Spatial and temporal dynamics of gene expression across wheat inflorescence development

Katie A. Long

A thesis submitted to the University of East Anglia

for the degree of Doctor of Philosophy

John Innes Centre

September 2025

Abstract

Inflorescence architecture is central to reproductive success in grasses and strongly influences yield-related traits in crops such as wheat (*Triticum aestivum L.*). Inflorescence architecture is patterned by gene expression, which influences the initiation and arrangement of spikelets —the repeating units of the grass inflorescence. Characterising these expression patterns across space and time is essential to understanding inflorescence development.

In this thesis, I apply spatial transcriptomics to characterise gene expression across wheat inflorescence development. Using Multiplexed Error-Robust Fluorescence *In Situ* Hybridisation (MERFISH), we mapped the expression patterns of 200 genes to cellular resolution across four developmental stages, within their native tissue context. Analysis of ~50,000 cells identified 18 expression domains and their enriched genes, revealing the spatio-temporal organisation of spikelet and floral development, and characterising tissue-level gene markers.

Using MERFISH, we investigated gene expression patterning along the apical–basal axis of the wheat spike. In wheat, the lanceolate-shaped inflorescence is defined by rudimentary spikelets at the base, which form as a result of delayed spikelet and floral development. Using domain- and cell-level maps, we identified distinct, spatially coordinated expression patterns that distinguish axillary meristems and their subtending leaf ridges across the apical-basal axis before visible spikelet formation, highlighting novel factors that pattern meristem identity and transition.

Given the novelty of spatial transcriptomics techniques to plant sciences, I document the optimisation of both imaging-based and sequencing-based approaches. In addition to MERFISH, I present the first application of Spatio-Temporal Enhanced Resolution Omics sequencing (Stereo-seq) to wheat inflorescence tissue, critically assessing its performance and limitations. Together, this work establishes spatial transcriptomics as a powerful technique for characterising developmental programs in complex plant tissues and provides new insights into the genetic regulation of wheat inflorescence development.

Acknowledgements

How lucky am I to have so many people to thank in so many ways! A small acknowledgements section could never begin to capture the way I feel, but I will give it try.

To my supervisor, Cristóbal, where can I even begin! Your mentorship has sent me on such a positive path in both my career and in life. I started in the lab straight out from my undergrad, new to the world of science, and perhaps a bit unprepared for it all. You always made me feel like I was capable of great things, and you gave me the opportunities to reflect that. You've taught me not only great science, but also the value in community. For this, I could never thank you enough.

To my PhD buddies Isa & Max – oh how the time has flown by! The idea of spending four intense years of PhD life with two random strangers seems like quite the gamble, but every day with you two throughout this journey has felt like winning the lottery. I can see the ways in which we have grown and changed, and I feel so honoured to have shared this intense (but magical) time of our lives together. I feel that this bond we have built together will last a lifetime. I can't wait to see all of the beautiful things you two will accomplish in the years to come.

To Nikolai, I know how much you hate getting a compliment, but you can't stop me this time. I hope you keep reading! Thank you for all of the fun times we shared- every opportunity to have a laugh with you was a highlight of my time at JIC. Through every step of the way, I felt encouraged by you, not only in the good times, but in the bad times too. You are a cheerleader for all of the students in the lab; I can't imagine this journey without you there supporting me and showing me the way. Thank you.

To rest of the Uauy lab, thank you for the ways in which you have taught me. To Anna, you've been a role model to me not only in science, but also in life. Thank you for sending me down a great path for my project, and for being a wonderful friend. To Sophie, days spent having a gab or a giggle in the tearoom with you was a true gift. You light up every room you enter, and I am so grateful for you. Our lab is full of great people with kind hearts, thank you to Simmo, Pam, Sophie, Marina, Vale, Tobin, Ella, Bernice, Andy, Jesus, Aura, Yunchuan, Lucy, Yuanrong, Ricardo, and to many others that came to visit us at the lab. I am forever grateful for you all!

To everyone in the tearoom, thank you for the fun times we shared. A great scholar once said, "I wish there was a way to know you're in the good old days, before you actually left them". I always knew that I would look back at times we spent together as the 'good old days'. You all filled my life with joy and levity, thank you.

I couldn't have made it through my PhD without all of my loved ones outside the lab supporting me along the way. Thank you to Mel, Mac, & Cheese for giving me a beautiful place to call home, and for taking care of me during my most vulnerable moments. Thank you to Issy, who has taught me to have an open heart to everyone and a gentle kindness towards myself. When I am with you, the future feels warm and bright. To Lu, our walks together helped me re-centre myself. You remind me of the way I want to show up in the world, and the type of person I want to be. To Roo, you are my other half. Through all the ways we have grown and changed, you have always made feel understood.

Most importantly, thank you to my family. To my mom, dad, and sister- all of the best parts of myself have come from you. You've given me everything and more. Thank you.

Table of contents

Table of figures

Table of tables

Supplementary data statement

Supplementary data referenced in the text are made available at:

doi.org/10.5281/zenodo.15720855

Code for original analysis is made available via Github:

https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics

Abbreviations

AM	Axillary meristem
BM	Branch meristem
CID	Coordinate Identity Sequence
DNB	DNA nanoball
ED	Expression domain
FM	Floral meristem
FISH	Fluorescence *in situ* hybridisation
ISS	*In situ* sequencing
IM	Inflorescence meristem
LR	Leaf ridge
LBM	Long branch meristem
MERFISH	Multiplexed error robust *in situ* hybridisation
O.C.T.	Optimal Cutting Temperature
PBM	Primary branch meristem
PC	Principal component
RBS	Rudimentary basal spikelet
SBM	Secondary branch meristem
SAM	Shoot apical meristem
smFISH	Single molecule fluorescence in situ hybridisation
scRNA-seq	Single-cell RNA sequencing
snRNA-seq	Single-nuclei RNA sequencing
Stereo-Seq	SpaTial Enhanced REsolution Omics-sequencing
SM	Spikelet meristem
SPM	Spikelet pair meristem
SR	Spikelet ridge
SAW	Stereo-Seq Analysis Workflow
TSM	Triple spikelet meristem
UMI	Unique molecular identifier
VA	Vegetative apex
$P1^{WT}$	wildtype *VRT-A2a* allele
$P1^{POL}$	misexpression *VRT-A2b* allele

Chapter 1 – General Introduction

1.1 - Cereals & humans – our past, present, and future

The history of our crops is entwined with the history of humanity; to trace one is inevitably to trace the other. This connection is especially clear in cereals, grass species cultivated for their edible grains. During the Neolithic period, cereals played a central role in the shift from nomadic to agricultural societies (Salamini et al. 2002; Hebelstrup et al. 2023), yet their importance predates domestication. Excavations from late Neanderthal sites reveal that wild *Triticeae* grains were cooked and eaten around 40–50 thousand years ago across Europe, the Near East, and Africa (Henry et al. 2014), and 20 thousand years ago, groups in the Levant had begun to settle more permanently, transforming wild cereals into staple foods analogous to modern-day bread (Maher et al. 2012; Arranz-Otaegui et al. 2018).

Agriculture arose independently in multiple regions (Fuller 2010; Fuller et al. 2011), spreading across most continents (Ellis et al. 2013), with the earliest centres of crop domestication emerging around 11,000 years ago (Dillehay et al. 2007; Bettinger et al. 2010; Clement et al. 2010; Fuller et al. 2011). Archaeological remains dated to 8,000–10,000 years ago provide evidence of domesticated cereals, including non-brittle spikes of proto-domesticated wheat (diploid, tetraploid and hexaploid *Triticum* species) and barley (*Hordeum vulgare spp. vulgare*) in the Fertile Crescent (Zeder 2011), non-shattering panicles of rice (*Oryza sativa)* in the Yangtze River valley, China (Fuller et al. 2009; Zheng et al. 2016) , and enlarged seeds of maize (*Zea mays* spp. *mays)* from the central Balsas River valley, Mexico (Piperno et al. 2009; Ranere et al. 2009; Pankin and von Korff 2017).

This study focuses on wheat, a crop with a complex history of domestication. Domesticated wheats occur at multiple ploidy levels, from diploids (2n = 2x = 14) to hexaploids (2n = 6x = 42) (Sax 1921, 1922), with representatives at each level domesticated at different times (Salamini et al. 2002). Polyploid wheats are allopolyploids, carrying sub-genomes (designated as A, B, and D) derived from distinct progenitor species. Their origins reflect successive hybridisation events between closely related wild grasses (Sax 1921, 1922; Glémin et al. 2019; Levy and Feldman 2022).

Tetraploid emmer wheat (*Triticum turgidum*, 2n = 4x = 28, AABB) arose approximately 800,000 years ago from a hybridisation between a diploid species similar to *Triticum urartu*

(AA; diverging from *T. urartu* ~1.3 MYA) and a close relative of *Aegilops speltoides* (BB; diverging from *Ae. speltoides* ~4.3 MYA) (Figure 1.1; Marcussen et al. 2014; Levy and Feldman 2022; Li et al. 2022). Although the precise origins of its domestication are uncertain (Weide 2015), selection for a non-brittle rachis produced the first domesticated form, *T. turgidum* spp. *dicoccon*. This was soon followed by a free-threshing, non-brittle variety, *T. turgidum* spp. *parvicoccum* (Kislev 1979; Nesbitt, M. 2001; Schultze-Motel 2019). Emmer wheat became the dominant cultivated cereal across the Fertile Crescent until the early Bronze Age (Zohary and Hopf 2000). Today it survives as a relict crop in regions such as Ethiopia (Salamini et al. 2002), while its free-threshing derivative, *T. turgidum ssp. durum* (durum or pasta wheat), remains a globally important crop (Dubcovsky and Dvorak 2007).

Hexaploid bread wheat (*T. aestivum*, 2n = 6x = 42, AABBDD) arose more recently, following hybridisation between tetraploid emmer wheat and *Aegilops tauschii* (2n = 2x = 14, DD), which contributed the D genome and established the allohexaploid ~9,000–8,500 years ago (Marcussen et al. 2014). This event likely followed the expansion of cultivated emmer into the distribution range of *Ae. tauschii* (Zohary and Hopf 2000). Shortly after its formation, bread wheat spread rapidly beyond the Fertile Crescent into diverse environments, becoming a global staple (Zohary and Hopf 2000; Levy and Feldman 2022). *T. aestivum* includes five subspecies —ssp. *spelta,* ssp. *macha* (hulled), and the free-threshing ssp. ae*stivum*, ssp. *compactum*, and ssp. *sphaerococcum* —though only ssp. *aestivum* is of major global significance (Levy and Feldman 2022).

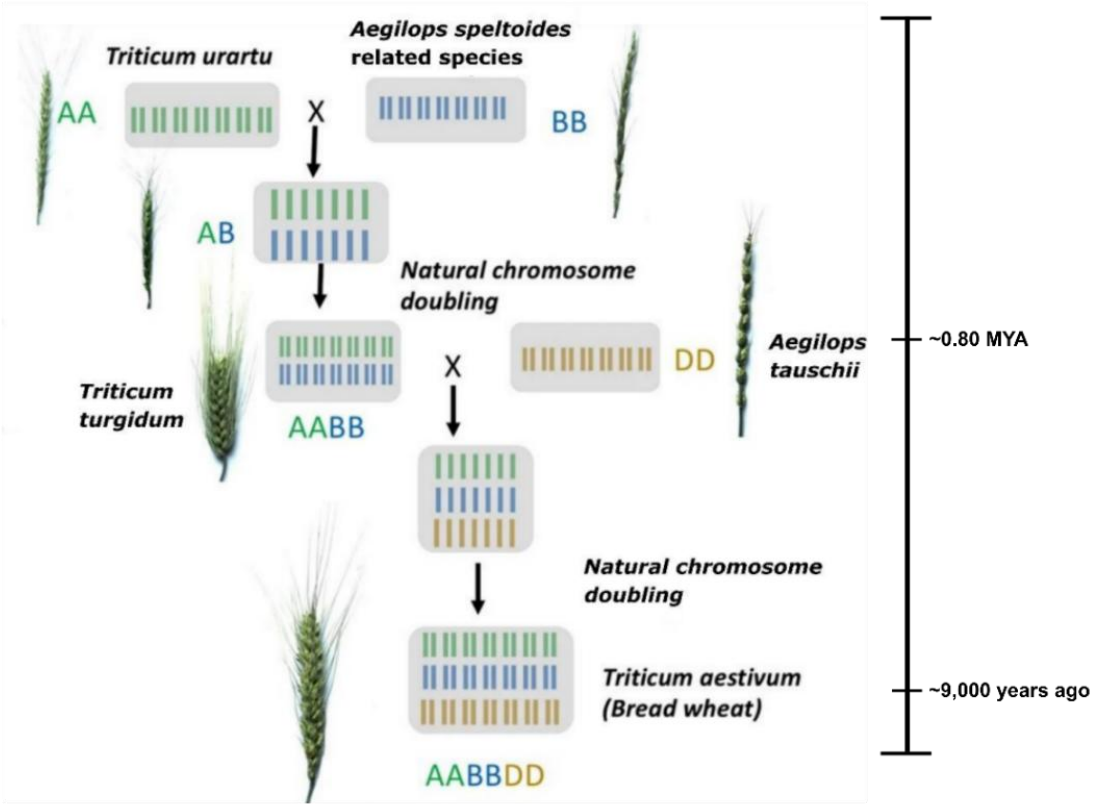


Figure 1.1 – The evolution of tetraploid and hexaploid wheat
Adapted from (Rosyara et al. 2019).

Semi-domesticated grains formed the basis of early agriculture, which in turn transformed humanity's relationship with the planet (Lewis & Maslin, 2015). Agriculture enabled the establishment of sedentary communities and supported higher population densities. It also facilitated the expansion of domesticated crops across wider geographical ranges. This global shift in land use fundamentally reshaped ecosystems: although net primary production has remained relatively constant (Running 2012). An estimated 25–38% of this productivity is now appropriated for food, fuel, fibre, and fodder (Running 2012; Krausmann et al. 2013), reducing what is available for other species. As of 2024, cropland (excluding pastures) occupies 1.8 billion hectares of Earth's 13 billion hectares of land, with wheat, maize, and rice dominating global production (FAO 2024).

Against this backdrop, agriculture not only tells an intimate story of our past but remains central to our future. At present, cereal crops underpin global food security, with wheat alone providing around 20% of dietary calories worldwide (Brinton and Uauy 2019). Yet population growth and dietary demand are intensifying pressure on production. The Food and Agriculture Organisation (FAO) projects a global population of 10 billion by 2050, requiring an additional 165 million hectares of new arable land to meet the requirements of food, feed, fibre, and biofuel (FAO 2018). Meeting this demand is further complicated by climate change, underscoring the need to enhance wheat resilience while minimising environmental impacts (Li et al. 2021c). To meet global demands, it is estimated that crop yields will need to increase by 2-3% annually (Hawkesford et al. 2013). To date, traditional breeding efforts have enhanced wheat yields and improved yield stability through various mechanisms. These include enhancing resistance to pathogens and environmental stresses, in addition to optimising physiological traits that influence energy capture (e.g., photosynthesis, nitrogen use efficiency) and its conversion into yield (e.g., spike size, tiller number). These advances have driven a steady global yield increase of ~1.16% per year over the past six decades (Fischer 2022), yet this rate remains insufficient to meet projected demand.

1.2 – Developmental biology for crops

Over the course of crop domestication, humans selected for traits that enhanced harvestability and palatability, which often involved modifying developmental characteristics of the plant (Boden and Østergaard 2019). In cereals including wheat, barley, rice, and maize, domestication selected for a characteristic suite of traits including non-shattering seeds, increased seed size, apical dominance, and reduced seed dormancy, which are collectively referred to as "Domestication Syndrome" (Hammer 1984; Preece et al. 2017; Alam and Purugganan 2024). Beyond initial domestication, subsequent breeding efforts further increased productivity by targeting developmental traits. A landmark example came in the 1950s, when Dr Norman Borlaug developed semi-dwarf wheat varieties that resisted lodging and allocated more resources to grain production, an innovation that drove the 'Green Revolution' (Khush 2001). While the main genes underlying Domestication Syndrome and the Green Revolution have been identified (reviewed by Boden & Østergaard, 2019), a central challenge moving forward is to elucidate the developmental mechanisms behind these traits and determine their potential for further modification. Addressing these questions is particularly urgent, as future crop improvement in the context of population growth and climate change will depend in part on optimising developmental traits to enhance productivity and efficiency.

The concept of the crop "ideotype," first introduced by C.M. Donald in 1968, defines a theoretical 'ideal' plant type designed to outperform existing cultivars and overcome limitations traditionally experienced by breeders during selection. This model emphasises the selection of morphological, physiological, and biochemical characteristics that enhance crop productivity under a specific set of environmental conditions (Donald 1968; Carbajal-Friedrich and Burgess 2024; Li et al. 2025). The first wheat ideotype proposed by Donald emphasised seven features: short, strong stems; erect, few and small leaves; large, erect ears with awns; and a single culm (Donald 1968). Building on this foundation, the Wheat Yield Consortium in 2011 refined the wheat ideotype by highlighting physiological traits such as increased photosynthetic capacity, maximised spike fertility, improved grain filling and size, optimised partitioning of assimilates to the grain, and enhanced lodging resistance (Reynolds et al. 2011; Carbajal-Friedrich and Burgess 2024). Central to these features are developmental traits, and achieving such targets requires a deep understanding of the genetic and developmental pathways that shape organ growth and architecture. Linking developmental biology with breeding will therefore be critical for

engineering new ideotypes that can meet the demands of modern agriculture (Reynolds et al. 2022).

Arabidopsis thaliana has long served as a powerful model system in plant developmental biology, offering a foundation of knowledge that can be applied to crop research. However, translating findings from *Arabidopsis* into crops remains challenging due to fundamental differences in anatomy, physiology, and genetics (Uauy et al. 2025). Moreover, developmental biology in crops has traditionally been limited by the availability of tools for functional characterisation. These difficulties are particularly pronounced in hexaploid wheat, where research is complicated by a large (16 Gb), highly repetitive genome (~85% repetitive elements; IWGSC et al. 2018; Walkowiak et al. 2020), long generation times (4-6 months), and extensive functional redundancy across its three homoeologous genomes. Such redundancy frequently obscures mutant phenotypes, as all gene copies must be disrupted to generate a clear loss-of-function, an effort that can take up to 18 months (Borrill et al. 2015, 2019; Adamski et al. 2020).

However, recent technological advances are helping to overcome these barriers, making it increasingly feasible to study developmental processes directly in crops, including wheat (Adamski et al. 2020; Uauy et al. 2025). A significant breakthrough occurred in 2017 with the release of EMS-mutagenized tetraploid and hexaploid TILLING lines, which were made available as a public collection of mutant seed stock. Exome capture sequencing revealed 23–24 missense or truncation alleles per gene, with over 90% of captured genes carrying at least one truncation or deleterious mutation. This combination of sequence data and seed stocks enables researchers to identify alleles of interest and accelerate functional characterisation rapidly (Krasileva et al. 2017). Genomic resources soon followed: in 2018, a chromosome-level genome assembly and annotation of *Triticum aestivum* cv. *Chinese Spring* was published (IWGSC et al. 2018), followed in 2019 by tetraploid *Triticum turgidum* ssp. *durum* cv. *Svevo* (Maccaferri et al. 2019). Together with variation data from TILLING lines, these resources are accessible through Ensembl Plants (Bolser et al. 2015; Howe et al. 2020). Complementary transcriptomic tools include platforms for visualising, sorting, and filtering RNA-seq datasets (Borrill et al. 2016), alongside a 70-tissue/timepoint developmental expression atlas generated in *T. aestivum* cv. *Azhurnaya* (Ramírez-González et al. 2018). Importantly, improvements in Agrobacterium-mediated transformation have made stable transformation more accessible, enabling rapid functional validation of candidate genes via CRISPR editing (Howells et al. 2018; Hayta et

al. 2019; Smedley et al. 2021). Together, these resources have transformed wheat into a feasible system for developmental biology.

1.3 – Basics of the grass inflorescence

In this work, I focus on the development of the grass inflorescence. Inflorescences form a diverse range of structures across grass species, and their form plays a central role in reproductive success and fitness in natural populations (Wyatt 1982; Friedman and Harder 2004; Kellogg et al. 2013). In cereals, inflorescence architecture directly influences seed number and size, making it a key determinant of yield and a repeated target of both domestication and modern breeding (Glémin and Bataillon 2009).

In grasses, much like other angiosperms, the inflorescence develops from the inflorescence meristem (IM), which is derived from the shoot apical meristem (SAM) during the floral transition (Kellogg et al. 2013; Bommert and Whipple 2018). In simple inflorescences, such as those found in *Arabidopsis thaliana*, the IM directly produces floral meristems (FMs) along its central axis. In contrast, grass inflorescences display more complex architectures that cannot be explained by only IM and FM identities alone (Kellogg 2007).

In all but four species of *Anomochlooideae,* grass inflorescences share the basic organisation of flowers within spikelets (Clifford 1987; Judziewicz and Soderstrom 1989; Clayton, W. D. 1990). Each spikelet contains sterile bracts called glumes, and one or more florets initiated on an axis known as rachilla (Figure 1.2; Clifford 1987; Kellogg 2022). While spikelets are considered the terminal units in grasses, spikelets themselves are an inflorescence. For this reason, the grass inflorescence is better described as a synflorescence- a compound structure formed of separate inflorescence structures (Weberling 1989; Vegetti and Weberling 1996; Endress 2010; Kellogg 2022). Following suit with a review by Kellogg (2022), I use the term "inflorescence" throughout to maintain commonly used nomenclature; however, this is noted to be somewhat inaccurate.

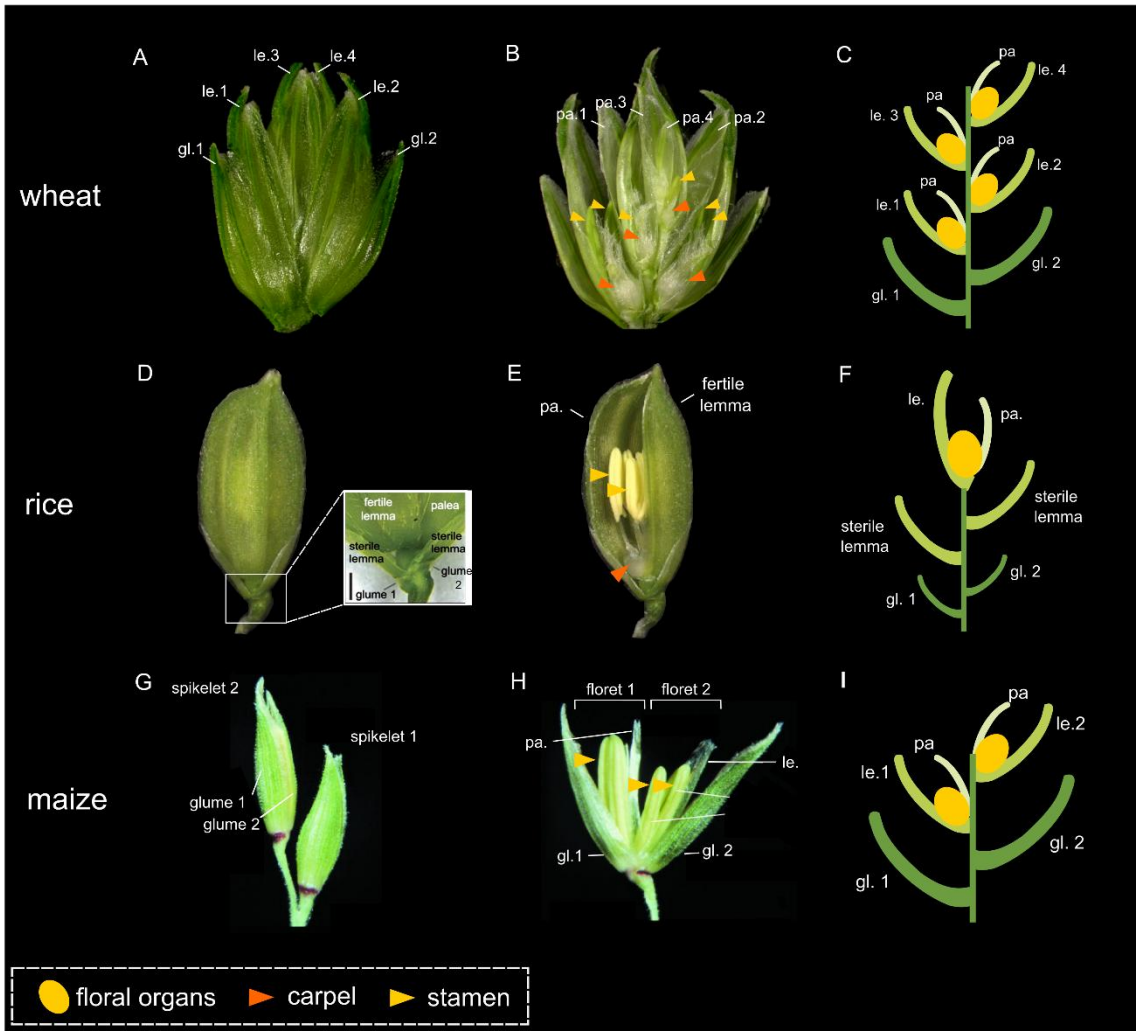


Figure 1.2 - Spikelets share characteristic features across grass species

Spikelet morphology of wheat **(A-C),** rice **(D-F),** and maize (tassel spikelet) **(G-I)** differs in its overall architecture, but shares key characteristics, including a pair of sterile bracts (glumes) and one or more florets initiated from a rachilla. In dissected spikelets in **B)**, **E)**, and **H)**, carpels are marked with an orange triangle, and stamens are marked with a yellow triangle. In diagrammatic spikelets in **C)**, **F)**, and **I)**, yellow ovals indicate floral organs. Images not to scale. Image of rice spikelets from Yoshida 2012. Inset image from Kellogg 2022. Images of maize spikelets from McSteen et al. 2000.

