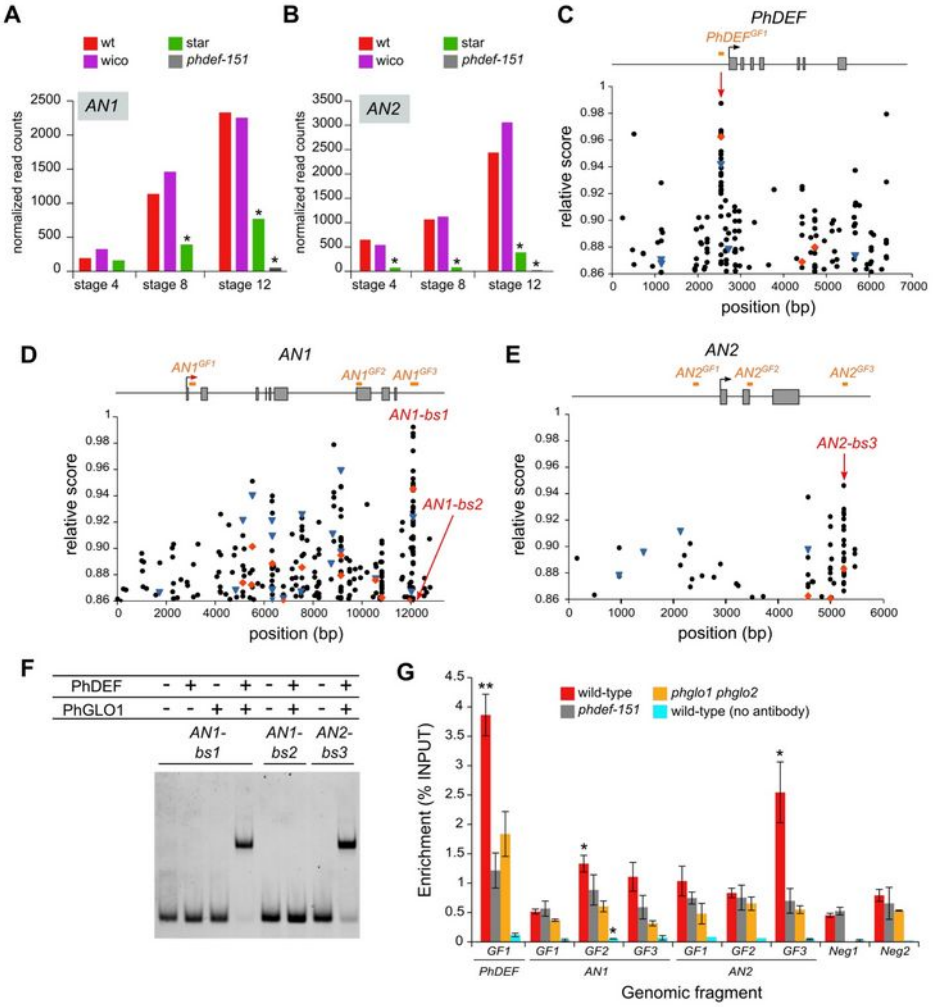
866 **Figure 4.** Epidermal and mesophyll cell identities in wt petals and sepals, and star and wico petals.
867 **(A)** From left to right: wt petals, wt sepals, star petals and wico petals cut open longitudinally to
868 show areas used for scanning electron microscopy and cross-sections. Petals were subdivided into

limb and tube area, and sepals were subdivided into a distal and a proximal part, as shown by the dotted white rectangles. **(B)** Representative scanning electron micrographs from the adaxial side of a wt petal, wt sepal, star petal and wico petal (from left to right). The red arrowhead points to a stomata and the white arrowhead points to a trichome. Scale bar: 30 µm. **(C)** Representative cross-sections from wt petals, wt sepals, star petals and wico petals (from left to right) stained with toluidine blue. Scale bar: 100 µm. **(D)** Average limb cell area from the adaxial side of wt, star and wico petals (n = 30 cells). Student's t test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$). **(E)** Average tube cell length from the adaxial side of wt, star and wico petals (n = 45 cells). Wilcoxon rank sum test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$). Error bars represent ± s.e.m. **(F)** Limb area from wt (top) and wico (bottom) petals, after their adaxial epidermis was manually peeled. For wt, the upper half of the picture shows the white underlying mesophyll. For wico, the green triangular area shows the green (chloroplastic) underlying mesophyll.



881 **Figure 5.** Gene deregulation in star and wico petals.

882 **(A)** Flowers from wt, star, wico and *phdef-151* at stages 4, 8 and 12 (only stage 12 for *phdef-151*),
883 whose petals or sepals were harvested for transcriptome sequencing. Flowers at anthesis are shown
884 for comparison. Scale bar: 1 cm. **(B)** Principal Component Analysis plot of the samples after
885 analysis of variance with DESeq2, showing that the first principal component corresponds to the
886 developmental stage and the second principal component corresponds to the genotype. **(C)** Number
887 of upregulated and downregulated genes in star, wico and *phdef-151*, as compared to wt at the
888 corresponding stages. **(D)** Venn diagram recapitulating the number of deregulated genes in star,
889 wico and *phdef-151* petal samples at stage 12, as compared to wt, and their different intersections.
890 Each sector contains the number of deregulated genes, and between parenthesis is the percentage of
891 genes that it represents from the total number of deregulated genes in the corresponding sample,
892 with a colour code (red = percentage of deregulated genes from star samples / blue = from wico
893 samples / black = from *phdef-151* samples).



894 **Figure 6.** PhDEF binds to *AN2* regulatory region *in vitro* and *in vivo*.

895 **(A, B)** Expression (as normalized read counts calculated by DESeq2) of *AN1* (A) and *AN2* (B) in

896 wt, star, wico and *phdef-151* second whorl organs at stages 4, 8 or 12. Stars indicate significant

897 down-regulation (log2FC < -1 and adjusted p-value < 0.01). **(C-E)** Relative score profiles for AP3

898 (red diamond), PI (blue triangle) and all other MADS-box TFs (black dots) available on Jaspar, on

899 the genomic sequences of *PhDEF* (C), *AN1* (D) and *AN2* (E). The relative score is computed using

900 the position weight matrix of each TF and is between 0 and 1; only relative scores higher than 0.86

901 are shown here. The gene model is represented above the score profile with exons as grey

902 rectangles, the transcription start site as an arrow, and the gene model is aligned with the position of

the predicted binding sites. For *PhDEF*, the position of a high-confidence CArG box, as explained in the main text, is indicated by a red arrow. In red, are indicated the positions of the sites tested by gel shift in (F): putative PhDEF binding sites (*AN1-bs1* and *AN2-bs3*) and a negative control with a low predicted binding score (*AN1-bs2)*. In orange, are depicted the genomic fragments tested by chromatin immunoprecipitation in (G). **(F)** Representative electrophoretic mobility shift assay (EMSA) gel performed with a combination of *in vitro*-translated PhDEF and/or PhGLO1 proteins, and Cy5-labelled *AN1-bs1*, *AN1-bs2* or *AN2-bs3* DNA fragments, whose position is depicted in (C-E). **(G)** Enrichment (as percentage of INPUT) of binding of PhDEF to different genomic regions of the chromatin purified from wt, *phdef-151* or *phglo1 phglo2* second whorl organs at stage 8, after immunoprecipitation with an anti-PhDEF directed antibody. The control without antibody was performed on chromatin isolated from wt petals. The position of the genomic fragments tested is depicted in (C-E). Neg1 and Neg2 represent two negative control fragments located in the promoter region of genes not deregulated in the *phdef-151* mutant, and present on different chromosomes than *PhDEF*, *AN1* and *AN2.* For unknown reasons, the Neg1 control region could not be amplified in the *phglo1 phglo2* samples. Stars indicate a significant enrichment of test regions over the average of the two negative control regions for each chromatin sample (one-tailed t-test, * $p<0.05$, ** $p<0.005$; n = 3 biological replicates for wt and *phdef-151*, 2 biological replicates for *phglo1 phglo2* and the control without antibody).

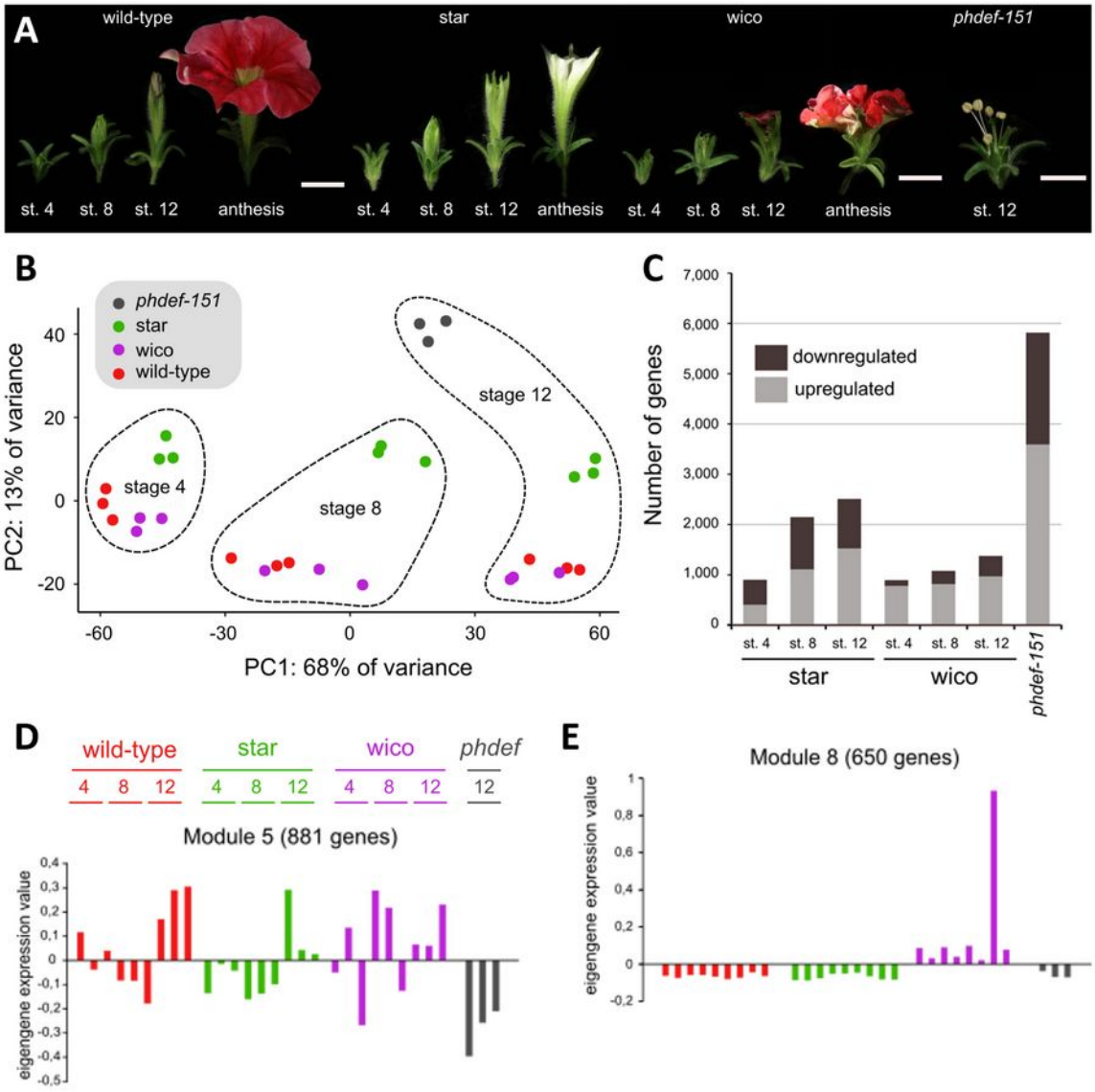


Figure 1.1: Gene deregulation in star and wico petals, and modules of gene co-expression

(A) Flowers from wild-type, *star*, *wico* and *phdef-151* at stages 4, 8 and 12 (only stage 12 for *phdef-151*), whose petals or sepals were harvested for transcriptome sequencing. Flowers at anthesis are shown for comparison. Scale bar: 1 cm. **(B)** Principal Component Analysis plot of the samples after analysis of variance with DESeq2 (Love et al., 2014), showing that the first principal component corresponds to the developmental stage and the second principal component corresponds to the genotype. **(C)** Number of downregulated and upregulated genes in *star*, *wico* and *phdef-151*, as compared to wild-type at the corresponding stages. **(D-E)** Modules 5 and 8 built by WGCNA on our transcriptomic dataset. For each module, the eigengene expression value is shown for the three biological replicates at each stage (indicated as numbers in module 1, similar order in the other modules). Wild-type expression values are in red, *star* is in green, *wico* is in purple and *phdef-151* is in gray.

I.2 – RNA sequencing of *star* and *wico* petals yields functional co-expression modules

To better understand the molecular basis for the *star* and *wico* phenotypes, bulk RNA-Seq was performed on petal tissue at three developmental stages, including wild-type and *phdef-151* samples. 3 developmental stages of study were chosen, an early stage (stage 4 as defined in (Reale et al., 2002)) when no major difference between genotypes is visible, an intermediate stage (stage 8) when tube length is at half its final size in wild-type flowers, suggesting that tube growth is active, and a late stage (stage 12) before limbs are fully expanded, suggesting that limb growth is active (Fig. 1.1, A). For *phdef-151* only the second-whorl sepal tissue at the late stage was sequenced. Principal component analysis showed that developmental stage is the first contributor to variation in gene expression, while genotype is the second axis of variation (Fig. 1.1, B). All samples clustered separately except *wico* and wild-type samples which were highly similar at all stages. One-to-one differential gene expression between samples was analyzed with R software and the package DESeq2 (Love et al., 2014) and samples showed on average 5,818 differentially expressed genes (DEGs) in *phdef-151*, as compared to 1,854 and 1,115 DEGs in *star* and *wico* respectively, when averaging for all stages (Fig. 1.1, C). There were generally more upregulated genes than downregulated ones in mutant genotypes, and the number of DEGs increased with aging of the petal in both *star* and *wico* (Fig. 1.1, C).

To identify candidate genes involved in tube and limb development, gene co-expression modules were built using Weighted Gene Correlation Network Analysis (WGCNA) (Zhang and Horvath, 2005; Langfelder and Horvath, 2008) to identify groups of genes similarly expressed between all samples. WGCNA classifies genes into different modules, each represented by an eigengene (the first principal component of the expression matrix of the module, or simply put, an ideal gene that best represents the expression profile of the module). With selected thresholds, WGCNA built 45 co-expression modules. However 32 of these modules display an expression profile where no regular pattern between replicates can be found (module 5 for instance), or where one biological replicate clearly stands out from the others (module 8 for instance), hence they do not strike as biologically meaningful (Fig. 1.1, D-E). The 13 remaining modules were analyzed closer and we applied a module membership cut-off, to select only genes that best fit the eigengene expression value of the module. Modules are ordered from the largest to the smallest, with module 1 containing 5,219 genes while module 45 contains only 8 genes.

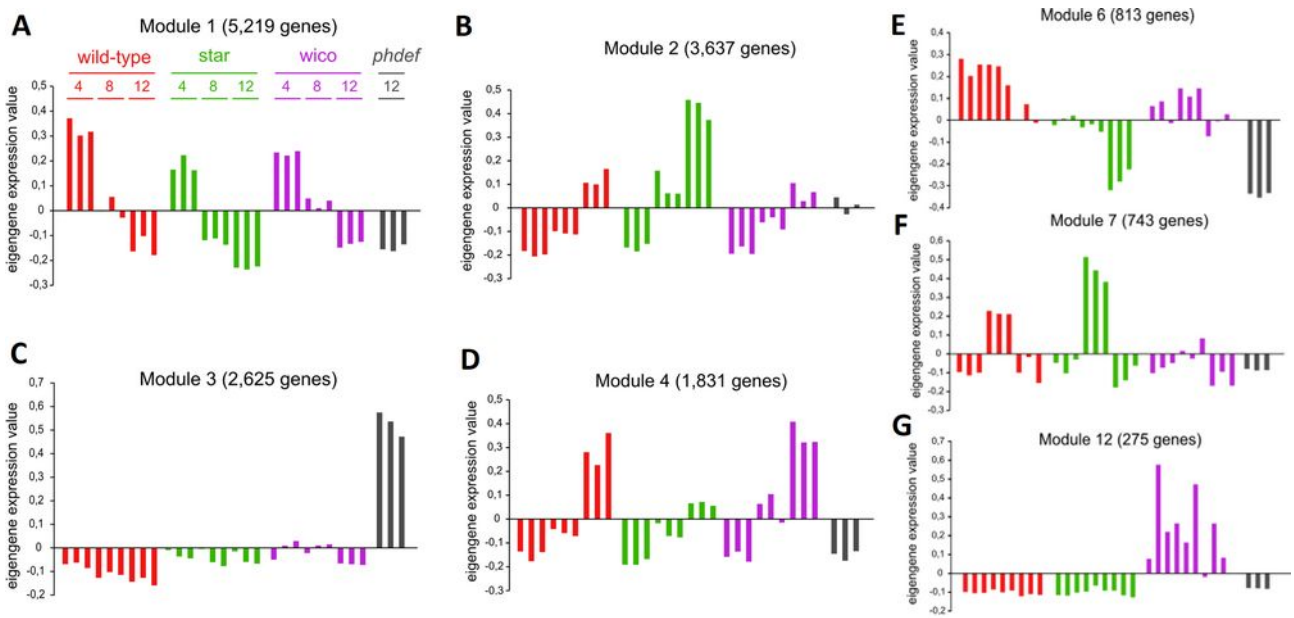


Figure 1.2: Modules of gene co-expression

(A) Module 1. **(B)** Module 2. **(C)** Module 3. **(D)** Module 4. **(E)** Module 6. **(F)** Module 7. **(G)** Module 12. For each module, the eigengene expression value is shown for the three biological replicates at each stage (indicated as numbers in module 1, similar order in the other modules). Wild-type expression values are in red, *star* is in green, *wico* is in purple and *phdef-151* is in gray.

Genes from module 1 display decreasing expression during petal development from stage 4 to 12, but show no pronounced difference between wild-type, *star* and *wico* samples, suggesting that the genes from this module are not likely to be involved in tube nor limb development specifically (Fig. 1.2, A). The 2,625 genes from module 3 are similarly lowly expressed between wild-type, *star* and *wico* samples at all stages of development studied here, but upregulated in *phdef-151*, suggesting that these genes are involved in the general establishment and/or maintenance of sepal identity or repression of petal identity (Fig. 1.2, C). The B-class genes *PhDEF*, *PhGLO1* and *PhGLO2*, are regrouped in module 6 showing reduced expression in *star* and *phdef-151* samples mostly but also slightly reduced expression in *wico* samples, as could be expected (Fig. 1.2, E).

Other modules likely to contain activators or repressors of limb development are present. The rationale is that activators of limb development should be expressed similarly between wild-type and *wico* samples, with an increasing expression level as the petal develops, and downregulated in *star* samples and *phdef-151* samples, particularly at stages 8 and 12 when limbs grow the most. Module 4 nicely meets these criteria and the 1,831 genes that it contains could be good candidates as activators of limb development (Fig. 1.2, D). Consistently, in this module 27 genes known to participate in anthocyanin production (out of a total of 41 manually annotated genes involved in this process in the whole *P. hybrida* genome) are present, including the genes encoding the major regulators ANTHOCYANIN1 (AN1), AN2, AN4, PH4 and DEEP PURPLE (DPL) as well as several anthocyanin biosynthesis enzymes. Module 2 contains 3,637 genes whose expression increases during petal development, and with higher expression in *star* samples at stages 8 and 12 (Fig. 1.2, B). These genes might be repressors of limb development, however in that case it is unclear why those genes would have an increasing expression level as the petal develops. Genes that activate or repress tube growth are likely to be expressed similarly between wild-type and *star* samples and to be downregulated or upregulated in *wico* samples respectively. However no module strictly meets these criteria. Genes from module 7 might represent activators of tube growth, transiently expressed at stage 8 when tube growth is maximal (Fig. 1.2, F), while genes from module 12 are potential repressors of tube growth, expressed in *wico* samples (Fig. 1.2, G).

Since WGCNA only identifies similar expression profiles, but does not apply cut-offs for gene deregulation, selecting DEGs within the WGCNA modules should yield a list of high-confidence candidate genes to play a role in tube or limb development. Applying this strategy to module 4, whose functional interpretation (limb development promotion) is the least ambiguous in all our modules, yielded interesting results.

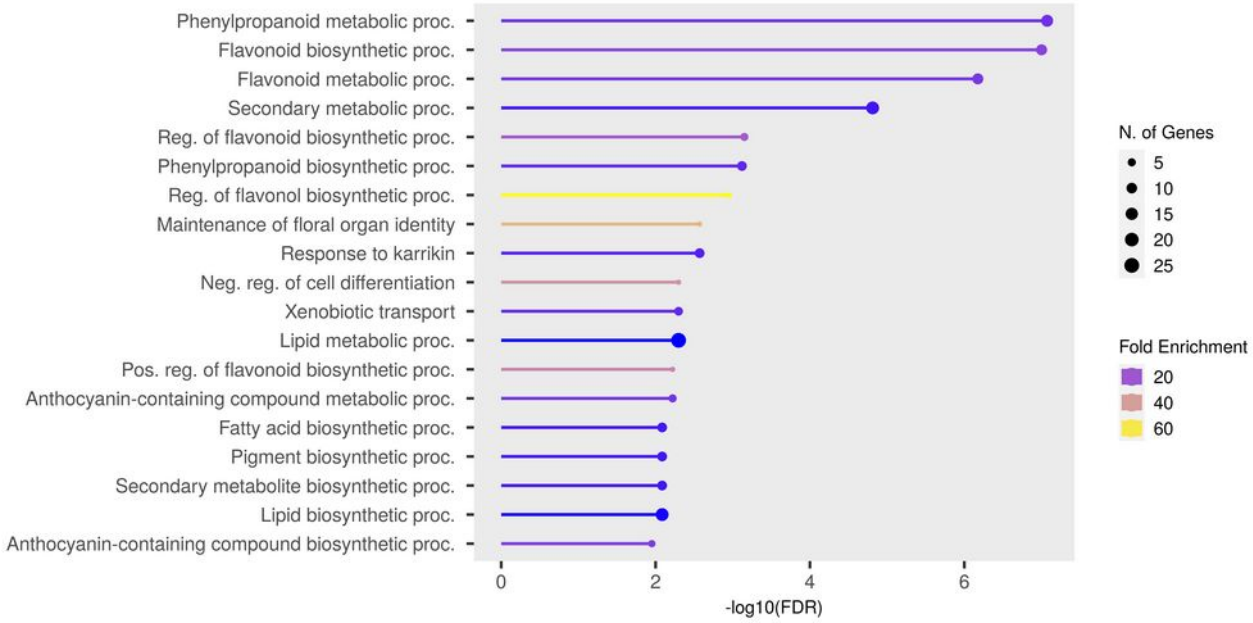


Figure 1.3: Gene Ontology (GO) term enrichment of significantly down-regulated genes in *star* at stage 12 within the module 4

Significantly down-regulated genes (Log2FC < -1.5 with p-value ≤ 0.05) in *star* at stage 12 within the module 4 were used to perform a GO analysis. The lollipop plot show the number of genes detected, their fold enrichment and the false discovery rate (proportion of genes associated with a GO term, divided by the corresponding proportion in the entire coding genome) of the top 20 significant GO biological processes.

Selecting genes significantly down-regulated in *star* samples at stage 12, resulted in a total of 292 genes, a strongly reduced number as compared to the initial 1,831 genes contained in module 4. In this reduced set of genes, we still found 21 anthocyanin-related genes, and Gene Ontology (GO) analysis show an enrichment in biological processes awaited to be down-regulated in *star* such as flavonoid-related processes or maintenance of floral identity (Fig. 1.3), suggesting that this strategy allows to keep a large part of the biologically meaningful genes. Thus, within these 292 genes are putative good candidates to play a specific role in late limb development (growth, pigmentation or conical cell formation for instance).

Originally, this bulk RNA-Seq dataset was intended as a control for scRNA-Seq and therefore, whole petal were used. In order to more reliably interpret the data regarding limb and tube development, having samples of tube tissue separately from samples with limb tissue, which could have been done by manually dissecting petals would have been better. The aforementioned data confounds the effect of *PhDEF* action on limb and tube development and its action on cell-layer identity. Thus, the interpretation of this dataset seemed too speculative to be included in (Chopy et al., 2023). However, I believe this data to still be of use to explore transcriptional dynamics and valuable as a comparison with scRNA-Seq data, hence its presence in this manuscript.

II – *Petunia* petal protoplasts generation comes with challenges

II.1 – Preamble: major drawbacks from scRNA-Seq

Most scRNA-Seq techniques require the samples of interest to be available as a cell suspension (Shaw et al., 2021). The 10X Genomics Chromium workflow (Zheng et al., 2017a) that I used to perform my experiments does not differ in this regard and its droplets-based isolation approach demands the use of a suspension of well dissociated and good viability (above 70%) cells. An efficient dissociation is key to avoid isolating multiplets, *i.e.* droplets with more than two cells that would create artificial cells with no real corresponding cell type. Although it is possible to detect and remove such droplets in downstream analysis (Dahlin et al., 2018; Kang et al., 2018; Stoeckius et al., 2018), as always, the cleaner the upstream data, the better quality the results will be. A good protoplast viability is even more important to ensure good downstream analysis. Low viability cells will release mRNAs in solution that will be captured by all droplets and rise background noise which, when too high because too many cells died, will negatively affect downstream analysis (Janssen et al., 2023).

Plant cells are encased by their cell-wall, an elaborate extra-cellular matrix of pectin, cellulose, glycans and lignin (Wilson and Hunt, 2002). This structure links cells to one another and provide organs part of their rigidity and cohesion, but hinders easy plant cell isolation into suspension compared to animal ones. Isolation is still possible either mechanically or by enzymatic digestion, the last approach being the most common. The resulting protoplasts, plant cells striped from their cell wall and only surrounded by their plasma membrane, have been used as tools in diverse research topics for over 60 years now (Cocking, 1960; Faraco et al., 2011). More recently, the emergence of scRNA-Seq in plant science also requires the use of protoplasts in suspension (Efroni et al., 2016; Svensson et al., 2018b; Zhang et al., 2019, 2021b; Kang et al., 2022).

Obtaining protoplasts, by enzymatic digestion as I did, comes with several drawbacks, among which a major one is cell viability. As stated earlier, good cell viability is key to a successful scRNA-Seq experiment. However, protoplasts are notoriously fragile and prone to bursting if not handled correctly. They are especially unforgiving when suspended in hypo-osmotic conditions. This fragility makes scRNA-Seq experiment on protoplasts tricky and needing careful experiment-specific optimization to keep them alive, even more since protoplasts characteristics can greatly differ from one another (Faraco et al., 2011).

A second drawback is the isolation bias. Firstly, all cell walls are not equal, their composition can greatly vary between cell types and cell developmental stage (Wilson and Hunt, 2002). Hence, the efficiency of the enzymes mix will vary from cell to cell and therefore, the output suspension will likely be enriched in cells with easy-to-digest cell walls but depleted in some other cell types. These differences are for example known in maize where, although it is easy to get leaves mesophyll protoplasts, vasculature cells are a challenge to isolate (Kim et al., 2021). Secondly, the location of cells inside a tissue will also impact the digestion yields in specific cell types, for instance, centrally located cells are under-represented in most *Arabidopsis thaliana* root scRNA-Seq datasets (Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Zhang et al., 2019). Thirdly, depending on the tissue of interest, cells can have very different sizes and could therefore be differently screened during the several filtration steps of an enzymatic protoplasting protocol. These multiple isolation biases will render most protoplast suspensions less representative of the crude tissue they originate from in terms of cell type. In most applications it will not matter that much, but when working with rare or particularly difficult to digest cell types, the impact can be source of major issues.

Last but not least, the conditions of digestion, loss of cell-wall and loss of inter-cellular signaling once in suspension will inevitably lead to possibly major changes in transcriptional

activity within the isolated cells. There is evidence that within the few hours of the digestion process, hundreds (Birnbaum, 2003; Jean-Baptiste et al., 2019) if not thousands of genes (Denyer et al., 2019b) can be differentially expressed between crude and protoplasted tissue.

It is therefore important to optimize the protoplasting protocol to shorten it as much as possible while maintaining high viability when planning a scRNA-Seq workflow afterwards, as well as assessing the different biases presented above, to ensure that the protoplast suspension is close enough to the tissue of interest in terms of cell populations and transcriptional activity.

It is important to note that all the following results apply to *Petunia x hybrida* W138 cultivar petal digestion and might not be applicable to other tissues or species.

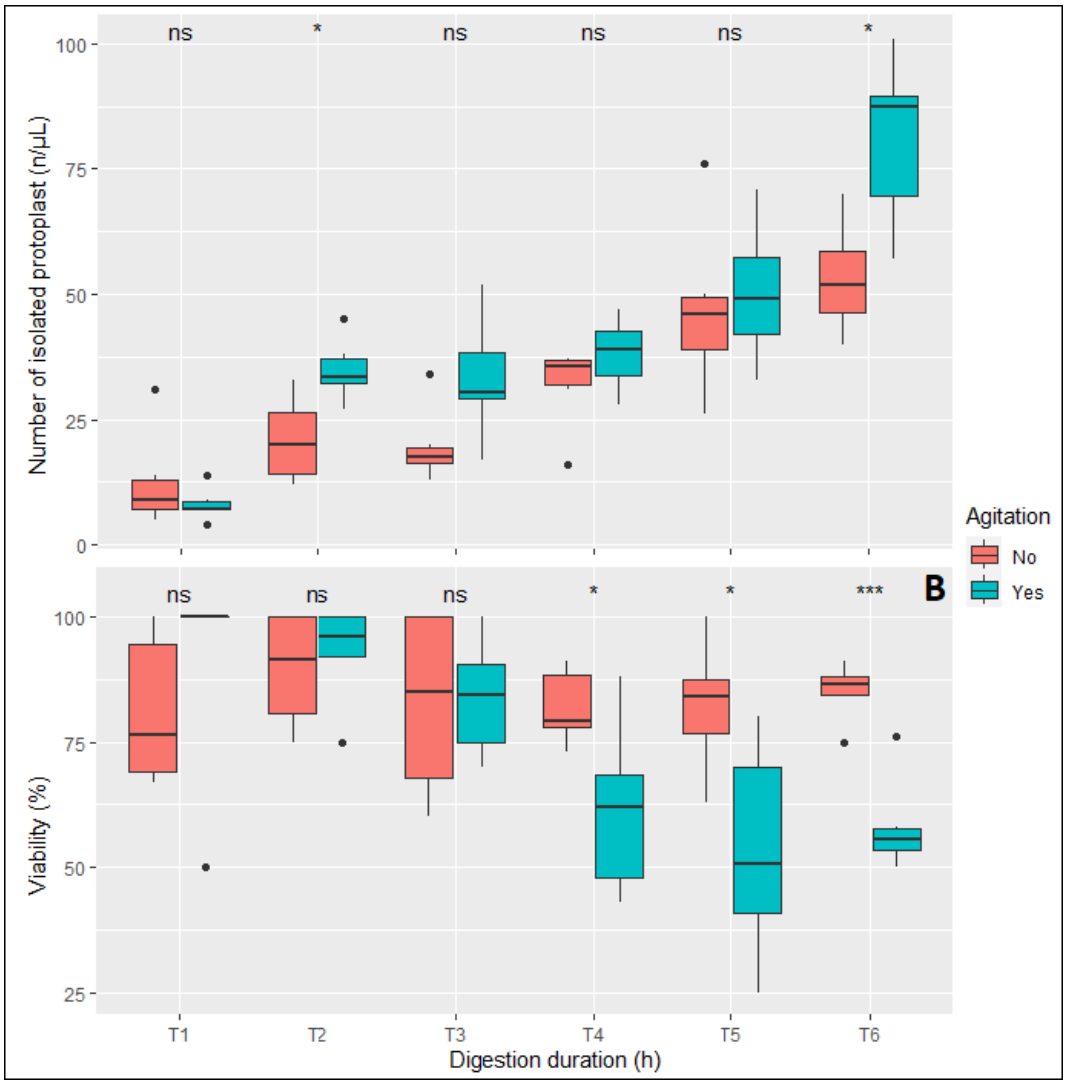


Figure 2.1: Isolated protoplast count and viability monitoring during enzymatic digestion

Six individual digestions where carried out on wild-type petals *(*as described later in materials and methods), of which three were gently agitated at 20 rpm on a small orbital agitator while the three other plates remained without agitation inside the same incubator. Every hour over the course of six hours a sample of each plate was made. The number of isolated protoplasts **(A)** and their viability **(B)** were directly estimated using Evans blue dye and a Kova slide under bright field microscopy. The results were plotted as a function of time for both conditions, without agitation in red and with agitation in blue. A Kruskal-Wallis test (Kruskal and Wallis, 1952) was performed between the two conditions at each time-point (ns: $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$).

II.2 – Agitation during isolation increases yields at the cost of cell viability

The protocol I used for protoplast isolation is derived from the one presented in (Faraco et al., 2011) used for *Petunia* petal protoplast transformation. However for the reasons I previously detailed, the overnight digestion time was not optimal. Since several other publications incubate their digestion mix under agitation for various tissues like root tips, fruits, leaves or petal, (Bargmann and Birnbaum, 2010; Titouh et al., 2015; Jean-Baptiste et al., 2019; Satterlee et al., 2020; Kang et al., 2022), I also tried to isolate protoplasts with or without agitation and followed both the number of isolated cells and their viability over the course of six hours of digestion.

Gentle agitation during digestion increases the speed of protoplast isolation. This trend is maintained throughout the follow-up, but is only significant at 4 and 5 hours of digestion (Fig. 2.1, A). The viability of the isolated protoplasts without agitation remained stable around the 80% mark, however when agitated, the viability significantly drops from the four hours mark and onward from approximately 80% to a little above 50% after six hours of digestion (Fig. 2.1, B).

Since viability is paramount when isolating protoplasts for a droplet-based scRNA-Seq approach, these results advocate against using 20 rpm orbital agitation to speed the process up for digestions longer than 3 hours. 10x Genomics recommend using a cell suspension between 700 and 1200 cells/µL for their Chromium workflow. According to my results, a four to five hour digestion without agitation seems right, with the four hours mark being on the low side and the five hours one on the high side of this range.

I therefore settled on a five hours long digestion without agitation which seemed to be most optimal considering my tissue of interest and the other fixed parameters.

II.3 – The protoplast suspension does not reflect exactly the cell type composition of the original tissue

Once the digestion parameters set, I decided to see how biased my protoplast suspension was compared to crude tissue. To this aim, I compared the number of abaxial epidermis cells, adaxial epidermis cells and mesophyll cells on the total cell number ratios in cross sections and a protoplast suspensions originating from the same *Petunia* petal limbs.