The diversity of grass inflorescence structures is often described by terminology borrowed from dicots and non-grass monocots, which requires us to describe the arrangement of spikelets, rather than flowers (Wayne 1982; Kellogg et al. 2013; Kellogg 2022). An inflorescence in which spikelets are sessile (directly attached) to the main inflorescence axis are referred to as “spikes”, which includes (but is not limited to) most members of tribe *Triticeae* (including wheat, barley and rye). Those in which the spikelets are attached by a pedicel, or stalk, to the central axis are “racemes”. These are relatively uncommon but are characteristic of the genus *Brachypodium* (including model species *Brachypodium distachyon*). For any inflorescence with a higher order of branching, whereby spikelets are not formed on the central axis, but instead are born on branches of higher orders of branching, they are termed “panicles” (Wayne 1982). This includes most grasses, including rice and the tassels of maize, which are commonly described as models of panicle-type inflorescences (Kellogg 2022).

To adequately summarise the broad scope of inflorescence architecture, it requires an understanding of changes in phyllotaxy, elongation of internodes, branching patterns and angles, and male/female differentiation of spikelets/florets (Bommert and Whipple 2018). However, here I refine my focus on how these architectures are built upon the repeated production of phytomers and how shifts in their developmental behaviour shape the variety of inflorescence structures in the grasses.

1.4 – Building the grass inflorescence through phytomers

Grass morphology is organised into repeating units called phytomers, each comprising an internode, a leaf, and an axillary meristem (AM). Phytomers are produced sequentially by the SAM, forming a modular architecture characteristic of grasses. Although the basic structure of each phytomer is repeated iteratively throughout vegetative and reproductive growth stages of the plant, its developmental fate can vary depending on position and developmental timing. This modularity allows for morphological diversity across species, as modifications in phytomer behaviour and AM fate generate the wide range of shoot and inflorescence forms observed in grasses (Briske 1991; Moore and Moser 1995).

Axillary meristems, which arise in the leaf axils of each phytomer, act as key developmental decision points. During vegetative growth, the SAM continuously initiates phytomers. Lateral leaf primordia expand, while AMs remain dormant. In a subset of phytomers, AMs may activate to form tillers, producing additional vegetative shoots (Pautler et al. 2013). This developmental program changes as the SAM transitions to an IM, marking the switch from vegetative to reproductive development. The apex elongates as the IM maintains a pool of undifferentiated meristematic cells at its apex while initiating lateral phytomer units on its flanks (Tanaka et al. 2013). In most grass species, growth of the leaves from each phytomer unit is suppressed, whereas the AMs grow immediately upon their formation (Kyozuka 2014; Kellogg 2022), producing an ebracteate (i.e. without bracts) inflorescence axis (Evans and Grover 1940; Latting 1972; Fraser and Kokko 1993; Chuck et al. 2010; Whipple et al. 2010).

Following their initiation by the IM, the lateral AMs of each phytomer unit transition through a series of meristem identity states, beginning as a branch meristem (BM), transitioning to a spikelet meristem (SM), and ultimately becoming a floral meristem (FM) (Tanaka et al. 2013; Bartlett and Thompson 2014). Each meristem identity is categorised by the type of lateral primordia a meristem produces. A BM produces additional phytomer units along its flanks, generating primary and secondary branches off the central axis of the inflorescence. A SM initiates two sterile bracts (glumes), followed by one or more phytomer units composed of a lemma (a modified bract) and a subtending FM. Each FM then gives rise to the floral organs (Bartlett and Thompson 2014).

This developmental framework, based on the fate of repeated phytomer units and shifts in meristem identity, enables us to break down an inflorescence into smaller developmental decisions. For example, in the panicle-type inflorescences of rice, spikelets are produced

after multiple rounds of branching. This occurs when the IM initiates AMs in a spiral phyllotaxis, which forms a type of BM termed 'primary branching meristems' (PBMs). These PBMs initiate additional AMs in a distichous phyllotaxis, termed 'secondary branch meristems' (SBMs), and both PBMs and SBMs elongate, subsequently producing SMs along their flanks, and terminating in a SM. Additionally, the IM will terminate in a SM. Each SM will produce a single FM (Figure 1.3A; Ikeda et al. 2004, 2019; Bommert et al. 2005).

Similarly, in the tassels of maize, the IM initiates several indeterminate BMs, termed long branch meristems (LBM), along its flanks in a spiral phyllotaxis. The LBMs subsequently initiate another type of BM, termed 'spikelet pair meristem' (SPMs) in a distichous phyllotaxy. Once several LBMs are produced, the IM shifts to producing distichously arranged SPMs along the central axis of the inflorescence. Each SPM will produce a pair of SMs: one pedicellate and one sessile. Each SM will produce two FMs on the axis of the rachilla (Figure 1.3B; Koppolu and Schnurbusch 2019).

In contrast, spike-type inflorescences such as those of wheat and barley (*Hordeum vulgare*) bypass the branching phase entirely, producing spikelets in a distichous phyllotaxy directly along the central inflorescence axis, the rachis (Koppolu and Schnurbusch 2019; Koppolu et al. 2022). In barley, AMs initially transition to a meristem identity referred to as the 'triple spikelet meristem' (TSM). Each TSM will produce three SMs per rachis node, and each SM will initiate a single FM (Figure 1.3C). In contrast, in wheat, AMs will form directly to SM identity, producing a single spikelet at each rachis node (Figure 1.3D; Koppolu and Schnurbusch 2019).

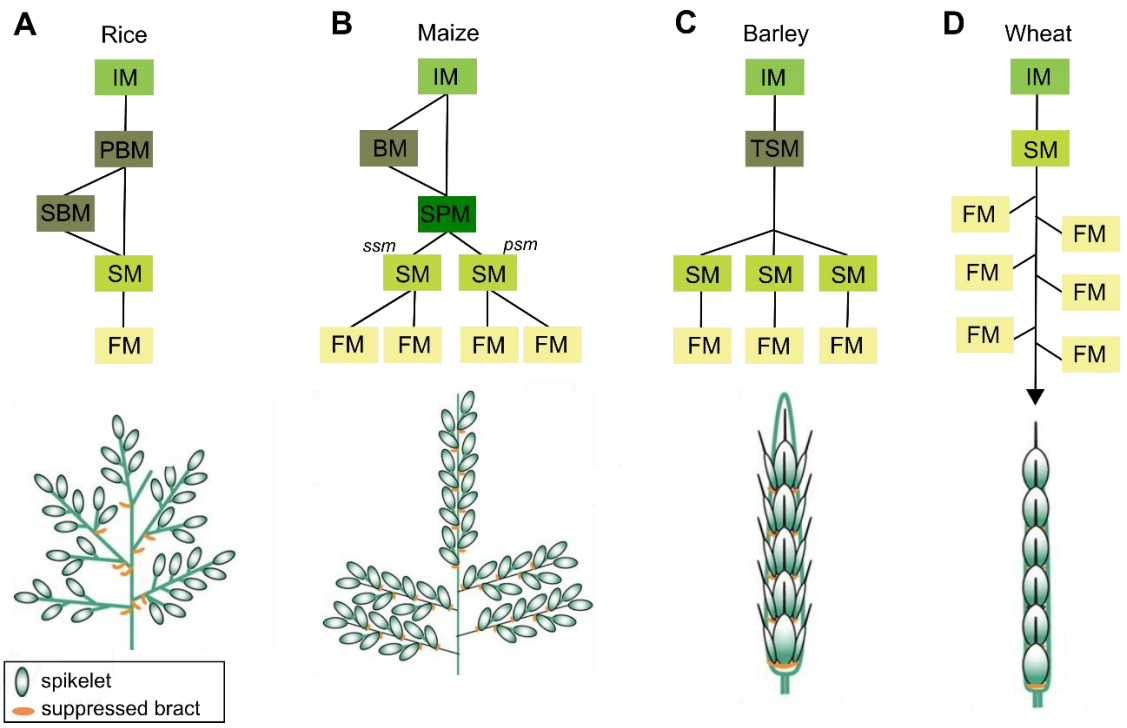


Figure 1.3 - Schematic of meristem identity transitions in major grass species and respective inflorescence architectures

A) Meristem identity transitions in rice. IM: inflorescence meristem; PBM: primary branch meristem; SBM: secondary branch meristem; SM: spikelet meristem; FM: floret meristem. **B)** Meristem identity transitions in maize. SPM: spikelet pair meristem; ssm: sessile spikelet meristem; psm: pedicellate spikelet meristem. **C)** Meristem identity transitions in barley. TSM: triple spikelet meristem. **D)** Meristem identity transitions in wheat. Schematics of meristem differentiation adapted from Koppolu and Schnurbusch (2019). Schematics of inflorescence architecture adapted from Kellogg (2022; not to scale).

1.5 – Examining the wheat inflorescence across developmental time

Given the primary focus on wheat inflorescence development in this body of work, I set out to describe the developmental progression of the wheat inflorescence in greater detail. To do so, I will examine the developmental stages described in the Waddington scale of spike development (Waddington et al. 1983), and at key stages, explore how morphology can be explained through individual phytomer units and their transition through meristem identity states.

To begin, we start at the transition of the SAM to the IM, which is characterised by elongation of the apex and the initiation of paired lateral ridges by the IM. During the double ridge stage, lateral ridges are arranged in a distichous phyllotaxis to the central axis (W2 to W2.5). Each double ridge consists of a smaller, suppressed bract primordium- the leaf ridge (LR)- and a larger axillary meristem- the spikelet ridge (SR; Figure 1.4A; Kirby and Appleyard 1984). The internodes of these phytomers will eventually elongate to form the rachis (Patil et al. 2019; McKim 2020). As each lateral AM will begin to develop a spikelet, the shift of the AM to SM identity is evident by the formation of the first two phytomer units along the secondary axis. These two phytomers form outgrown bracts (glumes), which form on the left and right flanks of the SM. Within the axils of each glume, axillary meristems are suppressed (Figure 1.4B; Kirby and Appleyard 1984; Kellogg 2022).

As a brief note, in the wheat literature, the term “spikelet ridge” is often used to describe the lateral meristems formed during early inflorescence development, reflecting their eventual role in initiating a spikelet. Here, I instead use the term “axillary meristem” for greater precision and to emphasise the distinction from a spikelet meristem, adopting terminology consistent with Kellogg (2022). This terminology also better reflects the developmental state at the double ridge stage, when the AM has not yet acquired spikelet identity or initiated spikelet patterning. Thus, the term “spikelet ridge” can be misleading in this context.

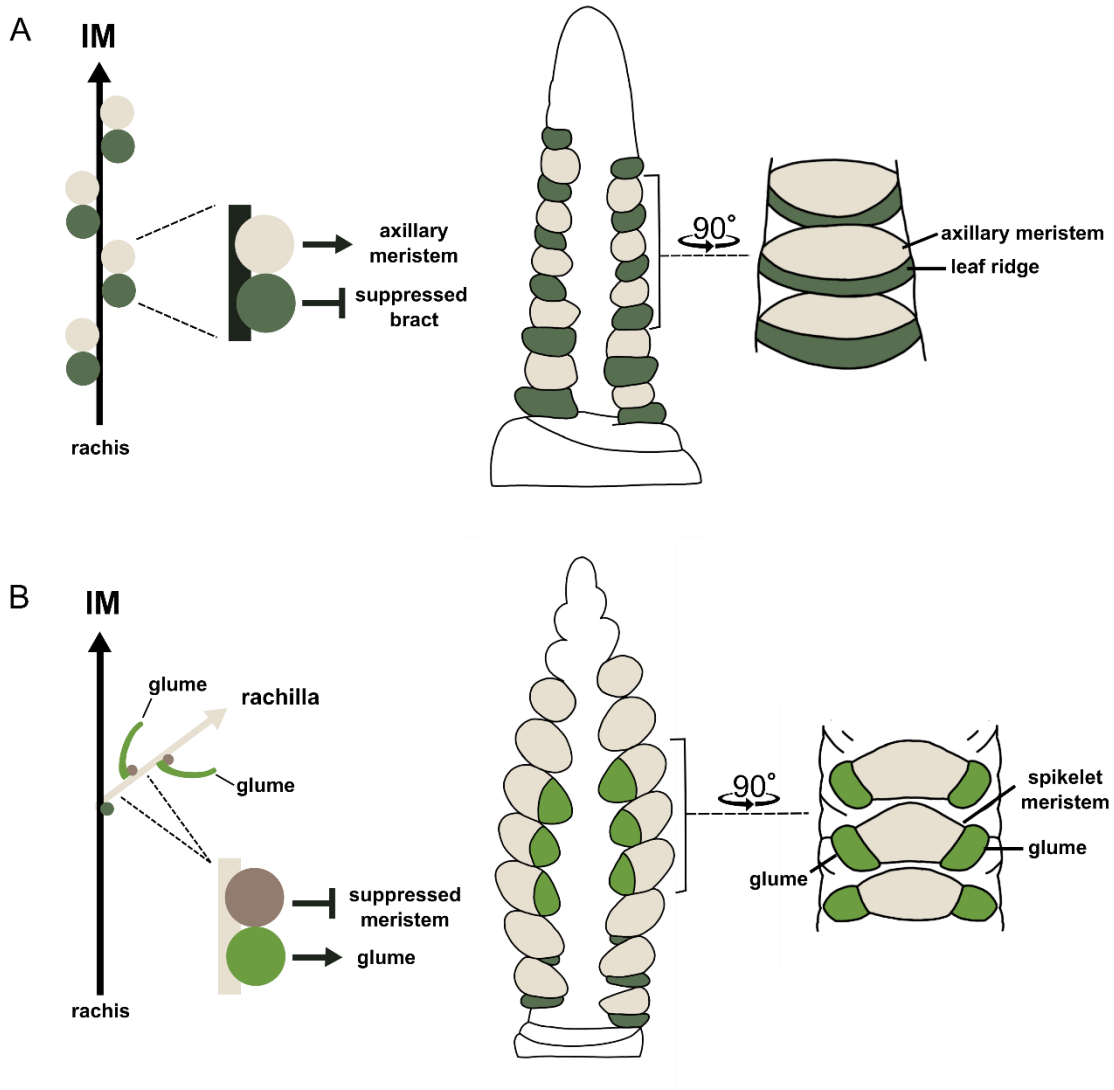


Figure 1.4 – Schematic of early wheat inflorescence development

A) Simplified model of phytomer units and diagrammatic representation of W2.5 spikes. **B)** Simplified model of phytomer units and diagrammatic representation of W3 spikes. Inflorescence diagrams not to scale, adapted from Kirby and Appleyard, 1984.

After the initiation of the glumes, the SM will initiate an indeterminate number of phytomers (10 - 12), alternating in a distichous manner. The rachilla, formed from the internodes of these spikelet phytomers, serves as the axis of the spikelet (Figure 1.5A; Kirby and Appleyard 1984; Koppolu and Schnurbusch 2019). While the SM may initiate over ten FMs, only four to six florets typically develop to maturity (Brinton & Uauy 2019). Each floral phytomer consists of a lemma, an outgrown bract that subtends the FM (Tanaka et al. 2013; Whipple 2017). By the floral primordia stage (W3.5), lemmas from floret positions 1 and 2 (the first initiated across developmental time) are visible as raised ridges formed in parallel to the glumes. FMs formed in the axil of florets one and two are visible as a rounded meristematic dome, and additional floral phytomers are initiated in floret positions 3, 4, and beyond (Figure 1.5B; Kirby and Appleyard 1984).

By stage terminal spikelet (W4), the final lateral AMs are initiated by the IM. The IM in wheat is determinate; a terminal spikelet marks the completion of spikelet initiation. Instead of the last initiated lateral AMs forming entire spikelets, they each form a component of the terminal spikelet- two glumes (glume + suppressed AM) and florets (lemma + FM), which form at a 90° angle to that of other spikelets (Figure 1.5C-D; Kirby and Appleyard 1984). At the terminal spikelet (W4) stage, and subsequent carpel extension round (W5 stages), the floral meristems differentiate into floral organ primordia. Each floret contains a palea (outer perianth; equivalent to sepals), two lodicules (inner perianth; equivalent to petals), three stamens and a gynoecium containing a single ovule (Figure 1.5E; Kirby and Appleyard 1984). Because florets within a spikelet are initiated sequentially, with floret one forming first, later florets are behind in developmental progression. For instance, at stage W5, florets in positions 5 and 6 have produced a lemma and established a floral meristem (FM) but have not yet initiated floral organ primordia (Figure 1.5D; Kirby and Appleyard 1984).

Although the phytomer concept defines each unit as a successive building block, its boundaries are less distinct in practice. Before internode elongation in the rachis and rachilla, phytomers remain highly compact, and internode boundaries are indistinct or absent during early inflorescence development (Patil et al. 2019; McKim 2020; Kellogg 2022). In floral structures, the model becomes more abstract than in vegetative shoots or inflorescence axes, as organ initiation occurs in quick succession and within a highly condensed space. This highlights the limitations of the phytomer model, yet it remains a valuable framework for understanding how simple components give rise to complex structures (Forster et al. 2007).

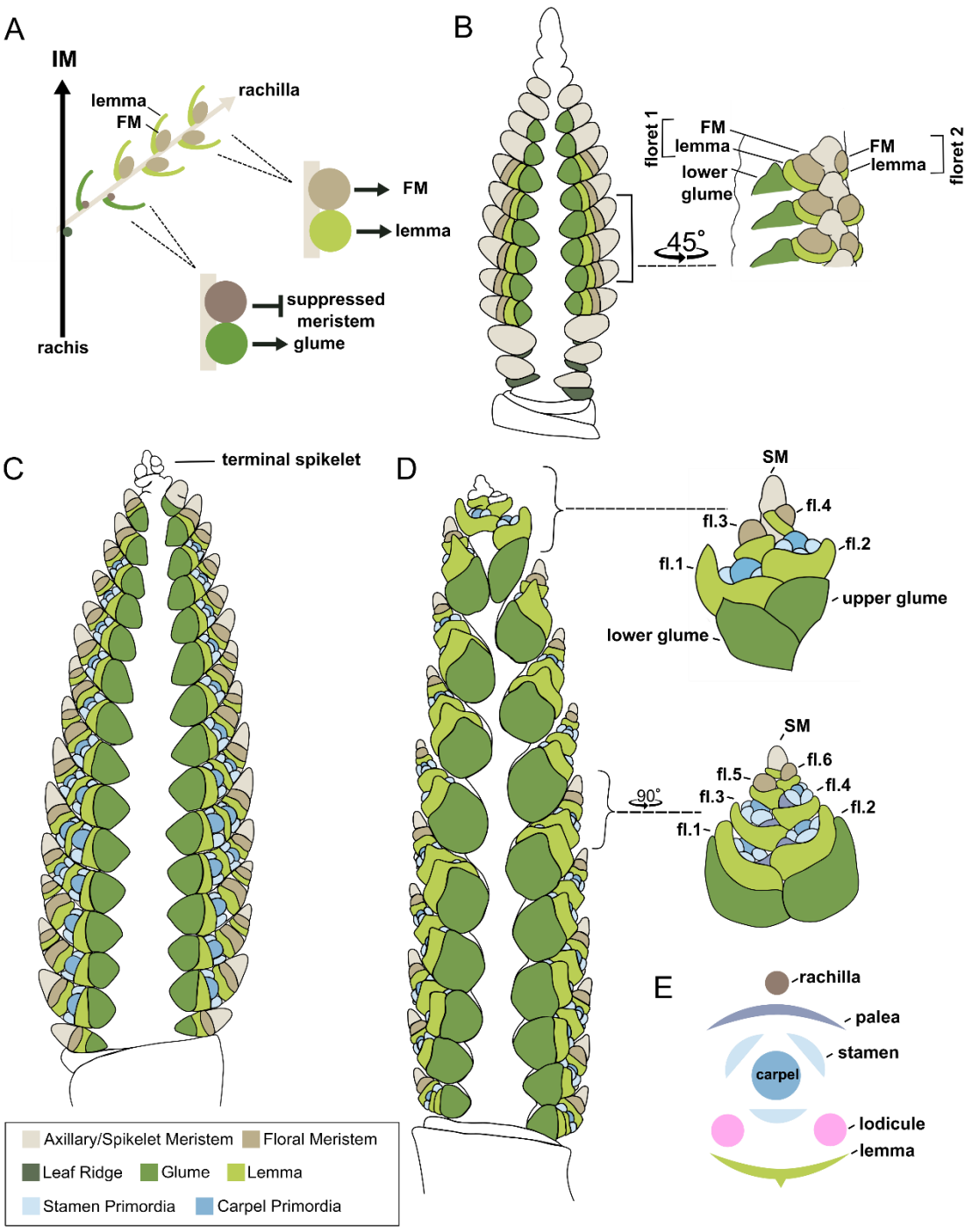


Figure 1.5 - Schematic representation of spikelet and floret development

A) Simplified model of phytomer units forming a wheat spikelet. Diagrammatic representations of **B)** W3.5 inflorescence, a side view of spikelets highlights the floral meristems developing in florets 1 and 2. **C)** W4 inflorescence, highlighting the formation of the terminal spikelet. **D)** W5 inflorescence, highlighting insets of the terminal spikelet and developing floral organs in central spikelets (fl: florets). Diagrams adapted from Kirby and Appleyard, 1987. **E)** Simplified representation of wheat spikelet, adapted from Debernardi et al. 2020 . Each spikelet consists of a lemma, a palea, two lodicules, three stamens and a carpel.

1.6 - Establishment of the lanceolate shape in wheat

As previously described, during early wheat inflorescence development, specifically at the so-called double ridge stage, the IM initiates phytomer units (LR + AM pairs). As these ridge pairs are initiated acropetally, a developmental gradient is established along the nascent inflorescence, with basal phytomers being the oldest (Bonnett 1966; Waddington et al. 1983). However, the timing of each SR to proceed through meristem identity transition (AM > SM > FM) does not follow this age gradient. Instead, central SRs are the first observed to form to glumes (the first indication of SM) and lemma primordia (indication of FM)—while basal SRs, despite being older, are delayed (Bonnett 1966). Due to this effect, it can be observed that within a single inflorescence, spikelet differentiation begins in the centre of the inflorescence axis, and continues bi-directionally, towards the top (apical), and bottom (basal) of the inflorescence (Kirby and Appleyard 1984).

While this variation in developmental progression across the apical-basal axis is evident early in development, it has lasting effects. In wheat, florets that are underdeveloped are aborted during a critical window, 10 – 20 days pre-anthesis. The survival of a floret depends on whether it has reached at least Waddington stage 5.5. This threshold is thought to narrow the developmental range of the remaining florets, contributing to the relatively synchronous anthesis observed across the spike. (Backhaus et al. 2023).

Floret abortion impacts several areas across the developing spike. At the apical end of the inflorescence, the spikelets are initiated last and have less time to mature, resulting in delayed development and likely floral abortion. Likewise, within each spikelet, 8–10 florets are typically initiated, but only the first four to six reach fertility and form a carpel, with the first initiated over developmental time in floret position 1, 2, and so on remaining (Evers and Millar 2002). In both these cases, this pattern reflects a temporal advantage - the earliest-formed florets have the longest period for growth and maturation. However, the effects of floral abortion are also observed in basal spikelets, which are initiated first in terms of developmental time. In typical field conditions, the basal-most spikelets are completely infertile and do not produce grain (Tamagno et al. 2024). This is thought to arise when no florets within a single spikelet surpass the developmental threshold to prevent floral abortion (Backhaus et al. 2023).

In this way, the number of surviving florets, and thereby the number of grains per spikelet, is a function of time (when a floret is initiated), as well as space (position of the spikelet across the apical-basal axis). This leads to the formation of the inflorescence in a

lanceolate shape. The central spikelets are larger and produce more grain when compared to apical and basal spikelets (Calderini and Ortiz-Monasterio 2003; Liu et al. 2006). This characteristic is not only observed in wheat, but also in other members of the *Triticum* genus, including *Triticum aestivum* (bread wheat), *Triticum spelta* (spelt), and *Triticum turgidum* ssp. *durum* (durum wheat), *Triticum turgidum* ssp. *dicoccum* (emmer), and *Triticum monococcum* (einkhorn).

The evolutionary and molecular reasons why a basal spikelet, which initiates first and has the most time to develop, falls behind, particularly at such an early stage of inflorescence development, are unclear. In a cascade of phytomers representing, in theory, identical and repeatable units forming across the apical-basal axis of the inflorescence, why do some behave differently than others? This requires us to frame the inflorescence as a compound structure and consider the context of each phytomer unit across both developmental time and space, such as a developmental axis. The fine dissection of these dimensions requires further enquiry.

1.7 - Implementation of semi-spatial transcriptomics in the grass inflorescence

The wheat inflorescence is a compound structure. As stated previously, the developmental progression of each phytomer unit is influenced by its relative time of initiation and its position along the apical-basal axis (i.e., spikelet position within the inflorescence). We observe this as early as the glume primordia stage (W3.25), where the progression to SM identity cannot be explained solely by the time of initiation. Each of these phytomers is formed within close proximity; ~1500 µm covers the entire length of the inflorescence.

Given this spatial complexity, the key transcriptional regulators determining the developmental progression/suppression of AMs across the apical-basal axis in wheat remain largely uncharacterised. Previous implementations of whole-tissue bulk RNA sequencing obscure spatial variation in gene expression. To address this limitation, semi-spatial transcriptomic techniques have been applied to grass inflorescences, including manual microdissection (Backhaus et al. 2022) and laser capture microdissection (Harrop et al. 2016; Thiel et al. 2021).

Low-input RNA-seq of apical, central and basal sections of micro-dissected wheat spikes revealed significant differences in gene expression profiles among them. For example, the MADS-box transcription factor *VEGETATIVE TO REPRODUCTIVE TRANSITION 2 (VRT2)* exhibited its highest expression in basal sections, with decreasing levels toward the apex, in a proposed gradient along the inflorescence (Backhaus et al. 2022). *VRT-A2* belongs to the *SHORT VEGETATIVE PHASE* (*SVP*) subfamily of MADS-box transcription factors, which are associated with vegetative growth and are downregulated after floral transition in *Arabidopsis* (Gregis et al. 2013), wheat (Adamski et al. 2021; Li et al. 2021b; Liu et al. 2021), rice (Sentoku et al. 2005; Lee et al. 2008), and barley (Trevaskis et al. 2007).

Genetic characterisation of *VRT2* reveals its connection to the lanceolate shape of wheat. In *Triticum turgidum* ssp. *polonicum* (Polish wheat), a natural variant at the *P1* locus, was traced to a partial deletion and rearrangement in intron 1 of *VRT-A2*. This allele ($P1^{POL}$) drives overexpression of *VRT-A2*, leading to elongated grains and floral organs (glumes and lemmas) characteristic of Polish wheat (Adamski et al. 2021). Introgression of $P1^{POL}$ into the *Triticum aestivum* cv. Paragon background produced near-isogenic lines (NILs) with a modest but significant increase in rudimentary basal spikelets (RBS) compared to the wild-type allele (+1.1 RBS under field conditions), suggesting that elevated *VRT-A2*

expression delays basal spikelet development (Adamski et al. 2021; Backhaus et al. 2022). Transgenic lines carrying additional copies of $P1^{POL}$ further supported this connection. *VRT-A2* expression scaled with allele copy number, and higher copy number lines exhibited more RBS than low copy number lines in a dosage-dependent manner (Backhaus et al. 2022).

Additionally, *SVP* transcription factors play a crucial role in the transition of meristem identity in wheat. Loss-of-function alleles in *vrt2* or its close paralog, svp1, cause delayed heading and an increase in the number of spikelets per spike. These effects are more pronounced in *vrt2 svp1* double mutants, indicating overlapping roles for these genes in promoting the transition from the SAM to the IM and in terminating the IM with a terminal spikelet. These genes also function as repressors of AM outgrowth in vegetative tissues. In *vrt2 svp1* double mutants, axillary spikelets or spikes subtended by leaves form at sub-peduncle nodes—structures absent in wild-type plants—while single *vrt2* or *svp1* mutants show a weaker effect. This phenotype suggests that *SVP* transcription factors restrict the transition of AMs to IM or SM identity during vegetative growth (Li et al. 2021b).

SVP function integrates with the activity of three SQUAMOSA-clade AP1/FUL-like MADS-box transcription factors (*VRN1*, *FUL2*, *FUL3*), which redundantly regulate meristem transitions throughout reproductive development. Mutations in these genes progressively delay the transition from the SAM to IM, with *vrn1*-null showing a mild delay, *vrn1 ful2* a more substantial delay, and *vrn1 ful2 ful3* the most severe, resulting in prolonged vegetative growth (Li et al., 2019). *VRN1*, *FUL2*, and *FUL3* also promote the transition from AM to SM identity, as displayed through their loss of function. In *vrn1 ful2* mutants, AMs in the inflorescence develop vegetative structures with occasional residual floral organs, while *vrn1 ful2 ful3* give rise to complete vegetative tillers subtended by de-repressed bracts, rather than spikelets (Li et al. 2019, 2021b).

Comparative transcriptomics of *vrn1 ful2* double mutants (with spikelets transformed into tillers) and *vrn1* single mutants (forming a typical inflorescence) identified three SVP genes (*SVP1*, *VRT2*, *SVP3*) strongly upregulated in the mutant background (Li et al. 2019), suggesting that the upregulation of these genes may prevent the normal meristem identity transition of AM > SM > FM. One hypothesised mechanism of action is that *SVPs* interfere with the formation of MADS-box SEPALLATA–SQUAMOSA protein complex required for normal spikelet development (Li et al. 2021b). Accordingly, coordinated *SVP*

downregulation alongside *SEPALLATA (SEP)* gene upregulation is essential for normal floral transition and spikelet development (Backhaus et al. 2022).

This downregulation of *SVPs* during the transition to inflorescence development is further refined in a semi-spatial laser capture micro-dissection experiment in barley. In this study, transcriptomes of the IM at the double ridge stage were compared with those of the vegetative apex (VA, encompassing the SAM and adjacent leaf ridges). *SVPs* were preferentially expressed in the VA, whereas *VRN1* and two *SEP* genes were specifically induced in the IM. These observations suggest that the regulatory balance between *SVPs* and *SQUAMOSA/SEP* genes during the shift from vegetative to reproductive growth of the apex is conserved between barley and wheat (Thiel et al. 2021).

In Backhaus et al., it is proposed that the gradient of *SVP* expression along the spike aligns with differences in AM activity: sub-peduncle axillary meristems are completely suppressed in a 'vegetative' identity, basal spikelet meristems experience partial suppression that limits their progression to floral development, and central to apical spikelets remain fully active, producing viable florets. Therefore, the spatial variation of *VRT-A2* across the apical-basal axis of the inflorescence may reflect the imperfect demarcation of developmental zones along the inflorescence, with rudimentary basal spikelets emerging as a developmental consequence of delayed or incomplete release from vegetative programs (Backhaus et al. 2022).

Additional studies have also characterised regulatory gradients formed across the apical-basal axis of the wheat inflorescence. In developing wheat spikes across three consecutive stages, expression of the AP2-like transcription factor *AP2-5* is highest in apical sections, in contrast to the elevated levels of its negative regulator miR172 in basal sections, together forming proposed opposing gradients along the spike. These patterns are consistent with mutant phenotypes. Enhanced miR172 activity (reducing *AP2-5* levels) produces empty florets, shortened awns, and incipient keels, most evident in basal spikelets, while inhibition of miR172 (increasing *AP2-5* levels) converts glumes into fertile florets, a phenotype restricted to apical spikelets (Debernardi et al. 2017). These findings support a model in which gene expression gradients establish distinct 'developmental zones' in basal, central, and apical phytomers.

Indeed, from the same semi-spatial dataset in wheat spikes generated by Backhaus et al., it was observed that the highest levels of differential expression were observed across the micro-dissections of apical, central, and basal sections of a single spike than between

any section belonging to consecutive developmental time points (Backhaus et al. 2022). Given the high levels of differential expression observed across the spike in this experiment, we hypothesised that other genetic factors, beyond *SVPs*, *SQUAMOSA/SEPs* and *AP2s*, contribute to apical-basal axis patterning that warrant further investigation.

However, the implementation of semi-spatial techniques does not come without limitations. Within each section of the wheat spike, divided into 'apical', 'central', and 'basal' zones, several tissue types are incorporated, obscuring precise gene expression patterns. For example, in the barley laser capture microdissection study, 64 genes were found to be differentially expressed between the AM and its subtending LR at the double ridge stage, underscoring the transcriptional distinctiveness of tissues within a single phytomer unit (Thiel et al. 2021). In wheat, 2,438 genes were identified as differentially expressed across apical, central, and basal sections at the double ridge stage (Backhaus et al. 2022), yet their spatial expression domains remain largely uncharacterized. Therefore, resolving these patterns will be essential to move beyond approximated gradients and towards a mechanistic understanding of inflorescence gene regulation.

1.8 - Spatial transcriptomic techniques move to plant tissues

Addressing *when* and *where* genes are expressed has been a central challenge in plant biology. Bulk tissue transcriptomics, a method most widely deployed to assess these questions, provides valuable insights but inevitably averages signals across heterogeneous cell populations (Nobori 2025). To further refine gene expression patterns, single gene *in situ* hybridisation has been used since the 1970s to define transcript localisation within a tissue context (Gall and Pardue 1969; John et al. 1969; Moses and Pachter 2022). However, this approach is limited in scalability.

As technology has advanced, through improvements in computing infrastructure, advanced automation and robotic techniques, and a decreasing cost to sequencing, a new set of techniques have emerged and moved into the mainstream: spatial transcriptomics (Moses and Pachter 2022). These techniques enable the quantification of mRNA expression for a large number of genes within the spatial context of tissues and cells (Giacomello 2021; Moses and Pachter 2022). Initially optimised in animal model systems, spatial transcriptomics was motivated by the need to identify cell-type-specific and spatially restricted genes, to link gene activity with developmental processes, and to uncover novel cell types not evident from morphology alone (Moses and Pachter 2022). Adaptation of these approaches to plants has served a similar purpose, providing a powerful means to map expression patterns of developmentally relevant candidate genes in a highly multiplexed manner (Giacomello et al. 2017; Giacomello 2021; Nobori 2025).

In plants, early applications used single-molecule fluorescence *in situ* hybridisation (smFISH), which improved upon traditional *in situ* methods by localising single RNA molecules to sub-cellular resolution (Duncan et al. 2016). More recent techniques can detect a larger set of genes and fall broadly into two categories: sequencing-based and imaging-based (reviewed in Nobori 2025), with both being applied recently to inflorescence tissues (Giacomello et al. 2017; Wang et al. 2024; Demesa-Arevalo et al. 2025; Liu et al. 2025; Xu et al. 2025, Chapter 2-4). Sequencing-based approaches enable unbiased transcriptome-wide studies; however, in practice, these techniques are limited in capturing low-abundance transcripts and often compromise cellular resolution. Imaging-based methods target a predefined set of genes, limiting the number detected per experiment, but allow for precise spatial and cellular resolution.

1.9 - Thesis aims

The overall aim of this thesis is to explore gene expression across the apical-basal axis of the developing wheat spike to novel resolution. Over the course of my PhD, I developed a strong interest in optimising emerging spatial transcriptomic methods, which, at the time, had rarely been applied to plant tissues. To this end, I report the development and adaptation of two complementary approaches: the imaging-based multiplexed error-robust in situ hybridisation (MERFISH) and the sequencing-based Stereo-seq.

First, I detail the implementation of MERFISH in wheat inflorescence tissues. In Chapter 2, I describe the preparation and optimisation steps, beginning with the analysis of a semi-spatial microdissection RNA-seq dataset to identify candidate regulators of apical–basal patterning. These genes were incorporated into a 200-gene panel. I then established protocols for plant tissue fixation, embedding, and sectioning, culminating in the first application of MERFISH to wheat tissues. Chapter 3 builds on this foundation by applying MERFISH across four developmental stages to examine tissue-specific transcriptional programmes and the differentiation of phytomer units along the apical–basal axis. In Chapter 4, I present the first application of Stereo-seq in wheat and critically assess its performance, highlighting both its potential and the technical limitations observed during these initial trials.

Chapter 2 – MERFISH resolves gene expression patterns in the wheat inflorescence to cellular resolution

This chapter includes results previously published in the following manuscript:

Spatial Transcriptomics Reveals Expression Gradients in Developing Wheat Inflorescences at Cellular Resolution

Katie A. Long, Ashleigh Lister, Maximillian R. W. Jones, Nikolai M. Adamski, Rob E. Ellis, Carole Chedid, Sophie J. Carpenter, Xuemei Liu, Anna E. Backhaus, Andrew Goldson, Vanda Knitlhoffer, Yuanrong Pei, Martin Vickers, Burkhard Steuernagel, Gemy G. Kaithakottil, Jun Xiao, Wilfried Haerty, Iain C. Macaulay, Cristobal Uauy

bioRxiv 2024.12.19.629411; https://doi.org/10.1101/2024.12.19.629411 (See Appendix 1)

This work was conducted in collaboration with Ashleigh Lister (Earlham Institute), Dr Nikolai Adamski (JIC), Dr Anna Backhaus (JIC), and Dr Maximillian Jones (JIC). For the bulk RNA-seq dataset, RNA extractions were performed by Nikolai, Anna, and Max. Anna contributed to the trimming and pseudo-mapping of RNA-seq data. Max conducted differential expression analysis and visualisation of the RNA-seq dataset. Ashleigh performed the MERSCOPE workflow for all steps following cryosectioning. I was assisted in the movement and storage of Vizgen data by Dr Burkhard Steuernagel (JIC), Dr Martin Vickers (JIC), and Gemy Kaithakottil (Earlham). The use of the MERSCOPE instrument was made possible through the Transformative Genomic platform (Earlham Institute), with contributions from Dr Iain Macaulay, Andrew Goldson, Vanda Knitlhoffer, and Ashleigh Lister. Rob Ellis and Dr Burkhard Steuernagel helped to implement the WebAtlas Interface for wheat-spatial.com. I am grateful to Dr Carole Chedid (Vizgen) for consultation on this project. Thank you to Tobin Florio for his help creating the wheat anatomical diagrams used in this chapter. I wish to thank the staff members of Vizgen for their technical support throughout this project. JIC Horticultural Services facilitated plant growth.

2.1 – Chapter summary

In this chapter, I present the development and optimisation of MERFISH (Multiplexed Error-Robust Fluorescence *In Situ* Hybridisation) for spatial transcriptomic profiling of wheat inflorescence tissues, representing the first application of this technique in wheat. Given the novelty of implementing MERFISH in plant tissues, I first detail the experimental preparation required for the Vizgen MERSCOPE platform, including validation of sample integrity using smFISH (single-molecule FISH) and the development of a robust embedding and cryosectioning protocol. These methods enabled the preparation of high-quality inflorescence sections, supporting the design and deployment of a 200-gene panel targeting genes implicated in apical–basal patterning of the spike. The resulting dataset, encompassing over 7 million transcripts and 50,731 high-quality segmented cells across four developmental stages and two genotypes, provides single-cell–resolution maps of gene expression in wheat spikes. Collectively, this chapter establishes a validated workflow for MERFISH in plant tissue, providing detailed protocols, performance benchmarks, and analytical approaches that will facilitate future spatial transcriptomic studies in diverse plant systems.

2.2 - Introduction

2.2.1 - Development of imaging-based techniques for spatial transcriptomics

As previously outlined in Chapter 1, a range of molecular techniques have been developed to profile gene expression while preserving spatial tissue integrity—collectively referred to as spatial transcriptomics. Among these, *imaging-based* approaches enable the quantitative detection of RNA transcripts in single cells using fluorescence microscopy, typically through the application of fluorophore-labelled, transcript-specific probes (Giacomello 2021). In plants, early implementations of these methods involved single-molecule fluorescence in situ hybridisation (smFISH), which offered a significant improvement over traditional *in situ* hybridisation by enabling the localisation of individual RNA molecules at subcellular resolution (Duncan et al. 2016). smFISH operates by hybridising single-labelled oligonucleotide probes, each around 20 nucleotides long, to a target mRNA molecule (Femino et al. 1998). To ensure a bright, detectable signal, a typical smFISH experiment employs multiple probes which hybridise across the length of the transcript, collectively producing discrete, diffraction-limited fluorescence spots visible under a microscope (Figure 2.1; Raj et al. 2008; Duncan et al. 2016). This technique can therefore quantify RNA copy number at the single-cell level, in addition to spatially mapping gene expression within a whole-tissue context.

Initial applications of smFISH in plants were conducted in *Arabidopsis thaliana* root cells, where Duncan et al. (2016) demonstrated robust RNA detection and quantification in individual cells isolated by physically squashing the roots between glass slides. This study established the feasibility of smFISH in plant systems. The technique was subsequently adapted for use in intact tissue sections, extending its utility to a whole-tissue context. Notably, Zhang et al. optimised smFISH for cryo-sectioned tissues (2024), including the inflorescence tissues of *Hordeum vulgare*, marking the first optimisations of smFISH techniques to grass inflorescence tissues. A key limitation of traditional smFISH is its restricted multiplexing capacity, as the number of transcripts detectable in a single experiment is constrained by the limited number of spectrally distinct fluorophores that can be used simultaneously. To overcome this bottleneck and dramatically increase the number of transcripts that can be profiled in one experiment, several advanced methods have been developed that modify the design of the oligo probes used in fluorescence *in situ*

hybridisation, which allow for the detection of hundreds or thousands of RNA species (Giacomello 2021).

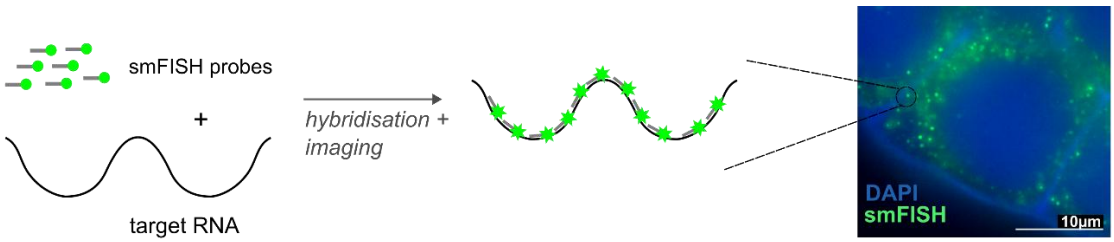


Figure 2.1 – Schematic of smFISH experimental design

2.2.2 - MERFISH expands multiplexing capacities of fluorescence *in situ* techniques

In this study, we focus on the application of MERFISH (Multiplexed Error-Robust Fluorescence *In Situ* Hybridisation), a method that extends the foundational principles of smFISH. Like smFISH, MERFISH uses fluorescently labelled probes to detect individual mRNA molecules. However, MERFISH introduces a combinatorial barcoding system and multiple rounds of hybridisation and imaging, which allows for the simultaneous profiling of hundreds to thousands of transcripts within a single sample. This approach not only enhances multiplexing capacity but also incorporates error-correction strategies to ensure robust and accurate transcript identification (Chen et al. 2015; Moffitt et al. 2016).

The MERFISH probe design differs from conventional smFISH by employing two types of probes: *encoding probes* and *readout probes*. Encoding probes are oligonucleotides designed to hybridise to target mRNAs and carry overhanging *readout sequences*, which serve as binding sites for fluorescently labelled readout probes (Figure 2.2A–B). To ensure high detection sensitivity, multiple encoding probes bind along the length of each RNA species (in this case, mRNA transcripts). While individual readout sequences can be reused across targets, each RNA species is identified by a unique combination of four readout sequences, which allow for its identification during imaging (Figure 2.2C). MERFISH relies on sequential cycles of hybridisation and imaging. In each round, fluorescently labelled readout probes are washed over the sample and hybridise to the complementary readout sequences found on the encoding probes. The sample is then imaged to detect the location of binding events. (Figure 2.2D; Chen et al. 2015).

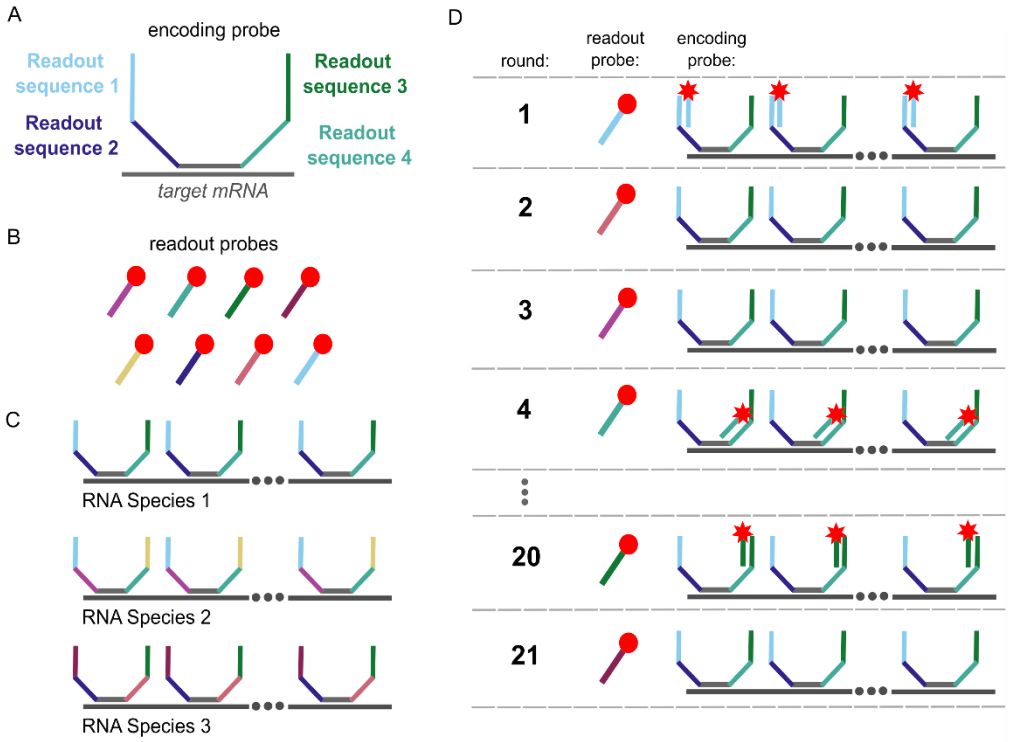


Figure 2.2- Encoding and readout probes enable sequential rounds of hybridisation and imaging
A) Encoding probes contain four readout sequences, which are complementary to B) readout probes- fluorescently labelled oligo probes. Each colour represents sequence complementarity. C) For each RNA species, a unique set of readout sequences is designed. D) Across sequential rounds of hybridisation and imaging, readout probes bind to the encoding probe, generating a fluorescent signal. The number of rounds is variable; here we display an example of n=21. Diagram adapted from Chen et al. 2015.

MERFISH utilises encoding probes and readout probes to produce a combinatorial label for each RNA species. Each transcript is assigned a unique binary barcode—a sequence of 1s and 0s—that encodes the presence or absence of fluorescence across imaging rounds. A signal detected at a specific location, in a given imaging round, is recorded as a '1'; the absence of a signal is recorded as a '0' (Figure 2.3A). Across all rounds, this sequential on/off pattern produces a binary barcode unique to each RNA species, enabling both its accurate identification and spatial localisation (Figure 2.3B; Chen et al. 2015). In this study, we employed the Vizgen MERSCOPE platform, which uses three imaging channels over seven rounds of imaging, generating a 21-bit barcode for each RNA species.

To ensure accurate RNA species identification in MERFISH, an error-robust barcode design is implemented. This design accounts for the increasing risk of false positives associated with additional imaging rounds. As described in the original MERFISH publication (Chen et al. 2015), each RNA species is assigned a unique binary codeword that is designed to tolerate and correct errors introduced during imaging. These codes are selected based on an error-correcting scheme that maintains a minimum Hamming distance of four between any two valid codewords. This means that at least four-bit errors would be required to misclassify one RNA as another, allowing for single-bit or even multi-bit errors to be identified and corrected during decoding. Notably, the design accounts for the asymmetry in error types common in fluorescence imaging: a missed hybridisation event (a 1 → 0 error) is much more likely than a false-positive signal (a 0 → 1 error). To mitigate this, each barcode is designed to contain a small number of '1' bits. In the Chen et al. (2015) implementation, each barcode contains only four '1's, minimising the chance that a missed signal would lead to an incorrect identification.

Each experimental design includes a codebook that maps each 21-bit barcode to its corresponding transcript identity (Figure 2.3C). After decoding, each detected barcode is matched against the codebook, with error-correction algorithms applied to resolve any discrepancies. This process generates a "dot" or transcript call, containing both the assigned RNA identity and its spatial coordinates within the image (Figure 2.3D). Critically, tissue structure and placement must be maintained throughout the experimental protocol to assign transcript identity accurately. Subsequently, DAPI and PolyT stains are used to segment the image into individual cells. The detected transcripts are then assigned to these segmented cells based on their spatial location, resulting in a cell-by-gene expression matrix (Chen et al. 2015).

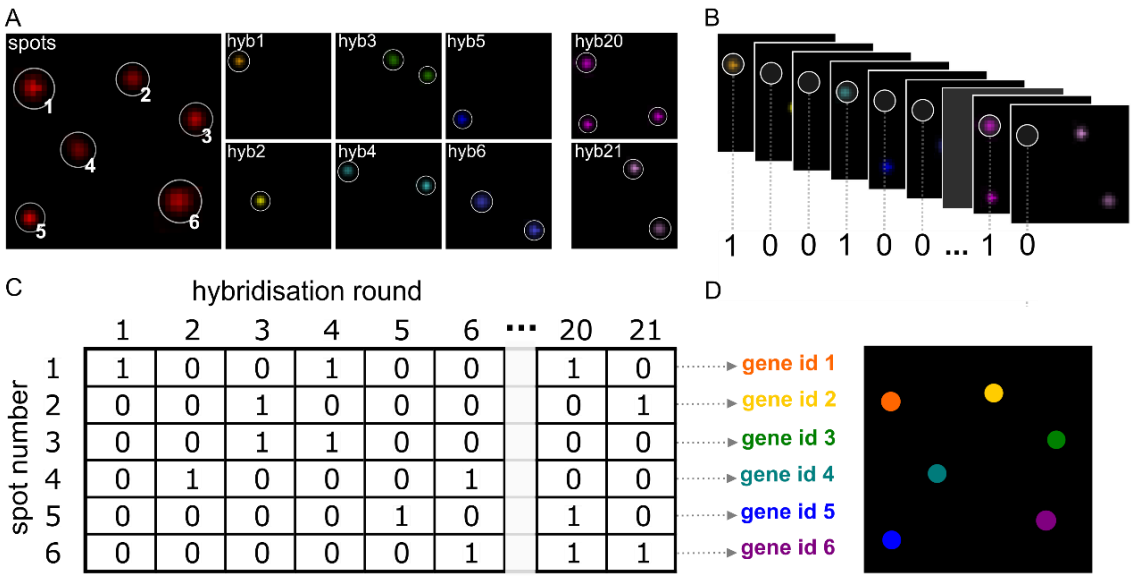


Figure 2.3 - MERFISH 21-bit barcoding enables RNA species identification

A) Simulated field of view showing the location of six RNA transcripts detected through the MERFISH experiment. B) Simulated MERFISH results showing fluorescence images from sequential rounds of hybridisation; white circles indicate the location of a single detected RNA transcript, as called by fluorescent signal. C) For each RNA transcript, a 21-bit binary barcode is constructed based on signal detection: a '1' indicates fluorescence detected in that round; a '0' indicates no signal detected. D) Barcodes are generated for all detectable RNA transcripts in the field of view and are matched to their corresponding gene identity, and a composite image with false dots representing each gene ID is generated. Diagram adapted from Chen et al. 2015.