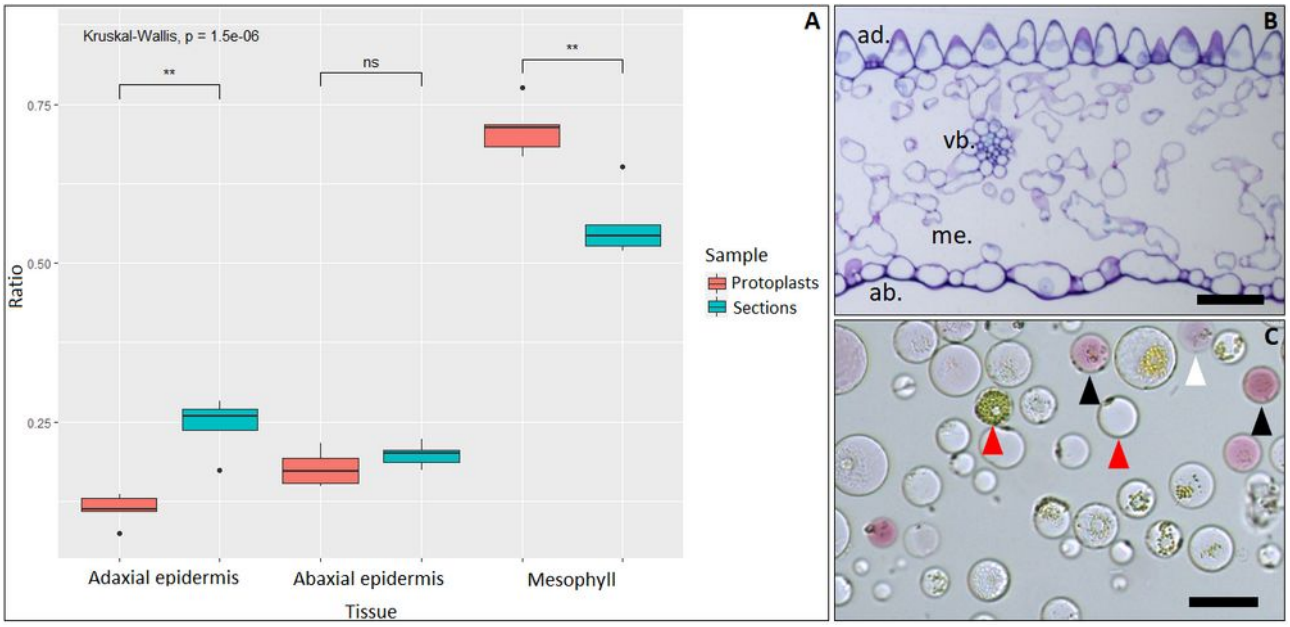


Figure 2.2: Ratio of adaxial epidermis, abaxial epidermis and mesophyll on total number of cells in petal limb, cross sections of petal limb and petal protoplast suspension

(A) Three independent digestions where carried out on wild-type petals and a small sample of each of the to-be digested petal limb was cross-sectioned. A total of nine protoplast suspension samples (three per plate) and 15 cross sections (one per flower) were observed under bright field microscopy. The number of adaxial epidermis, abaxial epidermis and mesophyll cells were counted (using a Kova slide for the suspension samples). The adaxial epidermis on total cells, abaxial epidermis on total cells and mesophyll on total cells ratios where calculated and plotted for each type of observation, protoplast suspension in red and cross section in blue. A Kruskal-Wallis test (Kruskal and Wallis, 1952) was performed between the two conditions for each type of observation (ns: $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$).

(B) Cross section of wild-type *Petunia x hybrida* W138 petal limb stained by Toluidine Blue and observed in bright field microscopy (ad.: adaxial epidermis; me.: mesophyll; vb.: vasculature bundle, ab.: abaxial epidermis; scale-bar: 50 µm).

(C) Suspension of wild-type *Petunia x hybrida* W138 petal limb protoplasts observed in bright field microscopy (black arrowheads: heavily anthocyanated adaxial epidermis cells; white arrowheads: lightly anthocyanated abaxial epidermis cells; red arrowheads: chloroplastic or non-pigmented mesophyll cells, scale-bar: 50 µm).

In petal limb cross sections these three different cell layers are easily distinguishable. The adaxial epidermis forms a single layer composed of typical conical cells (Cavallini-Speisser et al., 2021; Reed et al., 2022) (Fig. 1.2, B, ad.). The abaxial epidermis is a single layer of roundish cells at the opposite of the adaxial epidermis (Fig. 1.2, B, ab.). The mesophyll cells lie in between and form a very lacunous internal tissue (Fig. 1.2, B, me.) regularly crossed by vasculature bundles (Fig. 1.2, B, vb.).

In a protoplast suspension, differentiating unpigmented mesophyll cells from pigmented epidermal cells is not a problem. Discriminating between the two epidermises is more tricky, however the adaxial epidermis being much more pigmented than the abaxial epidermis in *Petunia x hybrida W138*, I assumed that the darker heavily anthocyanated cells were adaxial epidermis cells and the paler lightly anthocyanated cells were abaxial epidermis cells (Fig. 1.2, C, black and white arrows respectively) when identifying and counting the three cell groups.

The results show that the protoplast suspension is depleted in adaxial epidermis cells but enriched in mesophyll cells when compared to the cross sections. In contrast, the abaxial epidermis to total cells ratio is conserved between both conditions (Fig. 1.2, A).

These findings are not entirely surprising in regard of the literature. The depletion in adaxial epidermis cells could be explained by the thick cuticle at its surface (Liao et al., 2021; Ray et al., 2022) that could hinder cell wall digestion. On the contrary, the mesophyll cells enrichment in the protoplast suspension could be explained by its lacunous structure, allowing the digestion mix to more easily spread through the tissue, improving its efficiency. It is important to note that the various cell types composing vascular cells are indistinguishable by eye in a protoplast suspension and were therefore considered as mesophyll cells in this experiment. We expect some to be found in the protoplast suspension, however probably even more depleted than the adaxial epidermis since these cells are regrouped internally in compact bundles with thick secondary cell walls, as already shown in the literature (Denyer et al., 2019b; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Zhang et al., 2019).

No main cell type seems to be absent from the protoplast suspension in my case and the isolation bias exposed by these results shouldn’t be a problem to study cell-layer identity using the scRNA-Seq scope. Indeed since we are looking at transcriptomic data at the single-cell level for each isolated cell, the proportion of the different cell types matters little as long as every cell type of interest is isolated in sufficient quantities to be analyzed later, which is the case in the results presented above.

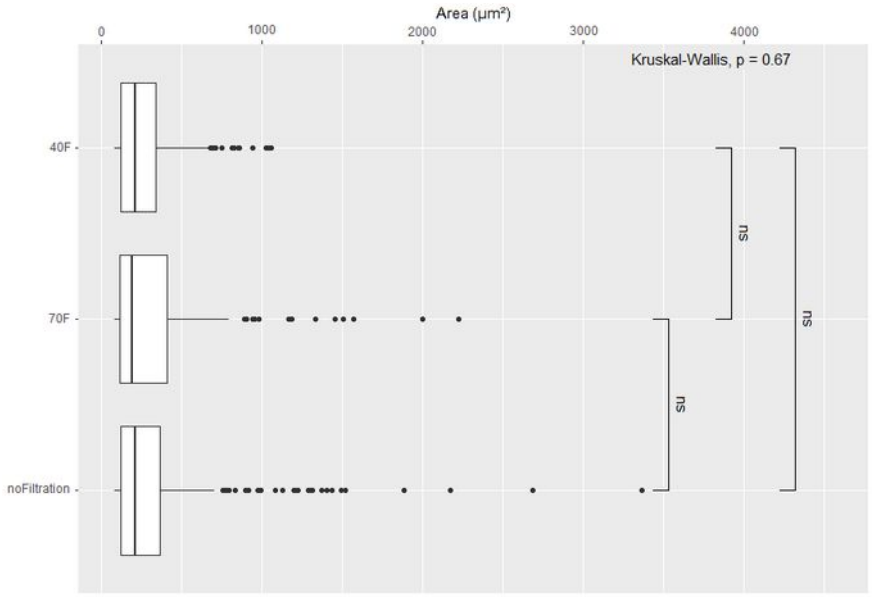


Figure 2.3: Distribution of the area of protoplasts after enzymatic isolation with or without filtration

A single digestion was carried out on wild-type petals, diluted, homogenized and split into three equal samples. One sample remained unfiltered, another was filtered on a 70 µm sieve (70F) and the last on a 40 µm one (40F). Each sample was then mounted between slide and coverslip, observed in bright field microscopy and analyzed. A Kruskal-Wallis test (Kruskal and Wallis, 1952) was performed to compare all samples one to another (ns: $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$).

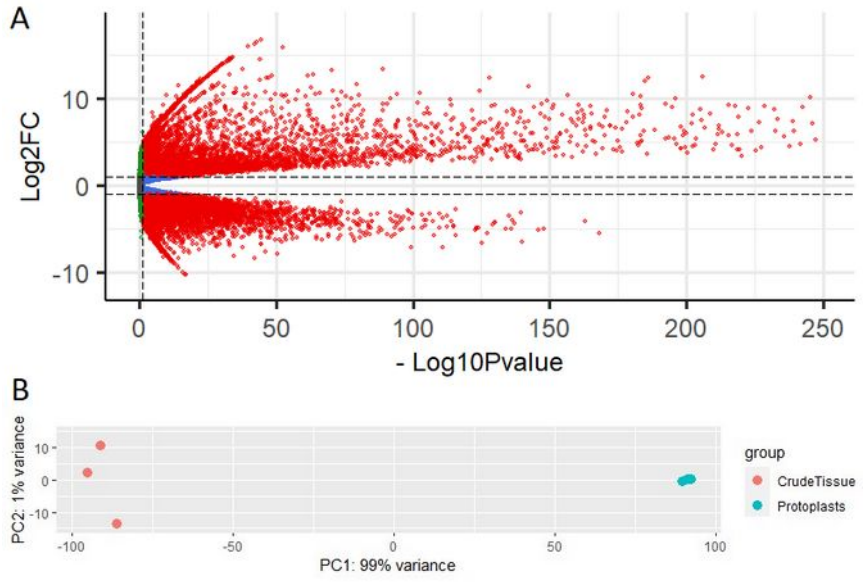


Figure 2.4: Key metrics of bulk RNA-Seq on crude and protoplasted *Petunia x hybrida* petals

(A) Volcano Plot of the detected genes expression levels. Considered as deregulated genes are in red. Non deregulated genes are in blue, green and grey with a Log2FC value between -1 and 1, or a pvalue > 0.05 or both, respectively. **(B)** Principal Component Analysis (PCA) projection of samples transcriptomic profiles.

II.4 – Lysate filtration does not introduce additional significant isolation bias

A key stage of isolating protoplasts to be used in droplet encapsulation is the filtration step. It is of crucial importance that as few debris as possible remain in suspension. The first reason is to avoid background noise. Debris that would be encapsulated with a cell could bring contaminant mRNAs and the droplet wouldn't be representing a single cell anymore. The second reason is purely technical, because debris bigger than a given limit, depending on the encapsulation technique (40 µm for 10x Genomics Chromium), could clog the encapsulation chip and ruin an entire run.

In some tissues, cells can be very diverse in size. Petunia petals are a good example and show cells ranging from smaller than 10 µm to over 50 µm in diameter (Fig. 2.2, B, C). It is therefore important to check whether filtrating debris could get rid of bigger cells and get rid of a meaningful cell population. To assess this in my conditions, I filtrated the same digested protoplasts either on a 70 µm or a 40 µm sieve and compared the results to the unfiltered lysate. There is no evidence of significant modification of the distribution of the area of the cells in either filtration condition (Fig. 2.3). The biggest outliers are removed but the impact on the cell suspension cells area is negligible from a statistical point of view. From a biological point of view, it is unknown whether these big cells are important for our biological question despite their small numbers. Although very minor, it is still an additional bias that should be kept in mind further on.

II.5 – Gene expression shift during protoplasting is massive

As stated earlier, it is known that protoplasts transcriptional state is different from its tissue of origin, even after a fast one hour digestion. Since I needed to digest *Petunia* petals for a minimum of five hours to isolate enough protoplasts, I expected the change to be even greater than previously shown in literature. In order to assess these awaited changes, I performed a differential gene expression analysis on bulk RNA-Seq data obtained from crude and protoplasted wild-type *Petunia* petal tissue.

This analysis showed that around 11,000 genes are deregulated between the two conditions. About 5,300 are upregulated (at least two-fold) and about 5,700 downregulated in the protoplast sample compared to the crude tissue sample (Fig. 2.3, A). The main driver of observed variations being the digestion treatment (Fig. 2.3, B). In order to further investigate what kind of biological process are affected by the protoplasting protocol, I performed a gene ontology (GO) analysis (Ashburner et al., 2000; Ebert et al., 2023) on the two sets of genes using their orthologs in *Arabidopsis* to link them to GO Biological Process terms.

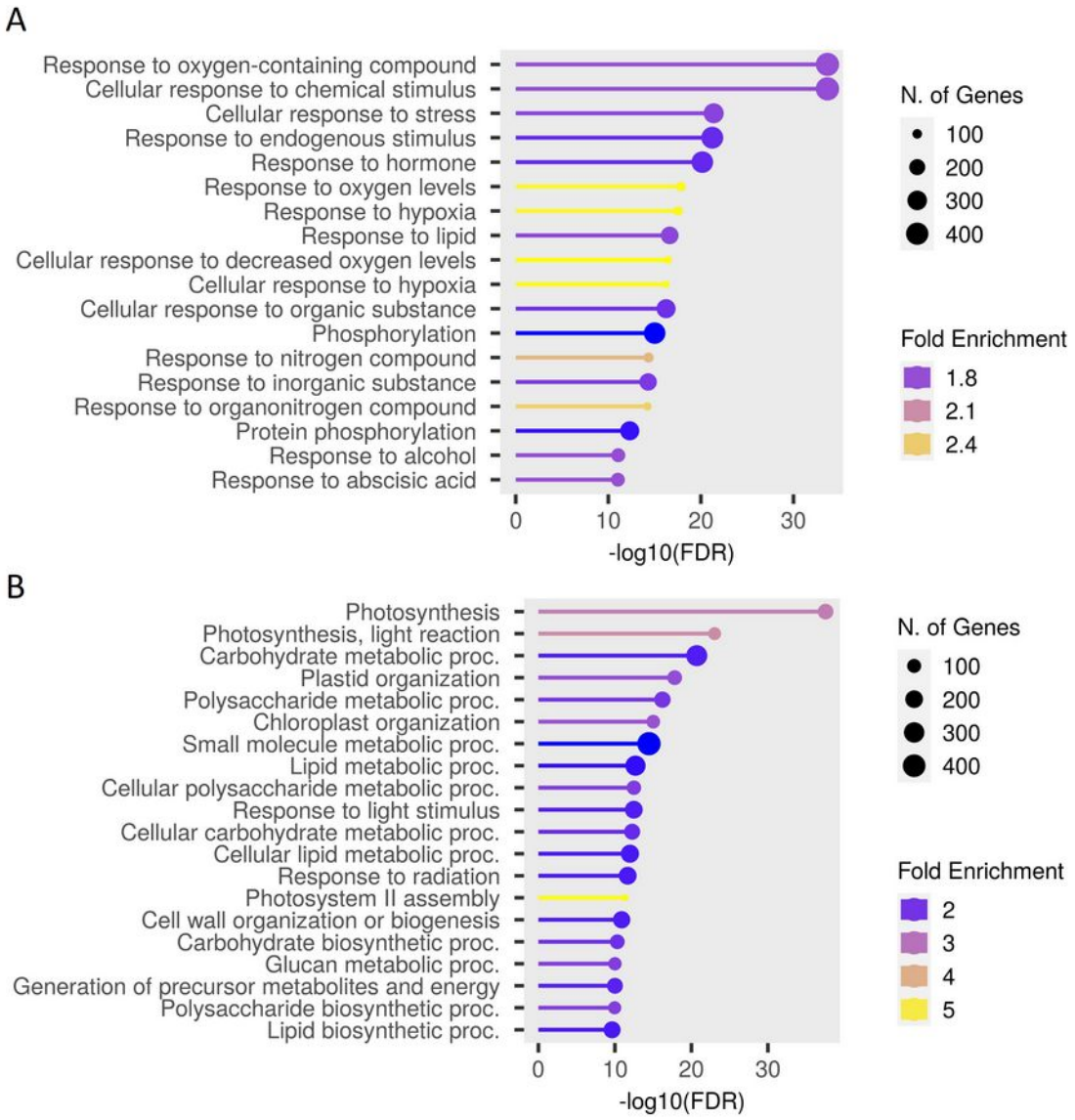


Figure 2.5: Gene Ontology (GO) term enrichment for upregulated (A) and downregulated (B) genes after protoplasting

Deregulated genes (Log2FC > 1 **(A)** and Log2FC < -1 **(B)** with p-value ≤ 0.05) after protoplasting were used to perform a GO analysis. Lollipop plots show the number of genes detected, their fold enrichment and the false discovery rate (proportion of genes associated with a GO term, divided by the corresponding proportion in the entire coding genome) of the top 20 GO biological processes of both genes lists.

To no surprise, most of the genes that are upregulated in the protoplasts compared to crude tissue are associated with GO terms related to a large variety of stress responses (Fig. 2.5, A). The most significantly enriched set of GO processes associated to these upregulated genes being reactive oxygen species (ROS), chemical stimulus, hypoxia and nitrogen compound response genes. These last two pathways have been shown to be implied in rapid ethylene-mediated plant response to flooding (Hartman et al., 2019), which is coherent with our digestion protocol. More globally the enriched set of GO processes associated to these upregulated genes as been described as plant adaptation to flooding (Pucciariello and Perata, 2017, 2021; Loreti and Perata, 2020). Response to abscisic acid (ABA) GO process related genes are also enriched in upregulated genes, once again in accord with the literature where this phytohormone is known as a key regulator of abiotic stress (Mehrotra et al., 2014).

The most significantly enriched set of GO processes associated to downregulated genes are involved in photosynthetic activity (Fig. 2.5, B). We also observe an enrichment in genes involved in lipid metabolism that are downregulated. Counteracting falling ATP concentrations by inhibiting lipid biosynthesis and activating membranes hydrolysis to generate energy is known as being part of the response to hypoxia (Xu et al., 2020).

Such results confirm that protoplasting effect is massive in my experimental conditions. As stated earlier, similar observations were made in most scRNA-Seq analysis using plant cell protoplasts (Birnbaum, 2003; Denyer et al., 2019b; Jean-Baptiste et al., 2019) but not in such large proportions. However, it was also shown that regardless of the inclusion or exclusion of known protoplasting-induced genes, pseudo-bulk (reconstituted bulk RNA-Seq from scRNA-Seq data) and single-cell RNA-Seq datasets highly correlated with each other (Birnbaum, 2003; Denyer et al., 2019b; Shulse et al., 2019). Upon close inspection of the deregulated genes it appeared that quite a few are involved in pathways that would be of interest for later data exploration such as anthocyanins biosynthesis related genes. Hence, and since literature does not advocate against it, as well as removing 11,000 genes from an analysis already known to detect low gene counts, I decided to still include them in further analysis while keeping in mind this massive deregulation.

There is also a positive aspect to such transcriptional deregulation between crude and protoplasted tissue. Most of the publications I used earlier to explain why finding the protoplasting-induced genes I observed in my experimental conditions is not a surprise were done *in planta*, not on protoplasts. Seeing protoplasts in suspension reacting the same way, as far as a GO analysis goes at least, as whole plants to protoplasting-induced stress is another confirmation, alongside countless studies using them, that they are a biologically valid model (Eeckhaut et al., 2013).

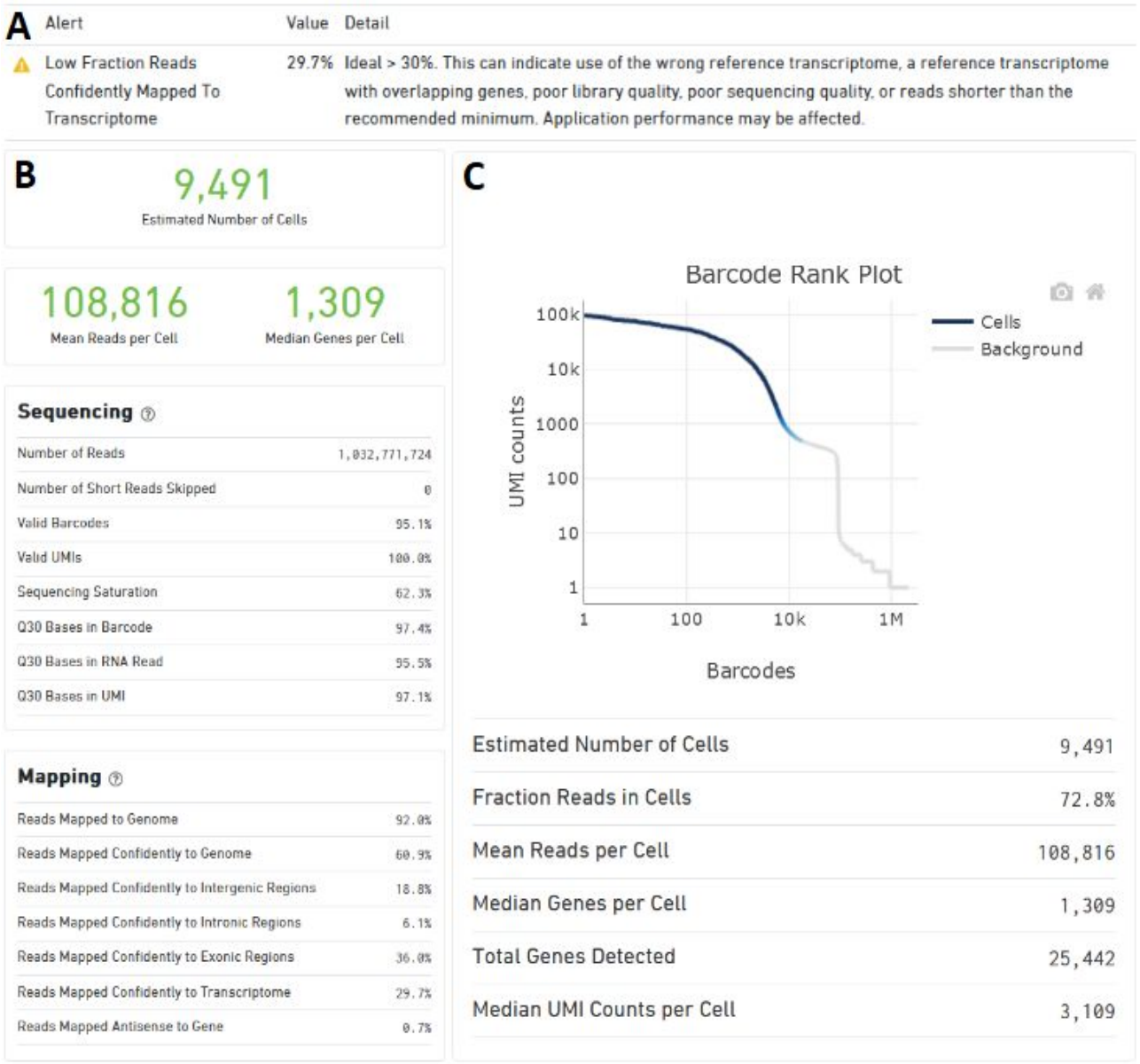


Figure 3.1: Cellranger HTML output showing key metrics of the scRNA-Seq pilot experiment

(A) Header showing errors and alerts regarding key metrics, here the low fraction of reads mapped confidently to the transcriptome of reference. **(B)** Three main metrics that are the isolated number of cells and the mean and median read count par cell. **(C)** Barcode rank plot (or simply put the signal/noise ratio) estimating encapsulation quality.

III – scRNA-Seq unveils key cell identity cues within wild-type *Petunia* petal

III.1 – scRNA-Seq is a valid approach using *Petunia* petal protoplasts

III.1.a – Pilot experiment design

As previously explained, the main goal of my thesis was to assemble a complete scRNA-Seq procedure, from protoplast isolation to bioinformatics analysis on *Petunia* petal tissue from our genotypes of interest; wild-type, *phdef-151*, *star* and *wico* at different developmental stages.

In order to test the feasibility of such a project, we designed a pilot experiment using a single sample, composed of protoplasts isolated from mature petals (or sepal-like second whorl organs for *phdef-151*) of all four genotypes of interest pooled together. Monitoring the output of this experiment allowed to check if cells from every sample were isolated as intended, if their characteristics allowed library generation and sequencing while still controlling costs since only a single sample was processed. This approach however meant that the obtained data would be challenging to analyze and draw biological conclusion from, since all genotypes are confounded.

Protoplasts of wild-type, *phdef-151*, *star* and *wico* petal were isolated, numbered and their viability checked before being pooled into a single suspension in equivalent proportions reaching a total of 560 cells/µL in the final suspension. The cells were loaded into 10x Genomics Chromium platform for droplet isolation, the library built and sequenced as previously described. Targeted metrics were 10,000 isolated cells and 100,000 reads per cells after sequencing.

III.1.b – Target metrics were met but an unexpected problem arose

Early bioinformatics analysis of the output of the pilot experiment showed that both isolated cell number and read count targets were quite closely met. Indeed, cellranger, 10x Genomics software (that maps the reads onto the reference genomes) identified 9,491 cells and 108,816 mean read count per cell (Fig. 3.1, B). Marginally less cells than expected were recovered, therefore increasing the mean read count per cell. The barcode rank plot showed the quality of the suspension was fine (Fig. 3.1, C). This graph plots the UMI counts (*i.e.* the number of unique mRNA molecules captured) as a function of each barcode (*i.e.* identified cell). A curve showing a plateau followed by steep drop around the cell count target, another plateau and yet again a steep drop, indicates the cells mostly remained intact during encapsulation and the noise of ambient mRNA background noise was low. Hence, the curve we obtained for this pilot experiment was promising regarding cells integrity (Fig. 3.1, C). However, this early analysis also show that only a small proportion, 27 %, of the reads map confidently to the transcriptome of reference (Fig. 3.1, A).

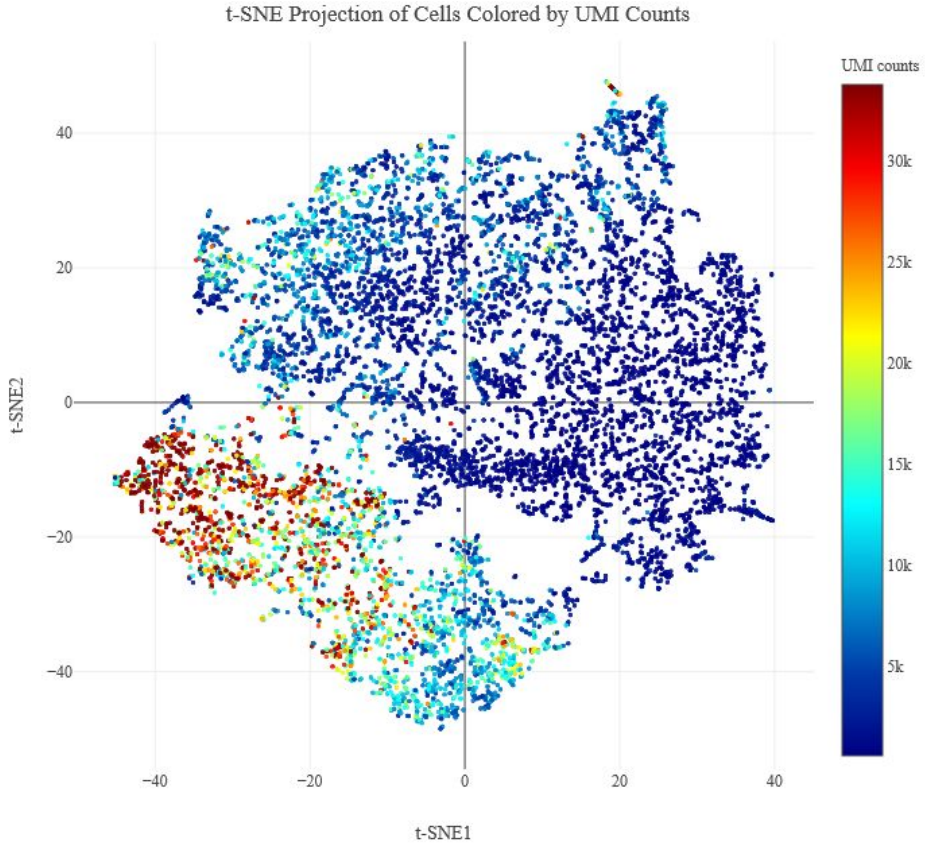


Figure 3.2: t-SNE projection of the pilot experiment UMI count

Each dot represents a cell. UMI counts are color-coded from high counts in red to low counts in blue.

Basically, only about 270,000 reads out of the 1,032,000,000 obtained will be exploitable for further analysis, which is low compared to the 70 % usually obtained and below the 30 % minimal threshold recommended by 10x Genomics. In order to try to mitigate this problem, I tried to improve the transcriptome annotation I was working with using bulk RNA-Seq datasets of *P. hybrida* flowers produced by the group between 2019 and 2022 and the genome assembler StringTie. The results gathered during this process will be presented in chapter 5. In short, although StringTie proved to efficiently improve our in-house *Petunia* genome annotation, the improvement was not sufficient to increase confidently mapped read counts in scRNA-Seq assays enough to justify its usage. To this day, I am still unsure of the reasons of such a poor mapping score, however, the use of a newer version of cellranger count increased this score by about 10 %, allowing the 30 % critical threshold to be met in all scRNA-Seq runs later on.

Another interesting metrics that cellranger outputs is a t-SNE projection (Maaten and Hinton, 2008) of the UMI count per cell (Fig. 3.2). This plot shows that our cells are very heterogeneous in terms of captured mRNA molecules. At first this worried me a bit but upon further reflection, such heterogeneity is not incoherent with the sample I used. Indeed, it makes sense that mature petals would be composed of a majority of cells with pretty low metabolism and consequently low mRNA counts, since not replicating anymore or at least at a slower rate than younger tissue.

The second main role of this pilot experiment was to provide data so I could test the bioinformatics pipeline I put together and keep improving it using data obtained in *Petunia*. In the next paragraph I will give an overview of its functioning and its main outputs.

III.2 – Custom bioinformatics pipeline overview

The pipeline that I built derives from the Seurat standard workflow (https://satijalab.org/seurat/articles/pbmc3k_tutorial), with custom improvements. This pipeline was made to run on the PSMN (*Pôle Scientifique de Modélisation Numérique*) computing clusters of the *ENS de Lyon* through the job scheduler SLURM (Yoo et al., 2003) and built around Bash and R programming languages. Bash (“run_pipeline.sh”) controls the overall pipeline workflow through SLURM job submissions, output files hierarchy, error handling and logs generation, while R scripts do the actual computing. The main R packages I used are Seurat (Hao et al., 2021) for scRNA-Seq specific computation, MultiK (Liu et al., 2021) to produce additional metrics helping with data exploration later-on, and Tidyverse (Wickham et al., 2019) for plot generation.

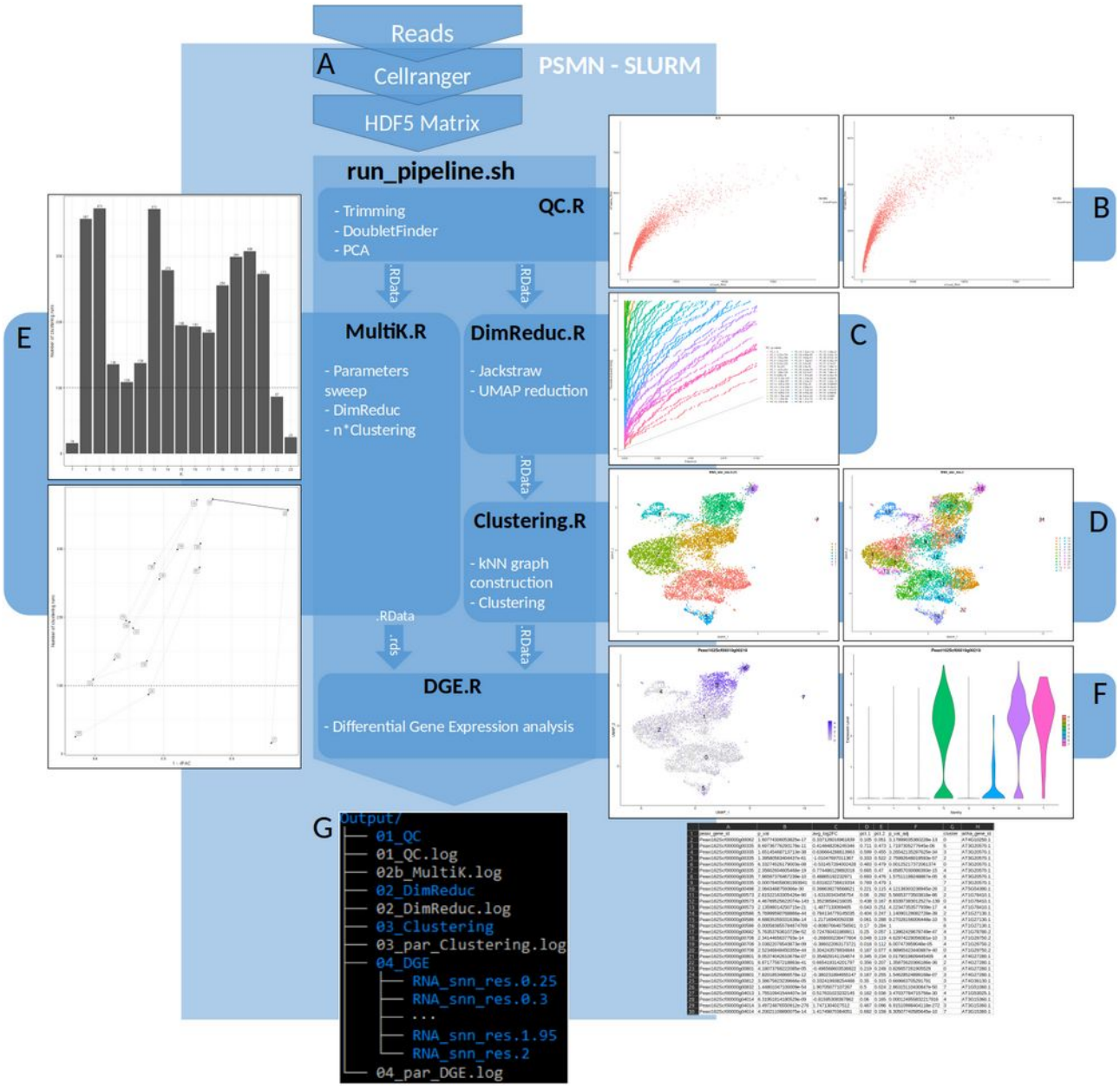


Figure 3.3: Bioinformatics scRNA-Seq unsupervised analysis pipeline overview

(A) The pipeline runs on the PSMN computing clusters and launches several computing scripts through SLURM job scheduler through “run_pipeline.sh”. **(B)** “QC.R” pre-process the data. **(C)** “DimReduc.R” reduces the dimensionality of the dataset. **(D)** “Clustering.R” clusters the dataset by regrouping cells of similar transcriptome together. **(E)** “MultiK.R” helps to choose an appropriate number of clusters for further analysis. **(F)** “DGE.R” performs differential gene expression analysis. **(G)** All along, multiple logs, plots and analysis datasets are produced and organized into a comprehensive file hierarchy.