2.2.3 - Implementation of imaging-based spatial transcriptomic techniques in plant tissues to date

In recent years, spatially resolved, imaging-based techniques have expanded rapidly, introducing novel strategies for interrogating gene expression in plant tissues. In this section, I highlight key publications that illustrate the development of these approaches in the field of plant biology and showcase the range of technologies now available to plant research.

In 2021, Laureyns et al. applied *in situ* sequencing (ISS) to the maize shoot apex, enabling spatial localisation of 90 developmental genes (2021). While ISS is often grouped with imaging-based approaches (Nobori 2025), some classify it separately from smFISH-based techniques (Moses and Pachter 2022). ISS relies on padlock probe hybridisation, rolling circle amplification, and sequence-by-ligation chemistry to detect individual mRNA molecules *in situ* (Ke et al. 2013). In this study, sections of the shoot apex were used to investigate the transcriptional dynamics underlying the formation of lateral organ primordia. A key finding was the spatial characterisation of *PLASTOCHRON1*, expressed at the boundary between indeterminate and determinate cells. Its expression partially overlapped with *ROUGHSHEATH1* and *OUTER CELL LAYER4*, an observation enabled by the high multiplexing capacity of ISS. Although this work represented a milestone in spatial profiling in plant tissues, the authors noted that the resolution of ISS at the time limited precise cellular localisation and quantification (Laureyns et al. 2021).

Subsequent advances have further refined imaging-based methods for plant tissues. In 2023, Nobori et al. introduced Phytomap (Plant Hybridisation-based Target Observation of Gene Expression Map), a protocol adapting multiplexed FISH for whole-mount plant samples. Building on *in situ* hybridisation and ISS, Phytomap employs DNA probes that hybridise to target mRNAs. After hybridisation, the probes are amplified *in situ* through rolling circle amplification. The amplified product contains a binding site for fluorescently labelled readout probes, enabling transcript detection via sequence-by-hybridisation chemistry. In their initial application, 28 genes were identified in a single experiment, validating cell-type marker genes detected through single-cell RNA-seq (scRNA-seq; Shahan et al. 2022; Nobori et al. 2023). A key advantage of Phytomap is its compatibility with whole-mount samples, circumventing the need for cryosectioning, particularly beneficial for small, delicate tissues such as *Arabidopsis* roots. The technique was subsequently applied to *Wolffia australiana*, a minute aquatic plant, where spatial localisation of 20 genes enabled the identification of four distinct

cell subpopulations, demonstrating its versatility across species and tissue types (Denyer et al. 2024). However, the number of genes that can be profiled with Phytomap is limited, as each imaging round captures four targets, with demonstrated effectiveness up to ten rounds (Nobori et al. 2023).

Another emerging technology, Molecular Cartography, developed by Resolve Biosciences, has gained traction in plant studies. This multiplexed *in situ* hybridisation technique does not rely on amplification, distinguishing it from ISS and Phytomap. Instead, it employs combinatorial smFISH across multiple imaging rounds, enabling the detection of up to 100 genes per sample—similar in principle to MERFISH (Groiss et al. 2021). Molecular Cartography was applied to maize roots to validate marker genes used in single-cell and single-nucleus RNA-seq-based cluster annotation (Guillotin et al. 2023). The method was subsequently used to map the spatial expression of 27 auxin-related genes and candidate regulators of vein patterning in maize leaves. Achieving single-cell resolution, this study performed cell segmentation and clustering across six tissue sections, revealing that distinct combinations of auxin influx and efflux transporters define major and minor leaf veins (Perico et al. 2024). Further applications also include the validation of cell type markers identified through scRNA-seq or single-nuclei RNA-seq (snRNA-seq). For example, Cervantes-Pérez et al. used Molecular Cartography to identify subpopulations of infected cells in mature soybean nodules, functionally annotating genes predicted to be cell-type specific (2024). Additionally, a study conducted large-scale single-cell profiling of maize and *Arabidopsis* meristems, identifying hundreds of cell type markers (Xu et al. 2024). Molecular Cartography was used to spatially map these markers in the developing maize ear, validating newly identified meristem marker candidates and revealing novel spatial domains in developing meristems (Xu et al. 2024).

The next major advance in plant spatial transcriptomics came with the integration of snRNA-seq data with imaging-based spatial platforms. This approach was first demonstrated by Nobori et al. (2025), who generated a time-resolved dataset of *Arabidopsis* leaves infected with *Pseudomonas syringae*. By combining single-nucleus multiome data (snRNA-seq and single-nuclei ATAC-seq) with a MERFISH panel targeting 500 genes, the authors validated cell-type markers and candidate genes involved in immunity, hormone signalling, and epigenetic regulation. Aligning spatial MERFISH profiles with snRNA-seq data further enabled inference of transcriptomic states beyond the 500 assayed genes (Nobori et al. 2025). This integrative strategy has since been applied to other plant systems, including barley and

wheat inflorescences, using Molecular Cartography (86 genes; Demesa-Arevalo et al. 2025; 99 genes; Xu et al. 2025). Most recently, a single-cell spatial transcriptomic atlas of the Arabidopsis life cycle was released (2025), profiling over 1,000 genes across ten developmental stages alongside a complementary snRNA-seq dataset of more than 400,000 nuclei (Lee et al. 2025). These studies illustrate both the growing multiplexing capacity of imaging-based spatial methods and their potential for cross-platform integration.

Together, these advances highlight the growing utility of imaging-based spatial methods for characterising gene expression in plant tissues. In 2024, I adapted MERFISH for use in wheat inflorescence tissues, extending the technique for the first time to a crop species. In this chapter, I describe the performance of MERFISH across multiple developmental stages of the wheat inflorescence and outline key sample preparation and analytical steps that are critical for establishing this method in a newly emerging area of plant biology.

2.3 - Results

2.3.1 - Generation of a 200 gene panel summarising wheat inflorescence development

To characterise gene expression along the apical–basal axis, we first selected genes for spatial profiling. Initial gene selection was informed by a microdissection RNA-seq dataset (Backhaus et al. 2022); however, given its limited developmental range and high variability, we conducted a more extensive analysis across spike development. We generated RNA-seq from central and basal spike sections across five development stages (Figure 2.4 A-B); Early and Late double ridge (EDR, LDR; Waddington stage W2, W2.5; respectively), Lemma Primordia (LP; W3.25), Terminal Spikelet (TS; W4), and Carpel Extension (CE; W5) (Waddington et al. 1983; Kirby and Appleyard 1984). Individual samples expressed, on average, 49,387 high-confidence genes, with 55,346 unique genes expressed across all samples. We identified 12,384 genes with significant differential expression between central and basal sections over time (Fig. 1B; See Methods), consistent with distinct spatial profiles along the apical-basal axis.

The MERFISH experimental workflow begins with the curation of a gene panel, which is used to design transcript-specific probes. We constructed a 200-gene panel capturing wheat inflorescence development (See Appendix 2), which incorporated genes spanning four broad categories: (1) genes differentially expressed along the apical–basal axis, (2) genes related to inflorescence development, (3) cell type marker genes identified from a single-nucleus RNA-seq study, and (4) housekeeping and control genes (Figure 2.4). In some cases, gene selection within each category was informed by prior genetic characterisation in wheat or in orthologs from related species, including *Oryza sativa*, *Zea mays*, and *Hordeum vulgare*. However, our selection also incorporated genes not yet characterised in grass species. Before final selections, we examined expression profiles from the microdissection RNA-seq dataset to ensure that each selected gene was expressed at one or more of the sampled developmental time points (significant expression considered >0.5 TPM).

Of the 200 gene panel, we selected 121 genes that are differentially expressed across the apical-basal axis of the inflorescence, as determined by the RNA-seq microdissection dataset. Our goal was to capture diverse transcriptional trajectories across developmental timepoints. For instance, we included genes with strong expression in basal regions of the inflorescence at early stages (Figure 2.4B - cluster 3; W2 - W2.5), such as *VEGETATIVE TO REPRODUCTIVE TRANSITION 2 (VRT2)*, *SHORT VEGETATIVE PHASE 1 (SVP1)*, and

VERNALIZATION 1 (VRN1). In contrast, we selected genes with elevated expression in apical regions at the same stages (Figure 2.4B - cluster 2), including *KNOTTED-LIKE HOMEOBOX 5 (KNOX5)* and *SEPALLATA 1–6 (SEP1–6)*. An additional 52 genes were chosen for their progressive upregulation during later floral development stages (Figure 2.4B, clusters 1, 4, 5; W3.25–W5).

An important component of the gene panel was the inclusion of housekeeping genes. Although the number of characterised housekeeping genes in wheat is limited, we supplemented our selection based on orthology with well-characterised housekeeping genes in other grasses. These genes are anticipated to be expressed across all cell types in the inflorescence and surrounding vegetative tissues, making them helpful in assessing transcript detection and expression uniformity across tissue samples (Figure 2.4C). The selected housekeeping genes span a range of expression levels, including low (e.g., *CycT1*, *mEF-G*, *ACT1–5*, *CAB*; average TPM 5.60–14.48), moderate (*ATG8g*, *eIF-4a*, *UBQ5*; average TPM 31.75–56.85), and high expression (*GAPC3*, *GAPC1*; average TPM 54.90 and 187.75, respectively).

In addition to transcription factors identified through differential expression analysis, we supplemented the panel with 39 transcription factors drawn from the literature for their general roles in inflorescence development across grasses. Together with those selected from the RNA-seq dataset, this brought the total to 127 transcription factors in the panel, representing 34 gene families. The most significant contributions came from the MADS-MIKC (23), HB-HD-ZIP (8), bHLH (8), bZIP (7), NAC (6), and AP2/ERF-AP2 (6) families, based on the classification system of Evans et al. (2022).

In addition, we incorporated putative cell-type marker genes identified in a recent ssnRNA-seq study of wheat inflorescence (Liu et al. 2025), with candidate gene selection guided by our collaborators. These genes are hypothesised to exhibit cell-type-specific enrichment, although the spatial distribution of the corresponding cell types within the wheat inflorescence remains uncharacterized. They serve a dual role: facilitating the annotation of snRNA-seq clusters and contributing to spatial analyses of MERFISH data to delineate major cell-type domains (Liu et al. 2025).

The final gene panel comprises 172 genes related to inflorescence development, 100 genes involved in grain development (for a separate project), and 28 additional genes used for housekeeping, controls, or miscellaneous purposes.

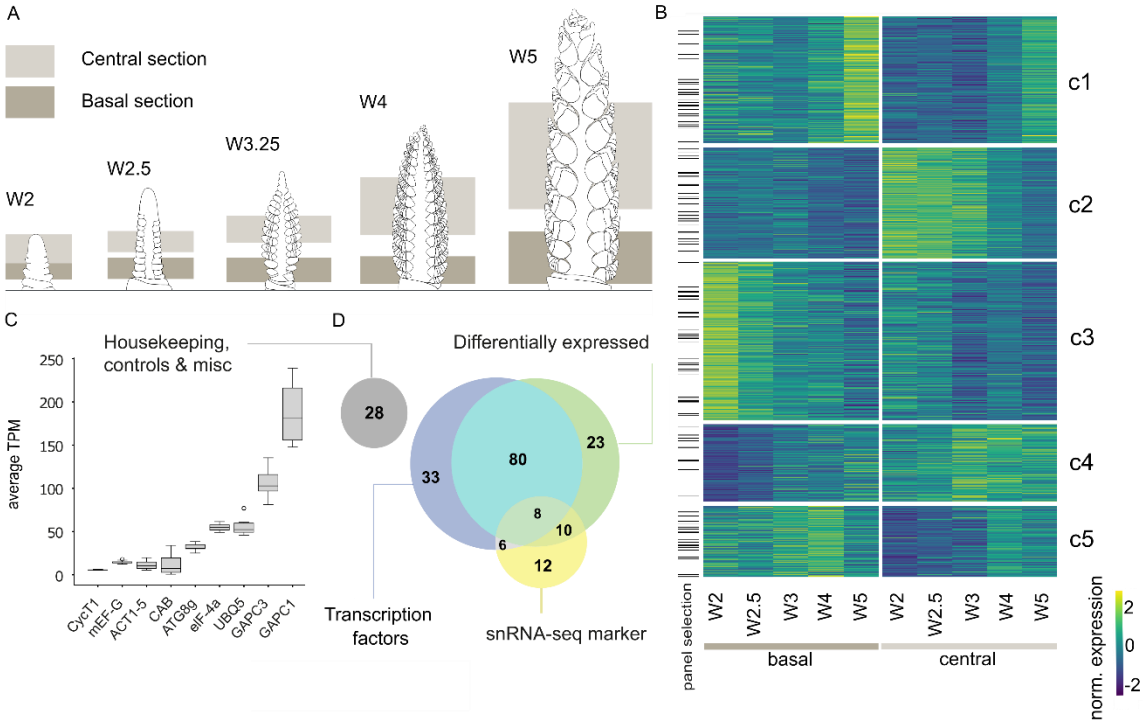


Figure 2.4 - Curation of a 200-gene panel used to investigate wheat inflorescence development

A) Schematic illustrating the central and basal spike regions sampled for pooled-tissue RNA-sequencing across five developmental stages (W2, W2.5, W3.25, W4, W5; not to scale). **B)** 12,384 genes were differentially expressed between central and basal regions of wheat inflorescence microdissections. The 121 genes selected for inclusion in the MERFISH panel are indicated by black bars. K-means clustering determines five (c1-c5) transcriptional trajectories identified through differential expression analysis. **C)** Average expression (transcripts per million; TPM) values for nine housekeeping genes in the micro dissected RNA-seq dataset (timepoints W2-W5), incorporated into the MERFISH panel. **D)** Composition of 200-gene MERFISH panel, including differentially expressed genes selected from microdissection dataset, single-nuclei RNA-seq markers (Liu et al. 2025), additional transcription factors, and housekeeping/control genes.

2.3.2 - Optimisation of smFISH in wheat inflorescence tissues

MERFISH sample preparation involves tissue dissection, fixation, embedding, and subsequent cryosectioning. A critical first step in optimising this workflow is to ensure that the preparation protocol preserves RNA integrity (particularly during fixation), and that the processed tissue yields high-quality MERFISH results. To assess this, we tested prepared samples using smFISH probes and a rapid imaging protocol on the MERSCOPE instrument. Briefly, encoding probes with two unique readout bits were imaged using two fluorescence channels. High colocalisation of signals between these two channels is used as a proxy for RNA integrity, with greater overlap indicating better RNA preservation (Figure 2.5A).

We used validation probes targeting *TraesCS6B02G144000*, the wheat ortholog of rice *EUKARYOTIC INITIATION FACTOR 4A-2 (EIF4A2)*. This gene is expected to be broadly expressed, with average expression ranging from 48.9 to 61.4 TPM in the microdissection RNA-seq dataset. To quantify spot detection and colocalization, we utilised the smFISH image processing toolbox FISH-quant (Imbert et al. 2022). We selected a 50 μm^2 region of spikelet meristem tissue from a high-quality section of wheat inflorescence tissue (W3.25). From this region, we detected 554 spots in channel 1, of which 154 (27.8%) overlapped with spots in channel 2. Conversely, of 507 spots in channel 2, 30.4% overlapped with channel 1 (Figure 2.5B). This process was repeated in two additional spikelet meristem regions of the inflorescence (Figure 2.5C-E). These yielded similar results, with overlap ranging from 25.8%-45.6% in total (Table 2.1). These results indicate spatial co-localisation between the two channels. However, because this metric is primarily used as an internal quality check within the Vizgen MERSCOPE workflow, it was challenging to identify comparable benchmarks in the literature. Following consultation with Vizgen staff, we interpreted these results as evidence of sufficient RNA quality and successful sample preparation.

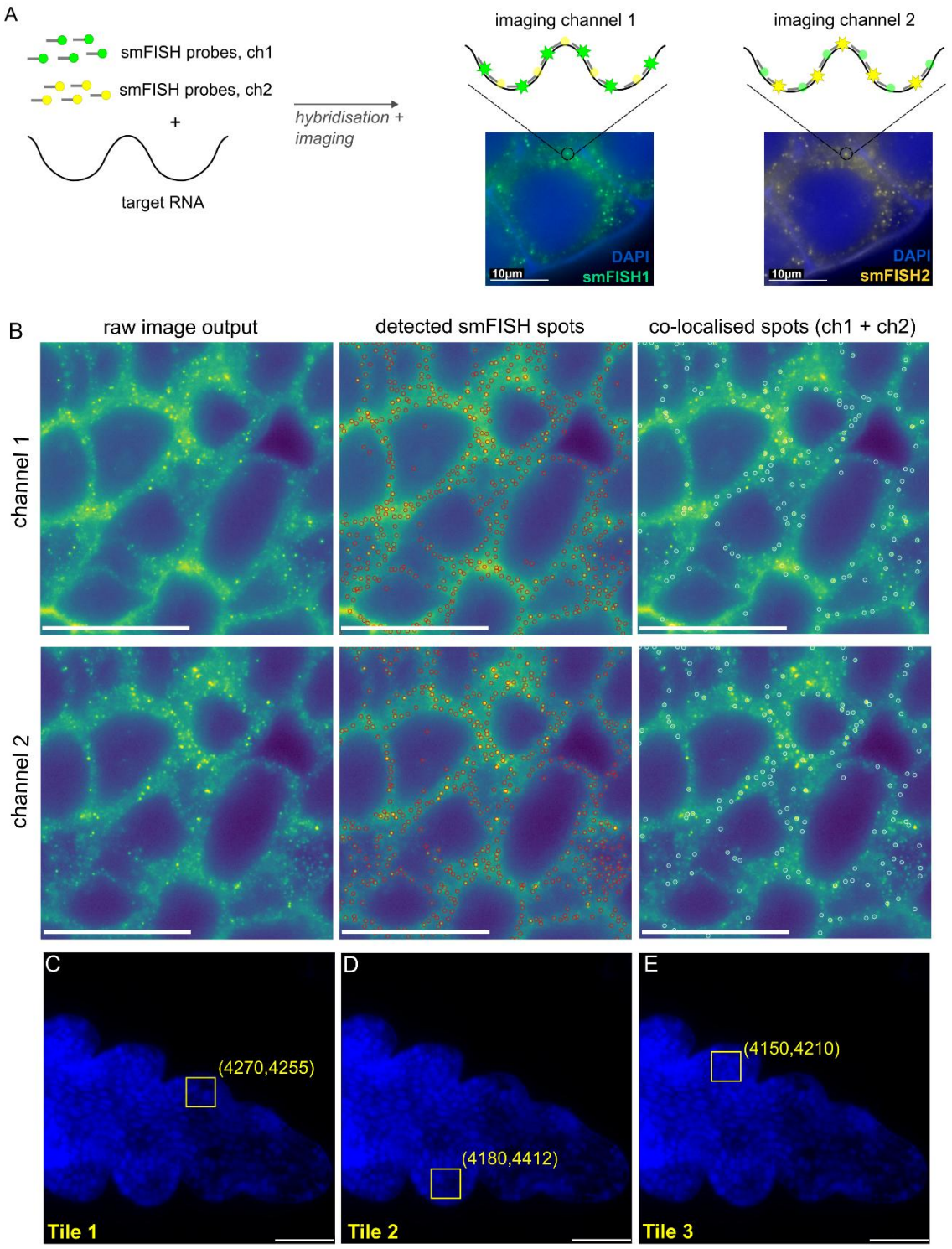


Figure 2.5 - Verification of smFISH signal co-localisation

A) Schematic of experimental design. smFISH probes were labelled with fluorophores detected in two imaging channels (647 nm and 541 nm). Microscopy images show probe hybridisation in both channels (smFISH1 and smFISH2), visualised using the MERSCOPE Visualizer Tool. DAPI = blue; smFISH1 = green; smFISH2 = yellow. Scale bar = 10 µm. **B)** Co-localisation of smFISH spots quantified with the smFISH image processing toolbox FISH-quant (Imbert et al., 2022). Example from a 50 µm ×

50 µm spikelet tissue tile (Tile1). All detected smFISH spots are shown in red; co-localised spots between channels are shown in white. Visualised using the BIG-FISH function plot_detection() with contrast enhancement. Scale bar = 25 µm. **(C-E)** Three 50 µm × 50 µm tiles were sampled for smFISH quantification. Images acquired on the MERSCOPE instrument. Scale bar = 100 µm; DAPI = blue.

Table 2.1 - smFISH spot co-localisation in 50 µm × 50 µm regions of meristem tissue.

Tile Name	Coordinates X, Y (µm)	smFISH channel	Spots Detected	Spots Co-localised	% Spots Co-localised
Tile 1	(4270, 4255)	1	554	154	27.80%
		2	507		30.40%
Tile 2	(4180, 4412)	1	544	173	31.80%
		2	379		45.60%
Tile 3	(4150, 4210)	1	414	107	25.80%
		2	392		27.30%

2.3.3 - Implementation of MERFISH workflow defines spatial gene expression patterns of a 200-gene panel

After designing and synthesising the 300-probe set, we performed the Vizgen MERSCOPE workflow (Figure 2.6). Briefly, developing wheat spikes were dissected at four stages (W2.5, W3.25, W4, W5; Waddington et al. 1983), embedded in optimal cutting temperature (OCT) compound, flash frozen, and cryo-sectioned (See Appendix 3). To accommodate for variable sectioning angle, we ensured the maximum amount of inflorescence tissues within a 1cm x 1cm area. Each OCT block included 5–36 spikes (depending on stage; Figure 2.7A-D) from two near-isogenic lines: one carrying the wildtype *VRT-A2a* allele ($P1^{WT}$) and the other the misexpression *VRT-A2b* allele ($P1^{POL}$) from *T. turgidum* ssp. *polonicum* (Adamski et al. 2021).

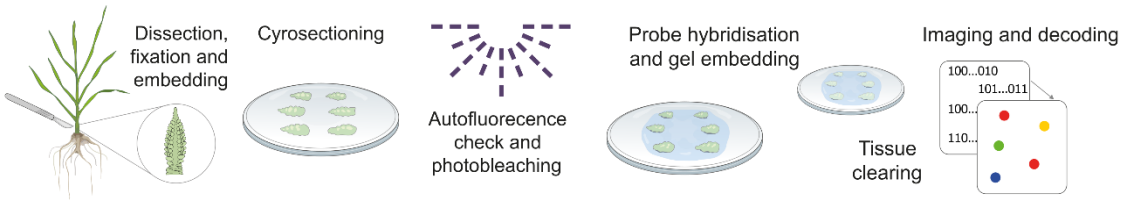


Figure 2.6 - MERSCOPE Workflow for plant tissues

To accommodate high auto-fluorescence in plant tissues, we performed autofluorescence checks and increased photobleaching times as required, totalling three hours. Additionally, we varied the amount of time left in clearing solution before imaging, ranging from 3 h to 4.5 days. Of the cryosectioned tissue included in the imageable area, some tissue experienced lifting from slides after tissue clearing steps. To assess the quality of tissue sections, we first imaged sections on the MERSCOPE instrument, using a rapid imaging round (~15 min) to generate a low-resolution mosaic of DAPI staining (Figure 2.7E-H). Using the MERSCOPE Instrument User Interface, we selected from the available sections a subset of tissues to image with the region of interest selection tool. For the datasets covered in this thesis, we imaged across four experimental runs in total, including a total of 36 imaged areas. The size of raw imaging data generated from each experimental run ranged from 460 GB to 1.52 TB total.