Reads are processed by cellranger, a 10x Genomics software that demultiplexes the reads into a HDF5 (The HDF Group, 2006) matrix (Fig. 3.3, A). Simply put, this matrix associates read counts, gene identity and cell identity. The matrix is loaded into “QC.R” which trims the dataset by removing genes detected in less than three cells and cells where less than 200 genes are detected. Doublets or multiplets are also removed using the package DoubletFinder (McGinnis et al., 2019) and a Principal Component Analysis (PCA) dimensional reduction is computed. Plots of the data before and after trimming are generated (Fig. 3.3, B). The output of “QC.R” is then processed by “DimReduc.R” which first uses the jackstraw approach to determine how many Principal Components (PCs) preserve most of the variability of the dataset (Chung and Storey, 2015; Chung, 2020) (Fig. 5.4, C). Then “DimReduc.R” uses this previous step result to compute the Uniform Manifold Approximation and Projection (UMAP, (McInnes et al., 2020)) dimensional reduction of the dataset (Fig. 5.4, B). The output of “DimReduc.R” is loaded into “Clustering.R” which builds the k Nearest Neighbor (kNN) graph of the dataset (Cover and Hart, 1967) and computes clustering runs across a range of parameters in order to produce a range of low-resolution to high-resolution clusterings (Fig. 5.4, D). In parallel to “DimReduc.R” and “Clustering.R”, the script “MultiK.R” is ran using the output of “QC.R” as input. The script uses the R package MultiK (Liu et al., 2021) which subsets 80% of the original dataset, runs the same dimensional reduction as described earlier and then computes 4,000 clustering runs. These clusterings are then used to produce several graphs giving a better idea of which clustering parameters best fit the dataset based on statistical cues (Fig. 3.3, E). Finally, the outputs of “Clustering.R” and “MultiK.R” are loaded into “DGE.R”. This last script performs unsupervised differential gene expression analysis on the statistically sound clustering run(s) suggested by MultiK in order to identify clusters marker genes that would allow to assign them a given cell identity. “DGE.R” also exports multiple plot types for these marker genes, as well as several sets of genes of interest (involved in petal identity, polarity, anthocyanins biosynthesis, epidermis identity, etc.), enabling easy data exploration, as well as of course outputting differential gene expression metrics (Log2 fold change (Log2FC), pvalue, cluster membership, etc.) tables for all genes and all clusters (Fig. 3.3, F). All outputs of the pipeline are organized into a comprehensive file hierarchy (Fig. 3.3, G).

Full pipeline with scripts used in analysis featured later and their logs are publicly available online in the supplementaries.

Figure 3.4: Cellranger summary of wild-type, *star* and *wico* samples from the second scRNA-Seq run

(A) Quality notifications of the wild-type sample analysis. **(WT)** Details of the wild-type sample analysis. **(*star*)** Details of the star sample analysis. **(*wico*)** Details of the wico sample analysis.

III.3 – Second scRNA-Seq run

The next scRNA-Seq run aimed at the isolation of 6,000 cells per sample of interest (wild-type, *phdef-151*, *star* and *wico*) and sequencing 100,000 reads per cell.

This run however was crippled by several technical difficulties. First, I was unable to obtain a concentrated enough *phdef-151* protoplast suspension. This problem was anticipated following the pilot experiment that already showed me that obtaining enough *phdef-151* protoplasts was going to be challenging, and as much *phdef-151* plants as possible were grown so I could dissect as much flowers as possible but even so, digestion rates were too low to allow isolation of even low suspension concentration limits recommended by 10x Genomics. Second, although protoplast viability was around 90 % prior to Chromium encapsulation, cellular integrity was heavily affected during this step, rendering subsequent data challenging if not impossible to analyze.

As previously stated, the cellranger summary shows that the gathered data is of poor quality. Cellranger summary illustrates this well for all three samples, although there is some heterogeneity in the results. All samples show even worse rates of read counts mapping confidently to the transcriptome of reference than for the pilot experiment (run 1) (Fig. 3.4, A). All samples also show low fraction reads in cells, basically meaning that a large fraction of the reads were not attributed to any cell (Fig. 3.4, A). This observation is usually the consequence of either the presence of a large cell population with very low mRNA counts that cannot be detected as cells by cellranger, or a high level of ambient mRNAs creating high background noise, making cells with lower mRNA counts to be detected. Considering the low read mapping rate and the fact that the pilot experiment did not show this second warning, I am more prone to suspect high level of ambient mRNA, meaning cell integrity problems during encapsulation. Indeed, ambient mRNA molecules would be degraded quicker by ambient RNases than mRNAs still inside intact cells, which would reduce their mapping rate (Gallego Romero et al., 2014) and explain the first alert observed. Barcode rank plots are very good representations of the metrics I just described. As opposed to the previous curve observed in run 1 (Fig. 3.1, C), the wild-type curve obtained in this run shows a gradual slope, the background noise blending in with the cell signal (Fig. 3.4, WT). *Star* and *wico* plots are cleaner since they show a clearer plateau, although perfectible, but the first steep drop is not as steep as it should be (Fig. 3.4, *star*, *wico*). This unclear separation between cell signal and background noise makes very difficult the identification of cells, hence the aberrant number (11,228) identified as such in the wild-type sample (Fig. 3.4, WT) which is obviously the most degraded. *Star* and *wico* samples are closer to the expected 6,000 isolated cells with respectively 6,300 and 4,290 identified cells (Fig. 3.4, *star*, *wico*).

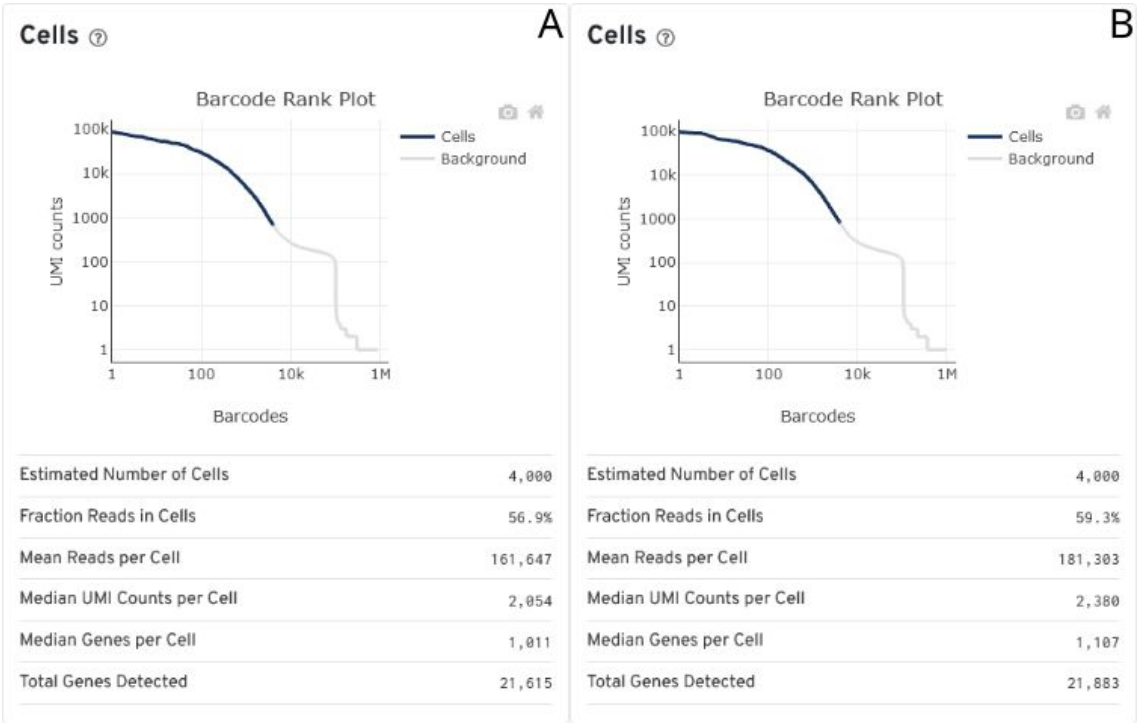


Figure 3.5: Barcode rank plots and key metrics of wild-type samples from the third scRNA-Seq run

(A) First replicate. **(B)** Second replicate.

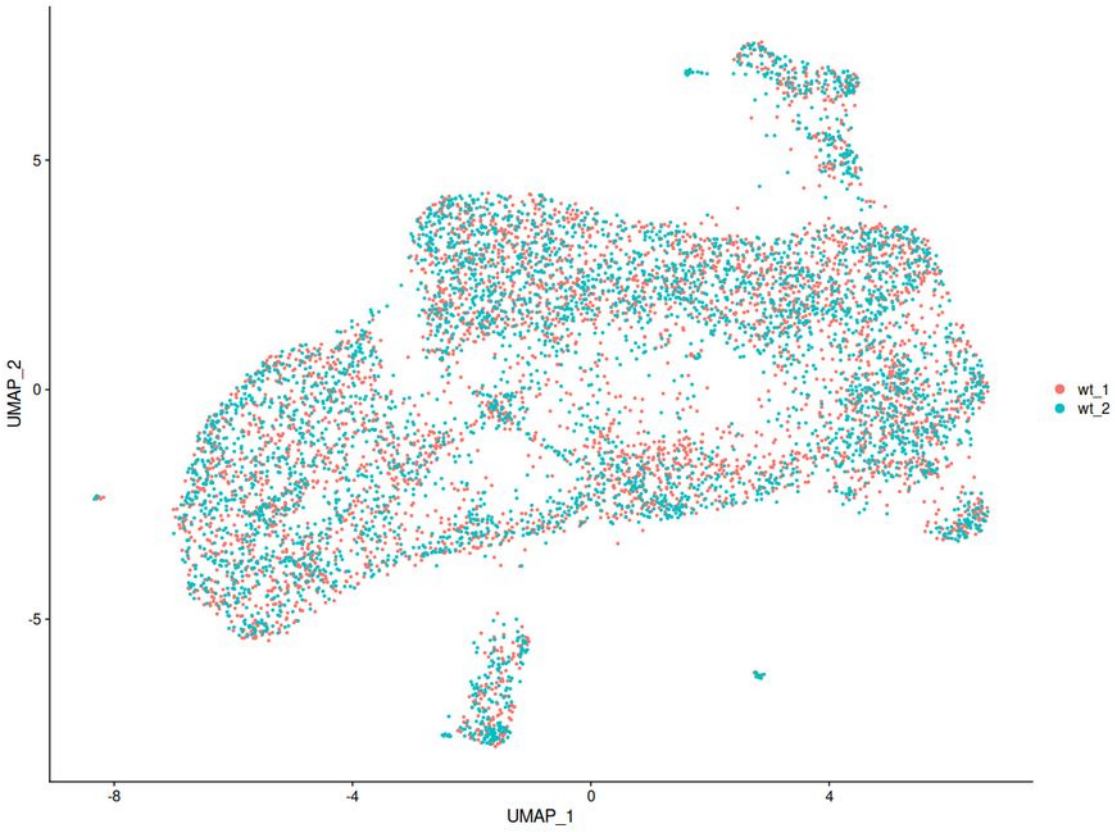


Figure 3.6: UMAP of wt_1 and wt_2 samples merged

All aforementioned considered, the wild-type sample does not look like it would be exploitable, however, *star* and *wico* samples could be. A way to gather exploitable data would be to force cellranger to identify a fixed number of cells, lower than the expected count, in order to capture mainly cells with high signal to noise ratio.

III.4 – Wild-type petal scRNA-Seq data expose key cell identity cues

III.4.a – Experimental design and run quality

Following these poor results, I planed a third and last scRNA-Seq run in order to try to obtain data from wild-type and *phdef-151* petals. The aim was to isolate 7,000 cells and sequence 100,000 reads per cell.

Yet again, I was unsuccessful at producing a concentrated enough *phdef-151* protoplast suspension. In consequence, I decided to run a technical replicate of the wild-type sample, in order to be able to integrate them together into a unique bigger dataset if need be. For both these samples, the encapsulation went better than during the second run although not as cleanly as I would have wanted. However, by fixing the number of cells cellranger identified to 4,000 (instead of the target 7,000).

III.4.b – Samples are comparable at the single-cell level

Since both datasets come from wild-type petal, I assumed they should be very similar. I checked whether it was the case and if they could be merged into a bigger dataset for further analysis. To this aim, I merged them into one while retaining sample identity, carried preprocessing and dimension reduction as described earlier and plotted the obtained UMAP (see supplementaries). As expected, both datasets perfectly clustered together (Fig. 3.6). Moreover, a differential expression analysis between the two datasets showed only 4 genes were significantly differentially expressed in wt_1 compared to wt_2, three of which were found to be upregulated by 0.53 to 0.81 Log2FC and one downregulated by -0.61 Log2FC (see supplementaries).

For further analysis, the two datasets were aggregated into a single one using “cellranger aggr” which first normalizes average read depth per cell and then combines the datasets together (Zheng et al., 2017b).

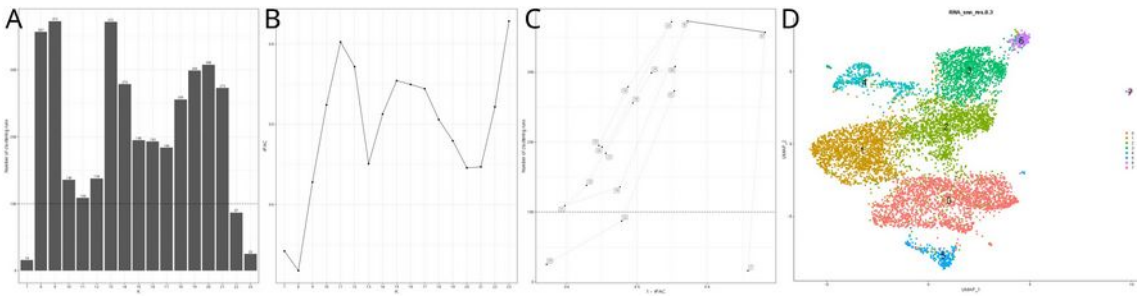


Figure 3.7: MultiK metrics and UMAP of wild-type aggregated dataset at the studied clustering resolution

(A) Bar plot of the frequency of runs for each K (cluster number) across all 4,000 subsampling runs, higher values mean they are statistically more reproducible cluster counts. **(B)** Plot of relative Proportion of Ambiguous Clustering (rPAC) score for each K. rPAC quantifies if the identity of a given cluster is well defined compared to other clusters. Local minimums of rPAC show good clustering quality. **(C)** Scatterplot of (1 – rPAC) and the frequency of K. Best values are in the upper right corner. Suggested values by MultiK are connected by a plain line. **(D)** UMAP of the clustering resolution eventually chosen for downstream analysis.

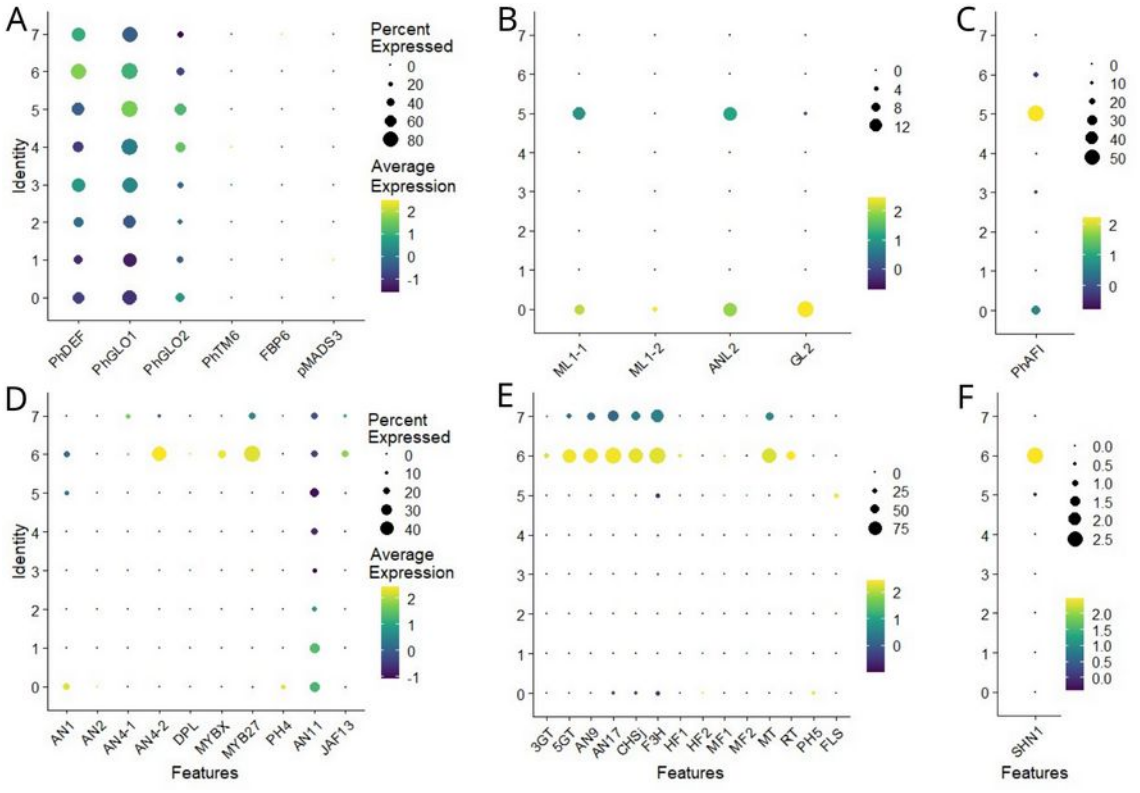


Figure 3.8: Dot plots representing differential gene expression analysis of key identity genes

(A) B- and C-class MADS-box homeotic genes. **(B-C)** Epidermal identity genes. **(D)** Anthocyanins biosynthesis regulator transcription factors. **(E)** Anthocyanins biosynthesis enzyme-encoding genes. **(F)** Conical cells marker.

III.4.c – Cluster-specific genes allow cluster identity deduction

Aggregated data was loaded into my analysis pipeline and its outputs further studied. Since the main goal of using scRNA-Seq data was to be able to differentiate between epidermal and inner cell layers, we chose to work on the best low resolution clustering proposed by MultiK, showing eight cell clusters (Fig. 3.7, A-C). The studied clustering UMAP is showed in (Fig. 3.7, D).

Since stamens and petals are fused in a portion of the tube in wild-type *Petunia* flowers, we assessed whether contamination of stamen tissue was present. All three B-class petal identity genes were detected in all clusters (Fig. 3.8, A; *PhDEF*, *PhGLO1*, *PhGLO2*). Their expression varies to some extent between the different clusters. Stamen markers such as the peculiar B-class gene *PhTM6* and both C-class genes are not detected in our sample (Fig. 3.8, A; *PhTM6*, *FBP6*, *pMADS3*). Hence, the sample is composed of petal cells only, or at least seems clean enough to not form stamen tissue specific cell clusters and further analysis can be performed without additional cleaning steps in this regard.

Known epidermal marker genes characterized as such in *Arabidopsis* such as ARABIDOPSIS THALIANA *MERISTEM LAYER1* (AT*ML1*), *ANTHOCYANINLESS2* (*ANL2*) and *GLABRA2* (*GL2*) show strong expression in cluster 0. *ML1* and *ANL2* are also expressed in cluster 5, also less strongly. In both clusters however, very few cells, 12 % at most, express these genes (Fig. 3.8, B). Another known epidermis marker described in *Antirrhinum*, *ANTIRRHINUM FIDDLEHEAD* (*AFI*) is also expressed in clusters 0 and 5, but its expression strength is interestingly inversed when compared to previously cited epidermal genes (Fig. 3.8, C). *AFI* also seems to be expressed in about 20% of the cells of cluster 6.

Petunia W138 line petal epidermis is pigmented while its mesophyll is not. Hence, its anthocyanins biosynthesis pathway is pretty well studied, so genes known to be involved should give valid cues to differentiate the epidermis from the mesophyll. Some of the known transcription factors indeed show a distinct expression pattern in clusters 0 and 6 on the strong side and 7 on the weaker side (Fig. 3.8, D). Interestingly, *ANTHOCYANIN 11* (*AN11*) is expressed throughout all clusters, which is coherent with the literature (Bombarely et al., 2016). The picture is even clearer when looking at genes encoding biosynthesis enzymes of the anthocyanins pathway (Fig. 3.8, E). More than half of the ones detected in our dataset are very strongly expressed in cluster 6. Interestingly, they are also strongly expressed in cluster 7, a cluster that did not show up before. Cluster 0 show some expression in a few of the genes such as *AN1*, but only in a small proportion of cells. Finally, *SHINE1* (*SHN1*), gene known to be expressed in conical cells also show strong expression in cluster 6 (Fig. 3.8, F), clearly identifying it as adaxial limb epidermis.

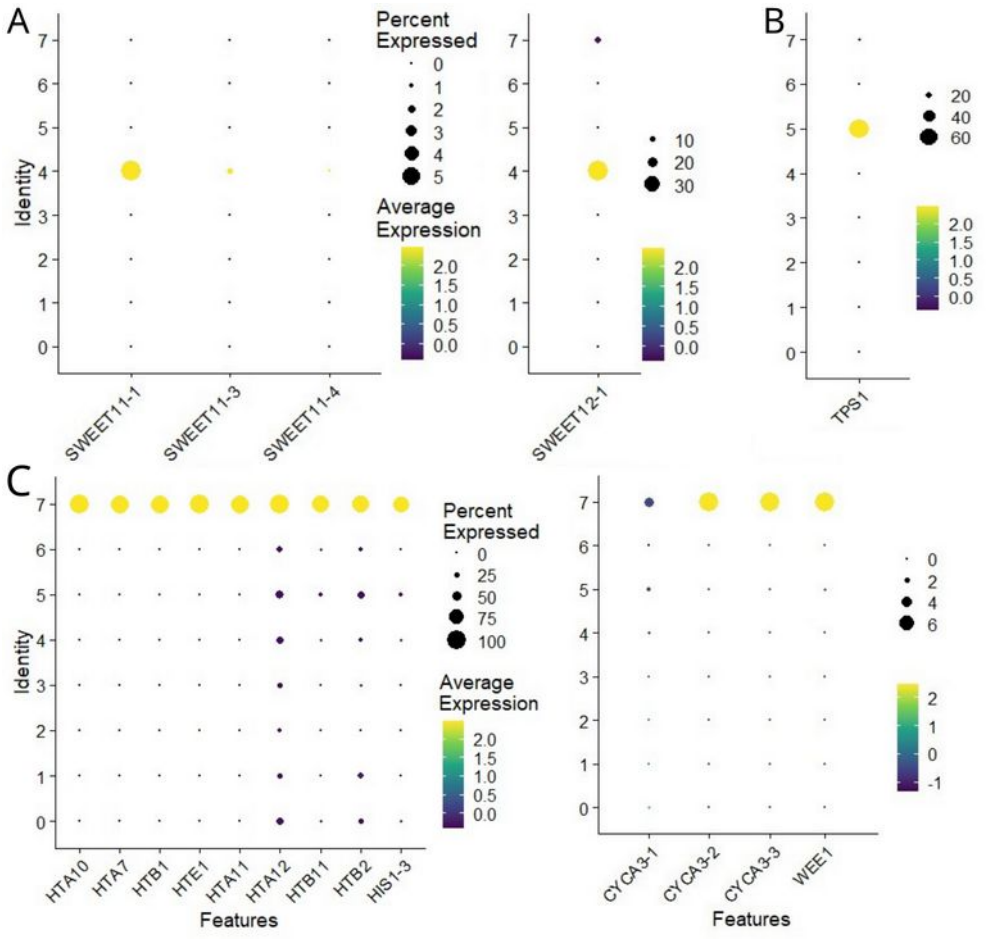


Figure 3.9: Dot plots of differential gene expression analysis of key cluster identity genes

(A) *SWEET11* and *SWEET12* sucrose efflux proteins encoding genes. **(B)** *TERPENE SYNTHASE 1* (*TPS1*) *Petunia* flower tube-specific gene. **(C)** Histone, *CYCLIN A3* and *WEE1* genes involved in cell-cyle regulation.

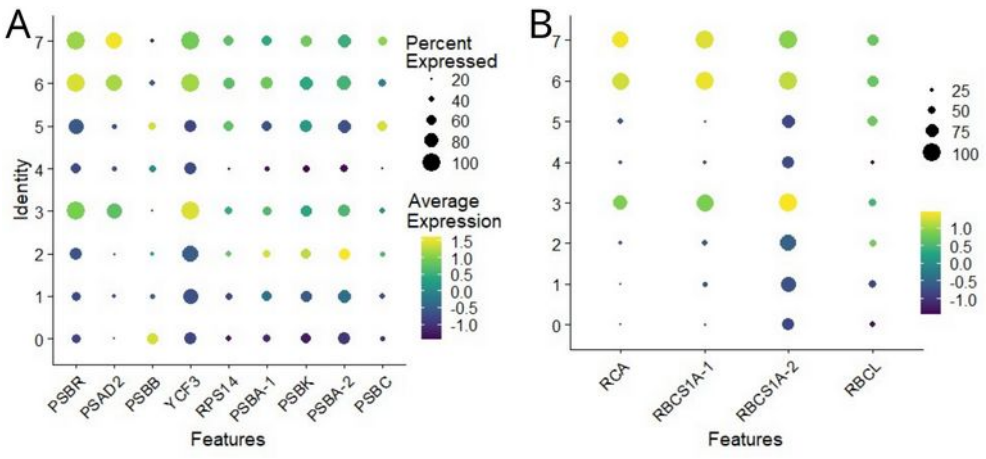


Figure 3.10: Dot plots of differential gene expression analysis of photosynthesis-related genes

(A) Genes encoding photosystems subunits. **(B)** Genes encoding RuBisCo subunits.

Surprisingly, this cluster did not appear as a clear epidermal cluster when looking at ML1, ANL2 or GL2 marker genes, suggesting that markers of epidermal identity can be quite different between epidermal cell types in a mature organ.

The cluster 4 seems to regroup vasculature cells. Indeed it shows strong expression of *SWEET11* and *SWEET12* genes encoding sucrose efflux transporter proteins involved in phloem loading (Fig. 3.9, A) (Fatima et al., 2022).

Aside from expressing known epidermis marker genes, cluster 5 also expresses a gene that was shown to be specifically expressed in *Petunia* tube. This cluster is therefore a good candidate for regrouping tube epidermal cells (Fig. 3.9, A) (Boachon et al., 2019). This would explain the lack of active anthocyanins biosynthesis while still expressing epidermis marker genes.

As stated earlier, cluster 7 stands out as active in anthocyanins biosynthesis, although not as strongly than cluster 6. Interestingly, cluster 7 shows histone-encoding genes highly overexpressed in most of its cells (Fig. 3.9, C). This suggested the presence of cells actively replicating DNA and so possibly in mitosis, and indeed, a small percentage of the cells within cluster 7 overexpress the cyclin-dependent protein activity regulator *CYCLIN A3* (*CYCA3*) (Fig. 3.9, C). The G2/M phase transition inhibitor *WEE1* is also overexpressed in some cells of cluster 7. These observations suggest that this cluster is mainly determined by its cell cycle state instead of its cell identity, which is a known problem in scRNA-Seq datasets, hence the various tools developed to correct this bias (Zheng et al., 2022). The fact that these same cells show strong anthocyanins biosynthesis activity suggests that they might constitute a small population of the epidermis actively replicating.

Petunia petal is slightly chloroplastic, especially in the tube, therefore some photosynthesis-related genes should be expressed and might exhibit cluster-specific patterns, and indeed they do. Although this plot is not very informative other than confirming most cells within *Petunia* petals are capable of photosynthesis, it is of help to confirm some of previous findings. For instance, the cluster 4 is the one showing the lowest expression and the lowest proportion of cells expressing genes encoding photosystem subunits (Fig. 3.10, A). It is also one of the clusters showing low expression for genes encoding Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo) subunits (Fig. 3.10, B). This is coherent with the suggestion made earlier that this cluster regroups vascular cells, which are very specialized internal cells, some of which are dead, and thus with no photosynthetic activity. Moreover, the striking very strong similarity in all photosynthesis-related gene expression between clusters 6 and 7 supports the idea that the cluster 7 regroups dividing epidermal cells that would otherwise belong in cluster 6 (Fig. 3.10, A-B).

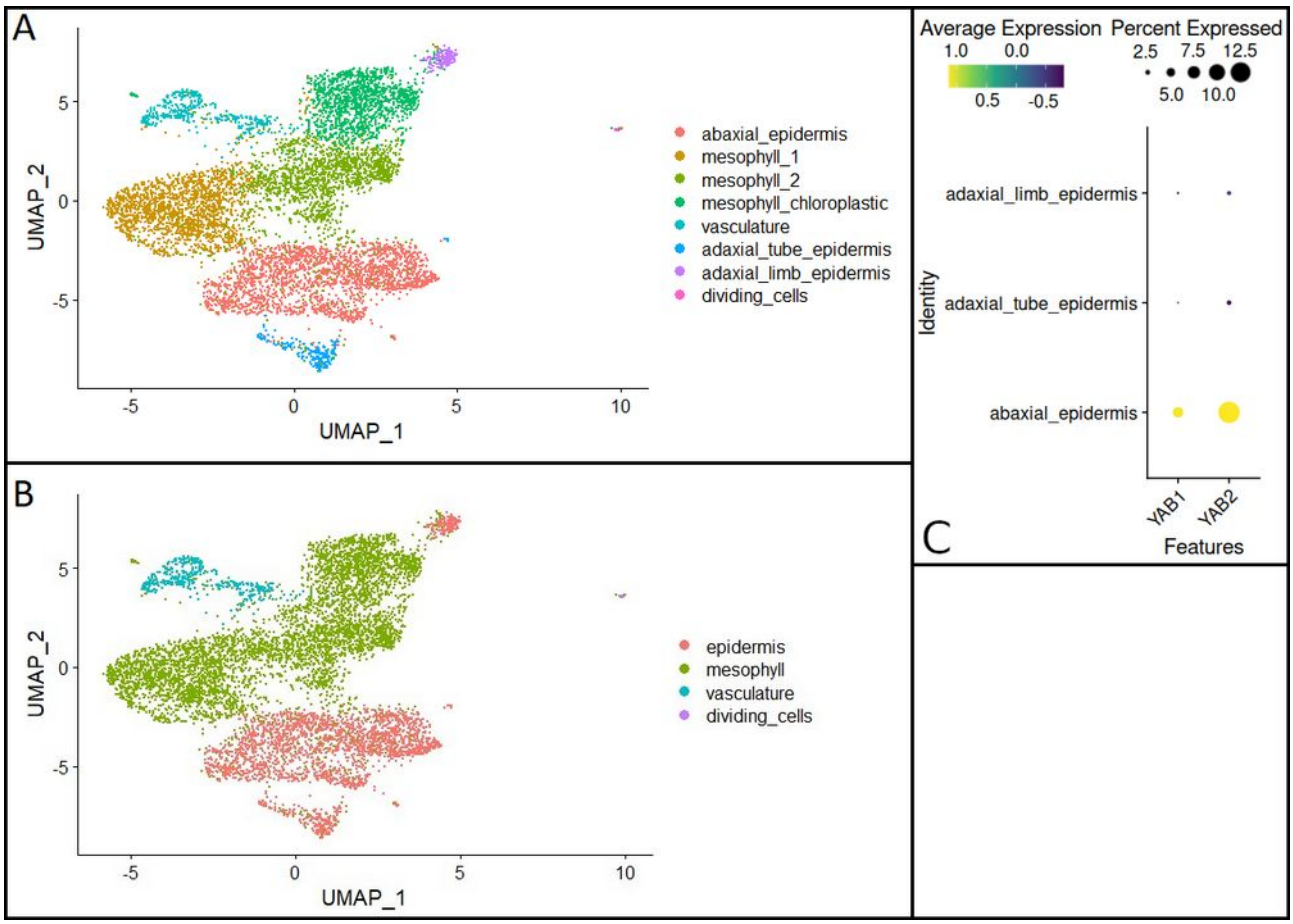


Figure 3.11: UMAP of the high-resolution and low-resolution clusterings of the dataset and dot plot of abaxial polarity marker genes

(A) High resolution clustering that allowed cluster identification. **(B)** Deduced lower resolution clustering. **(C)** Dot plot of the expression of abaxial polarity markers *YABBY1* and *YABBY2* in epidermis clusters.

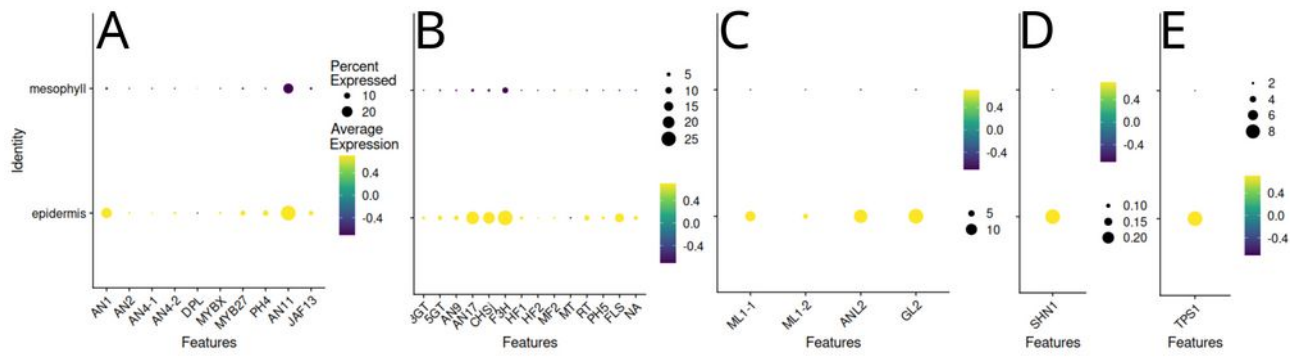


Figure 3.12: Dot plots of the expression levels of a set of key identity genes in mesophyll and epidermis.

(A) Anthocyanins biosynthesis regulator transcription factors. **(B)** Anthocyanins biosynthesis enzymes encoding genes. **(C)** Epidermal identity genes. **(D)** Conical cells marker. **(E)** Tube epidermis specific marker.

Interestingly, cluster 3 is also active photosynthesis-wise, but do not show up as epidermis. Hence, it might regroup heavy chloroplastic cells within the mesophyll (Fig. 3.10, A-B).