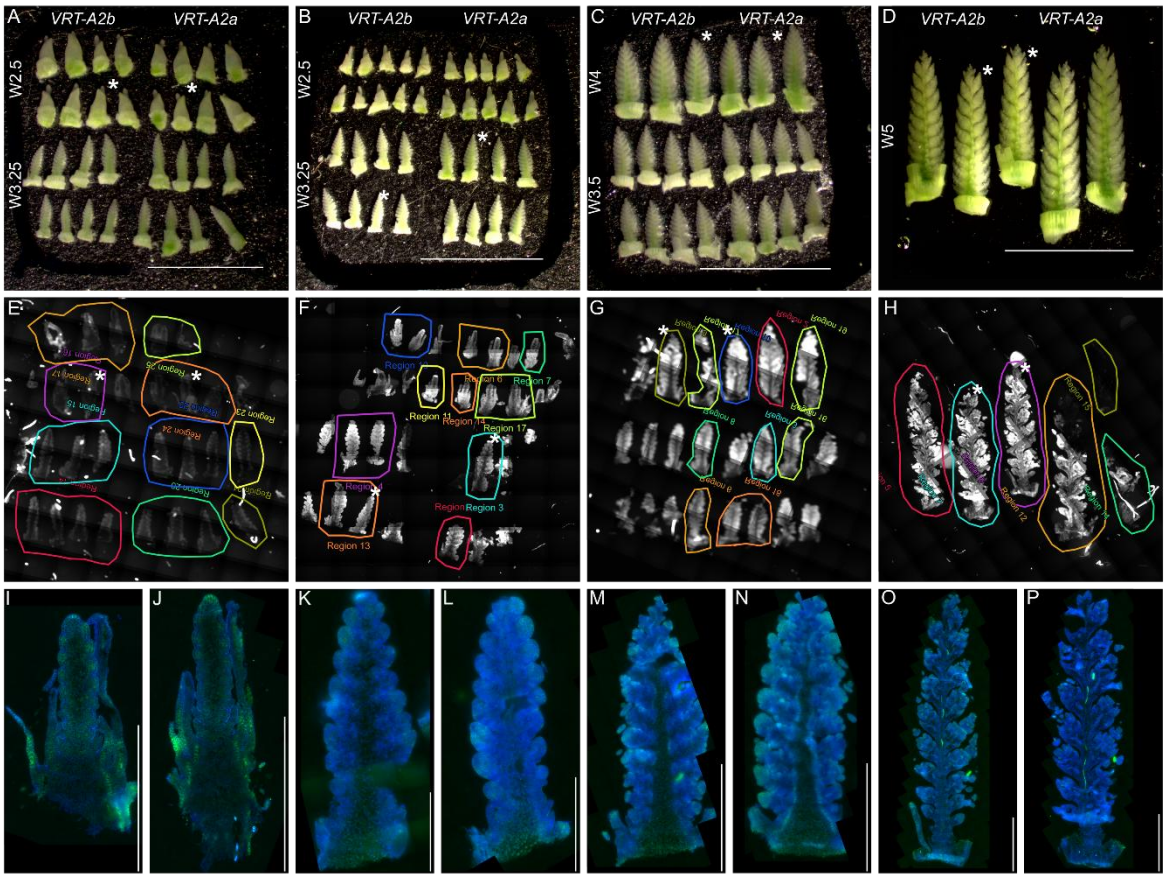


Figure 2.7 - Optimal cutting temperature (O.C.T.) block layout, 'Region of Interest' selections, and eight cryosections across four MERSCOPE experimental runs

A-D) Layout of wheat inflorescences in OCT blocks annotated with genotype and developmental time point annotations. Asterisks denote spikes selected for final analysis. Taken on Leica S9 stereomicroscope. Scale bar = 5mm. **E-H)** DAPI stain overview and experimental region selections from MERSCOPE Instrument output. **I-P)** Eight selected cryosections for onward analysis, stages W2.5 (I-J), W3.25 (K-L), W4 **(M-N)**, W5 **(O-P)** in genotypes $P1^{WT}$ **(I,K,M,O)** and $P1^{POL}$ **(J,L,N,P)**. Staining images taken on the MERSCOPE Instrument before the MERFISH experimental run (See Methods). DAPI = Blue, PolyT = Green.

For downstream analysis, we selected two representative samples from each developmental time point, one per genotype (*P1*^{WT} and *P1*^{POL}), resulting in a total of eight samples (Figure 3.7I-P). Selection was based on sectioning angle, tissue adherence, and the number of detected transcripts to ensure sample quality. Among these eight regions, the total tissue area captured on the slide ranged from 0.50 to 4.35 mm^2, and the number of transcripts detected within tissue boundaries ranged from 147,308 to 2,636,435 counts.

To summarise transcript density, we calculated the number of transcripts detected per 100µm^2, which ranged from 19.06 – 63.89 (Table 2.2). In total, we captured 7,308,224 transcripts across eight samples. The top five highest detected genes captured across eight experimental runs were *TraesCS7A02G313100* (739,501 counts, ortholog to rice *GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC*), *TraesCS3A02G155200* (422,689 counts, ortholog to rice *AUX/IAA PROTEIN 3*), *TraesCS2B02G260800* (393,558 counts, ortholog to rice *METHYLTRANSFERASE 1B*), *TraesCS3D02G284200* (341,756 counts, *AGAMOUS-LIKE 14*), and *TraesCS2A02G174300* (242,771 counts, *FRUITFULL3*).

Table 2.2 - Total capture area and detected transcript count across eight MERFISH samples

Genotype	Waddington Stage	Total detected transcripts	Tissue Area (mm^2)	Transcript counts / 100 µm^2
P1^{WT}	W2.5	295,537	0.7529	39.2516
P1^{POL}	W2.5	147,308	0.7728	19.0616
P1^{WT}	W3.25	401,150	0.8162	49.1483
P1^{POL}	W3.25	212,643	0.4997	42.5533
P1^{WT}	W4	575,126	1.5833	36.325
P1^{POL}	W4	990,962	1.8071	54.8371
P1^{WT}	W5	2,636,435	4.1266	63.8888
P1^{POL}	W5	2,049,063	4.3467	47.1407

2.3.4 - Cell segmentation and transcript assignment refine gene expression data to cellular resolution

A critical first step in spatial transcriptomic analysis is the segmentation of individual cells from stained tissue images. To achieve this, we used mosaic images of DAPI and PolyT stains and performed cell segmentation using Cellpose2 (Pachitariu and Stringer 2022). We optimised the segmentation parameters to maximise cell detection while minimising small segmentation artefacts, identifying an optimal flow threshold of 1.0 and a cell probability threshold of –3.0 (Figure 2.8).

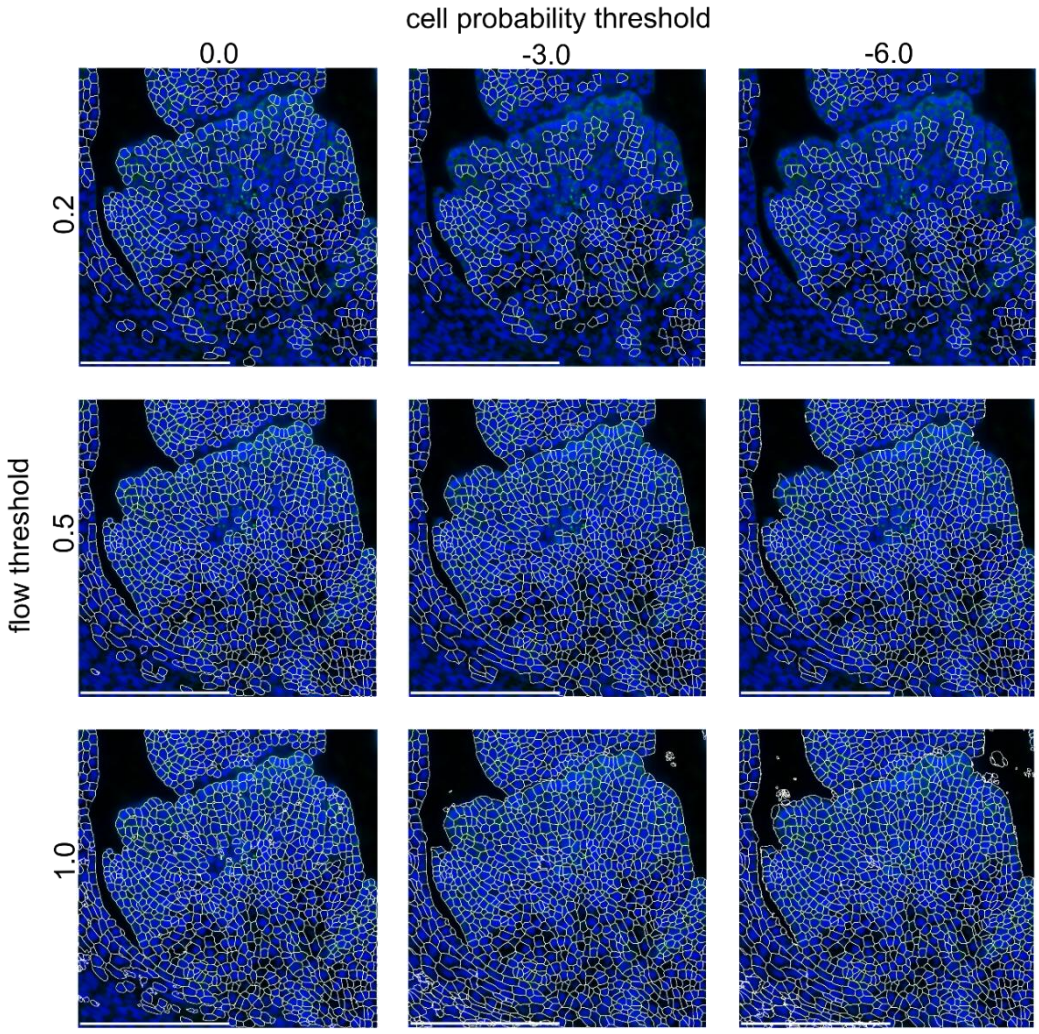


Figure 2.8 - Optimisation of cell segmentation parameters in Cellpose2 improves segmentation outputs

Segmentation outputs with parameters of flow threshold (0.2 to 1.0) and cell probability threshold (-6.0 to 0.0) in cellpose2 on a 500 µm x 500 µm region of wheat spikelet tissue (Pachitariu and Stringer 2022). Staining images taken on MERSCOPE Instrument prior to MERFISH experimental run (See Methods). DAPI = Blue. Segmentation lines = white. Scale bar = 250µm.

A particular challenge was accurate segmentation across stitching lines in the mosaic.tiff files generated by the MERSCOPE instrument. These stitching lines, which arise where adjacent image tiles are joined during processing, often introduce artificial boundaries that interfere with segmentation. To address this, we developed a preprocessing method in ImageJ to lighten the seam lines, enabling more accurate segmentation across tile boundaries (Figure 2.9; Schindelin et al. 2012; Schneider et al. 2012).

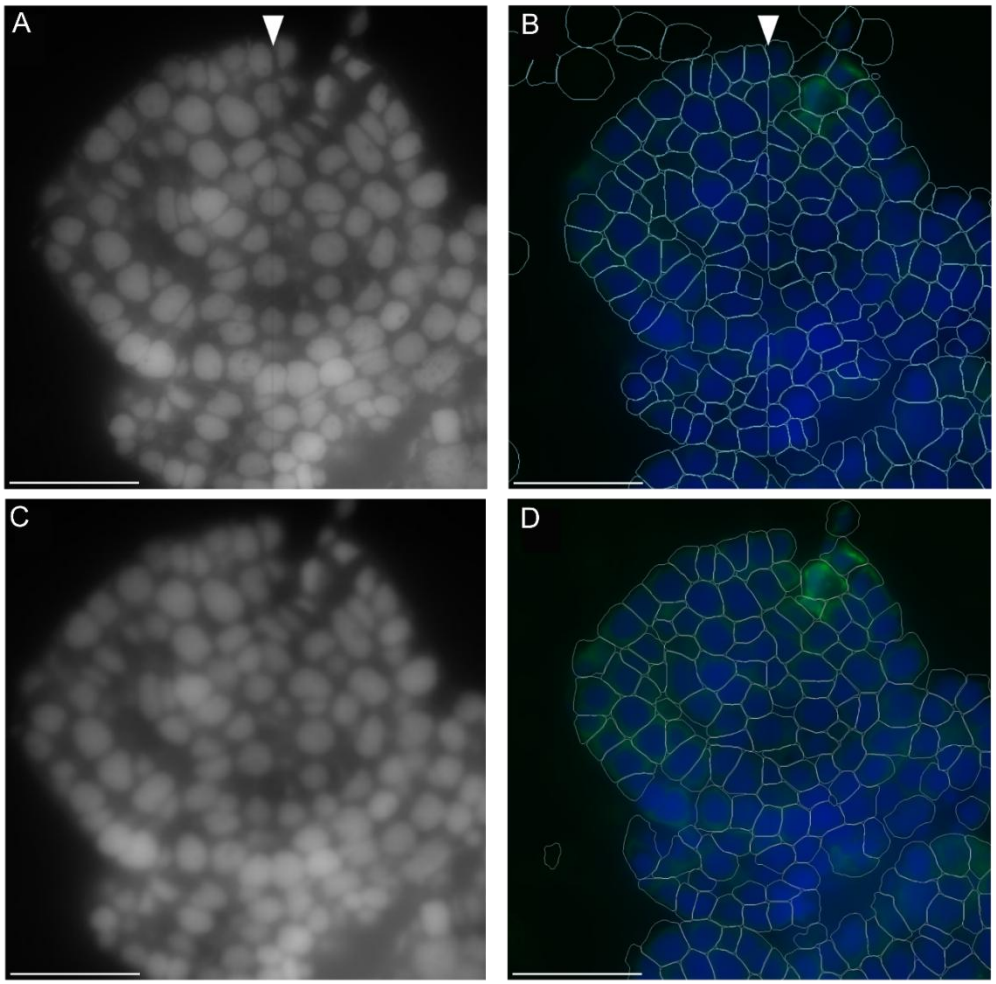


Figure 2.9 - Lightning of seam lines in staining images improves cellular segmentation outputs **A)** Raw DAPI stain, as output from MERSCOPE Instrument. **B)** Cellpose2 segmentation outputs prior to image edits, visualised in Vizgen MERSCOPE Visualiser Tool. White arrowhead denotes seam line detected through segmentation, resulting in false cell boundaries. **C)** Image J filters (Maximum Filter, x3 and Median Filter, x3, Gaussian blur, See Methods) applied to DAPI stain image. **D)** Cell segmentation of edited DAPI stain image visualised in Vizgen MERSCOPE Visualiser Tool. DAPI = blue, PolyT = green, segmentation lines = white. Scale bar = 50µm.

Cell segmentation (Pachitariu and Stringer 2022) and transcript assignment (Wiggin and Yu 2024) produced a cell-by-gene matrix detailing transcript counts per cell and gene counts per cell. This allowed for the filtering of segmentation artefacts (e.g., non-cellular objects) and low-quality cells with low transcript counts. Cells were removed from the dataset if they contained less than 25 transcript counts in total, and a volume threshold of >500 um^3, which reduced the number of segmented cells from 105,908 to 50,731 cells across eight samples (Figure 2.10).

After filtering low-quality cells, we evaluated sample quality based on total transcript abundance and the number of genes detected per cell. The number of cells captured per sample ranged from 1,569 (W2.5, $P1^{POL}$) to 15,185 (W5, $P1^{WT}$). Across all samples, average transcript counts per cell ranged from 77.9 to 152.4, while the average number of genes detected per cell ranged from 29.6 to 42.1 (Table 2.3).

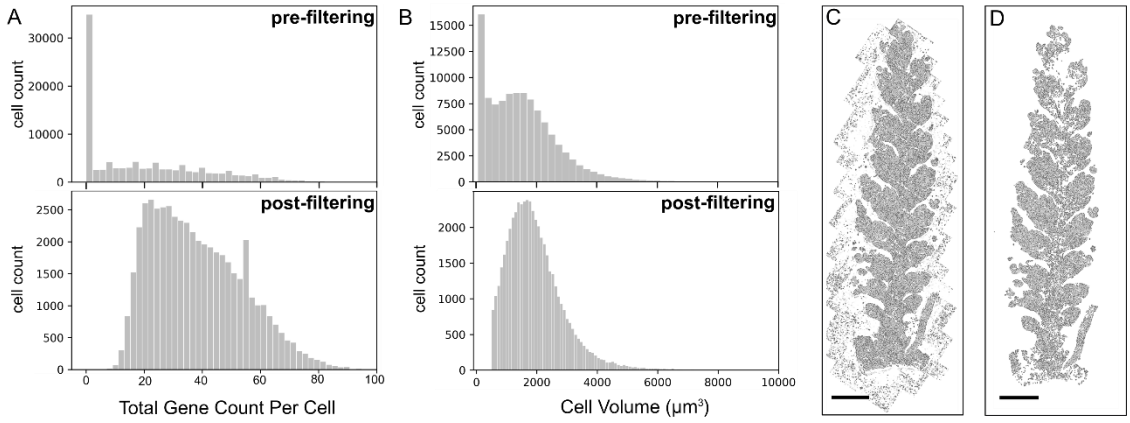


Figure 2.10 - Filtering of low-quality cells in eight selected MERFISH samples

A) Total gene count per cell before and after filtering low-quality cells (> 25 transcript counts per cell), combined across eight samples. **B)** Cell volume before and after filtering segmentation artefacts (> 500 μm^2). **C-D)** Cellular segmentation in cellpose2 before **(C)** and after **(D)** filtering of segmentation artefacts in sample W5, $P1^{WT}$.

Table 2.3 - Transcript count and gene count per cell across eight MERFISH samples

Genotype	Waddington Stage	Number of Cells	Metric	Mean	Median	Min	Max	Range
$P1^{WT}$	W2.5	2209	transcript count per cell	118.5	93	25	615	590
$P1^{POL}$	W2.5	1569	transcript count per cell	77.9	60	25	448	423
$P1^{WT}$	W3.25	2896	transcript count per cell	122.7	97.5	25	625	600
$P1^{POL}$	W3.25	2107	transcript count per cell	87.3	69	25	606	581
$P1^{WT}$	W4	5405	transcript count per cell	88.4	61	25	823	798
$P1^{POL}$	W4	6385	transcript count per cell	133.9	103	25	862	837
$P1^{WT}$	W5	15185	transcript count per cell	152.4	118	25	1092	1067
$P1^{POL}$	W5	14975	transcript count per cell	118.4	84	25	1009	984
$P1^{WT}$	W2.5	2209	number of genes per cell	36.9	36	9	81	72
$P1^{POL}$	W2.5	1569	number of genes per cell	29.6	27	3	80	77
$P1^{WT}$	W3.25	2896	number of genes per cell	39.6	38	11	97	86
$P1^{POL}$	W3.25	2107	number of genes per cell	34.4	32	9	87	78
$P1^{WT}$	W4	5405	number of genes per cell	32.2	29	7	99	92
$P1^{POL}$	W4	6385	number of genes per cell	39.6	38	8	93	85
$P1^{WT}$	W5	15185	number of genes per cell	42.1	41	5	105	100
$P1^{POL}$	W5	14975	number of genes per cell	37.6	35	5	99	94

After cell filtering, transcript counts per cell exhibited a non-normal distribution, with a pronounced left skew and a wide range (Figure 2.11A). This distribution pattern was consistent across all eight samples (Figure 2.11B). For instance, in the W2.5 spike sample ($P1^{WT}$), total transcript counts per cell ranged from 25 to 615 (Figure 2.11C). To mitigate the impact of technical variability and enable meaningful comparisons across cells and samples, transcript counts were normalised prior to downstream analysis (Figure 2.11D-F). To do so, we utilised standard normalisation functions in Scanpy (Wolf et al. 2018), which adjust the total number of counts to be uniform across all cells, followed by a log-transformation of the count matrix.

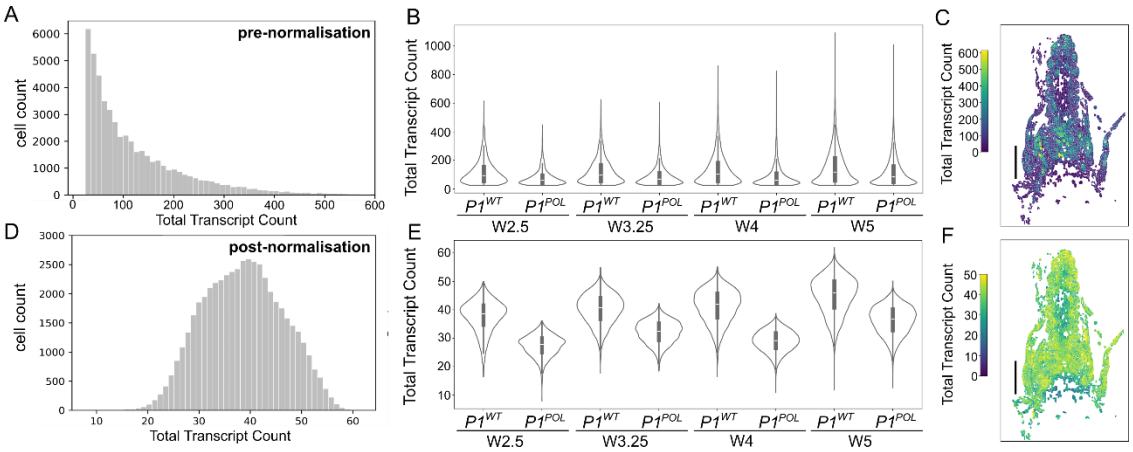


Figure 2.11 - Normalisation of transcript count in eight selected MERFISH samples

(A-C) Total transcript counts per cell before normalisation, displayed in **A)** frequency plot of across eight combined samples, B) violin plot of individual samples. **C)** heat plot of total transcripts per cell displayed over cellular segmentation in sample W2.5, $P1^{WT}$. **D-F)** Total transcript count per cell, after counts normalisation with Scanpy functions pp.normalize_total() and pp.log1p() (Wolf et al. 2018); displayed in D) frequency plot across eight samples combined, **E)** violin plot of individual samples, **F)** heat plot of total transcripts per cell displayed over cellular segmentation in sample W2.5, $P1^{WT}$.

Beyond results from cell segmentation, we observed gene expression patterns restricted to specific cell layers, demonstrating the cellular resolution achievable with MERFISH. For example, *TraesCS4D02G296400*, the ortholog to rice *ONION1,* showed highly localised transcript enrichment in the L1 and epidermal layers of the inflorescence, illustrating layer-specific gene expression (Figure 2.12A-E). This spatially resolved expression pattern highlights MERFISH's ability to detect transcriptional differences at the single-cell level.

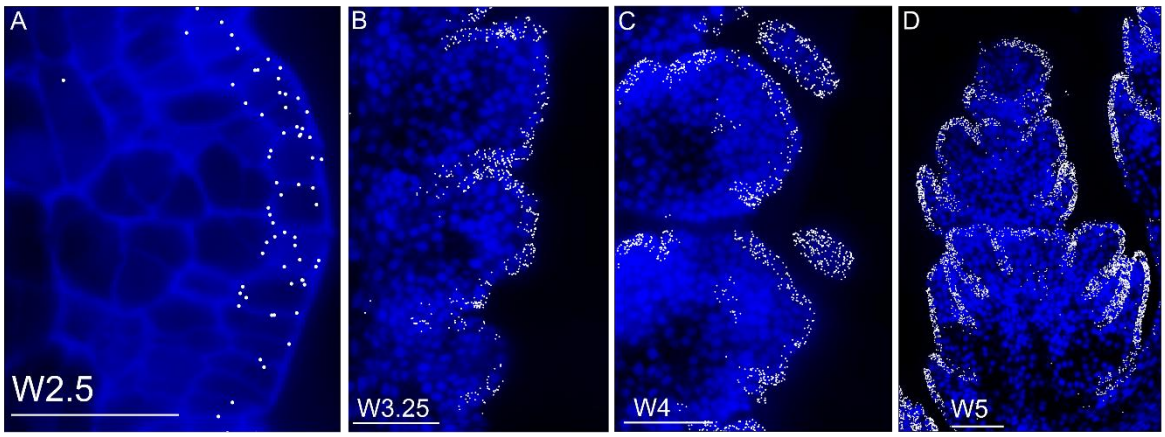


Figure 2.12- MERFISH refines gene expression to a single-cell layer

Transcript localisation of TraesCS4D02G296400 in four regions of *P1*^{WT} inflorescence. A-D) stages W2, W3.25, W4, and florets (E; W5). DAPI = blue. Detected transcripts = white. Scale bar = 100µm.

2.3.5 - MERFISH quality control and verification

We next implemented quality control measures to evaluate the off-target binding and false identification rate of MERFISH in wheat inflorescence tissue. In MERFISH, each transcript is identified through a unique barcode. To assess the percentage of errors in transcript detection, the MERSCOPE platform detects the presence of 15 'blank' barcodes not included in the experimental library (Chen et al. 2015). Across all samples, we detected minimal off-target hybridisation, with blank probe-derived transcripts accounting for less than 0.28% of total counts (range by sample: 0.24% to 0.34%; Table 2.4). This low background signal suggests high probe specificity and supports the robustness of our hybridisation conditions in a plant tissue context.

In some cases, gene probes were detected at low rates across all samples. To identify probes likely reflecting stochastic detection, we used the detection rates of blank barcodes as a baseline. For each sample, we normalised the per-cell count matrix (including both gene and blanks) and pooled the blanks to define a background distribution of per-cell counts associated with stochastic detection. Each of the 200 gene probes was then compared to this distribution using a two-sample Kolmogorov–Smirnov (KS) test (see Methods). This analysis identified eight gene probes that were consistently low confidence across all eight samples, suggesting their expression likely reflects non-specific or background detection. The affected genes include: *TraesCS2A02G306800, TraesCS1B02G274200, TraesCS5A02G405900, TraesCS7D02G233300, TraesCS2D02G019300, TraesCS5B02G560300, TraesCS3A02G093200,* and *TraesCS7A02G076500*.

Due to the hybridisation-based nature of MERFISH and the ~98% sequence identity among wheat homoeologs, we expected cross-hybridisation between homoeologous transcripts. To test this, we included a probe set designed to target a B-genome homoeolog with minimal expression (average 0.03 TPM; microdissection RNA-seq dataset, Figure 2.4), while the corresponding A and D copies are expressed (16.05 and 14.05 average TPM, respectively, Figure 2.13A). Despite being B-genome specific by design, this probe detected an average of 0.20 transcripts per cell, approximately tenfold higher than the signal observed in blank controls (Figure 2.13B). This signal likely reflects binding to the closely related A and D homoeologs (Figure 2.13C), highlighting a lack of homoeolog-specificity under the current hybridisation conditions.