All aforementioned considered, I attributed the following cell-identity to the clusters (Fig. 3.11, A). Cluster 0 is a good candidate for abaxial epidermis since it is expressing epidermal markers while only mildly expressing anthocyanins biosynthesis related genes and expressing abaxial polarity markers *YABBY1* (*YAB1*) and *YABBY2* (*YAB2*) (Fig. 3.11, C). Clusters 1 and 2 do not show any striking gene expression pattern and their cells would form the bulk of the mesophyll, while the cells regrouped into cluster 3 would constitute a chloroplastic population within the mesophyll. Cluster 4 is clearly a cluster of vascular cells since it is the only one to express known phloem loading genes. Cluster 5 is an epidermis cluster, and the presence of highly expressed tube-specific gene *TPS1* allows to identify it as adaxial tube epidermis. Cluster 6 is very well defined as the most active regarding anthocyanins and therefore is the right candidate for adaxial limb epidermis identity. Finally, cluster 7 regroups cells that are actively replicating, the presence of still highly expressed anthocyanins biosynthesis genes leading to classify them as adaxial epidermis also.

Since the main driver of my PhD is to precise cell-layer-specific regulation network involving PhDEF, I regrouped the different cell identities together in order to be able to check for differential gene expression between the epidermis and the mesophyll (Fig. 3.11, B). Clusters “mesophyll_1”, “mesophyll_2” and “mesophyll_chloroplastic” were therefore grouped into “mesophyll” identity while clusters “abaxial_epidermis”, “adaxial_tube_epidermis” and “adaxial_limb_epidermis” were merged as “mesophyll”. Both clusters “vasculature” and “dividing_cells” were left as is and will not be studied further.

III.4.d – Cell-layer clustering unveils candidate cell-layer-specific PhDEF partners

A new differential gene expression analysis was carried out between “mesophyll” and “epidermis” clusters. I obtained a list of 2,016 deregulated genes, of which 1,067 are upregulated in the epidermis (Log2FC > 0.25) and 949 are downregulated in the mesophyll (Log2FC < – 0.25).

First, I checked that known marker gene expression was coherent with this new clustering. Epidermis cluster showed upregulated anthocyanins biosynthesis regulators and biosynthesis enzyme-encoding genes expression, epidermal marker genes, conical marker gene *SHN1* and tube epidermis-specific gene *TPS1*. (Fig. 3.12).

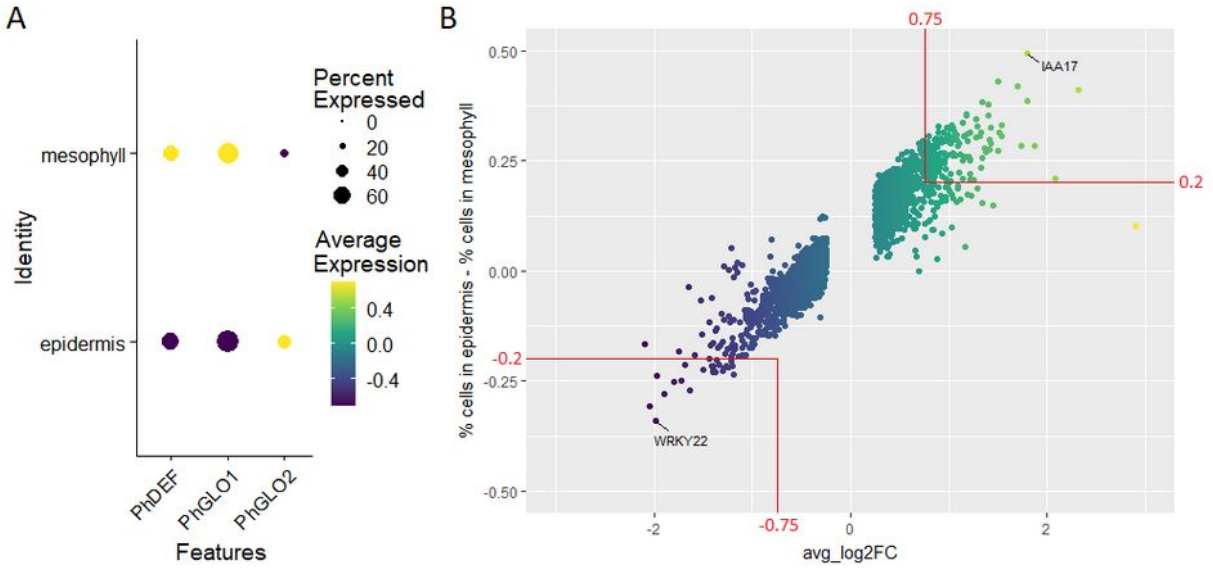


Figure 3.13: Petal-specific B-class genes expression levels in epidermis and mesophyll clusters and overall gene expression distribution in regard to their cluster specificity in epidermis and mesophyll clusters.

(A) Dot plot of the expression of the petal-specific B-class genes in mesophyll and epidermis clusters. **(B)** Scatter plot of the expression levels of all detected genes in regard to their cluster specificity. The x-axis indicates the Log2FC of gene expression in the epidermis vs. mesophyll clusters (a positive value indicates higher expression in the epidermis). The y-axis indicates the difference in the percentage of epidermis-expressing cells and the percentage of mesophyll-expressing cells, for each gene. In other words, a positive value indicates that there are more epidermal cells than mesophyll cells that express this gene.

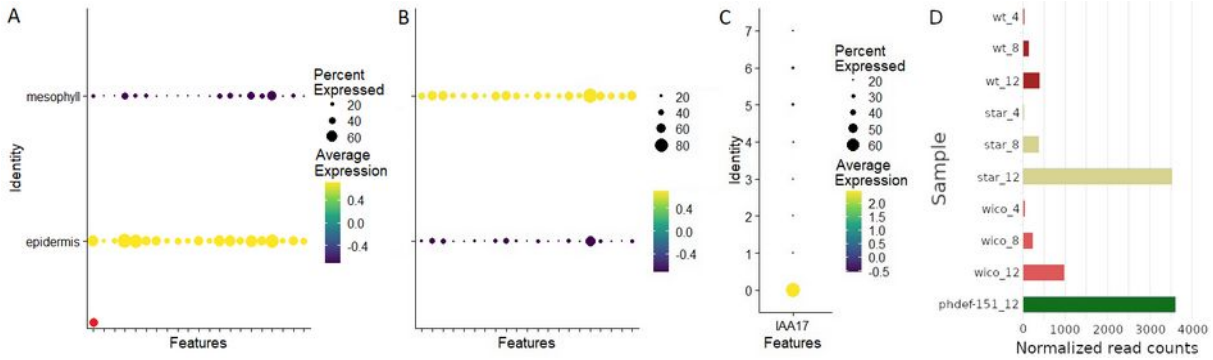


Figure 3.14: Dot plots representing epidermis- and mesophyll-specific genes expression levels in the epidermis and mesophyll clusters, and *IAA17* expression levels in high-resolution clustering and bulk RNA-Seq dataset

(A) Expression levels of the epidermis-specific genes. The red dot highlights *IAA17*. **(B)** Expression levels of the mesophyll-specific genes. **(C)** Expression levels of *IAA17* in high-resolution clustering and **(D)** bulk RNA-Seq in wild-type, *star*, *wico* and *phdef-151* petals at different stages of development (4, 8 and 12).

Interestingly, B-class genes are not uniformly expressed in mesophyll and epidermis clusters. Indeed, while *PhDEF* and *PhGLO1* are strongly expressed in the mesophyll, *PhGLO2* is not. In fact, *PhGLO2* shows the opposite expression pattern, with strong expression in epidermis cells and lower expression in the mesophyll (Fig. 3.13, A). This might be a first indication of cell-layer-specific differences in PhDEF regulation network. Indeed, one could hypothesize that petal development regulating molecular quartets containing either *PhDEF*/*PhGLO1* or *PhDEF*/*PhGLO2* would regulate different sets of genes. If *PhGLO1* is preferentially expressed in the mesophyll while *PhGLO2* is preferentially expressed in the epidermis, quartets could preferentially contain *PhGLO1* in the mesophyll and preferentially *PhGLO2* in the epidermis leading to different genes being preferentially expressed in either tissue.

In order to reduce the number of genes of interest for further study, I plotted the distribution of their expression levels in regard to their cluster specificity (Fig. 3.13, B). II arbitrarily set limits to isolate genes to be strongly cluster-specific and deregulated: the Log2FC threshold (log2 fold change of gene expression between epidermis and mesophyll cluster) was set at – 0.75 or 0.75. The difference between the percentage of cells expressing a given gene in the epidermis and the mesophyll was set to 0.2 (*i.e.* 20 % of cells). Applying these thresholds allowed the identification of 161 epidermis-specific and 21 mesophyll-specific genes. The 11 best and 10 worst epidermis marker expression are shown in (Fig. 3.14, A). All 21 mesophyll markers are shown in (Fig. 3.14, B). It is striking that much more epidermis-specific genes are found than mesophyll-specific genes, and indeed the distribution of the scatter plot (Fig. 3.13, B) is skewed towards higher values for the y-axis, meaning that there are more genes specific to the epidermis than to the mesophyll.

The best petal epidermal markers constitute interesting candidates to further investigate as playing a role in the identity or growth of the petal epidermis. As an example, the gene *IAA17* is of particular interest (Fig. 3.14, A, red dot). Indeed, in previously obtained bulk RNA-Seq data on wild-type, *phdef-151*, *star* and *wico* flowers it showed to be strongly upregulated in *phdef-151*, late *star* flower development as well as being lightly upregulated in late *wico* flower development compared to wild-type flowers (Fig. 3.14, D). It is also one of the most upregulated genes in the epidermal cluster in our scRNA-Seq data (Log2FC at 1.79), and 75% of epidermal cells vs. 26% of mesophyll cells express this gene, which is the highest difference (49%) we find in our dataset. Finally, it is also directly bound by PhDEF in our ChIP-Seq assay (see chapter 6). IAA proteins are transcription factors repressing auxin response genes at low auxin concentrations. This gene is a good candidate to play a role in petal limb expansion in late development, that is likely to be impaired in *star* and *phdef-151* when compared to wild-type considering their phenotype.

When looking at its expression pattern in the high-resolution clustering, it shows to be upregulated in the cluster 0, which has been identified as the abaxial epidermis (Fig. 3.14, C). Finding an auxin response regulator, potentially involved in cell expansion in the epidermis only, is also coherent with the idea that the epidermis is driving limb development, which we observe in the star and wico flowers (Chopy et al., 2023).

Other genes within these two groups have not yet been looked into, but they constitute good candidates for driving epidermis and mesophyll identities and therefore be affected by PhDEF in a cell-layer-specific manner. Exploring *star* and *wico* datasets from the second scRNA-Seq run is the next logical step to explore this hypothesis. However, this has not yet been accomplished yet due to lack of time, and data of lower quality.

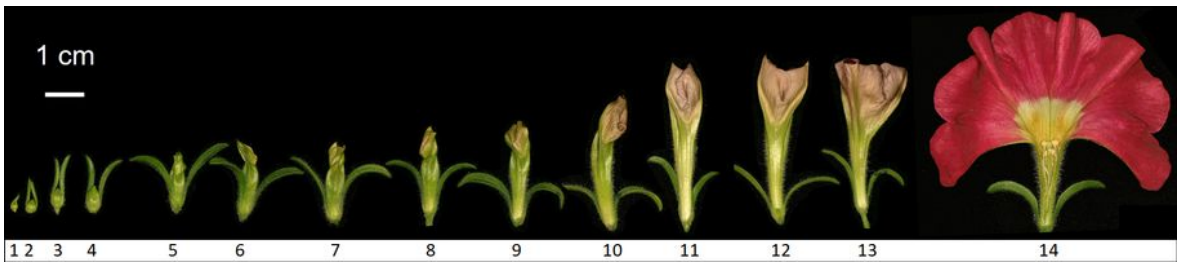


Figure 4.1: Wild-type Petunia hybrida flower development timeline

Three sepals out of five were removed. The final mature stage has been dissected open to expose internal organs. Picture provided by the courtesy of Mathilde Chopy.

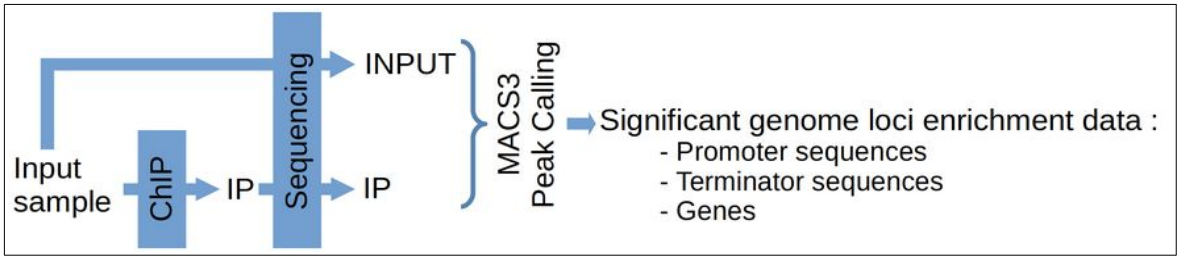


Figure 4.2: ChIP-Seq crude workflow

Chromatin is sequenced before (INPUT) and after (IP) ChIP experiment. IP and INPUT datasets are loaded into MACS3 which identifies significantly enriched genome regions in the IP sample.

Sample	WT_1	WT_2	phdef-151_1	phdef-151_2	phglo1;phglo2_1	phglo1;phglo2_1
Peak counts	8,978	5,151	7,827	1,261	29,124	67,882
After samples intersection	2,370		44		905	
In regulatory regions	1,142		0		222	

Table 6.1: Detected peaks in ChIP-Seq samples

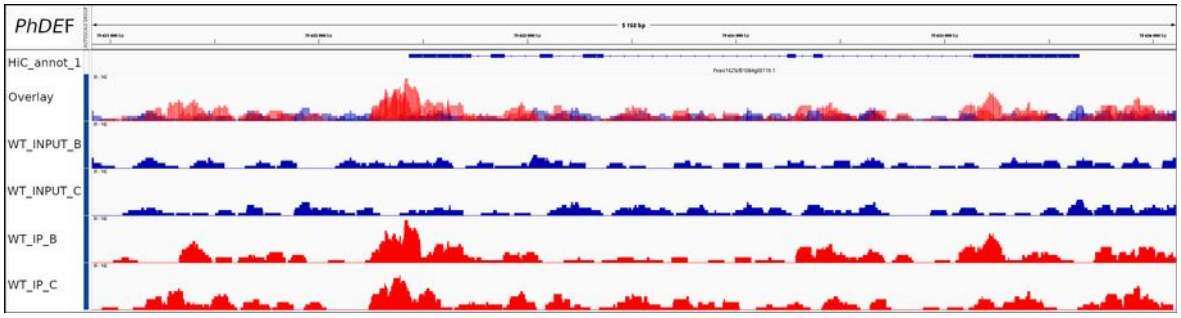


Figure 4.3: ChIP-Seq peaks visualization at *PhDEF* loci

Inputs: blue. IPs: red. Visualized with IGV (Thorvaldsdottir et al., 2013).

IV – ChIP-Seq reveals additional cues of *PhDEF* layer-specific target genes

IV.1 – Preamble and experimental design

We recently proposed that during Petunia petal development, limb growth is driven by the epidermis while tube growth is driven by the mesophyll, under a cell-layer-specific regulation network involving the B-class gene *PhDEF* (Chopy et al., 2023). The fact that PhDEF controls very different phenotypic traits when expressed in different layers, suggests that it regulates a different set of genes in these layers. As an example, we found by ChIP-qPCR that PhDEF binds to the regulatory sequence of *ANTHOCYANIN2* (*AN2*), a major regulator of petal pigmentation in the epidermis (Chopy et al., 2023). In order to identify *PhDEF* target genes at a genome-wide scale and putative differences between our mutants of interest *phdef*-151, *star* and *wico* and wild-type petal, I participated in setting up a chromatin immunoprecipitation followed by next generation sequencing (ChIP-Seq) experiment using anti-PhDEF antibody AB#1 previously validated. For each genotype of interest, two biological replicates were used as well as an additional wild-type sample not-immunoprecipitated that served as negative control. Stage 8 flower petals were used in order to capture genes that might be involved both in tube and limb growth (Fig. 4.1). Results are still very preliminary, the analysis was performed a few weeks ago and technical issues prevent me to draw conclusions in *star* and *wico* samples, but collected data is promising considering wild-type, *def-151* and *phglo1;phglo2* samples, which I will discuss now.

IV.2 – Peak calling yields coherent PhDEF target loci

This preliminary analysis was done by Brice Letcher and Carine Rey from the Laboratoire de Biologie et de Modélisation de la Cellule (LBMC, ENS de Lyon), in the context of a M2 practical course on Next Generation Sequencing. Peak calling using MACS3 (MACS3 project team, 2020) (Fig. 4.2) identified 8,978 and 5,151 peaks in wild-type replicates, 7,827 and 1,261 peaks in *phdef-151* replicates and 29,124 and 67,882 in *phglo1;phglo2* replicates as significantly enriched genome regions after anti-PhDEF ChIP. Intersecting both replicates for each genotype yielded 2,370 peaks in wild-type, 44 peaks in *phdef-151* and 905 peaks in *phglo1;phglo2*, most of which are in intergenic regions. When removing peaks detected in such regions (3 kb before transcription start sites (TSS) and 1 kb after transcription termination site (TTS)) 1,142 peaks are left in wild-type, none in *phdef-151*, and 222 in *phglo1;phglo2* (Table. 6.1). These values are coherent with one could await for the experimental design of this ChIP experiment. As illustration purposes, Fig. 4.3 shows detected peaks at *PhDEF* promoter, as per its known auto-activation (Vandenbussche et al., 2004).

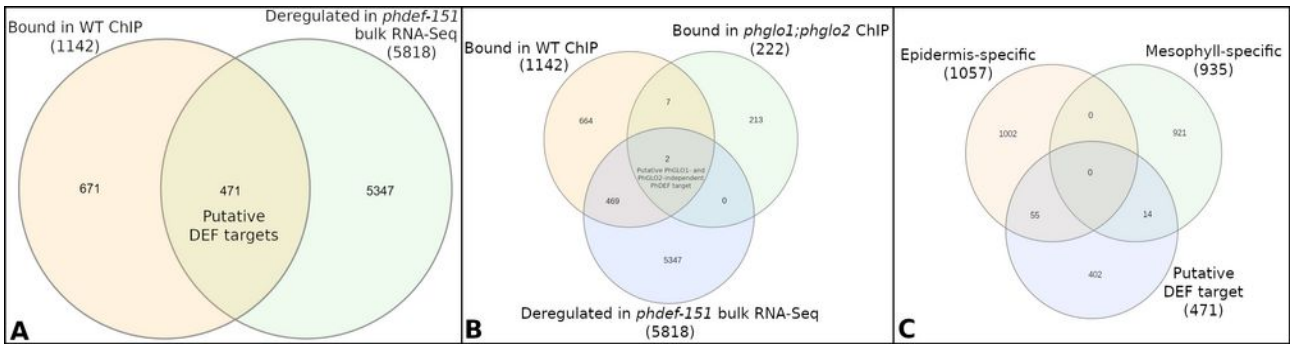


Figure 4.4: Venn diagrams of gene counts detected in ChIP-Seq, bulk RNA-Seq and scRNA-Seq assays

(A) Number of genes bound by PhDEF in wild-type petals (ChIP-Seq), number of genes deregulated in*phdef-151* (bulk RNA-Seq), and their intersection. **(B)** Same as A, but the intersection with the number of genes bound by PhDEF in *phglo1;phglo2* samples (ChIP-Seq) is shown. **(C)** Intersection between putative PhDEF targets (*i.e.* bound in ChIP in WT, and deregulated in *phdef-151* transcriptome) and epidermis- or mesophyll-specific genes detected in scRNA-Seq.

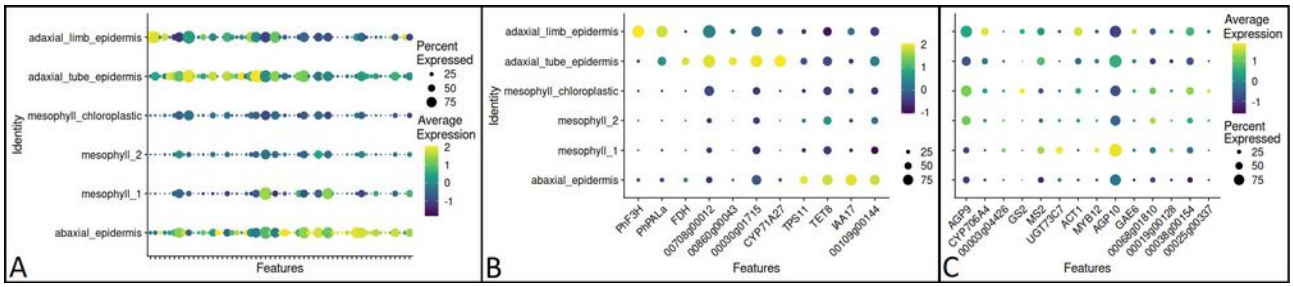


Figure 4.5: Dot plots of the tissue-specific expression of putative epidermis- and mesophyll-specific PhDEF target genes

(A) All 55 epidermal-specific genes. **(B)** A selection of 11 genes showing strong tissue specificity. **(C)** All 14 mesophyll-specific genes.

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details	Motif File
1		1e-269	-6.213e+02	27.09%	5.09%	54.0bp (70.3bp)	TCP4/MA1035.1/Jaspar(0.975) More Information | Similar Motifs Found	motif file (matrix)
2		1e-109	-2.514e+02	33.08%	14.79%	51.5bp (67.6bp)	SOC1/MA0554.1/Jaspar(0.904) More Information | Similar Motifs Found	motif file (matrix)
3		1e-74	-1.722e+02	33.50%	17.77%	54.6bp (63.9bp)	EDT1/MA0990.1/Jaspar(0.919) More Information | Similar Motifs Found	motif file (matrix)
4		1e-62	-1.448e+02	20.68%	9.24%	55.3bp (68.2bp)	An_0287(RRM)/Aspergillus_nidulans-RNCMPT00287-PBM/HughesRNA(0.773) More Information | Similar Motifs Found	motif file (matrix)
5		1e-52	-1.213e+02	24.77%	13.05%	53.6bp (65.6bp)	BIM2(bHLH)/col-BIM2-DAP-Seq(GSE60143)/Homer(0.962) More Information | Similar Motifs Found	motif file (matrix)
6		1e-47	-1.096e+02	33.04%	20.23%	53.4bp (64.3bp)	RAMOSA1/MA1416.1/Jaspar(0.942) More Information | Similar Motifs Found	motif file (matrix)

Figure 4.6: Main DNA motifs enriched in ChIP-Seq of wild-type samples

1: TCP. 2: MADS. 3: HDG. 4: SPL. 5: bHLH. 6: most likely artefactual.

We recently showed that 5,818 genes are deregulated in *phdef-151* petals (Chopy et al., 2023). When intersecting this data with the newly obtained ChIP-Seq data we find 471 genes that would be good candidates as PhDEF targets (Fig. 4.4, A). No gene is detected at the intersection between deregulated genes in *phdef-151* and *phglo1;phglo2* ChIP-Seq data (Fig. 4.4, B), suggesting that peaks detected in the *phglo1;phglo2* ChIP-Seq are artefactual, without partners, PhDEF might lose specificity and bind regions it does not in wild-type background. Finally, connecting scRNA-Seq data and ChIP-Seq data allows to identify 55 epidermis-specific and 14 mesophyll-specific putative PhDEF targets (Fig. 4.4, C). Interestingly, we find more epidermis-specific PhDEF targets than mesophyll-specific ones, suggesting that PhDEF actively specifies or reinforces epidermal identity in the petal.

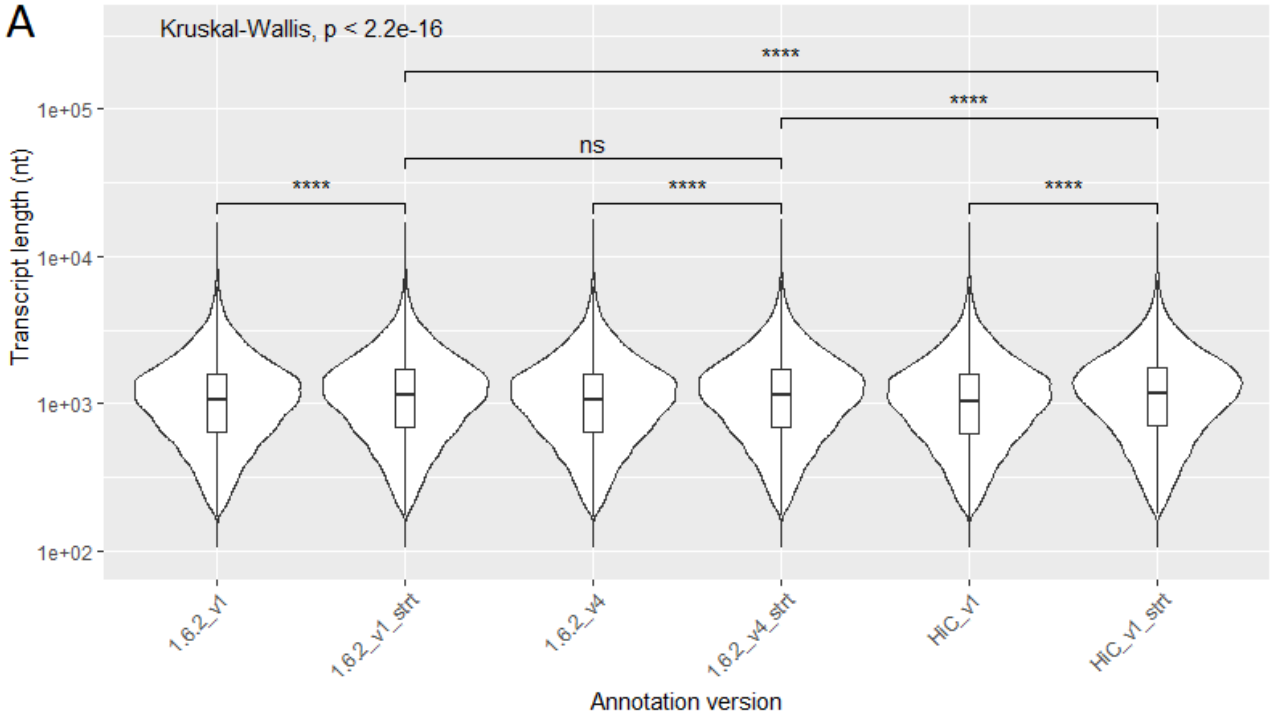
Using the high-resolution clustering from the scRNA-Seq data, we can even find cell-layer-specific genes among these putative DEF targets. Overall the expression pattern of these genes is clearly enriched in epidermal tissues, (Fig. 4.5, A), and some genes even show clear cell type-specific expression patterns (Fig. 4.5, B). The strong expression of *PhF3H*, encoding the flavanone 3-hydroxylase and *PhPALa*, encoding the phenylalanine ammonia-lyase in the limb epidermis is coherent with their known function upstream of the anthocyanins biosynthesis pathway in *Petunia*, however to my knowledge no regulators of such early steps of the pathway has been proposed so far (Bombarely et al., 2016). Although again these results are very preliminary, this finding might suggest that *PhDEF* is able to activate the anthocyanins biosynthesis pathway even more upstream than we recently proposed in (Chopy et al., 2023), which needs to be assessed further. Interestingly, *IAA17* of which we already talked about when presenting the scRNA-Seq results, is still present in the putative targets of PhDEF, supporting yet again the idea that *PhDEF* might influence cell growth through AUX/IAA signaling in a cell-layer-specific manner. The expression of the 14 mesophyll-specific putative PhDEF targets genes on the other hand, is less clear (Fig. 4.5, C). I would argue that deeper analysis is needed to unveil more subtle differentiation cues specifying the mesophyll as opposed to the obvious one of the epidermis.

Last but not least, the analysis of the ChIP-Seq data allowed to precise what type of DNA motif PhDEF might be preferentially binding to using JASPAR (Fornes et al., 2020). With no surprise, one can find among them CarG boxes, direct targets of PhDEF and other MADS-box proteins, but also motifs targets of TEOSYNTE BRANCHED/CYCLOIDEA/PROLIFERATING CELL FACTOR (TCP), HOMEODOMAIN-GLABROUS (HDG) and SQUAMOSA Promoter-Binding Protein-Like (SPL) transcription factors (Fig. 4.6).

TCP regulators have been described as MADS-box protein partners and are involved in *Arabidopsis thaliana* petal growth (Guo et al., 2013; Huang and Irish, 2016). HDG proteins, also known as HD-Zip class IV proteins, are known for their role in specifying epidermal identity (Schrick et al., 2023a) but have not been described as putative interactors for MADS proteins so far. Finally, SPL proteins have been reported to interact with MADS-box and TCP proteins and they also play a role in floral organ elongation (Egea-Cortines, 1999; Wang et al., 2016). The motifs identified are therefore coherent with the literature, but this deserves further investigation.

As previously stated, aforementioned data is still very fresh and the analysis I showed quite crude. However, I think it demonstrates that it contains valuable information, especially when crossed with bulk and single-cell RNA-Seq data, to provide deeper insights into PhDEF cell-layer-specific regulation networks.

A



B

	1.6.2_v1	1.6.2_v1_strt	1.6.2_v4	1.6.2_v4_strt	HiC_v1	HiC_v1_strt
Mean	1256	1336	1257	1337	1236	1366
Median	1056	1131	1056	1131	1039	1167
0%th quantile	105	105	105	105	105	105
25%th quantile	645	683	645	682	630	702
75%th quantile	1591	1710	1591	1709	1574	1760
100%th quantile	17082	17082	17856	17856	17087	17087

Figure 5.1: Transcripts length distribution of *P. axillaris* annotations

(A) Transcripts length distribution of annotations V1 (noted 1.6.2_v1), V4 (noted 1.6.2_v4) and in-house transcriptome (noted HiC_v1) before and after (noted strt) StringTie improvement. Kruskal-Wallis tests (Kruskal and Wallis, 1952) were performed to compare annotations before and after StringTie improvement and annotations after StringTie improvement against each other (ns: $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$). **(B)** Various metrics of transcript length distribution of annotations V1 (noted 1.6.2_v1), V4 (noted 1.6.2_v4) and in-house transcriptome (noted HiC_v1) before and after (noted strt) StringTie improvement.

V – StringTie improves predicted genome structural annotation coverage

V.1 – Preamble

As stated in chapter 3, upon browsing through preliminary scRNA-Seq results a striking problem arose, the reads mapped poorly to the *P. axillaris* reference transcriptome in use. According to 10x Genomics at least 30% of the reads should map for further analysis to be carried on effectively and the results were barely meeting this limit. Since reads showed good quality and all other metrics were fine, I wondered if the quality of the transcriptome I used was part of the problem. In order to explore this possibility, I gathered multiple datasets of bulk RNA-Seq of *P. hybrida* flowers produced by the team between 2019 and 2022 and used them with the assembler StringTie to try to improve 3 *P. axillaris* transcriptomes: annotations v1 and v4 of the published assembly (Bombarely et al., 2016) and our in-house transcriptome (Chopy et al., 2023).

V.2 – StringTie improved the predicted structural annotation coverage

This approach yielded good results considering the transcripts length distribution. Indeed, transcripts length is significantly modified in all newly constructed transcriptomes compared to their respective reference transcriptome although the overall distribution shape remains very similar (Fig. 5.1, A). In more details, the improved transcriptomes show average transcript length between 80 and 130 nucleotide (nt) longer than their initial transcriptomes, for 1.6.2_v1, v4 and HiC_v1 respectively (Fig. 5.1, B). As expected, since we removed any newly predicted transcripts not present in the reference after the use of StringTie, stringtied and non-stringtied transcriptomes show the exact same number of transcripts (Fig. 5.1, A). This last approach was decided to avoid predicting false positive transcripts since I assumed the quality of the preexisting predicted annotations would surpass that of mines.

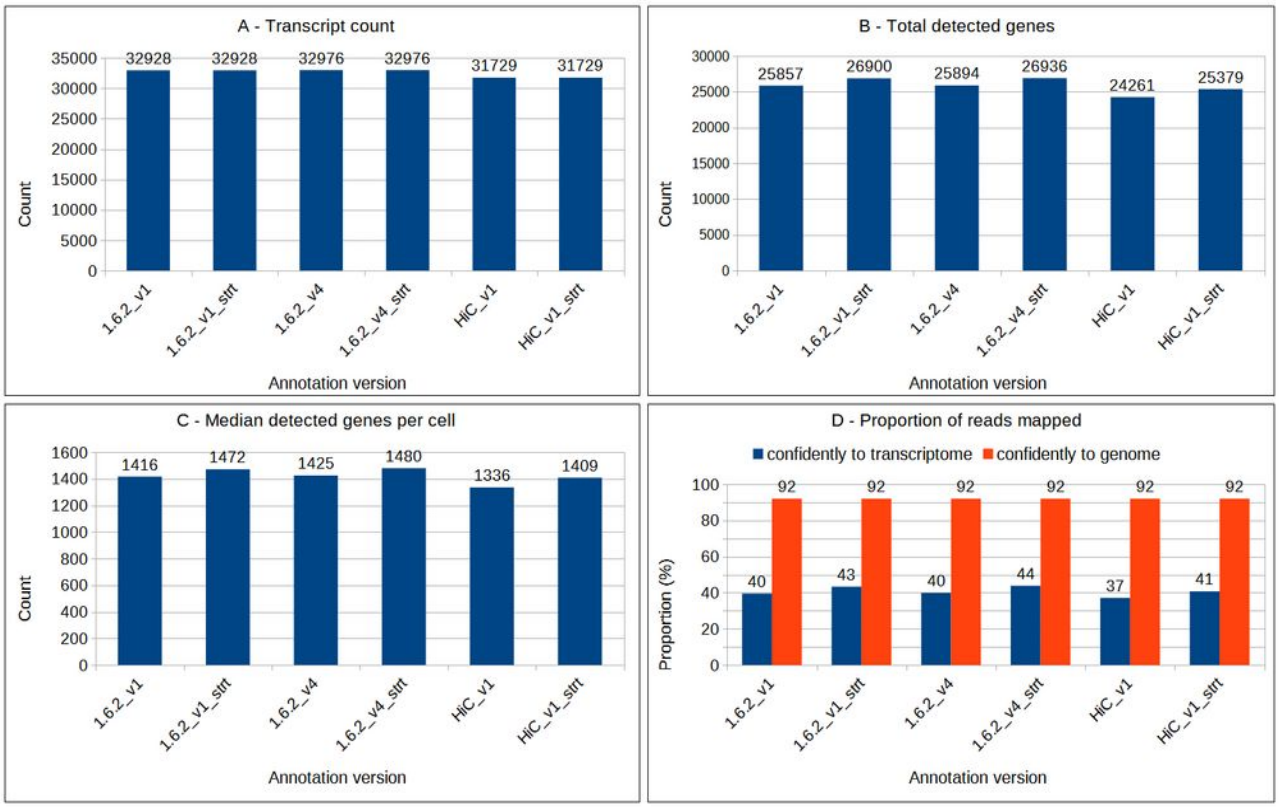


Figure 5.2: Cellranger metrics, before and after StringTie improvement.

(A) Number of predicted transcripts annotated on 6 reference annotations. From left to right, published annotation V1 of the published assembly before and after StringTie improvement, annotation V4 of the published assembly before and after StringTie improvement, in-house transcriptome before and after StringTie improvement. **(B-D)** The same scRNA-Seq run was mapped onto 6 reference transcriptomes using cellranger run and their outputs metrics compared. From left to right, published annotation V1 of the published assembly before and after StringTie improvement, annotation V4 of the published assembly before and after StringTie improvement, in-house transcriptome before and after StringTie improvement. **(B)** Total number of detected genes for the given assay for each reference transcriptome used. **(C)** Median number of detected genes per cell for the given assay for each reference transcriptome used. **(D)** Proportion of reads confidently mapped to the reference genome (FASTA file) and transcriptome (GTF file) for the given assay for each reference transcriptome used.

Once I showed that StringTie indeed improved the coverage of predicted transcripts, I assessed whether it would help with scRNA-Seq data mapping. To this aim, I generated 6 reference transcriptomes usable by “cellranger count” using “cellranger mkref” on the 6 annotations I had decided to work with: 1.6.2_v1, 1.6.2_v4, HiC_v1 and their stringtied counterparts 1.6.2_v1_strt, 1.6.2_v4_strt, HiC_v1_strt. Then, for each of these 6 references and the option --force-cells to force the detection of 4000 cells to ensure maximum reproducibility between runs, I ran “cellranger count” on the same scRNA-Seq run. The main metrics of interest that depend on read mapping are gathered in Fig. 5.2. and were gathered from the output summary of each run.

First of all, the number of total detected genes (*i.e.* in all cells) is increased in stringtied annotations compared to their original counterparts by about 4% (Fig. 5.2, B). The median number of genes detected in each cell is also slightly up by 4% using stringtied reference annotations (Fig. 5.2, C). Last but not least, the proportion of reads mapping with good confidence to the reference transcriptome is also up between 3 and 4% (Fig. 5.2, D).

All these results show that I managed to improve scRNA-Seq reads mapping onto my reference transcriptomes but that the improvement is underwhelming. The 30% minimum mapping threshold recommended by 10x Genomics is met, since all references show around 40% mapping rates, but the 50% to 80% usually awaited is not. These results show that the quality of the reference transcriptome in use, although still improvable as are all predicted annotations of any genome assembly, is not to blame regarding the poor reads mapping, or at least not its main driver. To this date, it is still unclear to me why a low fraction of my reads map to the transcriptome.

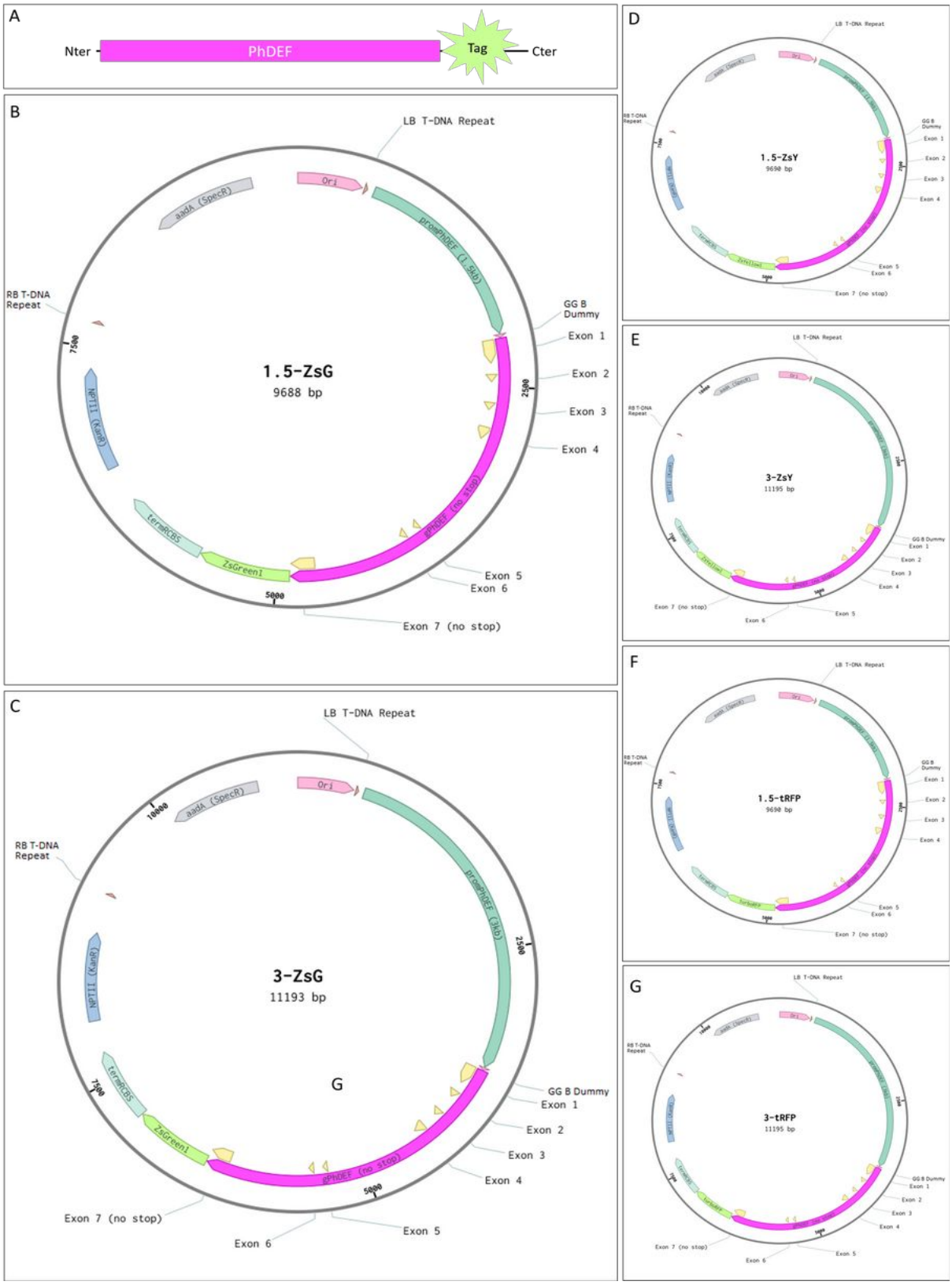


Figure 6.1: Schematic view of tagged PhDEF and designed plasmid maps

(A) Schematic view of PhDEF and a tag fused to its C-terminal region. **(B-G)** Maps of all designed plasmids (see supplementaries for GenBank format maps (Benson et al., 2013)).

VI – Assessing PhDEF protein localization in the petal

VI.1 – Preamble

As described in (Chopy et al., 2023), *PhDEF* displays a typical B-class gene expression pattern in all cell layers throughout petal development, its expression being strong at all developmental stages. However, PhDEF protein localization is still unknown in *Petunia*. This is an important point to address, since transcripts and protein levels are often uncorrelated (Maier et al., 2009). Moreover, we do not know if the PhDEF protein is equally represented in all cell layers of the wt flower; and in star and wico flowers, the presence of the PhDEF protein in layers that do not express the *PhDEF* gene (by protein movement between layers) could strongly change our interpretations of phenotypes. In order to gather more insights in this regard, I designed two approaches. First, I generated a set of *Petunia* reporter lines expressing PhDEF fused to a fluorescent marker in C-terminal (Fig. 6.1, A). Second, I performed immunolocalization assays on flower buds cross sections using a custom anti-PhDEF antibody. The first approach aimed at exploring PhDEF protein localization *in vivo* in wild-type plants and the second at visualizing its localization in mutants also, especially *star* and *wico*. Unfortunately, none of these two approaches or their many variations I tried yielded any exploitable results.

VI.2 – PhDEF reporter lines

VI.2.a – Plasmids design

Six different plasmid designs were assembled using Green Gate cloning (Lampropoulos et al., 2013), a derivative of Golden Gate cloning (Engler et al., 2008), that allows the assembly of up to 6 modules and a destination vector in a single reaction. As fluorescent tag proteins I settled on using turboRFP, ZsGreen1 or ZsYellow1 since all three were proven to work well in *Petunia* Mitchell line (Cho et al., 2019). The transposon line W138 is practically un-transformable for reasons still unclear to this day (Vandenbussche et al., 2016), hence the use of the Mitchell line, which usually show good results. The second and last variable part of the different constructs is the promoter sequence I used, corresponding either to the 1.5 kb or 3.0 kb sequence upstream of *PhDEF* loci in *Petunia x hybrida* genome. Maps of all designed constructs are showed in Fig. 6.1, B-G. The two only changes between the different constructs are the use of three different fluorescent proteins and two promoter sequences.

Figure 6.2: *Petunia hybrida* Mitchell flowers from of selection of the transformed lines showing petal development defects indicative of *PhDEF* silencing

(A-C) Recombinant B-class mutant-like flowers with poorly developed petals. **(D)** Recombinant wild-type-like flower. Scale bars: 1 cm.

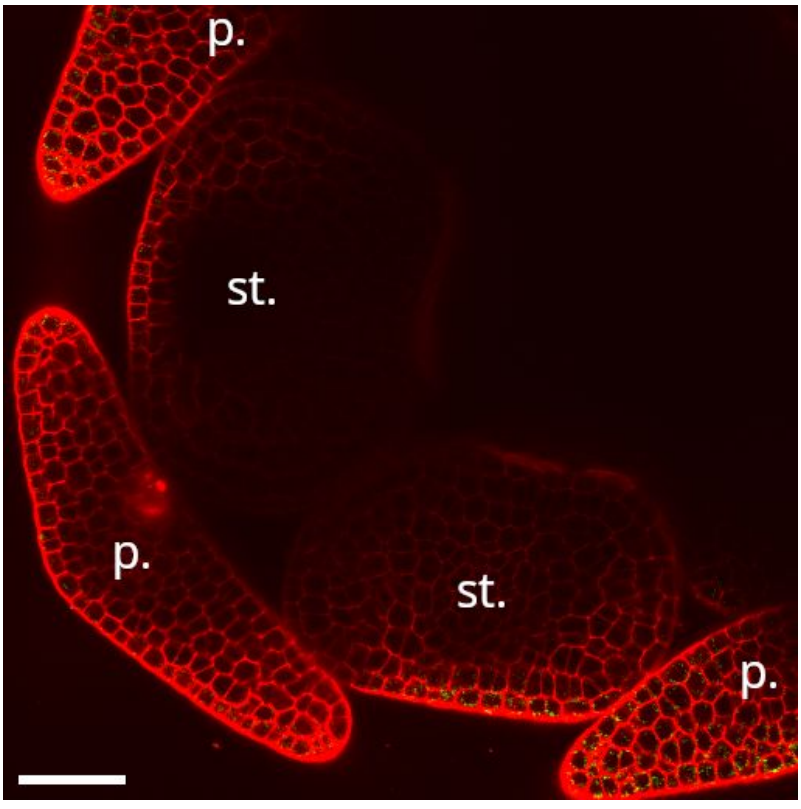


Figure 6.3: Top view composite image of a *Petunia hybrida* Mitchell early flower bud of PhDEF-ZsGreen1 recombinant line imaged in confocal microscopy

Cell walls are stained with Propidium Iodide and showed in red fluorescence. PhDEF-ZsGreen1 signal is in green (nothing but slight noise is visible since no signal was detected). P.: petals. St. Stamens. Scale bar: 50 µm.

VI.2.b – No signal for either construct was detected in transformed plants

Approximately 30 plantlets of each transformation batch were transferred into rooting medium during *in vitro* culture, of which a total of 48 plants were PCR screened as positive for construct insertion and later grew on soil. Specifically, fifteen 1.5-PhDEF-ZsGreen1, five 3.0-PhDEF-ZsGreen1, sixteen 1.5-PhDEF-ZsYellow1, eight 3.0-PhDEF-ZsYellow1, two 1.5-PhDEF-tRFP and two 3.0-PhDEF-tRFP plants were validated by genotyping.

Interestingly, a dozen plants across PhDEF-ZsGreen1 and PhDEF-ZsYellow1 transformed lines displayed flowers with petal development defects ranging from B-class mutant-like homeotic conversion of petals into sepals (Fig. 6.2, A) to wild-type-like flowers (Fig. 6.2, D) with a large variety of intermediate phenotypes (Fig. 6.2, B, C). Such observations suggest that the presence of the transgenic construct triggered *PhDEF* silencing strongly enough to degrade even endogenous *PhDEF* mRNAs. Seeing such strong silencing in a set of the plants is ambivalent. On the one hand, it means the inserted reporter gene is transcribed into mRNAs, which is good, but on the other hand if silencing is strong enough to even silence endogenous *PhDEF*, it would surely prevent the synthesis of PhDEF fusion proteins preventing imaging.

Unfortunately, the latter seems to have occurred. All plants were checked for fluorescence signal in very early flower buds under confocal microscopy and none showed neither ZsGreen1, nor ZsYellow1 nor turboRFP signal under appropriate observation parameters. There was no difference between plants showing signs of heavy *PhDEF* silencing and the ones not showing it. A representative image of the results systematically obtained while imaging the different lines is shown in Fig. 6.3. Red fluorescence shows cell walls stained with Propidium Iodide a few minutes prior imaging. In the featured image, ZsGreen1 signal would be green but only slight noise, due to the use of high gain parameters in order to gather as much signal as possible, is visible, indicating that no PhDEF-ZsGreen1 protein is present in the observed flower bud.

Previous attempts at creating reporter lines for PhDEF localization, with slightly different constructs, did not yield good results either although better than the ones I just presented. At the time, among all transformed plants, only two showed fluorescence and PhDEF localization was spotty and confined to the L1 (and therefore not heritable in the next generation).

Whether it is indeed silencing or other mechanisms, creating PhDEF reporter lines has proven to be an unexpected challenge. The use of a strong terminator (tRBCS terminator from pea) in my constructs might result in high transgene expression, and replacing it with the endogenous *PhDEF* terminator might mitigate this effect. In any case, immunolocalization seemed like another

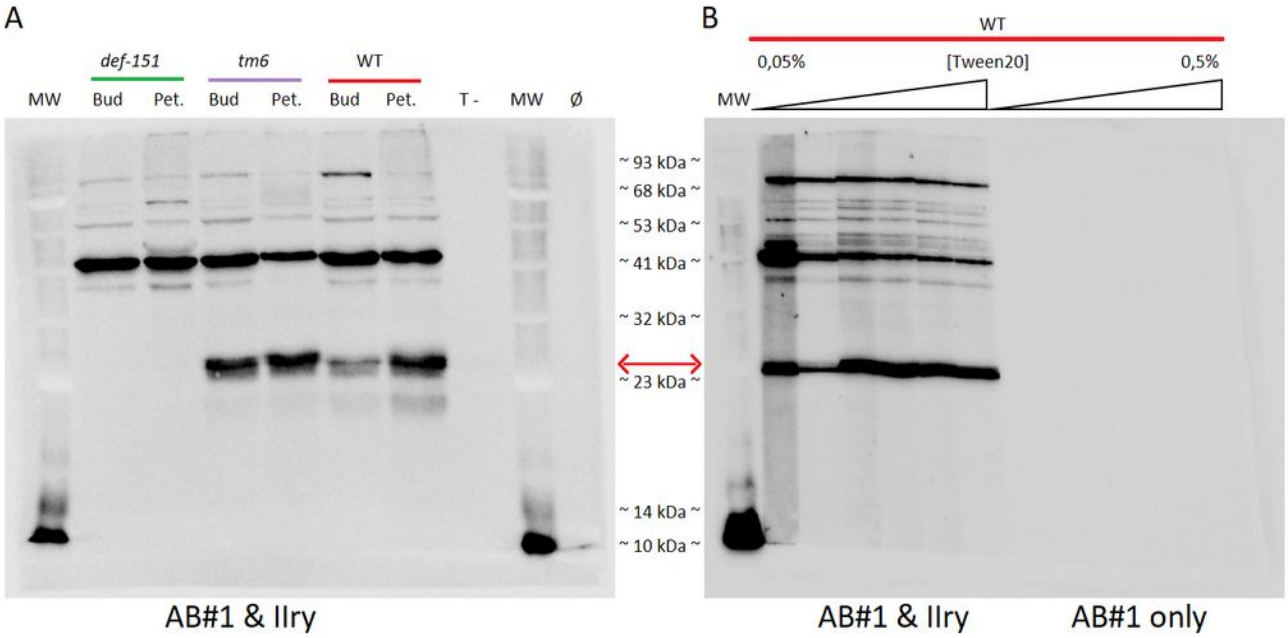


Fig. 6.4: Western Blot (WB) used for anti-PhDEF antibody AB#1 validation

(A) WB on wild-type, *phdef-151* and *phtm6* mutants flower bud and petal protein extract using AB#1 primary and HRP-conjugated anti-rabbit secondary antibodies. **(B)** WB on wild-type flower petal protein extract using AB#1 primary and HRP-conjugated anti-rabbit secondary antibodies on the left half (AB#1 & IIry) and only secondary antibody on the right half (IIry only). For each blot, and from left to right, six tracks are visible corresponding to six Tween-20 concentrations (0.05, 0.10, 0.20, 0.30, 0.40 and 0.50 %) applied during washing steps. MW: molecular weight ladder. Bud: flower bud. Pet.: flower petal. T-: no antibody negative control. Ø: empty gel well. PhDEF bands are shown by the red double arrow.

valid approach to gather valuable insights regarding PhDEF localization in wild-type as well as *star* and *wico* mutants.

VI.3 – PhDEF immunolocalization

VI.3.a – Anti-PhDEF custom antibody production and validation

Two custom rabbit anti-PhDEF polyclonal antibodies were produced by Proteogenix using a modified PhDEF protein as antigen. The MADS domain of the protein was removed in order to prevent antibody aspecificity due to the conserved nature of this epitope. Both antibodies (AB#1 and AB#2) were tested in Western Blot (WB) using wild-type, *phdef-151* or *phtm6* total proteins extracts of early petals or flower buds to check their sensitivity and specificity. The presence of *phtm6* mutant is to test whether the antibodies would aspecifically recognize PhTM6 protein since it is the closest paralog of PhDEF. In WB, AB#2 signal was extremely weak and is not showed here. AB#1 on the other hand shows a strong band in wild-type flower buds and petals at the awaited molecular weight (Fig. 6.4, A, WT, red arrow). The same band is absent in *phdef-151* mutant, showing that it is indeed detecting PhDEF protein (Fig. 6.4, A, *def-151*). Finally the same band of similar intensity is still present in *phtm6* mutant, showing AB#1 is not preferentially binding PhTM6 protein (Fig. 6.4, A, *phtm6*, red arrow). However, some aspecific high molecular weight bands are present, the most severe one spotted at around 45 kDa (Fig. 6.4, A).

In order to further assess AB#1 characteristics in regard of the aspecific bands previously described, I tried to increase the stringency of the washing buffer during WB preparation by using increasing Tween-20 concentration (Fig. 6.4, B, left). The second lane corresponding to 0.10 % Tween-20 concentration shows reduced intensity for all bands indicative of lower overall antibody binding, suggesting an artefactual effect. Lanes corresponding to all other Tween-20 concentrations however show a reduction of the intensity of all aspecific bands proportional to the concentration of Tween-20, the PhDEF band remaining essentially unaffected. I also checked if the aspecific bands were possibly caused by the secondary antibody, but the total absence of signal when using only the secondary antibody without AB#1 as a primary showed the aspecific bands were indeed caused by AB#1 and not the HRP-conjugated secondary one (Fig. 6.4, B, right).

The fact that increasing stringency reduced aspecific bands intensity in WB, a technique for which denaturant conditions and very high protein concentration originating from all cellular compartments tend to favor aspecificity, suggested that AB#1 is a valid antibody for immunolocalization and ChIP-Seq assays, while still keeping in mind the aspecificity it showed in WB. Since AB#2 wasn’t able to detect PhDEF in WB it was not used in later experiments.

Section thickness (µm)	Blocking	Rinse 1	Ier AB (@PhDEF)	Rinse 2	Rinse 3	Rinse 4	IIer AB (@rabbit IgG)	Rinse 5	Rinse 6	Rinse 7	Staining	Rinses 8, 9, 10
2-4	60 min. TBS-BSA 2%	5 min. TBS	1/10e in TBS-BSA 0.1% Overnight 4°C 1/10e in TBS-BSA 0.1% 2 h RT°	5 min. TBS	5 min. TBST 0.1%	5 min. TBS	1/1000e in TBS-BSA 0.1% 1 h RT°	5 min. TBS	5 min. TBST 0.1%	5 min. TBS	0.35% CFW In qqH2O 10 min.	3 x 5 min. in TBS
2-4					10 min TBST 0.1%				10 min TBST 0.1%			
2-4					15 min TBST 0.1%				15 min TBST 0.1%			
2-4					5 min. TBST 0.1%				5 min. TBST 0.1%			
2-4					10 min TBST 0.1%				10 min TBST 0.1%			
2-4					15 min TBST 0.1%				15 min TBST 0.1%			
2-4	60 min. PBS-BSA 2%	5 min. PBS	1/10e in PBS-BSA 0.1% Overnight 4°C 1/10e in PBS-BSA 0.1% 2 h RT°	5 min. PBS	5 min. PBST 0.1%	5 min. PBS	1/1000e in PBS-BSA 0.1% 1 h RT°	5 min. PBS	5 min. PBST 0.1%	5 min. PBS	0.35% CFW In qqH2O 10 min.	3 x 5 min. in PBS
2-4					10 min PBST 0.1%				10 min PBST 0.1%			
2-4					15 min PBST 0.1%				15 min PBST 0.1%			
2-4					5 min. PBST 0.1%				5 min. PBST 0.1%			
2-4					10 min PBST 0.1%				10 min PBST 0.1%			
2-4					15 min PBST 0.1%				15 min PBST 0.1%			
6-10	60 min. TBS-BSA	5 min. TBS	1/10e in TBS-BSA 0.1% Overnight 4°C 1/10e in TBS-BSA 0.1% 2 h RT°	5 min. TBS	5 min. TBST 0.1%	5 min. TBS	1/1000e in TBS-BSA 0.1% 1 h RT°	5 min. TBS	5 min. TBST 0.1%	5 min. TBS	0.35% CFW In qqH2O 10 min.	3 x 5 min. in TBS
6-10					10 min TBST 0.1%				10 min TBST 0.1%			
6-10					15 min TBST 0.1%				15 min TBST 0.1%			
6-10					5 min. TBST 0.1%				5 min. TBST 0.1%			
6-10					10 min TBST 0.1%				10 min TBST 0.1%			
6-10					15 min TBST 0.1%				15 min TBST 0.1%			
6-10	60 min. PBS-BSA 2%	5 min. PBS	1/10e in PBS-BSA 0.1% Overnight 4°C 1/10e in PBS-BSA 0.1% 2 h RT°	5 min. PBS	5 min. PBST 0.1%	5 min. PBS	1/1000e in PBS-BSA 0.1% 1 h RT°	5 min. PBS	5 min. PBST 0.1%	5 min. PBS	0.35% CFW In qqH2O 10 min.	3 x 5 min. in PBS
6-10					10 min PBST 0.1%				10 min PBST 0.1%			
6-10					15 min PBST 0.1%				15 min PBST 0.1%			
6-10					5 min. PBST 0.1%				5 min. PBST 0.1%			
6-10					10 min PBST 0.1%				10 min PBST 0.1%			
6-10					15 min PBST 0.1%				15 min PBST 0.1%			

Table 5.1: Summary of the different immunolocalization protocols tested

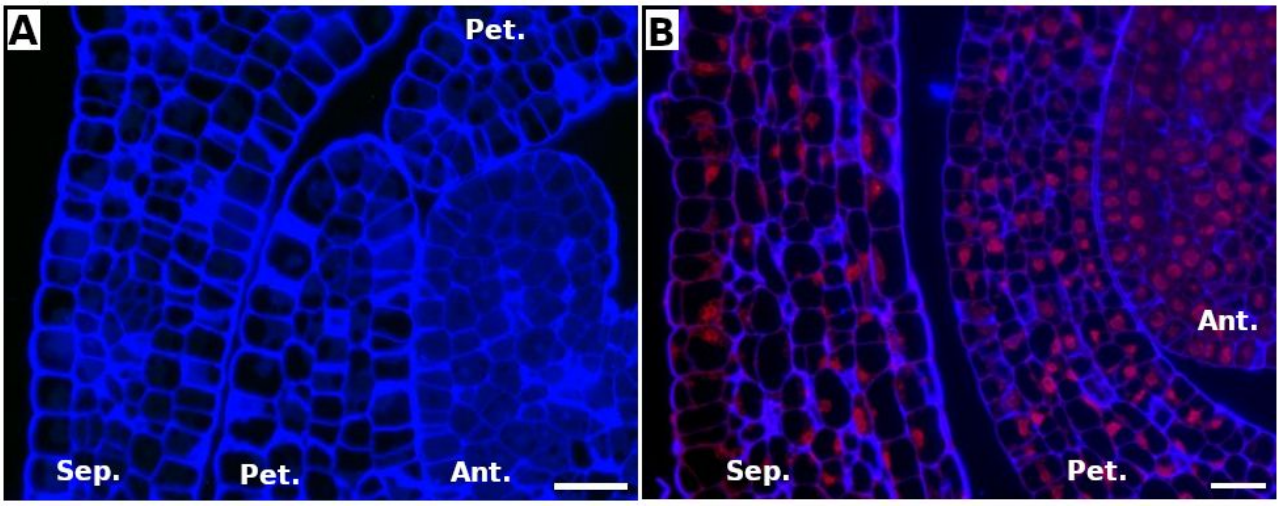


Fig. 6.5: Epifluorescence microscopy of wild-type early flower bud sections after immunolocalization

(A) Immunolocalization against PhDEF. PhDEF signal is in green (nothing but slight noise is visible since no signal was detected). **(B)** Immunolocalization against H3 histones (positive control) using a commercial antibody. H3 histones signal is in red. Cell walls are stained by Calcofluor White and showed in blue fluorescence in both images. Sep.: Sepal. Pet.: Petal. Ant.: Anther. Scale bars: 20 µm.

VI.3.b – Immunolocalization assays

I tested several immunolocalization protocols in order to visualize PhDEF in wild-type, *star* and *wico* young flowers resin-embedded cross sections (Table 5.1). I found that using Tris-HCl-Buffered Saline (TBS) instead of Phosphate-Buffered saline (PBS) yielded cleaner results and stronger nuclear signal when using the positive control anti-H3 histone antibody, so most tests were done using TBS. However, in none of the protocols I tried was I able to see any anti-PhDEF signal. Positive controls anti-H3 histone immunolocalization on test samples were successful using TBS and expected typical nuclear fluorescence pattern was observed (Fig. 6.5, B), but no signal was ever detected when using AB#1 anti-PhDEF antibody in either protocol I tested under epifluorescence (Fig. 6.5, A) or confocal microscopy.

Since both AB#1 and the anti-H3 histones antibodies I used were rabbit antibodies, the same secondary AF-568-conjugated anti-rabbit antibody was used for both PhDEF and H3 histones immunolocalization, although on different sections, showing that the absence of PhDEF signal is not caused by the secondary antibody. Different explanations can be imagined regarding such results and compared to the positive results obtained in WB. First, AB#1 target epitopes of PhDEF protein could be inaccessible outside of denaturant conditions either because of protein interactions involving PhDEF, or more simply because of the 3D conformation of PhDEF *in vivo*. Second, the chemical fixation could degrade PhDEF too much for it to be recognized by AB#1. Third, the problem could be the embedding resin I used, although I tested two different types, both of which were shown to work well for immunohistochemistry in plants.

Additional tests could be performed. To test wether it is a problem of protein conformation, one could try native WB that would allow to test AB#1 in non-denaturant conditions as proposed in (Sakuma et al., 2022). One could also try more variations of immunolocalization protocols, for instance using different blocking buffers (milk or glycine based, etc.) or different washing buffers (ASE (Rosas-Arellano et al., 2016), MTSB (Pasternak et al., 2015), etc.). Other fixation methods could also be tested, such as methanol fixation or cryofixation such as what was performed in Petunia (Wittich et al., 1999). I actually made a few trials of light fixation using acetic acid followed by cryofixation into liquid nitrogen, embedding in OCT resin and cryomicrotomy cross sectioning. But the structure of cut organs was not preserved well enough so I did not try immunolocalization afterwards.

DISCUSSION AND PERSPECTIVES

I – Conclusions

In this manuscript, I showed how the novel approach that is scRNA-Seq can be applied to study *Petunia* petal cell-layer- and cell-identity cues as well as to lay the groundwork for the study of petal development. Some technical aspects however, still need some adjustment to broaden its use to smaller-size samples, such as earlier flower developmental stages or *phdef-151* mutant.

I was able to identify key cell types and cell layers within the wild-type petal and perform cell-layer-specific differential gene expression (DGE) analysis between the mesophyll and the epidermis, or even between individual cell types in higher level clustering resolutions.

Crossing scRNA-Seq, ChIP-Seq and bulk RNA-Seq data, I identified cell-layer-specific gene expression patterns which constitute a solid resource for further exploration and provide a better understanding of *PhDEF* cell-layer-specific regulation networks.

In the following few pages, I will provide some ideas for technical and analytical improvement of the scRNA-Seq pipeline, and present some of the future work that should be undertaken in order to complete the results presented in this manuscript. I will also try to discuss how key findings of my work integrate in the existing literature as well as show their limitations.

II – A putative link between *PhDEF*, limb opening and pigmentation: *IAA17*

Most of the developmental aspects my thesis originally aimed to explore (*i.e.* the gene regulatory networks involved in tube *vs*. limb development) needed to be set aside due to technical difficulties. Nevertheless, I identified the Auxin/Indole-3-Acetic Acid (AUX/IAA) protein encoding gene *IAA17* as specifically expressed in petal abaxial epidermis at flower anthesis, as well as a putative target of PhDEF in ChIP-Seq, therefore as a potential actor of *Petunia* petal limb development shown to be driven by the epidermis (Chopy et al., 2023).

It is important to keep in mind that, as most genes in *Petunia*, *IAA17* has not yet been characterized in *Petunia* and its identification in our dataset is based on orthology with *Arabidopsis thaliana*. As a matter of fact, depending on the annotation version, *Peaxi162Scf00328g00118* is either annotated as *IAA14* or *IAA17* ortholog in *Arabidopsis*. A proper phylogeny of *IAA* genes in *Petunia* and *Arabidopsis* would be needed to identify its closest ortholog with better certainty.

AUX/IAA proteins are repressors of Auxin Response Factors (ARFs) which regulate, among others, cell elongation and anthocyanins biosynthesis (Li et al., 2016; Wang et al., 2018). *IAA17* has been described as involved in cell elongation and anthocyanins biosynthesis repression (Hou et al.,

2022; Jing et al., 2023). The presence of high expression levels of *IAA17* in the abaxial epidermis of *Petunia* flower could therefore be a part of the regulation pathway responsible for its weaker pigmentation compared to the adaxial epidermis. Moreover since *IAA17* represses cell elongation, it is also a good candidate to explain how the flower opens at anthesis. Repressing cell elongation in the abaxial epidermis but not in the adaxial one should indeed bend the limbs open, due to differences in resulting growth between the two tissues. Moreover, I found the expression of *IAA17* to be highly upregulated in *star* flowers at late developmental stages compared to wild-type flower. Its higher expression could be a driver for the total loss of pigmentation observed in *star* as well as the very poorly developed limbs. The *star* flower remains mostly closed at anthesis, which could be explained by the absence of cell elongation, especially in the abaxial epidermis.

IAA17 is therefore a good target gene candidate of PhDEF which could repress its expression to allow cell elongation and pigmentation. This hypothesis could be tested by identifying an *iaa17* mutant in our transposition mutant collection and observe both petal development and pigmentation. Although it is important to note that *IAA*s are very redundant and a single mutant might not show any phenotype.

III – Putative distinct regulation levels of the anthocyanin pathway by *PhDEF*

PhPALa encodes the phenylalanine ammonia-lyase, the enzyme that catalyzes the transformation of phenylalanine into cinnamic acid, which is the very first step of the phenylpropanoids pathway (Liu et al., 2018). This pathway branches into several others, among which the pathways producing volatile compounds, lignins and anthocyanins (Bombarely et al., 2016). I showed that *PhPALa* is a putative PhDEF target and is upregulated specifically in the adaxial *Petunia* petal epidermis. Hence, although this remains speculative, one could hypothesize that *PhDEF* is able to regulate the phenylpropanoids pathway at its very beginning and influence all downstream pathways, especially supporting the anthocyanins and volatile compounds biosynthesis activity which is very active in *Petunia* epidermis.

Among these pathways, PhDEF directly activates the expression of the major regulator *AN2* controlling the anthocyanin biosynthesis pathway (Chopy et al., 2023). My work suggests that PhDEF might also be capable of directly activating the expression of *PhF3H*, encoding the flavanone 3-hydroxylase which catalyzes the transformation of naringenin into dihydrokaempferol, considered to be the entry point into the anthocyanins biosynthesis pathway.

All things considered, one can hypothesize that *PhDEF* is regulating the anthocyanins pathway at multiple levels: indirectly via *PhPALa* expression promotion but also directly through *AN2* and *PhF3H* activation.

IV – Different interactors for PhDEF in the two cell layers (*PhGLO1*/*PhGLO2*)

I also identified the B-class genes *PhGLO1* to be preferentially expressed in the petal mesophyll and *PhGLO2* in the petal epidermis. PhGLO1 and PhGLO2 are known to be necessary interactors of PhDEF, either PhGLO1 or PhGLO2 being part of the molecular quartet driving petal identity (Vandenbussche et al., 2004). A cell-layer-specific expression pattern of *PhGLO1* and *PhGLO2* could be a cue driving *PhDEF* cell-layer-specific regulation networks, with either one of the two proteins being the preferred interactor of PhDEF in a given cell-layer. This suggestion is coherent with what was observed in *phglo1* single mutants. As briefly glanced upon in the Introduction, *phglo1* single mutants show a mostly wild-type-like petal phenotype. However, a subtle phenotype does exist: the petal is more chloroplastic, especially along the petals central veins, and more importantly the tube is smaller, with a shorten section were stamens and petals are fused (Vandenbussche et al., 2004). Coupled with the proposed model for tube and limb growth promotion in (Chopy et al., 2023), these differences advocate for a cell-layer-specificity of *PhGLO1*, the expression of *PhGLO2* alone not fully restoring PhGLO1 function in all layers, leading to a cell-layer-specific regulation network implicating *PhDEF*.