Table 2.4- Total Counts per sample of gene probes compared to blank barcode detection

Genotype	Waddington Stage	Total Counts (Including Blanks)	Blank Probe Counts	Percentage of Blank Probe in Total Counts	Gene Probe Counts	Percentage of Gene Probe in Total Counts
P1^{WT}	W2.5	262,562	713	0.27%	261,849	99.73%
P1^{POL}	W2.5	122,641	418	0.34%	122,223	99.66%
P1^{WT}	W3.25	356,417	973	0.27%	355,444	99.73%
P1^{POL}	W3.25	184,436	442	0.24%	183,994	99.76%
P1^{WT}	W4	478,986	1270	0.27%	477,716	99.73%
P1^{POL}	W4	857,615	2569	0.30%	855,046	99.70%
P1^{WT}	W5	2,320,062	6115	0.26%	2,313,947	99.74%
P1^{POL}	W5	17,77,816	5404	0.30%	1,772,412	99.70%

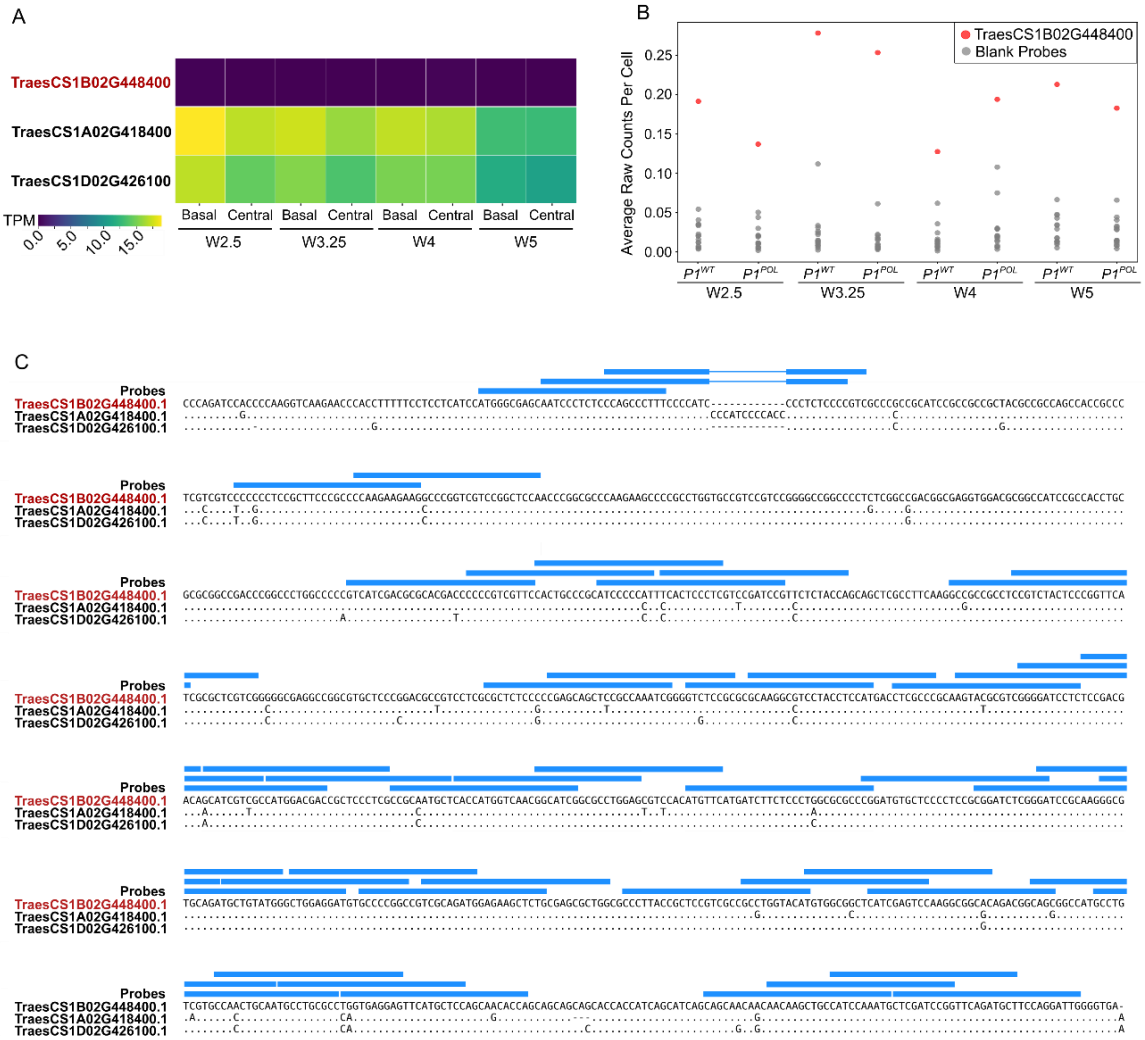


Figure 2.13 - MERFISH probes display non-homoeologous binding activity

A) Transcript per million (TPM) of the homoeologous triad from RNA-seq of micro-dissected basal and central spike regions at stages W2.5–W5. B) Average counts per cell (pre-normalisation) of TraesCS1B02G448400 compared to 15 blank probes across eight samples. C) Sequence alignment of homoeologous triad TraesCS1B02G448400, TraesCS1A02G418400, and TraesCS1D02G426100 using MUSCLE (v5, Edgar 2022). Probes were designed to TraesCS1B02G448400; blue bars indicate individual probe target sites along the transcript. Note that probes are purposely targeting polymorphic regions among homoeologs.

2.3.6 - Spatial gene expression patterns are consistent with single gene *in situ* hybridisation and bulk RNA-seq

To further validate the accuracy and biological relevance of our MERFISH data, we generated *in silico* sections that correspond to the physical microdissections used in a previous bulk RNA-seq experiment (Figure 2.4). We then compared the normalised gene expression data within the *in silico* sections with those obtained from bulk RNA-seq of manually dissected tissue. Across eight samples spanning four developmental time points, the average Spearman's correlation coefficient between the two datasets was 0.66 (range: 0.62–0.73; Figure 2.14), indicating moderate to strong agreement. Spearman's rank correlation was chosen for its robustness to non-linear relationships and differences in expression scale between technologies. This level of concordance supports the consistency between approaches.

Additionally, we compared MERFISH data with published *in situ* hybridisation studies in inflorescence tissues. Given the limited number of studies in wheat, we broadened the comparison to include other cereals (*Zea mays, Oryza sativa, Hordeum vulgare) and* compared results from $P1^{WT}$ samples to analogous stage and section areas between species (Figure 2.15). We observed the consistent expression patterns comparing genes in wheat, in *PISTILLATA1 (PI1), SHORT VEGETATIVE PHASE 1 (SVP1), FRUITFULL 2 (FUL2), AGAMOUS-LIKE 6 (AGL6), VEGETATIVE TO REPRODUCTIVE TRANSITION 2 (VRT2), SEPALLATA 1-6 (SEP1-6),* and *BARREN STALK1 (BA1).* In genes orthologous to maize, we observe consistent expression patterns in wheat to *OUTER CELL LAYER 4 (ZmOCL4), KNOTTED 1 (ZmKN1), TASSEL SHEATH 1 (ZmTSH1),* as well as the orthologs to rice *TONGARI-BOUSHI3* (*TOB3)* and barley *MANY-NODED DWARF1 (MND1).*

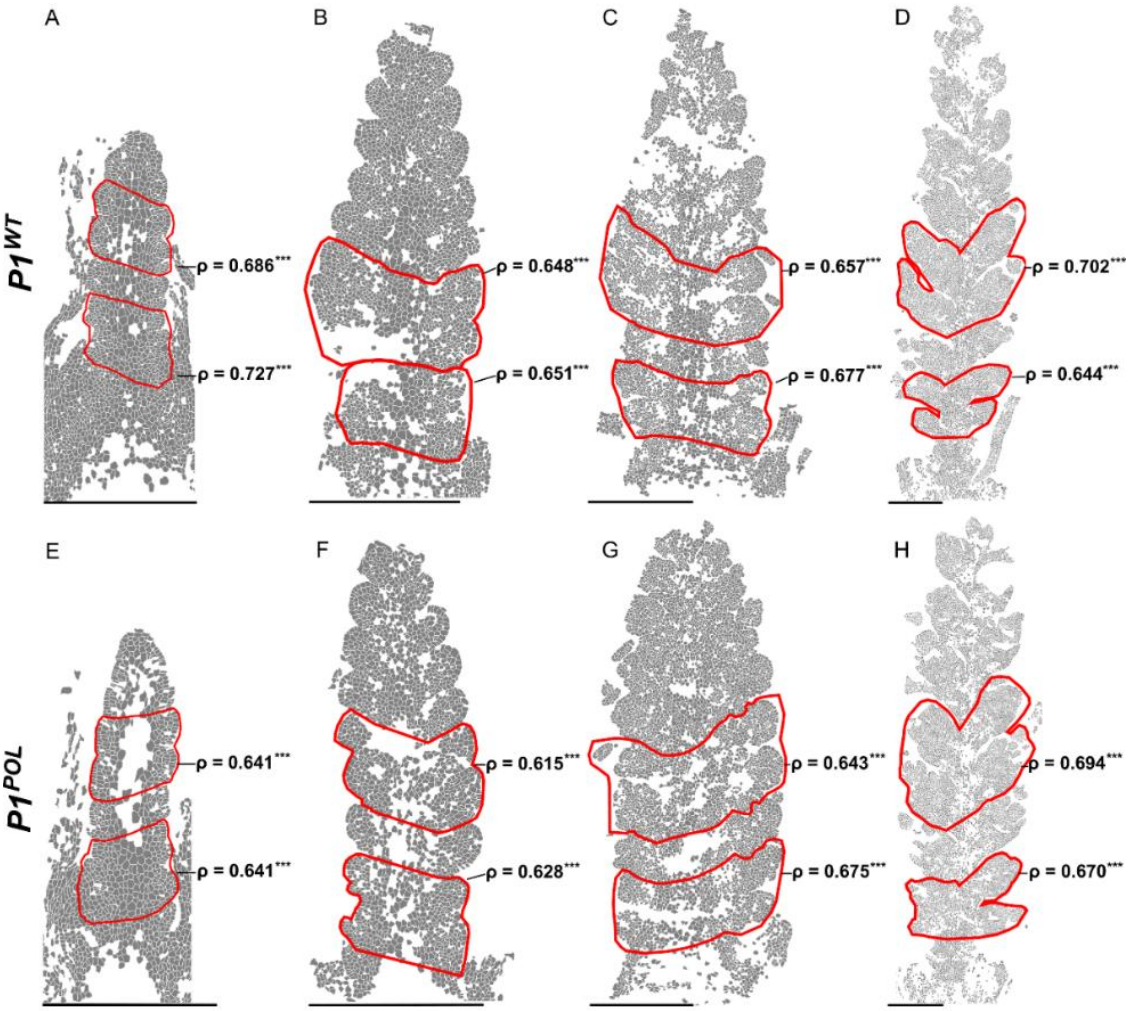


Figure 2.14- Normalised counts from *in silico* dissections of MERFISH regions show strong correlation with TPM expression values from corresponding RNA-seq microdissection regions

(A-H) Cell segmentations (grey) overlaid with in silico dissection regions (red), as selected using the MERSCOPE Visualizer tool. Spearman's rank correlation coefficients (ρ) are shown for each genotype, developmental stage, and section; significance is indicated (*** $P \leq 0.001$)

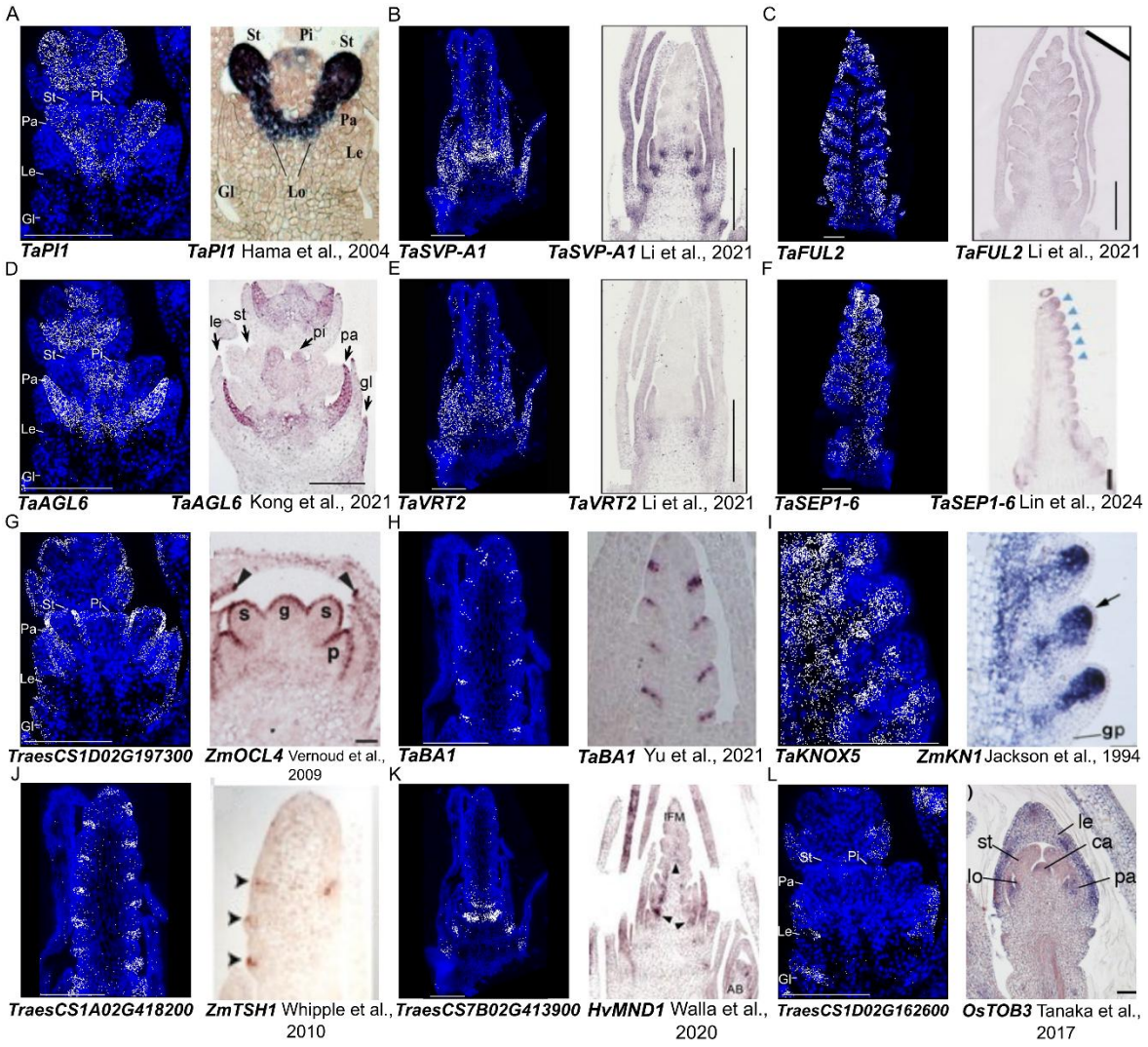


Figure 2.15 - Single gene *in situ* hybridisation in cereals from equivalent tissues and time points as those used for MERFISH

(A-L) Transcript localisation of wheat genes in MERFISH $P1^{WT}$ samples (left) compared to *in situ* hybridisation of wheat gene (*Ta*) or cereal ortholog (*Hv* = barley; *Os* = rice; *Zm* = maize) at equivalent inflorescence stage in published studies (right). Gl = glume, Le = lemma, Pa = palea, Lo = lodicule, St = stamen, Pi = Pistil.

2.3.7 - Inflorescence cell type markers verified with MERFISH data

We examined the spatial expression patterns of 36 putative cell-type marker genes from the 200-gene MERFISH panel. These genes were selected based on an analysis of an snRNA-seq dataset, which identified 20 transcriptionally distinct clusters and 7,211 significantly enriched marker genes across wheat inflorescence development (W2.5–W5). Cell-type annotations were assigned based on known gene expression patterns and functional enrichment (Liu et al. 2025).

To assess concordance between snRNA-seq–inferred cell types and spatial localisation, we compared MERFISH expression patterns of these 36 genes to their snRNA-seq annotations using broad tissue categories in the developing inflorescence (Figure 2.16A). Eighteen genes exhibited spatial patterns fully or partially consistent with their snRNA-seq–based annotations (Table 2.5). Most were enriched in multiple snRNA-seq clusters and showed precise MERFISH localisation in vasculature, epidermis, cortex/pith, proliferative zones, or floral tissues, with representative examples shown in Figure 2.16 (all gene expression patterns available at www.wheat-spatial.com).

The remaining 14 markers displayed less definitive patterns: one inferred ovary-specific gene (*TraesCS4B02G084800*) showed expression in vegetative and non-ovary floral tissues; four genes were broadly expressed without clear restriction; and nine showed low, stochastic signal likely reflecting low abundance, limited representation of the corresponding cell types, or suboptimal probe performance.

Further discussion of how MERFISH results inform and refine snRNA-seq cluster annotations is addressed in Liu et al. (2025). The major tissue domains mentioned (*vasculature*, *cortex/pith*, *florets*, and *epidermis*) are examined in further detail in Chapter 4. Overall, this analysis demonstrates that MERFISH can effectively validate and spatially resolve snRNA-seq–derived cell type markers, reinforcing its value as a high-throughput platform for spatially contextualising gene expression in developing plant tissues.

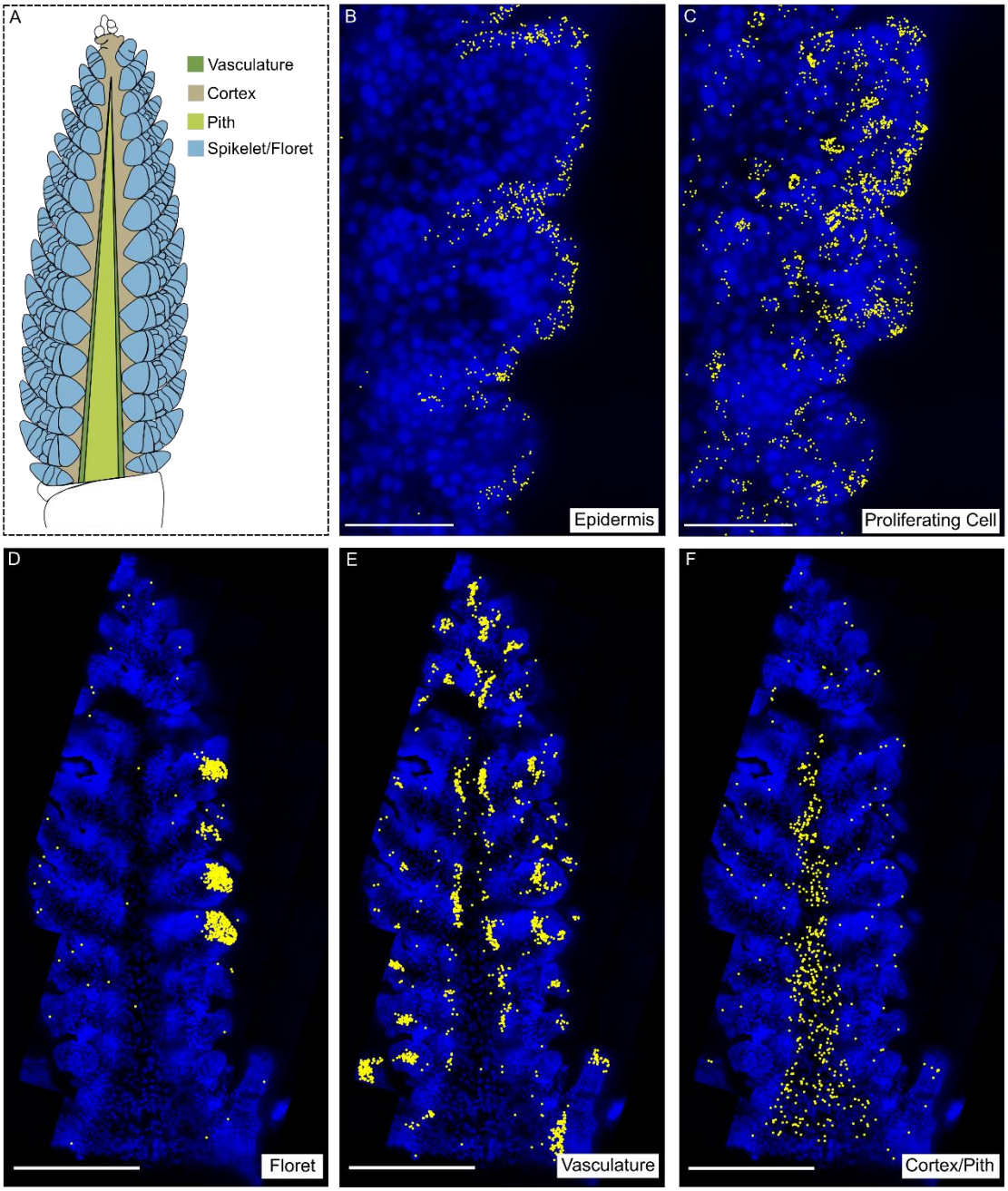


Figure 2.16 - Spatial validation of snRNA-seq–derived markers reveals tissue-specific expression patterns

A) Schematic showing the major tissue domains used for classification of MERFISH expression patterns in wheat inflorescence. Illustration adapted from original work by Xuemei Liu to represent a W4-stage spike. **(B-C)** Transcript detection in spikelet tissue (stage W3.25; $P1^{WT}$), **B)** *TraesCS4D02G296400* and **C)** *TraesCS2B02G260800*. Scale bar = 100µm. **(D-F)** Transcript detection in whole inflorescence tissue (stage W4 $P1^{WT}$), **D)** *TraesCS7D02G261600,* **E)** *TraesCS2A02G192600,* **F)** *TraesCS5B02G353200.* Scale bar = 500 µm. DAPI = blue. Yellow dots represent individual transcripts.

Table 2.5 - Comparison of snRNA-seq cell type annotations to MERFISH gene expression patterns

IWGSC RefSeq1.1 Gene ID	Gene Name	Cell Cluster, Liu et al. 2025	Cell Cluster Annotation, Liu et al. 2025	MERFISH Gene Expression Pattern	MERFISH agrees with snRNA-seq annotation?
TraesCS1A02G077800		R12, R17	Vasculature	Vasculature	Yes
TraesCS2A02G192600	*SHR*	R12, R14, R15, R17	Floret, Spikelet primordium, Vasculature	Vasculature, Spikelet (W3.25), Floret (W4, W5)	Yes
TraesCS5A02G230500		R12, R17	Vasculature	Vasculature	Yes
TraesCS5D02G385300		R12, R15, R17	Spikelet primordium, Vasculature	Vasculature, Spikelet (W3.25), Floret (W5)	Yes
TraesCS7D02G191600		R12, R15, R17, R5	Spikelet base, Spikelet primordium, Vasculature	Vasculature, Spikelet (W3.25, W4, W5)	Yes
TraesCS7A02G071700		R12, R15	Spikelet primordium, Vasculature	Vasculature (W5)	Partial
TraesCS1D02G343400		R14, R16, R3	Epidermis, Floret, Unknown	Epidermis	Yes
TraesCS6B02G251600		R10, R11, R14, R3	Epidermis, Floret, Glume/Lemma, Proliferating cell G1/S	Epidermis (W3.25, W4), Floret (W5)	Yes
TraesCS4D02G296400	*FDH*	R14, R16, R3	Epidermis, Floret, Unknown	Epidermis	Partial
TraesCS2B02G260800		R11, R14, R19, R4, R7	Floret, Ovary, Proliferating cell G1/S	Proliferating cell	Yes
TraesCS4D02G076900		R11, R19, R4, R7	Ovary, Proliferating cell G1/S	Proliferating cell	Yes

TraesCS6A02G259000	*AGL6*	R14	Floret	Floret (W4, W5)	Yes
TraesCS7B02G014500		R19	Ovary	Floret (W4, W5)	Yes
TraesCS7D02G261600	*SEP3-1*	R14, R19	Floret, Ovary	Floret (W4, W5)	Yes
TraesCS4B02G050200		R10, R14, R15, R19, R8	Floret, Glume/Lemma, Ovary, Pith, Spikelet primordium	Pith/Cortex (W3.25, W4), Floret (W4, W5), Vasculature (W5)	Yes
TraesCS1A02G156100		R10, R15, R17	Glume/Lemma, Spikelet primordium, Vasculature	Pith/Cortex (W3.25, W4), Vasculature (W5), Spikelet (W5)	Yes
TraesCS7D02G339600		R1, R13, R5	Cortex, Spikelet base	Pith/Cortex (W3.25, W4), Spikelet (W3.25, W4, W5)	Yes
TraesCS5A02G185600		R1, R13, R16, R18, R5	Cortex, Spikelet base, Unknown	Pith/Cortex (W3.25, W4), Spikelet (W3.25), Floret (W5), Leaf (W2.5)	Partial
TraesCS5B02G353200		R10, R13, R6, R8	Cortex, Glume/Lemma, Pith	Pith/Cortex (W2.5, W3.25, W4), Vasculature (W5)	Partial
TraesCS4B02G084800		R19	Ovary	Vasculature (W2.5-W5), Floret (W5)	No
TraesCS1D02G373800		R19, R4, R7	Ovary, Proliferating cell G1/S	Pattern unclear, low expression	No
TraesCS2A02G306800		R15	Spikelet primordium	Pattern unclear - low expression	No
TraesCS3A02G251500		R17	Vasculature	Pattern unclear - low expression	No
TraesCS3D02G401200	*nac6D*	R16, R18	Unknown	Pattern unclear - low expression	No
TraesCS5A02G098300	*HTA3*	R11, R19, R4, R7	Ovary, Proliferating cell G1/S	Pattern unclear - low expression	No

TraesCS5A02G165400		R10, R13, R6, R8	Cortex, Glume/Lemma, Pith	Pattern unclear - low expression	No
TraesCS5D02G136300	*TFL1*	R13, R15	Cortex, Spikelet primordium	Pattern unclear - low expression	No
TraesCS6A02G171800		R15, R17	Spikelet primordium, Vasculature	Pattern unclear - low expression	No
TraesCS7D02G008500		R17	Vasculature	Pattern unclear - low expression	No
TraesCS3A02G155200		R10, R12, R15, R17, R9	Glume/Lemma, Spikelet base, Spikelet primordium, Vasculature	Pattern unclear - expressed throughout sample	No
TraesCS5D02G133600		R11, R14, R19	Floret, Ovary, Proliferating cell G1/S	Pattern unclear - expressed throughout sample	No
TraesCS5D02G232900		R16, R18, R5	Spikelet base, Unknown	Pattern unclear - expressed throughout sample	No
TraesCS5D02G449200	*BHLH007*	R12, R15, R17	Spikelet primordium, Vasculature	Pattern unclear - expressed throughout sample	No

2.3.8 - MERFISH captures ectopic gene expression in *VRT-A2b* mutant

We next asked if MERFISH could quantify gene expression differences between genotypes. We used this to investigate how a *cis*-regulatory mutation in *VRT2* ($P1^{POL}$ allele), previously shown to increase *VRT2* expression (Adamski et al. 2021; Backhaus et al. 2022), alters its spatial expression domain. Prior genetic studies have demonstrated that the *VRT-A2b* allele leads to elevated *VRT2* expression, resulting in subtle developmental delays in basal spikelet formation and elongation of spikelet organs such as glumes and lemmas (Adamski et al. 2021; Backhaus et al. 2022). However, the spatial consequences of this elevated expression —specifically, whether it reflects ectopic expression or enhanced expression within native domains —have remained uncharacterised.