It is important to note though that scRNA-Seq data was obtained from fully developed flowers. Observed expression levels of B-class genes might not be representative of earlier developmental stages. Again, to further assess this differential expression pattern of *PhGLO1* and *PhGLO2* in the mesophyll and the epidermis and conclude on their putative role in cell-layer-specific PhDEF-mediated regulation networks, it would be necessary to obtain scRNA-Seq data at different flower developmental stages.

V – HDG proteins as putative epidermis-specific interactors of PhDEF

HOMEODOMAIN GLABROUS proteins, also known as HDG proteins, are class IV homeodomain leucine-zipper (HD-Zip IV) transcription factors well known to drive epidermal cell types differentiation (Schrick et al., 2023b). In my PhD I showed that one of the motifs enriched in anti-PhDEF ChIP-Seq on wild-type *Petunia* petal is a typical target of HDG proteins, suggesting that PhDEF might interact with such proteins to bind DNA and regulate gene expression.

Star and *wico* flower phenotype suggest that *PhDEF* plays a cell-layer-specific role in petal development, including considering its epidermis identity (Chopy et al., 2023). The fact that PhDEF could interact with HDG proteins known to be involved in epidermis identity suggests that PhDEF could directly control epidermal-specific targets by the means of this specific protein interaction. Indeed, the presence of PhDEF could facilitate, or straight-out allow, the binding of HDG proteins on specific targets which activation would define petal epidermis identity.

VI – About scRNA-Seq technical aspects

VI.1 – Protoplast isolation

Multiple technical challenges remain unresolved as of yet regarding scRNA-Seq on *Petunia* petal. The first being my inability to obtain sufficiently concentrated protoplast suspensions from little amounts of biological material. This caused me to abandon the idea of gathering temporal data across different key early developmental stages as well as gathering any data at all from *phdef-151* mutants. As stated earlier though, gathering such data is mandatory to tackle any developmental questions. A few unexplored possibilities could be tried to increase yields. First, it seems to me that the protoplast isolation efficiency exponentially dropped when less and less starting material was used. I explain this, based on what I observed during isolation runs, by the fact that protoplasts tend to aggregate together. After centrifugation steps they form a quite cohesive and dense layer, so the more protoplasts you have and the thicker the layer, the best chance this layer will remain cohesive. This considered, isolating protoplasts in narrower tubes (15 mL and not 50 mL for instance) could help increase yields. Still to try to increase the concentration of the initial lysate, one could also try to digest in less buffer, enabling to pool more plates in a single tube and again, thicken the protoplast layer after the initial centrifugation step. Second, the way petals were cut could also be modified to increase yields. I did a quick trial comparing small (ca. 25 mm^2) and big (ca. 100 mm^2) petal bits and smaller bits worked better, but I never pushed it further. Cutting the petals in very narrow strips as proposed in (Pan et al., 2022) could be a solution, although it might impact protoplast viability by activating wounding stress apoptosis-inducing pathways in much more cells. Finally, changing the enzymatic digestion buffer might also be a valid trial, especially considering *phdef-151* mutant. The protocol I used was optimized to obtain petal protoplasts (Faraco et al., 2011) and might not be fit to isolate protoplasts from sepal-like tissue, a tissue that is much thicker and potentially quite different in molecular composition. Using an enzyme mix developed towards leaf protoplasts isolation might yield better results.

Another technical issue is the drop in cell viability during encapsulation that I witnessed during the second run of scRNA-Seq. Protoplast viability was systematically assessed multiple times during the isolation process and was systematically over 90 % prior to encapsulation. IAll buffers I used in scRNA-Seq experiment were validated the day before by doing a small isolation batch and an osmolarity check, 490 to 500 $mOsm.kg^{-1}$ H_2O being the target. I also tested protoplast viability in Mannitol-BSA 0.1 % over a few hours and I did not see any meaningful viability shifts. However, when going over my notes about scRNA-Seq run 2, I noticed that instead of adjusting protoplast concentration using first 0.40 M Sucrose and then 0.44 M Mannitol as in run 1, I did it directly using 0.44 M Mannitol. So the final concentration of Mannitol was far higher during run 2 encapsulation than during run 1, which might be the reason for lower viability of the protoplasts, although the osmolarity was maintained. Moreover, using 0.40 M Sucrose for dilution purposes in run 3 yielded better results than in run 2, but still showed noticeable viability problems. Hence, I don't really have any strong explanation for this issue to this day. But it should be explored further and ideally addressed to guarantee cleaner scRNA-Seq experiments.

VI.2 – Poor scRNA-Seq reads mapping rates

As stated earlier, only between 30 and 50 % of the collected reads in the various scRNA-Seq experiments I performed during my PhD map to the available *Petunia axillaris* genome annotation. Blasting all reads against NCBI databases did not really show striking problems, no contamination, nor viral nor animal was detected. As presented in the Results, improving the annotation using bulk RNA-Seq datasets did not meaningfully improve this mapping rate. Blasting only reads that don't map would probably help to better understand the situation. As far as I understand, extracting them from the BAM file provided by cellranger after reads alignment should be feasible using the good set of tags, but this is not an explicit part of the cellranger workflow and I was not able to do so. This aspect is yet another which I feel is important to resolve in order to better assess the quality of the data produced.

VI.3 – Possible scRNA-Seq analysis pipeline improvements

I started working on the analysis pipeline entering the second year of my thesis. At the time, I had very scarce knowledge of bioinformatics, yet alone pipeline construction, data handling or reproducibility tests. This results in a pipeline that works and produce data I am confident with, however I feel like the pipeline could improve a lot. First, its structure is managed in Bash. A welcomed first change would be to manage the pipeline using Nextflow scripts, language of which I

heard about a year ago and that was built to manage pipelines. Second, I realize now that I make poor use of the PSMN clusters and its job scheduler system SLURM. The pipeline should be cut in smaller pieces and multiple steps should be parallelized by simply launching them all together on available nodes instead of using a single node to execute them sequentially. Finally, the last change I would make would improve reproducibility as well as improve other users experience. I would bundle a singularity container (*i.e.* a docker container that can run without super user access, (Kurtzer et al., 2017)) containing all softwares and packages used, so no version change would break it or its reproducibility and so dependencies would not need to be installed manually.

VII – About scRNA-Seq analysis

In this manuscript I have presented wild-type *Petunia* petal scRNA-Seq results. In order to further explore the main biological question driving the project, studying *star* and *wico* datasets is the obvious next step. However this might prove challenging. First of all, both datasets show quite high background noise due to technical difficulties during encapsulation and a low number of cells (4,000 when the target was 7,000) to analyze, neither will facilitate the analysis. Of the two, *wico* should be the easiest. Indeed, as demonstrated with the wild-type experiment, the main cell-identity cues detected in our datasets are epidermal L1 markers and pigmentation-related genes. Since their epidermis is wild-type-like, *wico* flowers will most likely have conserved such markers allowing epidermis clusters identification. On the other hand, *star* biological characteristics might add a layer of complexity. Indeed, *star* flowers show mixed epidermis cell identity; unpigmented domed cells intermediate between conical petal epidermis cells and flat sepal epidermis cells. Pigmentation cues should be lost and L1 or conical cell markers lost or altered, rendering epidermis clusters identification more difficult. Interestingly, s*tar* flower petal limb usually have revertant pigmented sectors. If detected, these cells should form a distinct cluster since pigmentation-related gene expression should be restored. This cluster might help to identify non-pigmented epidermal cells and more importantly, it could be used to perform DGE analysis between wild-type-like epidermis and *star* epidermis while being in the same *star* background, which could be quite informative.

Once *star* and *wico* datasets clusterings are characterized, DGE analysis between identified cell layers of interest across samples should allow to detect cell layer-specific genes deregulated under the control of *PhDEF*. Although one could integrate all datasets together into a single scRNA-Seq dataset and perform the analysis as if it was a single dataset indeed, the method of choice to perform DGE analysis across samples would be to create pseudo-bulk RNA-Seq datasets of each identified cell layer and use them as different sample inputs in a more classical DGE

analysis pipeline. Indeed, comparing different biological samples while retaining the single-cell nature of the data artificially underestimates the variance, since each cell is considered a sample, leading to misleadingly small p-values.

Finally, in order to maybe address the lack of data for *phdef-151* mutant using data already on hand, I believe it should be possible to integrate run 1, run 2 *star* and *wico* and run 3 wild-type datasets together and check if a cluster or more stand out and do not overlay with identified wild-type, *star* and *wico* clusters. If so, the cells within are likely to originate from *phdef-151* sample and it would become possible to analyze them using pseudo-bulk RNA-Seq DGE.

MATERIALS AND METHODS

I – Plant material, growth conditions and lines maintenance

All plants were grown in a culture room in long day conditions (16h day at 22°C, 8h night at 18°C, 75-Valoya NS12 LED bars, light intensity: 130 µE, 60% humidity). The *phdef-151* plants were obtained from the *Petunia x hybrida* W138 line and *wico* and *star* flowers were repeatedly obtained from several different *phdef-151* individuals and were maintained by cuttings.

III – Protoplast isolation

All the following steps were performed under a horizontal laminar sterile flow-hood (Thermo Scientific™ Heraguard™ ECO Clean Bench) and using sterile material. *Petunia x hybrida* wild-type, *def-151*, *star*, or *wico*, flowers were harvested and dissected to keep only the corollas. The corollas where disinfected in a 1 sec. bath in 70% Ethanol (v/v in qqH_2O) followed by a 30 sec. bath in 0.5% active chlorine Bleach (v/v in qqH_2O) and rinsed 3 times 5 sec. in consecutive 500 mL qqH_2O independent baths. 4 to 5 (wild-type, *star*, or *wico*) or about 30 (*def-151*) corollas were transferred in a 100 mm Petri dish (Corning™) containing 2 mL of Digestion Mix (0.4% macerozyme R-10 (Duchefa Biochemie B.V.), 0.8% Cellulase Onozuka R-10 (Duchefa Biochemie B.V.) w/v in TEX Buffer (3.1 g/L Gamborg B5 salts, 500 mg/L MES, 750 mg/L $CaCl_2$*$2H_2O$, 250 mg/L NH_4NO_3, 136.9 g/L Sucrose, pH 5.7)) and cut in ca. 0.5 cm² pieces using a new scalpel blade for each corolla to reduce tissue wounding as much as possible. 10 mL of Digestion Mix were added to the Petri dish before it was closed and sealed with Parafilm™ (Heathrow Scientific). The Petri dishes were incubated 5 h at 26°C in the dark, gentle orbital agitation (20 rpm) was turned on during the last 15 min. of incubation.

The digested mixture was filtered through a 40 µm cell strainer (Falcon) inside a 50 mL tube, the tube volume was adjusted to 25 mL using 0.40 M Sucrose (in qqH_2O, 492 mOsm.kg^{-1} H_2O) and was centrifuged using a swing-out rotor 10 min. at 100 g with acceleration 2/9 and deceleration 0/9. The protoplasts form a layer on top of the buffer, using a peristaltic pump (Gilson MINIPULS™ Evolution) and a Pasteur pipette as much of the underlying buffer as possible was removed without perturbing the protoplast layer, at a rate of ca. 100 µL/sec. The tube volume was adjusted to 25 mL using TEX Buffer and the whole process repeated twice. The protoplast concentration and viability was assessed using a Kova slide (Dutscher) and 2% (w/v in 0.40 M Sucrose, usage at 1% final) Evans Blue dye solution as described in (Fernandez-Da Silva and Menendez-Yuffa, 2006).

IV – scRNA-Seq

IV.1 – Run 1

Protoplasts of wild-type, *phdef-151*, *star* and *wico* petals were purified and their viability and concentration controlled as previously described. Wild-type, *phdef-151*, *star* and *wico* cell suspensions concentrations were adjusted first to match 1,000 cells/µL using 0,40 M Sucrose (w/v in qqH_2O, 492 mOsm.kg^{-1} H_2O) and further diluted to match *phdef-151* sample concentration which was the lowest (560 cells/µL) using Mannitol-BSA (0.44 M Mannitol, 0.1 % BSA (w/v, Sigma-Aldrich, A8022), in qqH_2O, 498 mOsm.kg^{-1} H_2O)) and 25 µL of each sample pooled into a single 100 µL 560 cells/µL suspension. This unique pooled suspension was loaded in the 10x Genomics Chromium chip and apparatus as per manufacturer recommendations. Target cell recovery was 10,000.

IV.2 – Run 2

Protoplasts of wild-type, *phdef-151*, *star* and *wico* petals were purified and their viability and concentration controlled as previously described. Wild-type, *star* and *wico* suspensions were adjusted to 345, 480 and 590 cells/µL respectively (target was 500 cells/µL) using Mannitol-BSA, *phdef-151* didn't yield enough protoplasts for its suspension to be used. Each sample was then loaded in the 10x Genomics Chromium chip and apparatus as per manufacturer recommendations. Target cell recovery was 7,000 per sample.

IV.3 – Run 3

Protoplasts of wild-type and *phdef-151* petals were purified and their viability and concentration controlled as previously described. Wild-type suspension was adjusted to 1,200 cells/µL (target was 1,000 cells/µL) using 0.40 M Sucrose. Again, *phdef-151* didn't yield enough protoplasts for its suspension to be used. Therefore, a technical replicate of the wild-type protoplast suspension was loaded in the 10x Genomics Chromium chip and aparatus as per manufacturer recommendations. Target cell recovery was 6,000 per sample.

IV.4 – Library preparation

Libraries were prepared using Chromium Next GEM Single Cell 3′ Reagent Kits v3.1 (Dual Index) kit and as per manufacturer recommendations. More detailed explanation of the process is featured in the Introduction, chapter VI.1.c.

IV.5 – Bioinformatics pipelines

IV.5.a – Reads quality check

FASTQ files integrity was checked using a custom Perl script. Reads quality was checked with fastqc version 0.12.1 (Simon, 2010). The output files were aggregated into one HTML report using multiqc version 1.8 for easy visualization of the results.

IV.5.b – Reads alignment, filtration and cell identification

Reads were aligned on the reference transcriptome, filtered and their barcode and UMI counted using cellranger count version 7.0.1 (Zheng et al., 2017a) as per 10x Genomics recommendations. Results were collected into a Hierarchical Data Format 5 (HDF5, (The HDF Group, 2006)) sparse feature-barcode matrix.

IV.5.c – Reference transcriptome generation

Reference transcriptome usable by cellranger was generated using cellranger mkref version 7.0.1 (Zheng et al., 2017a) on our in-house *P. axillaris* genome annotation as per 10x Genomics recommendations.

IV.5.d – Count matrix analysis

The HDF5 matrix outputs of cellranger count were analyzed in R version 4.1.2-foss-2021b (R Core Team, 2022) mainly using the package Seurat version 4.2.0 (Hao et al., 2021). Unless stated otherwise default function parameters were used for each function call.

The data was preprocessed by removing cells with less than 200 genes detected and genes detected in less than 3 cells. Data was normalized using the function NormalizeData, highly variable genes identified using FindVariableFeatures and scaled using ScaleData. Data dimension was reduced by Principal Component Analysis (PCA, (Pearson, 1901; Jolliffe and Cadima, 2016)) using RunPCA. Doublets (droplets that captured more than one cell) were detected using the package DoubletFinder (McGinnis et al., 2019) and removed.

The number of significant Principal Component (PC) was determined using the jackstraw approach (Chung and Storey, 2015) by running the function JackStraw. The second dimension reduction was performed using the number of significant PCs determined as explained above by applying the Uniform Manifold Approximation and Projection (UMAP, (McInnes et al., 2020)) dimension reduction technique using the function RunUMAP.

The package MultiK (Liu et al., 2021) was used to determine optimal cluster number of clusters without supervision. The resolution parameter was set as a sequence starting at 0.25, ending at 2 and with 0.05 increments. In parallel, a nearest neighbor graph was constructed using FindNeighbors and clustering computed using FindClusters running a Louvain algorithm (Waltman and van Eck, 2013) on the same set of parameters used in MultiK.

For the significant clusterings according to MultiK a Differential Gene Expression analysis (DGE) was performed and marker genes for each cluster identified using the function FindMarkers. Known genes of interest (pigmentation, organ identity, polarity, cell-layer specific, vasculature and photosynthesis related) were also identified. The cluster-specific expression levels of all genes identified at previous steps was plotted using functions FeaturePlot and VlnPlot.

V – Bulk RNA-Seq

V.1 – For crude and protoplasted petal tissues comparison

Three independent digestions were carried out on wild-type petals and a small sample of each of the to-be digested petal limb was flash frozen in liquid N_2 in three crude tissue 2 mL tubes corresponding to the petals of each digestion run. After protoplasts isolation the pellet of cells was also flash frozen. The 3 tubes containing crude tissues were mechanically ground into a fine powder (TissueLyser II, Retsch, Quiagen) and all 6 tubes underwent RNA extraction using Protocol A from Sigma's Spectrum™ Plant Total RNA Kit and On-Column DNase I Digestion Set as per manufacturer recommendations. RNA integrity and quantity were determined using a Bioanalyzer RNA 6000 Nano assay (Agilent), RINs between 8.60 and 9.40 were obtained. Libraries were prepared with poly-A enrichment and single-end 84-bp sequencing was performed on a NextSeq 500 platform (Illumina). Between 15 and 19 millions reads where obtained per sample. Reads where mapped on *P. axillaris* transcriptome as described in (Chopy et al., 2021). The dataset was further analyzed using R version 4.1.2-foss-2021b (R Core Team, 2022) and the package DESeq2 version 1.34.0 (Love et al., 2014) to generate lists of differentially expressed genes between the 2 conditions. ShinyGO version 0.77 (Ge et al., 2020) was used to perform the GO enrichment analysis and generate associated metrics and plots.

V.2 – For WT, *def-151*, *star* and *wico* comparison using WGCNA

Samples were prepared, the sequencing done and the reads mapped as described in (Chopy et al., 2021). The dataset was further analyzed using R version 4.2.1 (R Core Team, 2022) and the

package WGCNA version 1.69 (Zhang and Horvath, 2005; Langfelder and Horvath, 2008). Modules were built with standard settings, using a soft-thresholding power of 15 and the option MergeCutHeight set at 0.30, and a module membership value cut-off of 0.6 was arbitrarily used to select genes best matching the modules.

VI – StringTie bioinformatics pipeline

Reads from various RNA-Seq project from the team using wild-type W138 *Petunia x hybrida* where gathered and aligned as described in (Chopy et al., 2021) on two distinct genome assemblies of *Petunia axillaris* and three different predicted structural annotations. 1.6.2_v1 and 1.6.2_v4 corresponding respectively to the originally published *Petunia axilaris* genome assembly and predicted structural annotation (Bombarely et al., 2016) and a newer, unpublished predicted structural annotation of the same genome assembly, both hosted by the Sol Genomics Network (SGN) (https://solgenomics.net/ftp/genomes/Petunia_axillaris/). HiC_annot_1 corresponding to the v1 annotation transferred onto a genome assembly further scaffolded by HiC by DNA-Zoo (Dudchenko et al., 2017, 2018) (https://www.dnazoo.org/assemblies/Petunia_axillaris) by the team as described in (Chopy et al., 2021). Resulting binary alignment map (BAM) files were used with StringTie (Pertea et al., 2015) to generate corresponding gene transfer format (GTF) files using the -G and --conservative parameters allowing the use of a reference annotation file as a guide, limit *de-novo* transcripts prediction and favor 5’ and 3’ ends elongation of pre-existing transcripts annotations. Resulting GTF files were blended into one using the --merge parameter, removing any duplicated annotations. Any *de-novo* predicted transcripts was removed as well as tRNAs, present in the reference annotation, if needed using the text manipulation utility awk. The final GTF file was finally compared to the original reference using GffCompare (Pertea and Pertea, 2020) and the differences easily visualized using IGV (https://igv.org/app/) (Robinson et al., 2011, 2023; Thorvaldsdottir et al., 2013). All scripts, logs and final outputs are available at https://gitbio.ens-lyon.fr/qcavalli/stringtie.

VII – Histological cross sections of petal

All the following fixation steps were done on ice unless stated otherwise. *Petunia x hybrida* wild-type petal samples were dissected and immediately put into ice-cold PEM-PFA buffer (10 mM EGTA, 1 mM MgSO4, 100 mM PIPES, 75 mM Sucrose, pH 6.9 adjusted with 1 M NaOH, 4 % (w/v) PFA). The samples were vacuum infiltrated at – 0.06 MPa in the same buffer for 60 min., the

buffer changed for a fresh batch and the samples fixed at 4°C overnight. The samples were rinsed twice with ice-cold PEM-PFA buffer before dehydration in 60 min. long baths of 50%, 60%, 70%, 80%, 90%, 100%, 100% and 100% Ethanol (v/v in H_2O) under vacuum at – 0.06 Mpa. The samples were left in 100 % Ethanol at 4°C overnight and infiltrated in 60 min. long baths of 30/70, 50/50 and 70/30 LR White / Ethanol mix ((v/v), LR White Medium Grade Resin, 14380, Electron Microscopy Science) under vacuum at – 0.06 Mpa. The samples were left in the last bath overnight and infiltrated twice in a fresh 100% LR White 60 min. long bath under vacuum at – 0.06 Mpa. The samples were infiltrated 3 days in 100% LR White at 4°C under gentle agitation (20 oscillations per minute). Gelatin capsules were filled with LR White, each sample placed inside a capsule, the capsules filled with LR White, closed and polymerized overnight at 60°C in a dry stove. 5 to 10 µm sections were cut using a Microm Microtech HM355S microtome, loaded in a drop of water onto coated slides (SuperFrost Plus Gold Adhesion Microscope Slides, ET09.2, Epredia) and dried at 37°C on a hot bench. Up to 4 spots of sections per slide were prepared, consecutive sections were placed in spot 1, 2, 3, 4, 1, 2, 3, 4, etc. Slides where stored unmounted at -80°C awaiting use.

For histological studies under bright field microscopy, cross sections were stained using 0.5% (m/v in H_2O) Toluidine Blue for 1 min. and rinsed 1 min. with running H_2O. Slides were mounted using a drop of water right before observation.

VIII– Anti-PhDEF Western Blot

VIII.1 – Protein extraction

Wild-type, *phdef-151* and *phtm6* young flower buds and petals were collected, flash frozen in liquid N_2 and mechanically ground into a fine powder (TissueLyser II, Retsch, Quiagen). 1.5 mL of ice-cold Methanol + protease inhibitors (1X, cOmplete™ Protease Inhibitor Cocktail, Roche)) was added and each tube vortexed for ca. 30 sec and incubated 5 min at -20°C. Tubes were centrifuged 5 min at 16,000 g and 4°C. Supernatent was discarded and the pellet resuspended into 1 mL of ice-cold Acetone, each tube vortexed for ca. 30 sec and incubated 5 min at -20°C. Tubes were centrifuged 5 min at 16,000 g and 4°C. Supernatent was discarded and the pellet dried at room-temperature (RT°) for ca. 15 min. 50 µL of extraction solution (Plant Total Protein Extraction Kit working solution 4 (Sigma), 1X cOmplete™ Protease Inhibitor Cocktail (Roche)) was added to the pellet and the tubes vortexed 15 min. Tubes were centrifuged 30 min at 24,000 g and RT°. Supernatent was sampled in a new 1.5 mL tube and stored at -80°C awaiting later use.

Protein concentration was assessed using Bradford assay (Biorad) and UV-Vis spectroscopy at 595 nm as per manufacturer recommendations.

VIII.2 – Western Blot

Protein extracts of wild-type, *phdef-151* and *phtm6* young flower buds and petals were thawed on ice. 25 µL samples containing 20 µg of total proteins were prepared (1X Laemmli Sample Buffer (Biorad), H_2O, protein extract) and denatured 5 min at 95°C. The samples were loaded into a 12 % Acrylamide/Bisacrylamide gel (stacking gel: 0.5 mL 40 % Acrylamide/Bisacrylamide (Biorad), 1.25 mL stacking buffer (0.5 M Tris-HCl, pH 6.8), 3.25 mL H_2O; resolving gel: 3 mL 40 % Acrylamide/Bisacrylamide, 2.5 mL resolving buffer (1.5 M Tris-HCl, pH 8.8), 4.5 mL H_2O) and migrated at 90 V and 80 mA into 1X TG-SDS Buffer (10X TG Buffer pH 8.5 (Euromedex), 1 g/L SDS) until migration front reached the end of the gel (ca. 1 h). Migrated proteins were transferred onto a PVDF membrane (Millipore) using the P0 program of an iBlot™ 2 Gel Transfer Device (Invitrogen) as per manufacturer recommendations. Membrane was blocked overnight at 4°C under gentle rocking using 1X TBS-Tween-Milk (0.1 % Tween-20 (v/v, Sigma, P2287), 5 % Milk (Milchpulver, blotting grade, fetarm (Carl Roth)) in 1X TBS (100 mM Tris HCl pH 8.0, 1.5 M NaCl in H_2O)). Membrane was incubated 2 h at RT° under gentle rocking with primary antibody (rabbit anti-PhDEF, 1/500 in 1X TBS-Tween-Milk). Membrane was rinsed 3 × 10 min in 1X TBS-Tween-Milk. Membrane was incubated 1 h at RT° under gentle rocking with HRP-conjugated secondary antibody (goat anti-rabbit IgG HRP-conjugated (Invitrogen, SA1-9510) 1/5000 in 1X TBS-Tween-Milk). Membrane was rinsed 3 × 10 min in 1X TBS-Tween-Milk. Membrane was incubated with 2 mL Clarity™ Western ECL Substrate mix (Biorad) between 2 sheets of alimentary film (Sarogold™ Pro, Saropack) in a dark box. Excess ECL mix was wiped off and the membrane imaged under UV light using a ChemiDoc™ Touch (Biorad).

For the stringency assay, a single membrane was cut into stripes 1 well-wide and each stripe was incubated in different sets of buffers of different stringency.

IX – Immunolocalization

Slides prepared as described above were equilibrated at room-temperature and the section patches surrounded using a hydrophobic pen (PAP pen, GeneTex, GTX22601). To each patch the following succession of buffers was applied, using 100-200 µL depending on patch size. Sections were blocked in TBS-BSA 2% buffer (2% BSA (w/v, Sigma-Aldrich, A8022) in 1X TBS (100 mM Tris HCl pH 8.0, 1.5 M NaCl in H_2O)) for 45 min. and rinsed with 1X TBS for 5 min. Sections were incubated with primary antibody (1/10 to 1/1000 dilution depending on the antibody and assay in TBS-BSA 0.1%) overnight at 4°C in a damp black box. Slides were equilibrated at room-

temperature and rinsed 5 min. with 1X TBS, 10 min. with TBS-Tween 0.1% (0.1% Tween-20 (v/v, Sigma, P2287) in 1X TBS), and 5 min. with 1X TBS. Sections were incubated with secondary antibody (1/1000 dilution in TBS-BSA 0.1%) 60 min. at room-temperature in a damp black box before three consecutive rinses as described before. If needed, section were stained using Acriflavin (0.1% Acriflavine (w/v, Sigma, 01673) in H_2O), Toluidine Blue (as previously described), Calcofluor White (0.35% CFW (w/v, Sigma, F3543) in H_2O) or DAPI (3 mg/mL in H_2O (Sigma, D9542)) for 10 min. at room-temperature before three 5 min. rinses with 1X TBS. Finally, section were mounted in Vectashield (Eurobio scientific, H-1000).

Primary antibodies used were a custom made anti-PhDEF (truncated to remove the most conserved part of the protein, see annex 1) rabbit polyclonal antibody (see annex 2) or a commercially available anti-H3 Histone rabbit polyclonal antibody (Agrisera, AS10710) used as positive control. The secondary antibody used was a commercially available anti-rabbit IgG goat polyclonal antibody conjugated to Alexa 568 fluorophore (ThermoFisher, A11011).

1X PBS (Dutscher, X0515) instead of 1X TBS was used in several assays for testing purposes as further detailed in the Results chapter VI.3.b.

X – ChIP-Seq

X.1 – Nuclei isolation

2 or 3 stage 8 wild-type, *def-151*, *star*, *wico* and *phglo1;phglo2* flowers were dissected and their petals flash frozen in liquid N_2 inside 2 mL sterilized tubes each containing 2 small glass beads and stored at -80°C awaiting usage. Right before use, each sample was equilibrated again in liquid N_2 and mechanically ground down into a fine powder (TissueLyser II, Retsch, Quiagen). 1 mL of Fixation Buffer (10 mM HEPES pH 7.6, 0.5 M Sucrose, 5 mM KCl, 5 mM $MgCL_2$, 5 mM EDTA, cOmplete™ Protease Inhibitor Cocktail 1X (Roche), 14mM 2-Mercapto-ethanol, 2.5 mM DSG) was added to each tube, the mixture well homogenized and transferred into into a 15 mL tubes containing 9 mL of Fixation Buffer. The tubes were incubated between 60 and 90 min. on a turning wheel (30 rpm) at room-temperature (RT°). 300 µL of 37% FAA was added to each tube reaching ca. 1% final concentration and the tubes incubated precisely 5 min. at RT°. 1 mL of 2 M Glycine was added to each tube, quenching the solution and ending cross-linking and the tubes were directly put on ice. All next steps were performed on ice. Cells were mechanically lysed using an ice cold Dounce homogenizer and its pestles (Kimble, 885300-0040). For each sample, the loose pestle was used 5 times and the tight one, 7, the douncer being carefully rinsed with H_2O between each sample. 300 µL of Triton X-100 were added directly to the lysate inside the douncer reaching a 0.6% final

concentration. The lysate was filtered twice using 100 µm and 40 µm cell strainers, the douncer and every intermediate tubes being rinsed once using 1 mL of Fixation Buffer. Samples were centrifuged using a swing-out rotor 10 min. at 1000 g and 4°C. The supernatent was discarded and the pellet resuspended into 300 µL of nuclear isolation buffer (10 mM HEPES pH 7.6, 0.5 M Sucrose, 5 mM KCl, 5 mM $MgCL_2$, 5 mM EDTA, cOmplete™ Protease Inhibitor Cocktail 1X (Roche)). The nuclei suspension was carefully transferred onto 600 µL of 15% Percoll solution and the tubes centrifuged using a swing-out rotor 5 min. at 2,000 g and 4°C.

X.2 – Chromatin extraction and sonication

The supernatent was discarded and 900 µL of nuclear Lysis Buffer was added to the pellet. Each tube was vortexed vigorously for ca. 1 min. A 10 µL aliquot of non-sonicated chromatin of each sample was sampled and conserved at -80°C. The rest of each sample was transferred into a 1 mL sonication glass vial (Covaris, milliTUBE 1 mL AFA Fiber, 520130) and sonicated 2 × 15 min. (Covaris S220 Focused-ultrasonicator, peak power 105, duty factor 5, cycles per burst 200, 4°C). A 10 µL aliquot of sonicated chromatin of each sample was sampled and conserved at -80°C for later use, the rest of the sonicated chromatin was stored a few days at -80°C.

X.3 – Chromatin sonication validation

Sonication efficiency was checked using the aliquots of non-sonicated and sonicated chromatin. For decrosslinking, 0.5 µL 5M NaCl was added and the tubes incubated overnight at 65°C. 81.5 µL H_2O, 2.5 µL 0.5 M EDTA, 5 µL Tris-Hcl pH 6.5 and 1 µL RNAse A (24 mg/mL, Sigma-Aldrich, R4642) was added and the tubes incubated 60 min. at 42°C. 0.5 µL proteinase K (20 mg/mL, Invitrogen, 59895) was added and the tubes incubated at 42°C 90 min. 100 µL H_2O was added before performing a phenol:chloroform:isopentanol (Sigma-Aldrich, P2069) DNA extraction as per manufacturer recommendations. DNA was migrated in 1% agarose gel and sonication checked. Sonication was considered successful if sonicated samples showed a smear between 200 and 500 base-pairs and no trace of higher molecular weight DNA fragments.

X.4 – Chromatin immunoprecipitation

Following steps were carried out on ice. Per sample, 75 µL of protein A Dynabeads™ (Invitrogen, 10001D) and protein G Dynabeads™ (Invitrogen, 10003D) were mixed into 300 µL Dilution Buffer (15 mM Tris-Hcl pH 7.5, 150 mM NaCl, 1% Triton X-100, 1 mM EDTA, final pH

7-8). The beads were washed 2 times total by placing the tubes on a magnetic rack, waiting for the beads to clump up, the Dilution Buffer refreshed and the beads resuspended by gently rocking the tubes. Beads were finally resuspended into 150 µL Dilution Buffer per sample. For each samples, duplicates were processed later on. Per tube, 40 µL Dynabeads™ mix, 1.5 µL (2.5 µg) of anti-PhDEF antibody (custom, rabbit, 1.7 mg/mL, see annex 2) and 1.8 mL Dilution Buffer were mixed and the tubes incubated 2 h at 4°C on a rotating wheel (10 rpm). No antibody was added for negative-control tubes. Sonicated chromatin was thawed and centrifuged at 15 min. at 20,000 g and 15°C. Supernatent was recovered as chromatin input, the pellet discarded and a 20 µL aliquot sampled and stored at -80°C. For wild-type and *wico* chromatin 25 µL was added to the Dynabeads™ mix, 50 µL for *phdef-151*, s*tar* or *phglo1;phglo2*. For negative-control tubes, 25 µL of wild-type chromatin was used. Tubes were then incubated overnight at 4°C on a rotating wheel (10 rpm). Beads were washed using the following sequence by using a magnetic rack between each buffer addition and retrieval: short wash with 1 mL low-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-Hcl pH 8, 150 mM NaCl), long wash (15 min. at 4°C on a rotating wheel (10 rpm)) with 1 mL low-salt buffer, short wash with 1 mL high-salt-buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-Hcl pH 8, 500 mM NaCl), long wash with 1 mL high-salt-buffer, short wash with 1 mL LiCl buffer (0.25 M LiCl, 1% NP40 Igepal, 1% deoxycholate, 1 mM EDTA, 20 mM Tris-Hcl pH 8), long wash with 1 mL LiCl buffer. The last LiCl buffer washing was removed and 250 µL Elution Buffer (0.1 M $NaHCO_3$, 1% SDS) pre-heated to 65°C was added, the beads gently resuspended and the elution carried-out for 15 min. at 65°C with regular gentle manual shaking. The first elution was collected in a fresh tube and a second elution carried-out as previously described, finally both were pooled in the same tube. For decrosslinking, 0.5 µL 5M NaCl was added to the elution tubes and input aliquots to which 500 µL Elution Buffer was previously added, before overnight incubation at 65°C. 10 µL 0.5 M EDTA pH 8, 20 µL 1 M Tris-Hcl pH 6.5 and 1 µL proteinase K (20 mg/mL, Invitrogen, 59895) was added and tubes incubated 2 h at 42°C. A phenol:chloroform:isopentanol (Sigma-Aldrich, P2069) DNA extraction was finally performed as per manufacturer recommendations, final DNA pellet resuspended into 50 µL TE Buffer (10mM Tris-Hcl pH 8, 1 mM EDTA). DNA final concentration of each sample was tested either using a NanoDrop™ 3300 (Thermo Fischer Scientific) or a Qubit 2.0 (Invitrogen) depending on sample concentration and as per manufacturers recommendations before sample storage at -80°C awaiting use.

X.5 – DNA library preparation and sequencing

DNA library was prepared by collaborators using MicroPlex Library Preparation Kit v3 as per manufacturer recommendations, using 166 to 10,000 pg of template DNA depending on the sample. The library was sequenced on NextSeq5000 (Illumina), paired-end (2×76 bp, dual indexing i5/i7 2×8) over two runs to reach a minimum of 40 millions reads per sample.

X.6 – Data analysis

Data analysis was performed by collaborators by the following steps. Reads quality was assessed using fastqc (Simon, 2010), cleaned (trimming of adapters and low quality reads) using fastp (Chen et al., 2018) and mapped on our in-house *P. axillaris* genome annotation using bowtie2 (very sensitive local mode) (Langmead and Salzberg, 2012). Output BAM files where filtered for PCR and optical reads duplicates using sambamba (Tarasov et al., 2015) and read coverage normalized using bedtools (Quinlan and Hall, 2010). Peaks calling was done using MACS3 (MACS3 project team, 2020), and motif discovery using HOMER (Heinz et al., 2010). Data visualization in IGV (Thorvaldsdottir et al., 2013) was performed in house, output BAM files from both IP and INPUT replicates were intersected using “bedtools intersect” to keep only common peaks. Resulting files were loaded into IGV for data visualization.

XI – Recombinant plants generation and observation

XI.1 – Plasmids construction

All constructs were build using the GreenGate system (Lampropoulos et al., 2013) as per recommendations. All plasmids maps and primers used for their construction and verification by Sanger sequencing are available in the supplementaries.

XI.1 – *Agrobacterium* transformation and culture

20 ng of the plasmid of interest was added to a 25 µL aliquot of electrocompetent *Agrobacterium* just thawed on ice. The aliquot was then transferred into an electroporation glass vial and the bacteria electroporated for 5 ms at 2.2 kV (Gene Pulser Xcell, Biorad). 950 µL of LB broth (Thermo Fischer Scientific, 11798842) was added to the aliquot and the bacteria suspension incubated in a 2 mL tube 2 h at 28°C under orbital agitation at 180 rpm. The suspension was centrifuged 5 min at 2,000 g and the pellet resuspended into 150 µL LB broth. *Agrobacterium* suspension was plated on pre-warmed LB-Agar plates with appropriate selection (Spectinomycin,

Streptomycin or Ampicillin at 100 µg/mL or Kanamycin at 50 µg/mL) and incubated 48 h at 28°C. A single colony of the plate was picked to inoculate 5 mL of fresh LB broth with appropriate selection before incubation 48 h at 28°C under orbital agitation at 180 rpm.

XI.2 – *Petunia* Mitchell leaves transformation

All the following steps were performed under sterile conditions. 15 young leaves from *Petunia* Mitchell plants were harvested and disinfected 10 min in 1 L of 0.5% active chlorine bleach solution under gentle rotary agitation. The leaves were rinsed 5 × 2 min in 1 L of sterile H_2O and transferred into a large Petri dish. Leaves were cut into square pieces, removing all leave margins. In the mean time, the Agrobacterium culture was centrifuged 15 min at 2,000 g and resuspended with 40 mL sterile H_2O to which 4.5 µL Acetosyringone (100 mM in DMSO, Sigma-Aldrich, D134406) was added (11,25 µM final concentration). Leaves pieces were added to the bacteria suspension and incubated 30 min under gentle rotary agitation. Leaves were transferred onto sterile Whatman paper to remove any excess bacteria suspension and transferred adaxial side down onto co-culture Petri dishes (30 g/L Sucrose, 4.4 g/L MS (Duchefa, M0221), 1X Gamborg B5 vitamins (Duchefa, G0415), 4 g/L Phytagel™ (Sigma-Adricht, P8169), 2 mg/L BAP (Duchefa, B0904-1), 100 µg/L NAA (Duchefa, N0903), 20 µM Acetosyringone) which were then sealed with parafilm and incubated 48 h in long days conditions (21°C, 18 h day, 6 h night, Osram biolux neon tubes) in a closed box. The box was then slightly opened to let some indirect light in for 48 h.

XI.3 – Transformants selection

After four days of co-culture, the leaves were transferred onto selection Petri dishes (30 g/L Sucrose, 4.4 g/L MS, 1X Gamborg B5 vitamins, 4 g/L Phytagel™, 2 mg BAP, 100 µg NAA, 250 mg/L Carbenicillin (Duchefa, C0109), 1 mL Plant Preservative Mixture (PPM™, Plant Cell Technology)) with appropriate selection (300 mg/L Kanamycin, 5 mg/L glufosinate or 10 mg/L Hygromycin), which were then sealed with parafilm and incubated in the same conditions as before, but fully exposed to light this time. Petri dishes were check once a day and explants transferred onto fresh ones every two weeks while callus developed. Once leaves formed from the callus, callus was dissected in order to separate well formed plantlets which were then transplanted onto rooting medium in tall boxes (30 g/L Sucrose, 4.4 g/L MS, 1X Gamborg B5 vitamins, 4 g/L Phytagel™, 250 mg/L Carbenicillin (Duchefa, C0109), 1 mL PPM™) with appropriate selection. Once roots appeared, plants were screened by PCR (primers MLY2691 and 2692, see supplementaries) and positive ones were transferred into soil and cultivated as described earlier.

XI.4 – Transformants confocal imaging

Inflorescences were dissected to keep only very early flower stages when petals just initiate. Dissections were place inside a 35 mm Petri dish of low melting agarose inside a slit cut with a scalpel blade awaiting imaging. 5 min before imaging, a drop of propidium iodide was added to stain cell walls and the dissections submerged in H_2O before confocal imaging using a Leica TCS SP8 confocal microscope.

XII – Microscopy and image analysis

Epifluorescence microscopy pictures were taken using a Zeiss AxioImager microscope, an Axiocam 705 color camera and the software Zen version 3.2.

Confocal microscopy pictures were taken using a Leica TCS SP8 microscope, and the software LAS X version 4.1.13. Microscopy images analysis and annotation were done using FiJi (Schindelin et al., 2012). The data presented in Fig. 2.3 was obtained as follows: ten pictures of each slide were taken at random locations and analyzed as follows. Gaussian blur was applied to the image to smooth it, the image was converted into a mask and the plugin Analyze Particles was ran to generate CSV tables containing the size of all detected particles matching a set of size and circularity parameters (see supplementaries). All tables were pooled into one and further analyzed.

XIIII – General data analysis and plot generation

Data analysis and plot generation was mainly performed using R (R Core Team, 2022) and the packages from the tidyverse (Wickham et al., 2019), mainly ggplot2 (Wickham, 2016).

XIV – Availability of informatics scripts, pipelines and supplementaries

All informatics scripts, pipelines, logs and supplementary information cited within this manuscript are publicly available for review at the following repository:

https://gitbio.ens-lyon.fr/qcavalli/manuscript_supplementaries

ANNEXES

Annex 1 – Truncated PhDEF protein production details

ProteoGenix
Antibodies and Proteins

ProteoGenix SAS
Espace Européen de l'Entreprise
15 rue de La Haye
67300 Schiltigheim, France
Tel. +33 (0)3 90 20 54 70 - Fax +33 (0)3 88 56 16 88
contact@proteogenix.fr
www.ProteoGenix.science

Ref. ATGX-11886-PhDEF

Expression and Purification Tests Report

***E.coli* Expression System**

1. Project Information

Service Type: Recombinant Protein Production

Protein Name: PhDEF

Protein Alias for the project: PhDEF_long
PhDEF_short

Custom Services:

-Gene synthesis

-Subcloning in expression vector

-Expression and purification Tests

2. Gene and Expression Vector

2.1 Gene coding for the target protein

The cDNAs coding for the target proteins were chemically-synthesized with optimization for expression in *E.coli*. The sequences are illustrated below.

> PhDEF_long_cDNA - 552 bp

```
ATGGGCAGTCATCATCATCATCACCATAGCGGTCCGAGTATTACCACCAAACAGCTGTTTGATCTGTATCAGAAAACC
GTTGGTGTTGATCTGTGGAATAGTCATTATGAAAAAATGCAGGAACAGCTGCGTAAACTGAAAGAAGTTAATCGTAATCT
GCGCAAAGAAATTCGTCAGCGTATGGGCGAAAGCCTGAATGATCTGAATTATGAACAGCTGGAAGAACTGATGGAAAATG
TGGATAATAGCCTGAAACTGATTCGCGAACGCAAATATAAAGTTATTGGTAATCAGATCGAGACCTTTAAAAAGAAAGTG
CGTAATGTGGAAGAAATTCATCGTAATCTGCTGCTGGAATTTGATGCCCGTCAGGAAGATCCGTATGGCCTGGTGGAACA
GGAAGGTGACTATAATAGTGTTCTGGGTTTTCCGAATGGTGGTCATCGCATTCTGGCACTGCGCCTGCAGCCGAATCATC
ATCAGCCGAATCACCATCATCATCTGCATAGCGGTGGCGGCAGCGATATTACCACCTTTGCACTGCTGGAATAA
```

> PhDEF_short_cDNA - 372 bp

```
ATGGGCAGTCATCATCATCATCACCATAGTGGTCCGAGTATTACCACCAAACAGCTGTTTGATCTGTATCAGAAAACC
GTGGGTGTGGATCTGTGGAATAGTCATTATGAAAAAATGCAGGAACAGCTGCGCAAACTGAAAGAAGTTAATCGTAATCT
GCGTAAAGAAATTCGCCAGCGCATGGGCGAAAGTCTGAATGATCTGAATTATGAACAGCTGGAAGAACTGATGGAAAATG
TGGATAATAGTCTGAAACTGATTCGCGAACGCAAATATAAAGTTATTGGTAATCAGATCGAGACCTTTAAAAAGAAAGTT
CGCAATGTGGAAGAAATTCATCGTAATCTGCTGCTGGAATTTGATGCACGCTAA
```

2.2 Expression vector

The cDNA sequences described in the §2.1 were cloned in a pT7 expression vector. The expected proteins produced are illustrated below.

> PhDEF_long - 183 AAs

```
MGSHHHHHHSGPSITTKQLFDLYQKTVGVDLWNSHYEKMQEQLRKLKEVNRNLRKEIRQRMGESLNDLNYEQLEELMENV
DNSLKLIRERKYKVIGNQIETFKKKVRNVEEIHRNLLLEFDARQEDPYGLVEQEGDYNSVLGFPNGGHRILALRLQPNHH
QPNHHHHLHSGGGSDITTFALLE

Features :
His-tag with linker   : [1 : 11]
```

> PhDEF_short - 123 AAs

```
MGSHHHHHHSGPSITTKQLFDLYQKTVGVDLWNSHYEKMQEQLRKLKEVNRNLRKEIRQRMGESLNDLNYEQLEELMENV
DNSLKLIRERKYKVIGNQIETFKKKVRNVEEIHRNLLLEFDAR

Features :
His-tag with linker   : [1 : 11]
```

3. Expression Tests

3.1 Aim

-Determination of optimal conditions for protein expression by evaluation of:

*Induction strategy

*Temperature and time for induction

**E.coli* strain used for protein production

-Identification of best extraction condition: native or denatured

3.2 Short protocol description

-Culture:

*Bacteria starter obtained by incubation at 37°C

*IPTG induction with different temperatures and incubation times (see details in §3.3).

-Protein samples preparation (for each expression condition tested):

*Bacteria harvested by centrifugation

*Lysis with native buffer
*Clarification (centrifugation C1): The supernatant is the native protein extract (NPE)
*Solubilization of the pellet from C1 with a denaturing buffer (Urea 8M)
*Clarification (centrifugation): The supernatant is the denatured protein extract (DPE)
*PAGE analysis of NPE and DPE.

3.3 Results of the expression tests

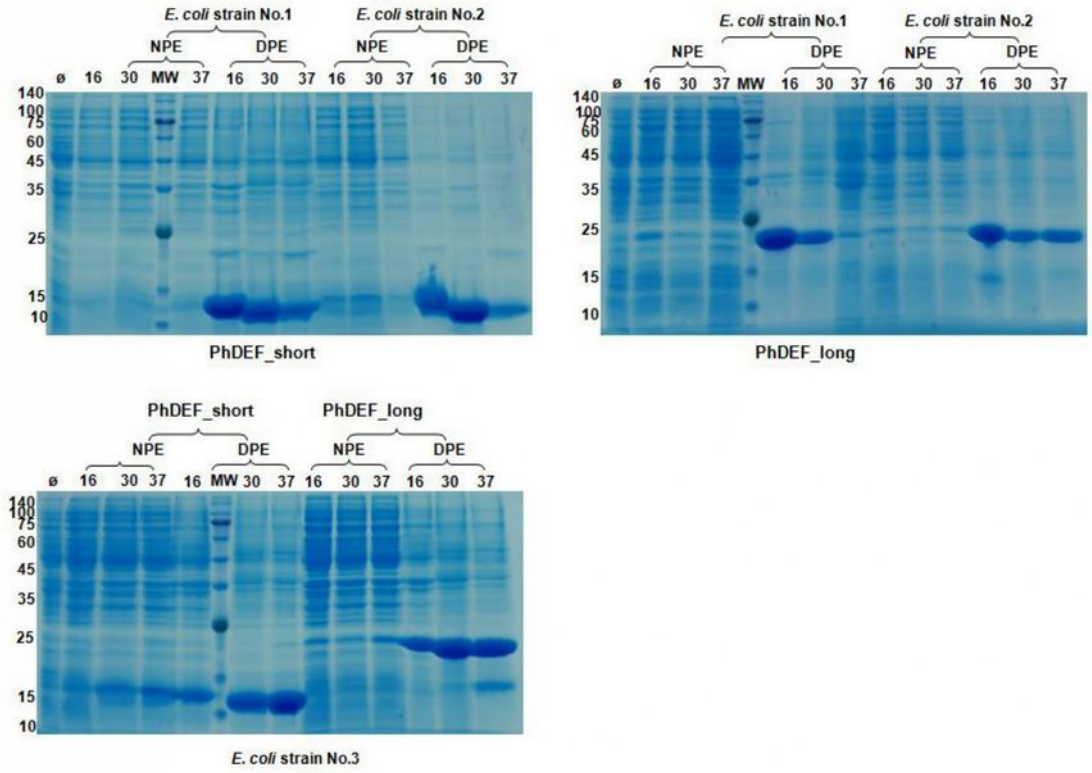


Figure 1. Expression tests of the target protein. *Reduced SDS-PAGE with Coomassie blue staining.*
Analysis of NPE and DPE prepared as described in §3.2.
MW. Molecular weight marker. **Ø.** Non-induced bacteria culture (negative control).
16, 30, and 37. Incubation temperature (°C) during induction with IPTG.
Induction with IPTG 1mM during 16h at 16°C, or during 4h for other temperatures.

3.4 Conclusion of the expression tests

Note: *E.coli strains tested are intended to overcome one or several expression issues such as (but not limited to) expression level, solubility or toxicity.*

The target protein is expressed in all *E.coli* strains tested at high levels, but present mainly in DPE. The optimal expression level in native conditions is indicated in the Table 1.

	Strain No.	Temperature	Induction Time
PhDEF_long (in NPE and DPE)	No.1	16°C	16h
PhDEF_short (in NPE and DPE)	No.3	37°C	4h

Table1. Optimal native expression conditions of target protein.

4. Small scale up purification tests.

4.1 Aim

Small-scale (200ml culture) purification tests in native and denatured condition. Starting material: NPE and DPE prepared after production in optimal native expression conditions described in the Table 1.

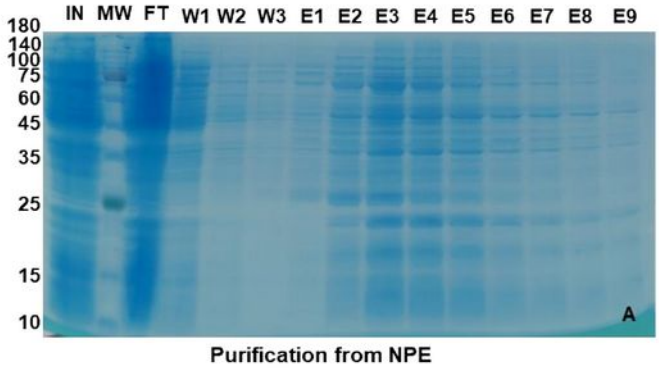
4.2 Short protocol description

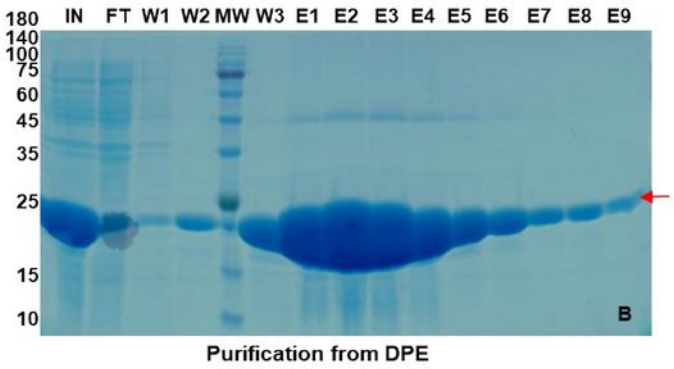
Purification by affinity vs. His-Tag on Nickel resin

-Equilibration and binding with PBS pH7.5 (add 8M urea for DPE purification)
-Washes and elution by imidazole shift
-Final sample QC: qualitative by SDS-PAGE, quantitative by Bradford method

After purification, the samples of interest are buffer exchanged vs. PBS pH7.5 (or PBS pH7.5, 4M urea for denatured condition).

4.3 Results of the purification tests





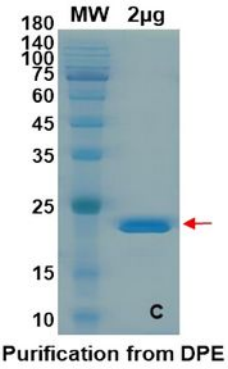


Figure 2. Purification tests of the PhDEF_long target protein. *Reduced PAGE with Coomassie blue staining.* **MW.** Molecular weight marker. **IN.** Input. **FT.** Flow through. **W1-W3.** Wash fractions. **E1-E9.** Eluted fractions. **A and B.** Purification profile (Reduced analysis). **C.** Final sample QC after buffer exchange.

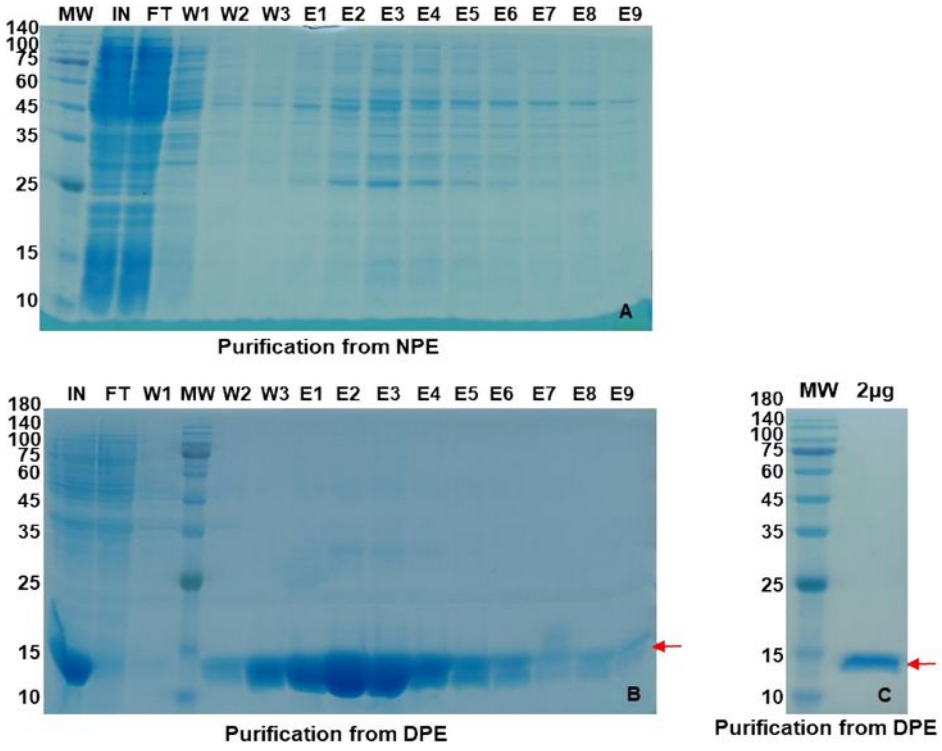


Figure 3. Purification tests of the PhDEF_short target proteins. *Reduced PAGE with Coomassie blue staining.* **MW.** Molecular weight marker. **IN.** Input. **FT.** Flow through. **W1-W3.** Wash fractions. **E1-E9.** Eluted fractions. **A and B.** Purification profile (Reduced analysis). **C.** Final sample QC after buffer exchange.

4.4 Conclusion of the purification tests

The target proteins can't be purified from NPE, but can be purified from DPE. The yield and purity obtained are indicated in the Table 2.

	Quantity* (mg)	Yield per liter** (mg)	Purity*** (%)
PhDEF_long (Purified from DPE)	26.4	132	90
PhDEF_short (Purified from DPE)	12.6	63	90

Table 2. Production and purification yields of the target proteins
*Obtained for the 200ml-culture test. **Estimate based on the 1L-culture test.
***Based on the 2µg loading in Figures 2C and 3C.

Final sample detail
PhDEF_long (Purified from DPE)
Final Conc: 2.2mg/ml
Volume: 12ml
Final product: 2.2mg/ml 1.5ml/vial 8vials
Storage buffer: PBS PH7.5, 4M urea

PhDEF_short (Purified from DPE)
Final Conc: 1.05mg/ml
Volume: 12ml
Final product: 1.05mg/ml 1.5ml/vial 8vials
Storage buffer: PBS PH7.5 ,4M urea

Annex 2 – Anti-PhDEF antibodies certificate of analysis

ProteoGenix SAS
Espace Européen de l'Entreprise
15 rue de La Haye
67300 Schiltigheim, France
Tel. +33 (0)3 90 20 54 70 - Fax +33 (0)3 88 56 16 88
contact@proteogenix.fr
www.ProteoGenix.science

Certificate Of Analysis - Custom Polyclonal antibodies
Project #: 12128

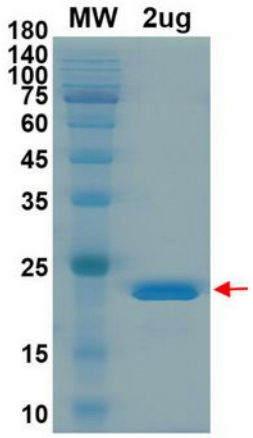
Antigen - Protein Production:

Protein name:PhDEF
Production Host: E.Coli
Molecular Weight:21.49kDa

Delivered protein sequence:
MGSHHHHHHSGPSITTKQLFDLYQKTVGVDLWNSHYEKMQEQLRKLKEVNRNLRKEIRQRMGESLNDLNYEQLEELMENV
DNSLKLIRERKYKVIGNQIETFKKKVRNVEEIHRNLLLEFDARQEDPYGLVEQEGDYNSVLGFPNGGHRILALRLQPNHH
QPNHHHHLHSGGGSDITTFALLE

Final buffer:PBS PH7.5, 0.02% NLS
Purity:95%
Concentration:2.2mg/ml

Protein Final QC:



Final sample QC. *Coomassie blue staining*

For research use only. Not suitable for in vitro diagnostic or human use

ProteoGenix SAS
Espace Européen de l'Entreprise
15 rue de La Haye
67300 Schiltigheim, France
Tel. +33 (0)3 90 20 54 70 - Fax +33 (0)3 88 56 16 88
contact@proteogenix.fr
www.ProteoGenix.science

Antibodies Production:

Species:Rabbit
Lot N°:12128-042320-A01
FINAL ELISA RESULTS:

Dilution	Blank	Negative	1:1000	1:4000	1:16000	1:32000	1:64000
Animal #1	0.032	0.045	4.24	4.05	3.81	3.45	2.87
Animal #2	0.032	0.045	4.42	4.24	4.13	4.05	3.34

NB: Negative serum: 1:1000 Animal #1 & Animal #2 <0.2 -Standard ELISA without optimization for this antigen

Antibodies Purification:

Purification Step: Affinity purification against the antigen
Final buffer: PBS, pH7.4

Final delivered samples:

Product	Animal #	Quantity	Concentration	Form
Pre-immune serum	1	0.5 ml		Liquid
Pre-immune serum	2	0.5 ml		Liquid
Immune serum	1	5 ml		Liquid
Immune serum	2	5 ml		Liquid
Purified Antibodies	1	3ml	1.7 mg/ml	Liquid
Purified Antibodies	2	2 ml	2.7 mg/ml	Liquid
The protein 12128- PhDEF		0.5 mg	7.2 mg/ml	

Storage advice:

ProteoGenix did not perform stability or storage tests for your specific end products.
Here are some **standard storage advice**:
- Store at 4°C for short term; Freeze and store at -20°C or -80°C for long term
- Freezing should be first tested on small aliquot(s); Glycerol up to 50% may be added for cryoprotection. Purified protein/antibodies sometimes require additives for optimal freezing and storage, and/or a specific freezing process (e.g., -20°C, -80°C, or flash freezing in liquid nitrogen).
- Purified proteins/antibodies should ideally be stored as 0.5 to 2mg/ml stock solutions.
- Avoid freeze/thaw cycles: aliquot the products according to your needs.

Special notes :

For research use only. Not suitable for in vitro diagnostic or human use

REFERENCES

Angenent, G. C., Franken, J., Busscher, M., Weiss, D., and van Tunen, A. J. (1994). Co-suppression of the petunia homeotic gene fbp2 affects the identity of the generative meristem. *Plant J.* 5, 33–44.

Asano, T., Masumura, T., Kusano, H., Kikuchi, S., Kurita, A., Shimada, H., et al. (2002). Construction of a specialized cDNA library from plant cells isolated by laser capture microdissection: toward comprehensive analysis of the genes expressed in the rice phloem. *The Plant Journal* 32, 401–408. doi: 10.1046/j.1365-313X.2002.01423.x.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: tool for the unification of biology. *Nat Genet* 25, 25–29. doi: 10.1038/75556.

Bargmann, B. O. R., and Birnbaum, K. D. (2010). Fluorescence Activated Cell Sorting of Plant Protoplasts. *JoVE*, 1673. doi: 10.3791/1673.

Barron, M., and Li, J. (2016). Identifying and removing the cell-cycle effect from single-cell RNA-Sequencing data. *Sci Rep* 6, 33892. doi: 10.1038/srep33892.

Bemer, M., Dijk, A. D. J. van, Immink, R. G. H., and Angenent, G. C. (2017). Cross-Family Transcription Factor Interactions: An Additional Layer of Gene Regulation. *Trends in Plant Science* 22, 66–80. doi: 10.1016/j.tplants.2016.10.007.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2013). GenBank. *Nucleic Acids Res* 41, D36-42. doi: 10.1093/nar/gks1195.

Bezrutczyk, M., Zöllner, N. R., Kruse, C. P. S., Hartwig, T., Lautwein, T., Köhrer, K., et al. (2021). Evidence for phloem loading via the abaxial bundle sheath cells in maize leaves. *The Plant Cell* 33, 531–547. doi: 10.1093/plcell/koaa055.

Bianchi, F., Cornelissen, P. T. J., Gerats, A. G. M., and Hogervorst, J. M. W. (1978). Regulation of gene action in Petunia hybrida: Unstable alleles of a gene for flower colour. *Theoret. Appl. Genetics* 53, 157–167. doi: 10.1007/BF00273576.

Biémont, C. (2010). A Brief History of the Status of Transposable Elements: From Junk DNA to Major Players in Evolution. *Genetics* 186, 1085–1093. doi: 10.1534/genetics.110.124180.

Birnbaum, K. (2003). A Gene Expression Map of the Arabidopsis Root. *Science* 302, 1956–1960. doi: 10.1126/science.1090022.

Boachon, B., Lynch, J. H., Ray, S., Yuan, J., Caldo, K. M. P., Junker, R. R., et al. (2019). Natural fumigation as a mechanism for volatile transport between flower organs. *Nat Chem Biol* 15, 583–588. doi: 10.1038/s41589-019-0287-5.

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C. S., et al. (2016). Insight into the evolution of the Solanaceae from the parental genomes of Petunia hybrida. *Nature Plants* 2, 1–9. doi: 10.1038/nplants.2016.74.

Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M., and Smyth, D. R. (1993). Control of flower development in Arabidopsis thaliana by APETALA1 and interacting genes. *Development* 119, 721–743. doi: 10.1242/dev.119.3.721.

Bowman, J. L., Smyth, D. R., and Meyerowitz, E. M. (1989). Genes directing flower development in Arabidopsis. *Plant Cell* 1, 37–52. doi: 10.1105/tpc.1.1.37.

Bowman, J. L., Smyth, D. R., and MEYEROWITZt, E. M. (1991). Genetic interactions among floral homeotic genes of Arabidopsis. 23.

Brennecke, P., Anders, S., Kim, J. K., Kołodziejczyk, A. A., Zhang, X., Proserpio, V., et al. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* 10, 1093–1095. doi: 10.1038/nmeth.2645.

Buettner, F., Natarajan, K. N., Casale, F. P., Proserpio, V., Scialdone, A., Theis, F. J., et al. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol* 33, 155–160. doi: 10.1038/nbt.3102.

Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 357, 661–667. doi: 10.1126/science.aam8940.

Causier, B., Schwarz-Sommer, Z., and Davies, B. (2010). Floral organ identity: 20 years of ABCs. *Semin Cell Dev Biol* 21, 73–79. doi: 10.1016/j.semcdb.2009.10.005.

Cavallini-Speisser, Q., Morel, P., and Monniaux, M. (2021). Petal Cellular Identities. *Frontiers in Plant Science* 12. doi: 10.3389/fpls.2021.745507.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560.

Cho, K. H., Kim, J. Y., Alvarez, M. I., Laux, V. Y., Valad, L. K., Tester, J. M., et al. (2019). Strong Fluorescence Expression of ZsGreen1 in Petunia Flowers by Agrobacterium tumefaciens–mediated Transformation. *J. Amer. Soc. Hort. Sci.* 144, 405–413. doi: 10.21273/JASHS04776-19.

Chopy, M., Cavallini-Speisser, Q., Chambrier, P., Morel, P., Just, J., Hugouvieux, V., et al. (2021). Cell layer-specific expression of the B-class MADS-box gene *PhDEF* drives petal tube or limb development in petunia flowers. Plant Biology doi: 10.1101/2021.04.03.438311.

Chopy, M., Cavallini-Speisser, Q., Chambrier, P., Morel, P., Just, J., Hugouvieux, V., et al. (2023). Cell layer-specific expression of the homeotic MADS-box transcription factor PhDEF contributes to modular petal morphogenesis in petunia. *The Plant Cell*, koad258. doi: 10.1093/plcell/koad258.

Chung, N. C. (2020). Statistical significance of cluster membership for unsupervised evaluation of cell identities. *Bioinformatics* 36, 3107–3114. doi: 10.1093/bioinformatics/btaa087.

Chung, N. C., and Storey, J. D. (2015). Statistical significance of variables driving systematic variation in high-dimensional data. *Bioinformatics* 31, 545–554. doi: 10.1093/bioinformatics/btu674.

Cocking, E. C. (1960). A Method for the Isolation of Plant Protoplasts and Vacuoles. *Nature* 187, 962–963. doi: 10.1038/187962a0.

Coen, E. S., and Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353, 31–37. doi: 10.1038/353031a0.

Colombo, L., Franken, J., Koetje, E., van Went, J., Dons, H. J., Angenent, G. C., et al. (1995). The petunia MADS box gene FBP11 determines ovule identity. *Plant Cell* 7, 1859–1868. doi: 10.1105/tpc.7.11.1859.

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inform. Theory* 13, 21–27. doi: 10.1109/TIT.1967.1053964.

Dahlin, J. S., Hamey, F. K., Pijuan-Sala, B., Shepherd, M., Lau, W. W. Y., Nestorowa, S., et al. (2018). A single-cell hematopoietic landscape resolves 8 lineage trajectories and defects in Kit mutant mice. *Blood* 131, e1–e11. doi: 10.1182/blood-2017-12-821413.

Davies, B., Egea-Cortines, M., de Andrade Silva, E., Saedler, H., and Sommer, H. (1996). Multiple interactions amongst floral homeotic MADS box proteins. *EMBO J* 15, 4330–4343.

Deal, R. B., and Henikoff, S. (2011). The INTACT method for cell type–specific gene expression and chromatin profiling in Arabidopsis thaliana. *Nat Protoc* 6, 56–68. doi: 10.1038/nprot.2010.175.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. *Science* 295, 1306–1311. doi: 10.1126/science.1067799.

Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K., and Timmermans, M. C. P. (2019a). Spatiotemporal Developmental Trajectories in the Arabidopsis Root Revealed Using High-Throughput Single-Cell RNA Sequencing. *Dev. Cell* 48, 840-852.e5. doi: 10.1016/j.devcel.2019.02.022.

Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K., and Timmermans, M. C. P. (2019b). Spatiotemporal Developmental Trajectories in the Arabidopsis Root Revealed Using High-Throughput Single-Cell RNA Sequencing. *Developmental Cell* 48, 840-852.e5. doi: 10.1016/j.devcel.2019.02.022.

Dolezel, J., Číhalíková, J., Weiserová, J., and Lucretti, S. (1999). Cell cycle synchronization in plant root meristems. *Methods Cell Sci* 21, 95–107. doi: 10.1023/a:1009876621187.

Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., et al. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* 356, 92–95. doi: 10.1126/science.aal3327.

Dudchenko, O., Shamim, M. S., Batra, S. S., Durand, N. C., Musial, N. T., Mostofa, R., et al. (2018). The Juicebox Assembly Tools module facilitates *de novo* assembly of mammalian genomes with chromosome-length scaffolds for under $1000. Genomics doi: 10.1101/254797.

Ebert, D., Feuermann, M., Gaudet, P., Harris, N. L., Hill, D. P., Lee, R., et al. (2023). The Gene Ontology knowledgebase in 2023.

Eeckhaut, T., Lakshmanan, P. S., Deryckere, D., Van Bockstaele, E., and Van Huylenbroeck, J. (2013). Progress in plant protoplast research. *Planta* 238, 991–1003. doi: 10.1007/s00425-013-1936-7.

Efremova, N., Schreiber, L., Bär, S., Heidmann, I., Huijser, P., Wellesen, K., et al. (2004). Functional conservation and maintenance of expression pattern of FIDDLEHEAD-like genes in Arabidopsis and Antirrhinum. *Plant Mol Biol* 56, 821–837. doi: 10.1007/s11103-004-5576-y.