In $P1^{WT}$ inflorescence tissues, *VRT2* expression is localised to the vegetative tissues below the inflorescence, and the developing cortex during early developmental stages (W2.5, W3.25, W4), and becomes expressed primarily in the leaves and peduncle below the inflorescence by stage W5 (Figure 2.17A-D). In contrast, MERFISH revealed that in $P1^{POL}$ inflorescence tissues, *VRT2* is expressed ectopically, expanding its expression domain into spikelet and floral tissues, with this expansion most pronounced at stages W4 and W5 (Figure 2.17E-H). These results demonstrate the ability of MERFISH to detect tissue-specific changes in a developmental mutant.

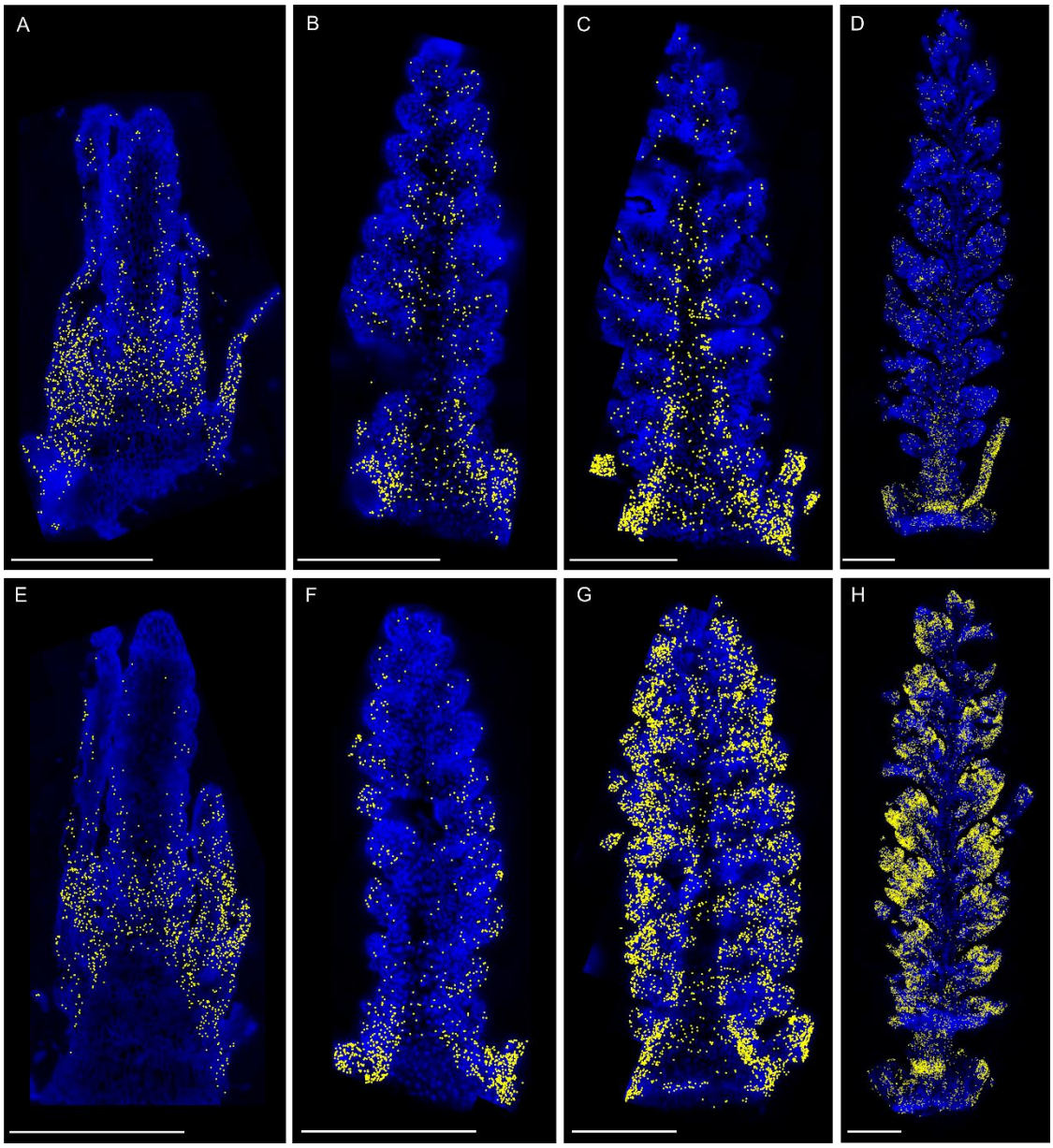


Figure 2.17- MERFISH detects ectopic expression of *VRT2* in the *P1^{POL}* genotype.

Transcript detection of *VRT2* (TraesCS7A02G175200) in genotypes *P1^{WT}* (A-D) and *P1^{POL}* (E-H). Stages W2.5 (A, E), W3.25 (B, F), W4 (C, G), W5 (D-H). Scale bar = 500 µm. DAPI = blue. Yellow dots represent individual transcripts.

2.4 - Discussion

2.4.1 – Development of an open-access spatial transcriptomic dataset with MERFISH

This chapter presents the development of a spatial transcriptomic dataset profiling the expression of 200 genes across four developmental stages of the wheat spike. To achieve this, we optimised and implemented Multiplexed Error-Robust Fluorescence *In Situ* Hybridisation (MERFISH) in wheat inflorescence tissue, representing the first implementation of MERFISH to a crop species.

Through MERFISH, we generated high-quality, reproducible, single-cell–resolution data. The gene expression patterns detected with MERFISH were consistent with expression patterns observed in single-gene *in situ* hybridisation, including *MND1* in barley (Walla et al. 2020), *KN1* in maize (Jackson et al. 1994), and *TOB3* in rice (Tanaka et al. 2017), underscoring its utility for cross-species comparisons of conserved developmental programs. In parallel, a related dataset of 99 genes was generated in tetraploid wheat using Molecular Cartography at developmental stages W1.5, W2.5, and W3.25 (Xu et al., 2025). Our panel included 34 overlapping genes, which showed strong concordance between the two platforms (detailed in Xu et al., 2025). This cross-validation across independent spatial transcriptomic approaches reinforces the robustness of the newly characterised spatial expression patterns.

As MERFISH is an emerging technique in plant sciences, we have established a workflow that spans from sample embedding to imaging, which can be applied to additional plant tissue types. (Open-source protocol available at Long et al. 2025). Together with an open-access resource for all generated data (doi.org/10.5281/zenodo.15720855), we aim to provide a foundation for future analyses and applications of imaging-based spatial transcriptomics in plants.

2.4.2 - Interactive web tool provides access to spatial transcriptomic data to the wider plant science community

A central goal in generating this dataset was to make it broadly accessible to the plant science community. Spatial transcriptomic datasets can be challenging to explore due to their large file sizes, complex formats, and, in the case of MERFISH, the need for specialised software to open the final data outputs. To overcome these barriers and maximise research impact, we developed an interactive online resource (www.wheat-spatial.com) that enables the exploration of the wheat inflorescence MERFISH dataset.

The platform is built on the open-source WebAtlas framework (Li et al. 2024), which supports user-friendly, browser-based navigation of large spatial datasets. Users can visualise transcript counts at single-cell resolution, expression domain assignments, and compare expression profiles across developmental stages without requiring local installation of MERFISH-specific software (Li et al. 2024). This resource provides an openly accessible reference for the wheat research community and a template for sharing spatial transcriptomic data in other plant species.

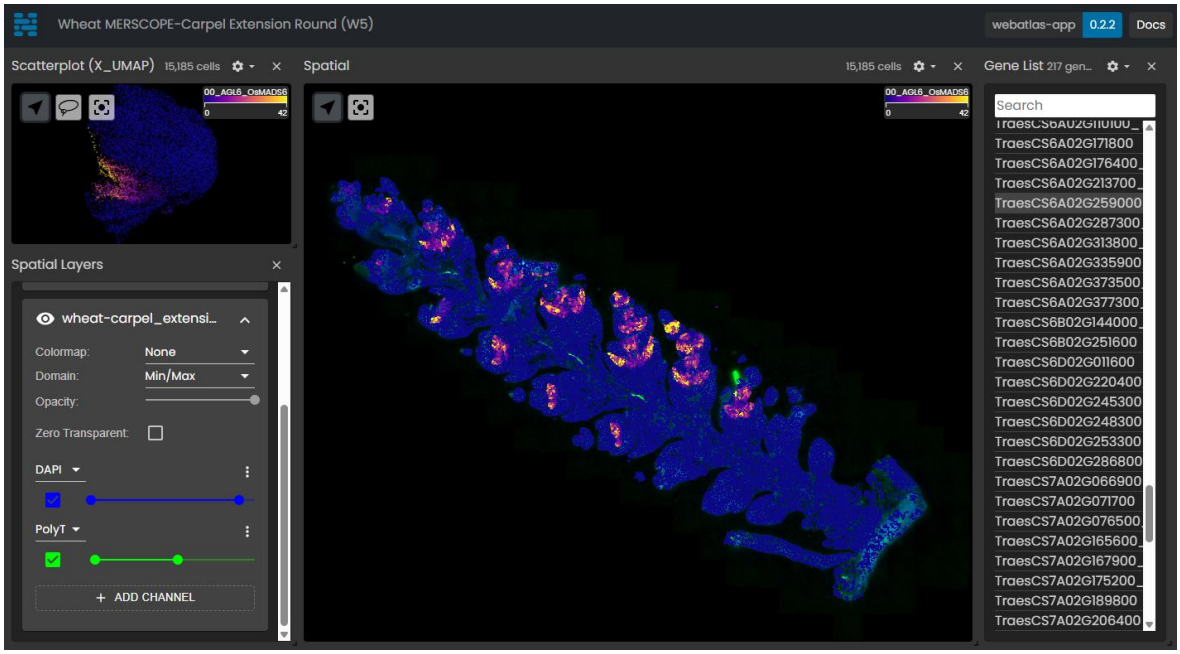


Figure 2.18 - www.wheat-spatial.com visualises cell-level gene expression information in an interactive web browser.

2.4.3 - Limitations of MERFISH technology

Both spatial transcriptomics and single-cell RNA sequencing technologies are emerging as powerful tools in plant biology (Nobori 2025). The wide variety of technologies on the market continues to expand, each presenting its own benefits and limitations. To discuss some of the limitations of MERFISH, I present questions I commonly encounter when presenting and discussing this work:

Can MERFISH easily be adapted to other plant tissue types?

Imaging-based spatial transcriptomic methods have been successfully applied to a wide range of plant tissues, demonstrating their adaptability to diverse tissue compositions (Nobori 2025). Notably, imaging-based techniques offer distinct advantages over scRNA-seq, particularly in plant systems. Unlike scRNA-seq, spatial transcriptomic approaches do not require protoplast isolation, enabling analysis of tissue types that are challenging to optimise for scRNA-seq due to factors such as rigid cell walls, large cell size, or the presence of secondary metabolites (Giacomello 2021).

Nonetheless, adapting MERFISH for plant tissues presents specific technical challenges. High and variable levels of autofluorescence across tissue types can interfere with probe detection, necessitating optimisation of clearing protocols and cell wall digestion (Duncan et al. 2022). While imaging-based techniques have been most commonly applied to relatively soft tissues Laureyns et al. 2021; inflorescences, Lee et al. 2025; Xu et al. 2025), further protocol refinement is required for tissues with substantial secondary cell wall deposition. Permeabilisation must balance probe accessibility with structural preservation; over-digestion can damage tissue architecture or cause detachment from slides, compromising imaging quality (Giacomello 2021).

Cryosectioning introduces additional complexity. In thin tissues, such as *Arabidopsis* roots, producing consistent longitudinal sections is technically challenging, making whole-mount strategies (PHYTOMap; Nobori et al. 2023) advantageous. In contrast, thicker tissues such as wheat inflorescences are more straightforward to section but present different obstacles: their intricate three-dimensional arrangement means that the sectioning angle strongly influences the anatomical regions captured. Achieving consistent sectioning angles across all samples and interpreting the anatomical context of resulting sections were significant challenges in our MERFISH experiments.

Will probe binding efficiency change vary between probes?

Probe binding efficiency in fluorescence *in situ* hybridisation (FISH)–based techniques, including MERFISH, is inherently variable. Differences arise from both probe design and the biological context of the target transcript. Factors such as partial RNA degradation, secondary structure, or protein binding can reduce hybridisation efficiency and lower fluorescence intensity to a level indistinguishable from background signal (Duncan et al. 2022).

Quantifying the proportion of poorly performing or "stochastic" probes in plant spatial transcriptomic studies remains challenging. However, an ISS study of the maize shoot apex documents the efficacy of their gene probe performance. 15 of 90 probes were classified as stochastic based on a threshold of <15 total counts detected (Laureyns et al., 2021). In our dataset, comparing counts per cell for blank barcodes and gene-targeting probes across developmental stages identified eight genes (out of 200) whose signals were indistinguishable from background noise, suggesting that they do not detect the transcript of interest at a biologically meaningful level. This highlights variability in probe performance, with a subset of probes failing to capture detectable expression.

Across species and tissue types, MERFISH generally shows a strong correlation with RNA-seq datasets (Chen et al. 2015; Wang et al. 2018; Liu et al. 2022; Choi et al. 2023; Cisar et al. 2023). However, transcript abundance measured by MERFISH does not always scale proportionally with expression values from RNA-seq. In a comparison of MERFISH counts to bulk RNA-seq counts in mouse liver and kidney, MERFISH detected systematically higher transcript counts, with fold-change increases ranging from ~10× to >1000×, indicating that relative gene abundance estimates do not scale uniformly across all targets (Liu et al. 2022).

Is this method truly quantitative? Could I use it for differential gene expression across different samples?

FISH–based methods, including MERFISH, are inherently quantitative: each fluorescent spot corresponds to a single RNA molecule, allowing for the absolute count of transcripts within individual cells (Duncan et al. 2022). Owing to its high sensitivity and specificity, smFISH is considered the gold standard for validating rare or variable gene expression and is often used to complement scRNA-seq, which captures only a fraction of the transcriptome and exhibits substantial cell-to-cell variability (Torre et al. 2018; Zhang et al. 2020; Duncan et al. 2022).

However, interpreting MERFISH data in the context of differential gene expression (DGE), particularly across biological samples, requires careful consideration of several technical factors. In our dataset, we detected a median of 85.7 transcripts per cell (range: 61–118) and 34.5 genes per cell (range: 27–41), comparable to a recent *Arabidopsis* MERFISH study using a 500-gene panel (median: 161 transcripts, 79 genes per cell; Nobori et al., 2025). Given our more targeted 200-gene panel, these values suggest that our sample quality and capture efficiency are comparable to those of other plant MERFISH datasets. However, total transcript counts varied between replicates and within tissue sections. For example, in two W5 stage replicates, the wild type ($P1^{WT}$) had a median of 188 transcripts and 118 genes per cell, while the ($P1^{POL}$) mutant had 103 transcripts and 84 genes per cell. In one W2.5 sample, high-quality cells ranged from 25 to over 600 transcripts. Such variability underscores the need for normalisation when comparing samples, and caution when interpreting DGE from MERFISH data alone.

Despite these caveats, MERFISH is powerful for validating and spatially contextualising DGE identified by other transcriptomic approaches. In our study, MERFISH confirmed elevated *VRT2* expression in the $P1^{POL}$ mutant (Adamski et al., 2021; Backhaus et al., 2022) and revealed ectopic expression in spikelet and floret tissues, providing spatial resolution to the regulatory effect of the *cis*-regulatory mutation. Similarly, Nobori et al. (2025) utilised MERFISH in Arabidopsis leaves to detect pathogen-induced transcriptional differences between two immune cell populations—an analysis that would not have been possible without single-cell spatial resolution.

Ultimately, whether MERFISH is suitable for DGE depends on the specific biological question being asked. Because MERFISH relies on a pre-designed gene panel, it is often used as a follow-up to transcriptomic studies, such as bulk RNA-seq or single-cell RNA-

seq. For instance, if prior transcriptomic data indicate that a gene is differentially expressed across conditions or genotypes, MERFISH can reveal whether an altered spatial distribution, such as emergence in a new tissue domain or cell type, a shift in spatial expression boundaries, or an association with cell identity transitions accompanies this change.

In summary, while MERFISH is quantitative and capable of supporting differential expression analyses, its real strength lies in resolving where changes in expression occur within the tissue. When integrated with broader transcriptomic datasets, it becomes a powerful tool for uncovering the spatial context behind gene regulation.

Does the curation of a gene panel incorporate bias into results?

In contrast to sequencing-based spatial transcriptomics, which can capture the transcriptome in an untargeted manner, FISH-based approaches, such as MERFISH, require a predefined set of target genes, inherently limiting the experiment to a small fraction of transcripts (Giacomello 2021). Panel design typically relies on prior transcriptomic data to select genes based on criteria like expression abundance, differential expression (across time, treatments, or cell types), and prior genetic characterisation. This process inevitably introduces selection bias, potentially excluding genes that are uncharacterized or exhibit complex expression patterns (Covert et al. 2023).

In our implementation of MERFISH, the 200-gene panel encompassed four broad categories: (1) genes differentially expressed along the apical–basal axis based on bulk RNA-seq, (2) characterised inflorescence development genes from the literature, (3) putative cell-type markers from a wheat snRNA-seq dataset, and (4) housekeeping and control genes. While this design introduces bias and may overlook novel or unexpected expression patterns, MERFISH's ability to resolve transcript localisation at single-cell resolution provides a powerful advantage over untargeted sequencing-based spatial methods, particularly for hypothesis-driven studies.

2.4.4 - Future perspectives

The application of spatial transcriptomics offers clear advantages over bulk RNA-seq and snRNA-seq for developmental biology. By preserving positional context and detecting transcript localisation to a single-cell resolution, these approaches can be applied to the precise delineation of sharp boundaries between developmental domains, reveal gradual expression gradients along developmental axes, and quantify cell–type–specific co-expression patterns *in situ*. Such spatially resolved insights can be used to understand how transcriptional programs coordinate tissue differentiation and organ patterning and can also aid in comparing tissue patterning between plant species.

Recent developments underscore the importance of integrating spatial transcriptomics with scRNA-seq or snRNA-seq to connect high-resolution spatial maps with broader transcriptomic coverage (Demesa-Arevalo et al. 2025; Nobori et al. 2025; Xu et al. 2025). Future integrative strategies, such as combining FISH-based techniques with protein co-detection (Tao et al. 2023) or co-applying MERFISH, ATAC-seq, and snRNA-seq (Nobori et al. 2025), holds promise for capturing transcriptional, epigenetic, and post-transcriptional regulation within its native spatial context, offering a more comprehensive understanding of gene regulatory networks in plant development.

As demonstrated in this chapter, MERFISH offers a unique opportunity to investigate developmental biology with single-cell spatial resolution, enabling precise mapping of transcriptional programs within complex plant tissues. My next chapter explores in greater depth how these data can be analysed for biological insights, revealing new perspectives on the spatio-temporal dynamics of wheat inflorescence development.

2.5 – Methods

2.5.1 - Plant materials and growth conditions

We used two BC_6 near-isogenic lines (NIL) differing for *VRT-A2* alleles in a hexaploid wheat (cv Paragon) background. One NIL carried the wildtype Paragon *VRT-A2a* allele, here named $P1^{WT}$, whereas the second NIL carried the *VRT-A2b* allele from *Triticum turgidum* ssp. *polonicum* (named $P1^{POL}$; Adamski et al. 2021) Plants were grown under a 16/8 h light/dark cycle at 20/15 °C, 65% relative humidity and bottom-watering irrigation (Simmonds et al. 2024).

2.5.2 - Tissue dissections and sample preparation

The VRT-A2a NIL was used for semi-spatial RNA-seq, whereas both NILs were used for MERFISH. For semi-spatial RNA-seq, we used a published dissection methodology (Faci et al. 2024) to produce basal and central/apical sections. At the Early Double Ridge stage (EDR, W2), spikes were bisected. In contrast, for the Late Double Ridge (LDR, W2.5), Lemma Primordia (LP, W3.25), Terminal Spikelet (TS, W4) and Carpel Extension (CE, W5) stages, the basal section consisted of the most basal four spikelets from each spike. Two spikelets were skipped, then the subsequent four spikelets were harvested to comprise the central section (Fig. 1A). Samples were stored at -70 °C until RNA was extracted from the pooled micro dissected spikes using Qiagen RNeasy Plant Mini and Zymo Direct-zol RNA Microprep kits as described in the manufacturer's manual. Total RNA (~1 µg) was sent to Novogene UK for PCR-free library preparation and Illumina sequencing (PE150; 50M reads per sample).

For MERFISH, we used a similar dissection protocol (Faci et al. 2024) but retained the youngest leaves surrounding meristems (Supplementary Fig. 2). After dissection, meristems were transferred using an RNase-free pipette tip into 4% paraformaldehyde (PFA) in 1× PBS (prepared from 6% formaldehyde [w/v], methanol-free; Pierce 28906) in 2 mL RNase-free Eppendorf tubes. Samples were vacuum infiltrated for 10 minutes or until tissue sank and incubated overnight at 4 ºC. The PFA solution was removed, and the samples were washed three times with 1× PBS. Tissue was then immersed in 15% sucrose in 1× PBS at 4 °C for 6 hours, followed by immersion in 30% sucrose in 1× PBS at 4 °C overnight.

2.5.3 - Analysis of semi-spatial RNA-seq data

We trimmed raw reads with cutadapt (v1.9.1, Martin 2011) and generated read counts and transcripts per million (TPM) values using Kallisto pseudo-alignment (v0.44.0, Bray et al. 2016) for all genes in the IWGSC RefSeq v1.1 annotation (IWGSC 2014). We conducted subsequent analyses for high-confidence gene models with non-zero counts in at least one sample. We transformed read counts (rlog function; DESeq2 v1.34.0; Love et al. 2014), and performed principal component (PC) analysis with prcomp (R Core Team 2018). We calculated differential expression ($p < 0.001$; Benjamini-Hochberg corrected) between central and basal sections across time using ImpulseDE2 (v3.6.1; Fischer et al. 2018), on genes with average > 0.5 TPM for at least one stage-section combination. We clustered the 12,384 differentially expressed genes using k-means (k1:10) and displayed with pheatmap (v1.0.12; Kolde 2019). RNA-seq sample metadata, TPM values, and ImpulseDE2 output, see archive folder 'RNAseq.tar.gz' at doi.org/10.5281/zenodo.14515926.

2.5.4 - Gene panel selection and design for MERFISH

We designed a 300-gene panel for MERFISH, comprising 200 genes associated with spike development, and 100 genes from a separate wheat grain project, which are not described here. We removed genes that could not accommodate at least 25 specific probes, based on Vizgen's probe design software, except for three genes targeted by between 20 and 25 probes. MERFISH probes were designed and synthesised by Vizgen.

2.5.5 - Meristem embedding and sectioning

We cleaned all surfaces and dissection tools with RNABlitz before use. We marked a 1 cm × 1 cm area on the back of a Tissue-Tek mold (25 × 20 × 5 mm; Thermo Fisher, AGG4580) and filled with Tissue-Plus OCT compound (Agar Scientific, AGR1180). We also filled a 60 mm Petri dish with OCT. We removed individual meristems from the 30% sucrose solution using clean dissection tools, aided by a drop of OCT on the tool to adhere the meristems during collection. We transferred meristems to the OCT-filled Petri dish, where they were mixed with OCT to remove residual sucrose and ensure complete coating. Using a stereomicroscope (Leica S9 with an HXCAM HiChrome HR4 Lite camera and a Photonic Optics light source), we inspected meristems for air bubbles, which were carefully

removed with a fine dissection tool. We trimmed excess vegetative tissue as needed. Meristems were then placed into the OCT-filled Tissue-Tek mold, arranged within the marked 1 cm^2 region according to genotype and developmental stage. Each OCT block contained 5–36 meristems, depending on the developmental stage and we imaged them using GX Capture-T. The OCT blocks were flash-frozen and stored at −70 °C.