Efroni, I., Ip, P.-L., Nawy, T., Mello, A., and Birnbaum, K. D. (2015). Quantification of cell identity from single-cell gene expression profiles. *Genome Biol* 16, 9. doi: 10.1186/s13059-015-0580-x.

Efroni, I., Mello, A., Nawy, T., Ip, P.-L., Rahni, R., DelRose, N., et al. (2016). Root Regeneration Triggers an Embryo-like Sequence Guided by Hormonal Interactions. *Cell* 165, 1721–1733. doi: 10.1016/j.cell.2016.04.046.

Egea-Cortines, M. (1999). Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus. *The EMBO Journal* 18, 5370–5379. doi: 10.1093/emboj/18.19.5370.

Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic Gene Expression in a Single Cell. *Science* 297, 1183–1186. doi: 10.1126/science.1070919.

Engler, C., Kandzia, R., and Marillonnet, S. (2008). A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS ONE* 3, e3647. doi: 10.1371/journal.pone.0003647.

Faraco, M., Di Sansebastiano, G. P., Spelt, K., Koes, R. E., and Quattrocchio, F. M. (2011). One Protoplast Is Not the Other! *Plant Physiology* 156, 474–478. doi: 10.1104/pp.111.173708.

Fatima, U., Balasubramaniam, D., Khan, W. A., Kandpal, M., Vadassery, J., Arockiasamy, A., et al. (2022). SWEET11 and SWEET12 transporters function in tandem to modulate sugar flux in Arabidopsis: An account of the underlying unique structure–function relationship. 2022.03.26.485957. doi: 10.1101/2022.03.26.485957.

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

Fernandez-Da Silva, R., and Menendez-Yuffa, A. (2006). Viability in protoplasts and cell suspensions of Coffea arabica cv. Catimor. *Electron. J. Biotechnol.* 9, 0–0. doi: 10.2225/vol9-issue5-fulltext-4.

Fernandez-Pozo, N., Menda, N., Edwards, J. D., Saha, S., Tecle, I. Y., Strickler, S. R., et al. (2015). The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. *Nucleic Acids Research* 43, D1036–D1041. doi: 10.1093/nar/gku1195.

Fornes, O., Castro-Mondragon, J. A., Khan, A., van der Lee, R., Zhang, X., Richmond, P. A., et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 48, D87–D92. doi: 10.1093/nar/gkz1001.

Frank, M. H., and Chitwood, D. H. (2016). Plant chimeras: The good, the bad, and the “Bizzaria.” *Dev. Biol.* 419, 41–53. doi: 10.1016/j.ydbio.2016.07.003.

Gala, H. P., Lanctot, A., Jean-Baptiste, K., Guiziou, S., Chu, J. C., Zemke, J. E., et al. (2021). A single-cell view of the transcriptome during lateral root initiation in Arabidopsis thaliana. *The Plant Cell* 33, 2197–2220. doi: 10.1093/plcell/koab101.

Gallego Romero, I., Pai, A. A., Tung, J., and Gilad, Y. (2014). RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biol* 12, 42. doi: 10.1186/1741-7007-12-42.

Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36, 2628–2629. doi: 10.1093/bioinformatics/btz931.

Gerats, A. G., Huits, H., Vrijlandt, E., Maraña, C., Souer, E., and Beld, M. (1990). Molecular characterization of a nonautonomous transposable element (dTph1) of petunia. *Plant Cell* 2, 1121–1128.

Gerats, T., and Strommer, J. eds. (2009). *Petunia: evolutionary, developmental and physiological genetics*. New York: Springer.

Gerdes, M. J., Sood, A., Sevinsky, C., Pris, A. D., Zavodszky, M. I., and Ginty, F. (2014). Emerging Understanding of Multiscale Tumor Heterogeneity. *Front. Oncol.* 4. doi: 10.3389/fonc.2014.00366.

Guo, G., Huss, M., Tong, G. Q., Wang, C., Li Sun, L., Clarke, N. D., et al. (2010). Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Developmental Cell* 18, 675–685. doi: 10.1016/j.devcel.2010.02.012.

Guo, S., Xu, Y., Liu, H., Mao, Z., Zhang, C., Ma, Y., et al. (2013). The interaction between OsMADS57 and OsTB1 modulates rice tillering via DWARF14. *Nat Commun* 4, 1566. doi: 10.1038/ncomms2542.

Hagemann-Jensen, M., Ziegenhain, C., and Sandberg, R. (2022). Scalable single-cell RNA sequencing from full transcripts with Smart-seq3xpress. *Nat Biotechnol* 40, 1452–1457. doi: 10.1038/s41587-022-01311-4.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573-3587.e29. doi: 10.1016/j.cell.2021.04.048.

Haque, A., Engel, J., Teichmann, S. A., and Lönnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med* 9, 75. doi: 10.1186/s13073-017-0467-4.

Hartman, S., Liu, Z., van Veen, H., Vicente, J., Reinen, E., Martopawiro, S., et al. (2019). Ethylene-mediated nitric oxide depletion pre-adapts plants to hypoxia stress. *Nat Commun* 10, 4020. doi: 10.1038/s41467-019-12045-4.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38, 576–589. doi: 10.1016/j.molcel.2010.05.004.

Honma, T., and Goto, K. (2000). The Arabidopsis floral homeotic gene PISTILLATA is regulated by discrete cis-elements responsive to induction and maintenance signals. *Development* 127, 2021–2030. doi: 10.1242/dev.127.10.2021.

Hou, Q., Zhao, W., Lu, L., Wang, L., Zhang, T., Hu, B., et al. (2022). Overexpression of HLH4 Inhibits Cell Elongation and Anthocyanin Biosynthesis in Arabidopsis thaliana. *Cells* 11, 1087. doi: 10.3390/cells11071087.

Huang, T., and Irish, V. F. (2016). Gene networks controlling petal organogenesis. *Journal of Experimental Botany* 67, 61–68. doi: 10.1093/jxb/erv444.

Iida, H., Mähönen, A. P., Jürgens, G., and Takada, S. (2023). Epidermal injury-induced derepression of key regulator ATML1 in newly exposed cells elicits epidermis regeneration. *Nat Commun* 14, 1031. doi: 10.1038/s41467-023-36731-6.

Iida, H., and Takada, S. (2021). A Quarter Century History of ATML1 Gene Research. *Plants (Basel)* 10, 290. doi: 10.3390/plants10020290.

Irish, V. (2017). The ABC model of floral development. *Current Biology* 27, R887–R890. doi: 10.1016/j.cub.2017.03.045.

Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J.-B., Lönnerberg, P., et al. (2011). Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* 21, 1160–1167. doi: 10.1101/gr.110882.110.

Jack, T. (2004). Molecular and Genetic Mechanisms of Floral Control. *The Plant Cell* 16, S1–S17. doi: 10.1105/tpc.017038.

Janssen, P., Kliesmete, Z., Vieth, B., Adiconis, X., Simmons, S., Marshall, J., et al. (2023). The effect of background noise and its removal on the analysis of single-cell expression data. *Genome Biol* 24, 140. doi: 10.1186/s13059-023-02978-x.

Jean-Baptiste, K., McFaline-Figueroa, J. L., Alexandre, C. M., Dorrity, M. W., Saunders, L., Bubb, K. L., et al. (2019). Dynamics of Gene Expression in Single Root Cells of *Arabidopsis thaliana*. *Plant Cell* 31, 993–1011. doi: 10.1105/tpc.18.00785.

Jing, H., Yang, X., Emenecker, R. J., Feng, J., Zhang, J., Figueiredo, M. R. A. de, et al. (2023). Nitric oxide-mediated S-nitrosylation of IAA17 protein in intrinsically disordered region represses auxin signaling. *Journal of Genetics and Genomics* 50, 473–485. doi: 10.1016/j.jgg.2023.05.001.

Jolliffe, I. T., and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A.* 374, 20150202. doi: 10.1098/rsta.2015.0202.

Jussieu, A.-L. (1803). Sur le Pétunia, genre nouveau de la famille des plantes solanées. *Annales du Muséum National d'Histoire Naturelle* 2, 214–216.

Kang, H. M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., et al. (2018). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol* 36, 89–94. doi: 10.1038/nbt.4042.

Kang, M., Choi, Y., Kim, H., and Kim, S. (2022). Single-cell RNA-sequencing of *Nicotiana attenuata* corolla cells reveals the biosynthetic pathway of a floral scent. *New Phytologist* 234, 527–544. doi: 10.1111/nph.17992.

Kim, J.-Y., Symeonidi, E., Pang, T. Y., Denyer, T., Weidauer, D., Bezrutczyk, M., et al. (2021). Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *The Plant Cell* 33, 511–530. doi: 10.1093/plcell/koaa060.

Koes, R., Verweij, W., and Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci* 10, 236–242. doi: 10.1016/j.tplants.2005.03.002.

Köster-Töpfer, M., Frommer, W. B., Rocha-Sosa, M., and Willmitzer, L. (1990). Presence of a transposon-like element in the promoter region of an inactive patatin gene in Solanum tuberosum L. *Plant Mol Biol* 14, 239–247. doi: 10.1007/BF00018564.

Kramer, E. M., Dorit, R. L., and Irish, V. F. (1998). Molecular Evolution of Genes Controlling Petal and Stamen Development: Duplication and Divergence Within the APETALA3 and PISTILLATA MADS-Box Gene Lineages. *Genetics* 149, 765–783. doi: 10.1093/genetics/149.2.765.

Kruskal, W. H., and Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association* 47, 583–621. doi: 10.1080/01621459.1952.10483441.

Kurtzer, G. M., Sochat, V., and Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PLOS ONE* 12, e0177459. doi: 10.1371/journal.pone.0177459.

Kusters, E., Della Pina, S., Castel, R., Souer, E., and Koes, R. (2015). Changes in *cis* -regulatory elements of a key floral regulator are associated with divergence of inflorescence architectures. *Development*, dev.121905. doi: 10.1242/dev.121905.

Lai, X., Vega-Léon, R., Hugouvieux, V., Blanc-Mathieu, R., Van Der Wal, F., Lucas, J., et al. (2021). The intervening domain is required for DNA-binding and functional identity of plant MADS transcription factors. *Nat Commun* 12, 4760. doi: 10.1038/s41467-021-24978-w.

Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J. U., and Forner, J. (2013). GreenGate - A Novel, Versatile, and Efficient Cloning System for Plant Transgenesis. *PLoS ONE* 8, e83043. doi: 10.1371/journal.pone.0083043.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559. doi: 10.1186/1471-2105-9-559.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357–359. doi: 10.1038/nmeth.1923.

Lenser, T., Theißen, G., and Dittrich, P. (2009). Developmental Robustness by Obligate Interaction of Class B Floral Homeotic Genes and Proteins. *PLOS Computational Biology* 5, e1000264. doi: 10.1371/journal.pcbi.1000264.

Li, S.-B., Xie, Z.-Z., Hu, C.-G., and Zhang, J.-Z. (2016). A Review of Auxin Response Factors (ARFs) in Plants. *Frontiers in Plant Science* 7. Available at:

https://www.frontiersin.org/articles/10.3389/fpls.2016.00047 [Accessed September 28, 2023].

Liao, P., Ray, S., Boachon, B., Lynch, J. H., Deshpande, A., McAdam, S., et al. (2021). Cuticle thickness affects dynamics of volatile emission from petunia flowers. *Nat Chem Biol* 17, 138–145. doi: 10.1038/s41589-020-00670-w.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* 326, 289–293. doi: 10.1126/science.1181369.

Lipniacki, T., Paszek, P., Marciniak-Czochra, A., Brasier, A. R., and Kimmel, M. (2006). Transcriptional stochasticity in gene expression. *Journal of Theoretical Biology* 238, 348–367. doi: 10.1016/j.jtbi.2005.05.032.

Liu, C., Wu, T., Fan, F., Liu, Y., Wu, L., Junkin, M., et al. (2019). A portable and cost-effective microfluidic system for massively parallel single-cell transcriptome profiling. Genomics doi: 10.1101/818450.

Liu, Q., Luo, L., and Zheng, L. (2018). Lignins: Biosynthesis and Biological Functions in Plants. *Int J Mol Sci* 19, 335. doi: 10.3390/ijms19020335.

Liu, S., Thennavan, A., Garay, J. P., Marron, J. S., and Perou, C. M. (2021). MultiK: an automated tool to determine optimal cluster numbers in single-cell RNA sequencing data. *Genome Biol* 22, 232. doi: 10.1186/s13059-021-02445-5.

Lohmann, J. U., Hong, R. L., Hobe, M., Busch, M. A., Parcy, F., Simon, R., et al. (2001). A Molecular Link between Stem Cell Regulation and Floral Patterning in Arabidopsis. *Cell* 105, 793–803. doi: 10.1016/S0092-8674(01)00384-1.

Long, Y., Cheddadi, I., Mosca, G., Mirabet, V., Dumond, M., Kiss, A., et al. (2020). Cellular Heterogeneity in Pressure and Growth Emerges from Tissue Topology and Geometry. *Curr Biol* 30, 1504-1516.e8. doi: 10.1016/j.cub.2020.02.027.

Long, Y., Stahl, Y., Weidtkamp-Peters, S., Postma, M., Zhou, W., Goedhart, J., et al. (2017). In vivo FRET-FLIM reveals cell-type-specific protein interactions in Arabidopsis roots. *Nature* 548, 97–102. doi: 10.1038/nature23317.

Loreti, E., and Perata, P. (2020). The Many Facets of Hypoxia in Plants. *Plants* 9, 745. doi: 10.3390/plants9060745.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. doi: 10.1186/s13059-014-0550-8.

Lyons, E., and Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *The Plant Journal* 53, 661–673. doi: 10.1111/j.1365-313X.2007.03326.x.

Maaten, L. van der, and Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 2579–2605.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161, 1202–1214. doi: 10.1016/j.cell.2015.05.002.

MACS3 project team (2020). MACS3. Available at: https://github.com/macs3-project/MACS.

Maier, T., Güell, M., and Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Letters* 583, 3966–3973. doi: 10.1016/j.febslet.2009.10.036.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380. doi: 10.1038/nature03959.

Martin, C., Bhatt, K., Baumann, K., Jin, H., Zachgo, S., Roberts, K., et al. (2002). The mechanics of cell fate determination in petals. *Phil. Trans. R. Soc. Lond. B* 357, 809–813. doi: 10.1098/rstb.2002.1089.

Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer* 12, 323–334. doi: 10.1038/nrc3261.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences* 36, 344–355. doi: 10.1073/pnas.36.6.344.

McClintock, B. (1965). The control of gene action in maize. *Brookhaven Symposia in Biology* 18, 162–184.

McClintock, B. (1987). *The discovery and characterization of transposable elements: the collected papers of Barbara McClintock*. New York: Garland Pub.

McGinnis, C. S., Murrow, L. M., and Gartner, Z. J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Systems* 8, 329-337.e4. doi: 10.1016/j.cels.2019.03.003.

McInnes, L., Healy, J., and Melville, J. (2020). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Available at: http://arxiv.org/abs/1802.03426 [Accessed January 26, 2023].

Mehrotra, R., Bhalothia, P., Bansal, P., Basantani, M. K., Bharti, V., and Mehrotra, S. (2014). Abscisic acid and abiotic stress tolerance – Different tiers of regulation. *Journal of Plant Physiology* 171, 486–496. doi: 10.1016/j.jplph.2013.12.007.

Melzer, R., Härter, A., Rümpler, F., Kim, S., Soltis, P. S., Soltis, D. E., et al. (2014). DEF- and GLO-like proteins may have lost most of their interaction partners during angiosperm evolution. *Ann Bot* 114, 1431–1443. doi: 10.1093/aob/mcu094.

Melzer, R., and Theissen, G. (2009). Reconstitution of “floral quartets” in vitro involving class B and class E floral homeotic proteins. *Nucleic Acids Res.* 37, 2723–2736. doi: 10.1093/nar/gkp129.

Melzer, R., and Theißen, G. (2009). Reconstitution of ‘floral quartets’ in vitro involving class B and class E floral homeotic proteins. *Nucleic Acids Res* 37, 2723–2736. doi: 10.1093/nar/gkp129.

Melzer, R., Verelst, W., and Theissen, G. (2009). The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes in vitro. *Nucleic Acids Res.* 37, 144–157. doi: 10.1093/nar/gkn900.

Mendes, M. A., Guerra, R. F., Berns, M. C., Manzo, C., Masiero, S., Finzi, L., et al. (2013). MADS domain transcription factors mediate short-range DNA looping that is essential for target gene expression in Arabidopsis. *Plant Cell* 25, 2560–2572. doi: 10.1105/tpc.112.108688.

Messenguy, F., and Dubois, E. (2003). Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene* 316, 1–21. doi: 10.1016/S0378-1119(03)00747-9.

Monniaux, M., and Vandenbussche, M. (2018). How to Evolve a Perianth: A Review of Cadastral Mechanisms for Perianth Identity. *Front Plant Sci* 9, 1573. doi: 10.3389/fpls.2018.01573.

Morel, P., Heijmans, K., Rozier, F., Zethof, J., Chamot, S., Bento, S. R., et al. (2017). Divergence of the Floral A-Function between an Asterid and a Rosid Species. *Plant Cell* 29, 1605–1621. doi: 10.1105/tpc.17.00098.

Natarajan, K. N. (2019). "Single-Cell Tagged Reverse Transcription (STRT-Seq)," in *Single Cell Methods: Sequencing and Proteomics* Methods in Molecular Biology., ed. V. Proserpio (New York, NY: Springer), 133–153. doi: 10.1007/978-1-4939-9240-9_9.

Norman, C., Runswick, M., Pollock, R., and Treisman, R. (1988). Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. *Cell* 55, 989–1003. doi: 10.1016/0092-8674(88)90244-9.

Pajoro, A., Madrigal, P., Muiño, J. M., Matus, J. T., Jin, J., Mecchia, M. A., et al. (2014). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome Biology* 15, R41. doi: 10.1186/gb-2014-15-3-r41.

Pan, Z., Hung, Y., Chen, T., Shih, Y., Lin, Y. J., and Wang, C. (2022). Development of a petal protoplast transfection system for Sinningia speciosa. *Appl Plant Sci* 10, e11476. doi: 10.1002/aps3.11476.

Parcy, F., Nilsson, O., Busch, M. A., Lee, I., and Weigel, D. (1998). A genetic framework for floral patterning. *Nature* 395, 561–566. doi: 10.1038/26903.

Passmore, S., Elble, R., and Tye, B. K. (1989). A protein involved in minichromosome maintenance in yeast binds a transcriptional enhancer conserved in eukaryotes. *Genes Dev.* 3, 921–935. doi: 10.1101/gad.3.7.921.

Pasternak, T., Tietz, O., Rapp, K., Begheldo, M., Nitschke, R., Ruperti, B., et al. (2015). Protocol: an improved and universal procedure for whole-mount immunolocalization in plants. *Plant Methods* 11, 50. doi: 10.1186/s13007-015-0094-2.

Pearson, K. (1901). LIII. *On lines and planes of closest fit to systems of points in space*. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2, 559–572. doi: 10.1080/14786440109462720.

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., and Yanofsky, M. F. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 405, 200–203. doi: 10.1038/35012103.

Perbal, M.-C., Haughn, G., Saedler, H., and Schwarz-Sommer, Z. (1996). Non-cell-autonomous function of the *Antirrhinum* floral homeotic proteins *DEFICIENS* and *GLOBOSA* is exerted by their polar cell-to-cell trafficking. *Development* 122, 3433–3441. doi: 10.1242/dev.122.11.3433.

Pertea, G., and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. *F1000Res* 9, 304. doi: 10.12688/f1000research.23297.1.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33, 290–295. doi: 10.1038/nbt.3122.

Peterson, T. ed. (2013). *Plant Transposable Elements: Methods and Protocols*. Totowa, NJ: Humana Press doi: 10.1007/978-1-62703-568-2.

Pnueli, L., Hareven, D., Broday, L., Hurwitz, C., and Lifschitz, E. (1994). The TM5 MADS Box Gene Mediates Organ Differentiation in the Three Inner Whorls of Tomato Flowers. *Plant Cell* 6, 175–186. doi: 10.1105/tpc.6.2.175.

Pucciariello, C., and Perata, P. (2017). New insights into reactive oxygen species and nitric oxide signalling under low oxygen in plants. *Plant, Cell & Environment* 40, 473–482. doi: 10.1111/pce.12715.

Pucciariello, C., and Perata, P. (2021). The Oxidative Paradox in Low Oxygen Stress in Plants. *Antioxidants (Basel)* 10, 332. doi: 10.3390/antiox10020332.

Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., et al. (2014). Structural Basis for the Oligomerization of the MADS Domain Transcription Factor SEPALLATA3 in *Arabidopsis*. *The Plant Cell* 26, 3603–3615. doi: 10.1105/tpc.114.127910.

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/.

Ray, S., Savoie, B. M., Dudareva, N., and Morgan, J. A. (2022). Diffusion of volatile organics and water in the epicuticular waxes of petunia petal epidermal cells. *The Plant Journal* 110, 658–672. doi: 10.1111/tpj.15693.

Reale, L., Porceddu, A., Lanfaloni, L., Moretti, C., Zenoni, S., Pezzotti, M., et al. (2002). Patterns of cell division and expansion in developing petals of Petunia hybrida. *Sexual Plant Reproduction* 15, 123–132. doi: 10.1007/s00497-002-0150-8.

Reed, A., Rudall, P. J., Brockington, S. F., and Glover, B. J. (2022). Conical petal epidermal cells, regulated by the MYB transcription factor MIXTA, have an ancient origin within the angiosperms. *Journal of Experimental Botany* 73, 5490–5502. doi: 10.1093/jxb/erac223.

Rijpkema, A. S., Royaert, S., Zethof, J., van der Weerden, G., Gerats, T., and Vandenbussche, M. (2006). Analysis of the *Petunia TM6* MADS Box Gene Reveals Functional Divergence within the *DEF* / *AP3* Lineage. *The Plant Cell* 18, 1819–1832. doi: 10.1105/tpc.106.042937.

Robinson, J. T., Thorvaldsdottir, H., Turner, D., and Mesirov, J. P. (2023). igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics* 39, btac830. doi: 10.1093/bioinformatics/btac830.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat Biotechnol* 29, 24–26. doi: 10.1038/nbt.1754.

Rosas-Arellano, A., Villalobos-González, J. B., Palma-Tirado, L., Beltrán, F. A., Cárabez-Trejo, A., Missirlis, F., et al. (2016). A simple solution for antibody signal enhancement in immunofluorescence and triple immunogold assays. *Histochem Cell Biol* 146, 421–430. doi: 10.1007/s00418-016-1447-2.

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 360, 176–182. doi: 10.1126/science.aam8999.

Ryu, K. H., Huang, L., Kang, H. M., and Schiefelbein, J. (2019). Single-Cell RNA Sequencing Resolves Molecular Relationships Among Individual Plant Cells. *Plant Physiol.* 179, 1444–1456. doi: 10.1104/pp.18.01482.

Sakuma, C., Nakagawa, M., Tomioka, Y., Maruyama, T., Entzminger, K., Fleming, J. K., et al. (2022). Western blotting of native proteins from agarose gels. *BioTechniques* 72, 207–218. doi: 10.2144/btn-2022-0012.

Satterlee, J. W., Strable, J., and Scanlon, M. J. (2020). Plant stem-cell organization and differentiation at single-cell resolution. *Proc Natl Acad Sci USA* 117, 33689–33699. doi: 10.1073/pnas.2018788117.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682. doi: 10.1038/nmeth.2019.

Schrick, K., Ahmad, B., and Nguyen, H. V. (2023a). HD-Zip IV transcription factors: Drivers of epidermal cell fate integrate metabolic signals. *Current Opinion in Plant Biology* 75, 102417. doi: 10.1016/j.pbi.2023.102417.

Schrick, K., Ahmad, B., and Nguyen, H. V. (2023b). HD-Zip IV transcription factors: Drivers of epidermal cell fate integrate metabolic signals. *Current Opinion in Plant Biology* 75, 102417. doi: 10.1016/j.pbi.2023.102417.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., and Sommer, H. (1990). Genetic Control of Flower Development by Homeotic Genes in *Antirrhinum majus*. *Science* 250, 931–936. doi: 10.1126/science.250.4983.931.

Shahan, R., Hsu, C.-W., Nolan, T. M., Cole, B. J., Taylor, I. W., Vlot, A. H. C., et al. (2020). A single cell *Arabidopsis* root atlas reveals developmental trajectories in wild type and cell identity mutants. Plant Biology doi: 10.1101/2020.06.29.178863.

Shaw, R., Tian, X., and Xu, J. (2021). Single-Cell Transcriptome Analysis in Plants: Advances and Challenges. *Molecular Plant* 14, 115–126. doi: 10.1016/j.molp.2020.10.012.

Shulse, C. N., Cole, B. J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M., et al. (2019). High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. *Cell Reports* 27, 2241-2247.e4. doi: 10.1016/j.celrep.2019.04.054.

Simon, A. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Sink, K. C. (1984). *Petunia: Monographs on Theoretical and Applied Genetics*. New York: Springer-Verlag.

Smaczniak, C., Immink, R. G. H., Angenent, G. C., and Kaufmann, K. (2012). Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development* 139, 3081–3098. doi: 10.1242/dev.074674.

Smaczniak, C., Muiño, J. M., Chen, D., Angenent, G. C., and Kaufmann, K. (2017). Differences in DNA Binding Specificity of Floral Homeotic Protein Complexes Predict Organ-Specific Target Genes. *The Plant Cell* 29, 1822–1835. doi: 10.1105/tpc.17.00145.

Smith, F. J., de Jong, J. H., and Oud, J. L. (1975). The use of primary trisomics for the localization of genes on the seven different chromosomes ofPetunia hybrida. I. Triplo V. *Genetica* 45, 361–370. doi: 10.1007/BF01508311.

Sommer, H., Beltrán, J. P., Huijser, P., Pape, H., Lönnig, W. E., Saedler, H., et al. (1990). Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. *The EMBO Journal* 9, 605–613. doi: 10.1002/j.1460-2075.1990.tb08152.x.

Spelt, C., Quattrocchio, F., Mol, J. N. M., and Koes, R. (2000). anthocyanin1 of Petunia Encodes a Basic Helix-Loop-Helix Protein That Directly Activates Transcription of Structural Anthocyanin Genes. 13.

Stehmann, J. R., Lorenz-Lemke, A. P., Freitas, L. B., and Semir, J. (2009). “The Genus Petunia,” in *Petunia*, eds. T. Gerats and J. Strommer (New York, NY: Springer New York), 1–28. doi: 10.1007/978-0-387-84796-2_1.

Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B. Z., Mauck, W. M., et al. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* 19, 224. doi: 10.1186/s13059-018-1603-1.

Svensson, V., Vento-Tormo, R., and Teichmann, S. A. (2018a). Exponential scaling of single-cell RNA- seq in the last decade. 13.

Svensson, V., Vento-Tormo, R., and Teichmann, S. A. (2018b). Exponential scaling of single-cell RNA-seq in the past decade. *Nat Protoc* 13, 599–604. doi: 10.1038/nprot.2017.149.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 6, 377–382. doi: 10.1038/nmeth.1315.

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. *Bioinformatics* 31, 2032–2034. doi: 10.1093/bioinformatics/btv098.

The HDF Group (2006). Hierarchical Data Format (HDF5). Available at: https://hdfgroup.github.io/hdf5/.

Theißen, G. (2001). Development of floral organ identity: stories from the MADS house. *Current Opinion in Plant Biology* 4, 75–85. doi: 10.1016/S1369-5266(00)00139-4.

Theissen, G., and Saedler, H. (2001). Plant biology. Floral quartets. *Nature* 409, 469–471. doi: 10.1038/35054172.

Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14, 178–192. doi: 10.1093/bib/bbs017.

Titouh, K., Khelifi, L., Slaoui, M., Boufis, N., Morsli, A., Titouh Hadj Moussa, K., et al. (2015). A simplified Protocol to Induce Callogenesis in Protoplasts of Date Palm (Phoenix dactylifera L.) Cultivars. *Iran J Biotech* 13, 26–35. doi: 10.15171/ijb.1054.

Tung, C.-C., Kuo, S.-C., Yang, C.-L., Yu, J.-H., Huang, C.-E., Liou, P.-C., et al. (2023). Single-cell transcriptomics unveils xylem cell development and evolution. *Genome Biol* 24, 3. doi: 10.1186/s13059-022-02845-1.

van Mourik, H., Chen, P., Smaczniak, C., Boeren, S., Kaufmann, K., Bemer, M., et al. (2023). Dual specificity and target gene selection by the MADS-domain protein FRUITFULL. *Nat. Plants* 9, 473–485. doi: 10.1038/s41477-023-01351-x.

Vandenbussche, M., Chambrier, P., Rodrigues Bento, S., and Morel, P. (2016). Petunia, Your Next Supermodel? *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00072.

Vandenbussche, M., Zethof, J., Royaert, S., Weterings, K., and Gerats, T. (2004). The Duplicated B-Class Heterodimer Model: Whorl-Specific Effects and Complex Genetic Interactions in *Petunia hybrida* Flower Development. *Plant Cell* 16, 741–754. doi: 10.1105/tpc.019166.

Waltman, L., and van Eck, N. J. (2013). A smart local moving algorithm for large-scale modularity-based community detection. *Eur. Phys. J. B* 86, 471. doi: 10.1140/epjb/e2013-40829-0.

Wang, W., Zhong, Y., Zhuang, Z., Xie, J., Lu, Y., Huang, C., et al. (2021a). Multiregion single-cell sequencing reveals the transcriptional landscape of the immune microenvironment of colorectal cancer. *Clinical & Translational Med* 11, e253. doi: 10.1002/ctm2.253.

Wang, X., He, Y., Zhang, Q., Ren, X., and Zhang, Z. (2021b). Direct Comparative Analyses of 10X Genomics Chromium and Smart-seq2. *Genomics, Proteomics & Bioinformatics* 19, 253–266. doi: 10.1016/j.gpb.2020.02.005.

Wang, Y., Wang, N., Xu, H., Jiang, S., Fang, H., Su, M., et al. (2018). Auxin regulates anthocyanin biosynthesis through the Aux/IAA–ARF signaling pathway in apple. *Hortic Res* 5, 1–11. doi: 10.1038/s41438-018-0068-4.

Wang, Z., Wang, Y., Kohalmi, S. E., Amyot, L., and Hannoufa, A. (2016). SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 2 controls floral organ development and plant fertility by activating ASYMMETRIC LEAVES 2 in Arabidopsis thaliana. *Plant Mol Biol* 92, 661–674. doi: 10.1007/s11103-016-0536-x.

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., and Meyerowitz, E. M. (1992). LEAFY controls floral meristem identity in Arabidopsis. *Cell* 69, 843–859.

Weigel, D., and Meyerowitz, E. M. (1994). The ABCs of floral homeotic genes. *Cell* 78, 203–209. doi: 10.1016/0092-8674(94)90291-7.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. 2nd ed. 2016. Cham: Springer International Publishing : Imprint: Springer doi: 10.1007/978-3-319-24277-4.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., et al. (2019). Welcome to the Tidyverse. *JOSS* 4, 1686. doi: 10.21105/joss.01686.

Wijsman, H. J. W. (1986). Evidence for transposition in Petunia. *Theoret. Appl. Genetics* 71, 791–796. doi: 10.1007/BF00276419.

Wilson, J. H., and Hunt, T. eds. (2002). *Molecular biology of the cell. prob, 2002: A problems approach / John Wilson & Tim Hunt*. 4. ed. New York: Garland.

Wittich, P. E., de Heer, R. F., Cheng, X.-F., Kieft, H., Colombo, L., Angenent, G. C., et al. (1999). Immunolocalization of the petunia floral binding proteins 7 and 11 during seed development in wild-type and expression mutants ofPetunia hybrida. *Protoplasma* 208, 224–229. doi: 10.1007/BF01279093.

Wuest, S. E., O'Maoileidigh, D. S., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., et al. (2012). Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proceedings of the National Academy of Sciences* 109, 13452–13457. doi: 10.1073/pnas.1207075109.

Xu, L., Pan, R., and Zhang, W. (2020). Membrane lipids are involved in plant response to oxygen deprivation. *Plant Signaling & Behavior* 15, 1771938. doi: 10.1080/15592324.2020.1771938.

Xu, X., Smaczniak, C., Muino, J. M., and Kaufmann, K. (2021). Cell identity specification in plants: lessons from flower development. *Journal of Experimental Botany* 72, 4202–4217. doi: 10.1093/jxb/erab110.

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A., and Meyerowitz, E. M. (1990). The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. *Nature* 346, 35–39. doi: 10.1038/346035a0.

Yephremov, A., Wisman, E., Huijser, P., Huijser, C., Wellesen, K., and Saedler, H. (1999). Characterization of the FIDDLEHEAD gene of Arabidopsis reveals a link between adhesion response and cell differentiation in the epidermis. *Plant Cell* 11, 2187–2201.

Yoo, A. B., Jette, M. A., and Grondona, M. (2003). SLURM: Simple Linux Utility for Resource Management. in *Job Scheduling Strategies for Parallel Processing* Lecture Notes in

Computer Science., eds. D. Feitelson, L. Rudolph, and U. Schwiegelshohn (Berlin, Heidelberg: Springer), 44–60. doi: 10.1007/10968987_3.

Zeisel, A., Muñoz-Manchado, A. B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347, 1138–1142. doi: 10.1126/science.aaa1934.

Zhang, B., and Horvath, S. (2005). A General Framework for Weighted Gene Co-Expression Network Analysis. *Statistical Applications in Genetics and Molecular Biology* 4. doi: 10.2202/1544-6115.1128.

Zhang, T.-Q., Chen, Y., and Wang, J.-W. (2021a). A single-cell analysis of the Arabidopsis vegetative shoot apex. *Dev Cell*. doi: 10.1016/j.devcel.2021.02.021.

Zhang, T.-Q., Chen, Y., and Wang, J.-W. (2021b). A single-cell analysis of the Arabidopsis vegetative shoot apex. *Developmental Cell* 56, 1056-1074.e8. doi: 10.1016/j.devcel.2021.02.021.

Zhang, T.-Q., Xu, Z.-G., Shang, G.-D., and Wang, J.-W. (2019). A Single-Cell RNA Sequencing Profiles the Developmental Landscape of Arabidopsis Root. *Molecular Plant* 12, 648–660. doi: 10.1016/j.molp.2019.04.004.

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., et al. (2017a). Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 8, 14049. doi: 10.1038/ncomms14049.

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., et al. (2017b). Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 8, 14049. doi: 10.1038/ncomms14049.

Zheng, H., Wu, F., Li, S., Zhang, X. S., and Sui, N. (2021). Single-cell profiling lights different cell trajectories in plants. *aBIOTECH* 2, 64–78. doi: 10.1007/s42994-021-00040-7.

Zheng, S. C., Stein-O'Brien, G., Augustin, J. J., Slosberg, J., Carosso, G. A., Winer, B., et al. (2022). Universal prediction of cell-cycle position using transfer learning. *Genome Biology* 23, 41. doi: 10.1186/s13059-021-02581-y.