We performed sectioning using a Leica CryoStar NX70. All inside surfaces and tools were cleaned with Blitz RNase Spray (Severn Biotech Ltd, 40-1735-05), and a fresh blade (MX35 Ultra™ Microtome Blade, 3053835) was used. We set the chuck temperature to −20 °C, and the blade temperature to −18 °C. We pre-warmed samples at the back of the cryostat for 30 minutes before sectioning, and brought the MERSCOPE Slides (Vizgen, 20400001) to room temperature. We trimmed OCT blocks to remove excess OCT, mounted to the chuck, and further trimmed until tissue was exposed; 10 µm sections were cut to inspect tissue regions on glass slides. Once we identified the region of interest at the optimal depth and angle, 10 µm sections were flattened with paintbrushes, flipped, and mounted onto room-temperature MERSCOPE slides, following the placement and technique outlined in the MERSCOPE user guide. After mounting, we placed the slides in 60 mm Petri dishes and incubated at the back of the cryostat for 30 minutes. We then fixed the slides in 4% methanol-free PFA in 1× PBS for 10 minutes. We washed the slides three times with 1× PBS, incubating for 5 minutes per wash. We aspirated residual PBS and air-dried the slides for 1 hour in a cell culture hood with the Petri dish lid closed. We then incubated the slides with 5 mL of 70% ethanol prepared in RNase-free water. Petri dishes were sealed with parafilm and stored at 4 °C, either overnight or for up to 7 days. For more details, see dx.doi.org/10.17504/protocols.io.rm7vzqwb4vx1/v1, or Appendix 3.

2.5.6 – Vizgen MERSCOPE workflow

Slide preparation for the Vizgen verification kit was followed using 91600004_MERSCOPE Sample Verification Kit User Guide_Rev D and 91600002_MERSCOPE Fresh and Fixed Frozen Tissue Sample Preparation User Guide_Rev E (Vizgen), following guidelines for non-resistant fixed frozen tissue clearing. Careful cleaning using both 70% ethanol and then RNAseZAP (Invitrogen, AM9782) or Blitz RNase Spray (Severn Biotech Ltd, 40-1735-05) was conducted on gloves, tweezers, scalpels, workbenches, incubators and hybridisation box (Brabantia, 203480). Slides were prepared using Vizgen sample preparation kit (Vizgen, 10400012). Autofluorescence levels were assessed using a EVOS FL2 Auto (Invitrogen by

Thermo Fisher Scientific, AMAFD2000), and found to be low at 10X. Slide was bleached for 3 hours and then reassessed using the microscope. Custom verification kit (Vizgen, 10400124, unique ID ZM1147) was applied and incubated at 37 °C for 24 hours. Nuclear and cytoplasmic staining were performed using a Primary Staining Solution (DAPI and PolyT, see Guide 91600002). After staining, checks for efficiency were also made. Clearing at 47 °C for 1 day with clearing buffer (Vizgen) with proteinase K (NEB, P81070S) addition. A further day of clearing at 37 °C, without proteinase K was performed. Care was taken to minimise smears and lint on the slide before running, by cleaning with 80% ethanol and lens cleaning tissue (2105-841, Whatman).

For a full MERSCOPE experimental run with a 300-gene panel, we performed slide preparation following guidelines for non-resistant fixed frozen tissue clearing (91600002_MERSCOPE Fresh and Fixed Frozen Tissue Sample Preparation User Guide_Rev E (Vizgen)). We prepared slides using Vizgen sample preparation kit (Vizgen, 10400012) with all instruments (including hybridisation box (Brabantia, 203480)) cleaned using both 70% ethanol and then RNAseZAP (Invitrogen, AM9782) or Blitz RNase Spray (Severn Biotech Ltd, 40-1735-05). We performed checks for autofluorescence at 10X using an EVOS FL2 microscope under a DAPI light cube, recording light intensity levels to decide on the reduction of autofluorescence before and after photobleaching (performed for between 3 and 8 hours in EtOH 70%, Vizgen 10100003). A 300-gene probe set (Vizgen product number 20300007) was applied and hybridised for 48 hours. On the days of a run, we re-checked autofluorescence levels and topped up using the photobleacher for 3 hours if necessary. After DAPI staining, we also made checks for the efficiency of staining. Clearing times varied depending on run slots. A standard clearing at 47 °C for 1 day with clearing buffer (Vizgen) containing proteinase K (NEB, P81070S) addition was always performed, whereas additional days (1 to 4 days) of clearing at 37 °C without proteinase K in the buffer were performed. Tissue is never fully cleared by eye, nor when using a light microscope, before MERSCOPE runs.

Upon imaging, care was taken to minimise smears and lint on the slide by cleaning with 80% ethanol and lens cleaning tissue (2105-841, Whatman). We outlined regions of interest around individual spikes on the slide overview using DAPI staining. Following 60x imaging, we decoded transcripts using the panel-specific MERSCOPE Codebook as detailed in the Introduction (Figure 2.3). We processed raw data with the MERSCOPE Instrument Software to generate and output file structures as described in the MERSCOPE instrument User Guide.

2.5.7 - Cell segmentation and processing

We performed cell segmentation on stitched images of DAPI and PolyT staining. Prior to segmentation and to minimise error, we lightened seam lines in the stitched images using FIJI (v 1.54f; Schindelin et al. 2012). Dark stitching lines were processed using the following steps: 'Process>Filters>Maximum' (radius = 2 pixels, applied twice) and 'Process>Filters>Median' (radius = 2 pixels, applied twice). We then applied three times across the entire image the following step: 'Process>Filters>Gaussian Blur' (radius = 4 pixels). Image edits and segmentation results are exemplified in Figure 2.8.

We performed cell segmentation and transcript assignment using the Vizgen Post-Processing Tools (v1.2.2; Wiggin and Yu 2024), within a Python virtual environment on Ubuntu 20.04. We used the Cellpose2 cyto2 model (Pachitariu and Stringer 2022) with DAPI (blue channel) as the nuclear marker and PolyT (green channel) as the cytoplasmic marker. For segmentation parameters, see logs (doi.org/10.5281/zenodo.14515926). We exported segmentation results as polygon geometry in both mosaic and micron space, assigned transcripts to cell boundaries using the partition-transcripts function in VPT, and generated cell metadata with the derive-entity-metadata function. Finally, we integrated cell boundaries into existing .vzg files for visualisation in the Vizgen MERSCOPE Visualizer Tool (MERSCOPE Vizualizer 2023) using the update-vzg function. All implementation scripts are available (https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics) with example commands. For a gene x cell matrix for all samples, see archive folder 'QC.tar.gz' at doi.org/10.5281/zenodo.14515926.

2.5.8 – smFISH spot detection

Regions of interest for smFISH spot detection were extracted from smFISH channel 1 and smFISH channel 2 (laser channels 647nm and 561nm, respectively) high-resolution TIFF mosaics using tifffile (v2024.5.10; Gohlke 2024). First, we used the affine transformation matrix to convert user-defined tile centre coordinates (in microns) to pixel coordinates with NumPy (v1.26.3, Harris et al. 2020). The specified tile width and height defined a rectangular crop area, which was clipped to image bounds and read directly from the high-resolution mosaic .tiff file (zarr v2.16.1). Crops were returned as 16-bit grayscale arrays,

and percentile-based intensity scaling was applied using NumPy and tifffile to generate display-ready TIFFs.

Spot detection and co-localisation was quantified with FISH-quant, using the python-based analysis package big-FISH (v0.6.2; Imbert et al. 2022). First, we processed each smFISH channel separately using the detection.detect_spots() function with a voxel size of 108 nm (X,Y) and an expected spot radius of 300 nm. The spot radius was chosen based on the anticipated size of a diffraction-limited smFISH spot (200-300nm; Lyubiomova et al. 2013). An initial automatic detection threshold was estimated for each channel and subsequently increased 2-fold to reduce false positives. Subpixel localisation of detected spots was refined using function detection.fit_subpixel(). Colocalization between channels was quantified using multistack.detect_spots_colocalization(), with a maximum pairing distance threshold of 324 nm (three pixels). The function returned the coordinates of colocalised spots for each channel, inter-spot distances, and the index positions of matched spots. The total number of detected spots and the proportion of colocalised spots were calculated separately for each channel. Plots were displayed with BIG-FISH function plot_detection().

2.5.9 - Quality checks and filtering

Raw stain images were imported with tifffile (v2024.5.10) and an affine transform matrix storing imaging metadata was read in with pandas (v2.2.0; McKinney 2010; The Pandas Development Team 2024) and converted to a NumPy array (read_affine_matrix; v1.26.3, Harris et al. 2020). The transformation matrix was converted to micron space and applied the raw image using Scikit-image (v0.23.3, Walt et al. 2014). After transformation, the image was normalised, and image contrast and brightness were adjusted with Scikit-image (v 0.23.2; Walt et al. 2014). The enhanced image was rendered as a blue-channel RGB for visualization. Tissue masks were generated from the blue channel with Scikit-image functions Otsu thresholding (skimage.filters.threshold_otsu), morphological closing and small-object removal (skimage.morphology.binary_closing; remove_small_objects). Transcript coordinates were loaded, filtered to include the 200-spike development and control genes, and classified as in-tissue by indexing the binary mask. Mask area (mm^2) was computed from the number of true pixels multiplied by the squared pixel size, and transcript density was reported per specified area unit (mm^2). For each sample, the

workflow returned the number of transcripts within the mask, mask area and density metrics.

We loaded the spatial transcriptomic data from the eight samples into AnnData objects (anndata v0.10.7; Virshup et al. 2021) and processed using Squidpy (v1.4.1; Palla et al. 2022) and Scanpy (v1.10.0; Wolf et al. 2018). We filtered expression data to include the 200 spike development and control genes, selected cells from a single inflorescence within the imaged area and excluded low-quality cell segmentations based on volume (>500 pixels) and transcript count (>25 counts). For each sample, we calculated quality control (QC) metrics (total counts per cell, number of genes detected per cell, percentage of counts from top-expressed genes).

We assessed off-target hybridisation using total counts per cell in 15 blank probes, and across wheat homoeologs. We normalised expression data for all samples using scanpy functions sc.pp.normalize_total and sc.pp.log1p (v1.10.0, Wolf et al. 2018). We identified low-detecting probes by pooling the normalised counts per cell of blank probes to define a background distribution of per-cell counts associated with stochastic detection. Each of the 200 gene probes was then compared to this distribution using a two-sample Kolmogorov–Smirnov (KS) test (SciPy v1.13.0; Virtanen et al. 2020). Probes that were not significantly different from the blank distribution ($p > 0.1$) were classified as low-confidence, unless they exhibited high expression (log1p-normalised value > 1.5) in at least 25 cells. Supplementary information including QC metrics per sample and total blank counts per sample, can be accessed in archive folder 'QC.tar.gz' at doi.org/10.5281/zenodo.14515926.

As a further quality control metric, we performed *in silico* spike dissections equivalent to those captured by physical microdissection for 8 high-quality MERFISH sections using the MERSCOPE Visualiser 'draw ROI polygon' tool (MERSCOPE Vizualizer 2023). We extracted cell ids within each selected region and summated the total normalised counts, yielding transcript count information on 200 genes total. We used Spearman's rank correlation coefficients between these values and the mean transcript per million (TPM) values from the relevant genotype-section-stage combination of the semi-spatial RNA-seq data (Supplementary Fig. 7). For information of ROI area and correlation results, see archive folder QC.tar.gz' at doi.org/10.5281/zenodo.14515926.

Additionally, we identified *in situ* hybridisation results in wheat, barley, rice, and maize from equivalent tissues and time points as those used for MERFISH and visualised them in

side-by-side comparisons as described in 'Staining, Segmentation, and Transcript Visualisation'.

2.5.10 - Staining, segmentation, and transcript visualisation

We processed the cell segmentation data as GeoDataFrames (Geopandas v0.14.4; Jordahl et al. 2020), and converted the transcript coordinates into a GeoDataFrame from global x and y coordinates. We performed a spatial join operation to assign transcripts to segmented cells, retaining only transcripts located within cell boundaries. We loaded the DAPI staining image as a .tiff file, alongside a transformation matrix enabling conversion between pixel space to physical (micron) space, as generated by the MERSCOPE Instrument Software. The transformation matrix converted to micron space and applied to the raw image using Scikit-image (v0.23.3, Walt et al. 2014). After transformation, the image was normalised, and image contrast and brightness were adjusted with Scikit-image. We next rotated segmented cell polygons and transcript coordinates using NumPy (v1.26.3, Harris et al. 2020), and visualised cell geometries as polygons using Matplotlib (v3.8.2, Hunter 2007) with polygon handling and transformations facilitated by Shapely (v2.0.4, Gillies et al. 2022). Transcripts were overlaid as point features. The corresponding image was rotated with Scipy.ngimage (Virtanen et al. 2020; v1.13.0). Full details in implementation scripts, see https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics.

Chapter 3 – MERFISH reveals the spatio-temporal dynamics of wheat inflorescence development

This chapter includes results previously published in the following manuscript:

Spatial Transcriptomics Reveals Expression Gradients in Developing Wheat Inflorescences at Cellular Resolution

Katie A. Long, Ashleigh Lister, Maximillian R. W. Jones, Nikolai M. Adamski, Rob E. Ellis, Carole Chedid, Sophie J. Carpenter, Xuemei Liu, Anna E. Backhaus, Andrew Goldson, Vanda Knitlhoffer, Yuanrong Pei, Martin Vickers, Burkhard Steuernagel, Gemy G. Kaithakottil, Jun Xiao, Wilfried Haerty, Iain C. Macaulay, Cristobal Uauy

bioRxiv 2024.12.19.629411; https://doi.org/10.1101/2024.12.19.629411 (See Appendix 1)

This work was conducted in collaboration with Ashleigh Lister (Earlham Institute) and Dr Nikolai Adamski (JIC). Ashleigh performed the MERSCOPE workflow for all steps following cryosectioning. Nikolai assisted in conducting a literature review of proposed ABCDE model genes in wheat and related grass species, as well as interpreting the findings of floral gene expression patterns in the MERFISH dataset. I was assisted in the movement and storage of Vizgen data by Dr Burkhard Steuernagel (JIC), Dr Martin Vickers (JIC), and Gemy Kaithakottil (Earlham). The use of the MERSCOPE instrument was made possible through the Transformative Genomic platform (Earlham Institute), with contributions from Dr Iain Macaulay, Andrew Goldson, Vanda Knitlhoffer, and Ashleigh Lister. I am grateful to Dr Carole Chedid (Vizgen) for consultation on this project. I wish to thank the staff members of Vizgen for their technical support throughout this project. JIC Horticultural Services facilitated plant growth.

3.1 - Chapter summary

In this chapter, I demonstrate how MERFISH can be applied to investigate the spatio-temporal regulation of spikelet and floral development in wheat. I first analysed spatial transcriptomic data by clustering ~50,000 cells into 18 expression domains, followed by the identification of domain-enriched genes that serve as tissue-level markers. Using these domain- and cell-level maps, I examine gene co-expression relationships within spikelet and floral tissues, including shifts in gene expression observed in the regulatory mutant *P1POL* NILs. I also investigate genes differentially expressed along the apical–basal axis, identifying distinct and spatially coordinated patterns that distinguish axillary meristems and subtending leaf ridges prior to visible spikelet formation across the apical-basal axis. Together, these analyses reveal novel regulatory factors that pattern meristem identity and guide the transition to spikelet development.

3.2 - Introduction

3.2.1 - MERFISH as a tool for candidate gene validation and hypothesis generation

Spatial transcriptomic approaches offer a powerful platform for gene function. Over recent years, their use has expanded considerably, complementing traditional transcriptomic analyses, such as bulk RNA-seq and single-cell RNA-seq (scRNA-seq), which primarily identify genes through differential expression across tissues or conditions. Spatial techniques, such as MERFISH, enable researchers to go a step further by mapping the distribution of these genes within intact tissue (example studies reviewed in Nobori, 2025), providing additional context on gene function at a scale not previously attainable through single-gene *in situ* hybridisation.

However, the utility of MERFISH extends well beyond candidate gene validation and can be harnessed as a platform for hypothesis generation in plant developmental biology. Rather than focusing exclusively on individual genes, MERFISH datasets can be interrogated at the scale of an entire gene panel, enabling analyses that draw on bioinformatic approaches commonly applied in scRNA-seq (Wolf et al. 2018; Palla et al. 2022). These methods allow the identification of groups of cells with transcriptionally similar profiles, thereby reframing the analysis from the spatial context of a single gene to the broader context of cells and tissues. Within this framework, cells can be organised into gene expression “domains” - clusters of cells that share a common transcriptional state (Shi et al. 2023b). Importantly, such domains can be defined across all samples in a dataset, offering a powerful basis for comparing transcriptional states across tissues, developmental stages, and genotypes (Hie et al. 2019). In this chapter, I apply this approach to explore how tissue identity is established and maintained over developmental time, to investigate tissue-specific patterns of gene co-expression, and to assess how positional context influences cell fate decisions within the developing wheat inflorescence.

3.2.2 - Genetic regulation of tissue identity in grasses

To interpret the spatial patterns revealed by MERFISH, it is necessary to place them within the established framework of developmental regulation in grasses. Many of these expression patterns are conserved among cereals, reflecting shared regulatory programs that underlie tissue identity (Kellogg 2022). Here, I explore the genetic characterisation of key inflorescence-related genes in grass species, including rice (*Oryza sativa*), maize (*Zea mays*), barley (*Hordeum vulgare*), and wheat (*Triticum spp*.), and their corresponding gene expression patterns. This is not intended to be a complete review of the genetic controls in inflorescence development, but instead highlights some of the known developmental genes with distinct expression patterns characterised across the grasses that will be explored further in this chapter. I structure this as a relatively chronological series of events: the initiation and maintenance of axillary meristems, the transition of axillary meristems (AM) to spikelet meristem (SM), and the formation of floral meristems (FM) with the elaboration of floral organs.

3.2.2.1 - Axillary meristem initiation and maintenance

As the inflorescence meristem (IM) elongates, it maintains a pool of undifferentiated meristematic cells at its apex, while initiating lateral AM + bract pairings on its flanks (Kirby and Appleyard 1984). In this process, AMs must not only be initiated but also maintained as they transition through successive developmental identities (Tanaka et al. 2013). In grasses, neighbouring regions adjacent to AMs function as signalling centres, regulating meristem growth in a non-cell-autonomous manner. These are often described as 'boundary genes', which direct cell fate at the meristem boundary, determining whether cells contribute to the bract, remain part of the meristem, or cease division to establish the boundary domain (Whipple 2017; Kellogg 2022; Xiao et al. 2022).

One key regulatory gene, basic helix–loop–helix (bHLH) transcription termed *BARREN STALK 1 (BA1)* in maize and *LAX PANICLE 1 (LAX1)* in rice, is a regulator of AM initiation and maintenance acting non-cell autonomously. In maize, *BA1* is expressed in a distinct cell layer above the initiating AM. In *ba1* loss-of-function mutants, plants fail to produce AMs in vegetative and reproductive structures, leading to an inability to form tillers, ears, and tassels (Gallavotti et al. 2004). In rice, the ortholog *LAX1* displays a conserved function similar to that of maize. *LAX1* is expressed in a few cell layers surrounding the axillary meristems of the inflorescence, including primary and secondary branch meristems. Strong loss-of-function

lax1 alleles suppress lateral spikelet initiation and reduce panicle branching. Unlike its ortholog in maize, *LAX1* is not expressed in the vegetative shoot apical meristem (SAM), indicating a more specialised role in inflorescence development (Komatsu et al. 2001, 2003). In wheat, *lax1-aabbdd* mutant lines produce compact spikes, suggesting potential neofunctionalization relative to maize and rice (He et al. 2021). A regulator of *BA1/LAX1* has also been identified in rice, *MONOCULM1* (*MOC1*), a GRAS family transcription factor. This gene acts upstream as a positive regulator of *LAX1; moc1* mutants exhibit reduced inflorescence branching (Li et al. 2003), highlighting its role as a positive regulator in AM formation (Kellogg 2022).

Disruption of additional genes involved in boundary specification often impairs AM initiation, underscoring the importance of a bract–meristem boundary. Among these, NAC-domain transcription factors of the *CUP-SHAPED COTYLEDON (CUC)* family have been extensively studied in eudicots (Souer et al. 1996; Aida et al. 1997, 1999; Weir et al. 2004). Grass species also contain *CUC* homologs with expression patterns restricted to boundary regions (Zimmermann and Werr 2005; Chang et al. 2021; Wang et al. 2021). In maize, orthologs of *CUC1* and *CUC3* are expressed at the boundaries between the bract and spikelet pair meristem (SPM), on the flanks of the SPM, and between the SM and glume primordia during floral development. Loss-of-function mutants phenocopy *ba1*, with defective axillary bud initiation and absence of ears. Notably, *BA1* expression persists in these mutants, suggesting that *CUC* genes act independently of *BA1* in axillary meristem initiation (Zhong et al. 2025). Similarly, in rice, loss-of-function mutants in *NO APICAL MERISTEM (NAM;* ortholog to *AtCUC1* and *AtCUC2*) display small panicles with reduced primary branch meristem (PBM), and rare secondary branch meristem (SBM) formation (Chang et al. 2021). However, the contribution of *CUC* genes to wheat inflorescence development remains uncharacterised.

Once an AM is initiated, the maintenance of meristematic identity is regulated by several genetic pathways; one key player is Class 1 KNOTTED 1-like homeobox (KNOX) proteins. In *Arabidopsis,* the class-I *KNOX* gene *SHOOT MERISTEMLESS (STM)* is required to maintain the SAM. Strong loss-of-function *stm-1* mutant alleles completely lack a SAM, while weaker *stm-2* alleles show abnormal SAM organisation and arrested floral development (Barton and Poethig 1993; Clark et al. 1996; Long et al. 1996). These phenotypes demonstrate that *STM* suppresses differentiation within the meristem, preventing meristematic cells from being recruited into organ primordia (Scofield et al. 2014, 2018). In maize, the *STM* homolog *KNOTTED1 (KN1)* fulfils a comparable role. Loss-of-function *kn1* mutants exhibit severe inflorescence and floral defects, including absent ears, reduced branching in tassels and

ears, and extra carpels and proliferating ovules in female flowers (Kerstetter et al. 1997). *KN1* is expressed in meristematic tissues, ground tissue, and developing vascular bundles, and is downregulated as leaves and floral organs are initiated (Smith et al. 1992, Jackson et al. 1994). Orthologs to *KN1,* the rice ortholog *ORYZA SATIVA HOMEOBOX1* (*OSH1*), and barley ortholog *KN1* display similar expression patterns (Sentoku et al. 1999; Demesa-Arevalo et al. 2025), suggesting a role in meristematic maintenance in inflorescence and floral tissues more broadly across the grasses.

3.2.2.2 - Bracts act as signalling centres to adjacent axillary meristems

In most grasses, the lateral phytomers initiated along the central axis of the inflorescence display a suppressed growth of leaf ridges (LRs), whereas the AMs grow immediately upon their formation. This developmental pattern is underpinned by a conserved genetic program across the grasses and depends on the precise coordination of developmental progression or suppression within a small region of the inflorescence (Kellogg et al. 2013; Kellogg 2022). Genetic characterisation has suggested that bracts act as local signalling centres, providing positional cues that influence the determinacy of adjacent meristems (Whipple 2017).

In grasses, LRs are characterised by a distinct gene expression program linked to their arrested growth, setting them apart from other lateral organs. Key genes restricted to the ridge act in bract repression, including the maize genes *TASSELSHEATH1* (a GATA domain zinc-finger transcription factor) and *TASSELSHEATH4* (a SQUAMOSA BINDING PROTEIN transcription factor). Loss-of-function mutations in either result in ectopic leaves forming in the inflorescence (Chuck et al. 2010; Whipple et al. 2010), and a double loss-of-function mutant in *tsh1* and *tsh4* results in a more severe bract outgrowth phenotype, indicating their redundancy (Xiao et al. 2022). This phenotype is explained clearly by its expression pattern: both *TSH1* and *TSH4* are localised to the bract primordium from the earliest stages of bract initiation (Chuck et al. 2010; Whipple et al. 2010).

Interestingly, these mutants also affect inflorescence branching. Loss-of-function in *tsh1* or *tsh4* separately both reduce the number of long tassel branches (Chuck et al. 2010; Whipple et al. 2010), and in *tsh1 tsh4* double mutant tassels, there is a heightened effect with no long tassel branches present (Xiao et al. 2022). In maize, in addition to *TSH1* and *TSH4*, two SBP transcription factors, *UNBRANCHED2* (*UB2*) and *UNBRANCHED3* (*UB3*), display a similar role. Maize *ub2/ub3* double mutants produce extra leaves at the base of the tassel and initiate fewer BMs, consistent with their co-localisation with *TSH4* in bracts subtending

axillary meristems. Triple mutant *ub2/ub3/tsh4* lines displayed a more severe phenotype, characterised by the complete absence of tassel branches, suggesting genetic redundancy (Chuck et al. 2014). Together, these findings highlight the connection between the bract and the adjacent meristem in the inflorescence. One interpretation is that de-repressed bract growth occurs at the expense of AM development, thereby reducing branching potential (Chuck and Bortiri 2010).

Evidence from other cereals suggests a broadly conserved role for these bract suppression genes. *TSH1*-mediated bract suppression appears to be broadly conserved across the grasses; comparable phenotypes are observed in orthologs of *TSH1*, including *NECK LEAF1* (*NL1*) in rice and *THIRD OUTER GLUME* (*TRD1*) in barley (Wang et al. 2009; Houston et al. 2012). In rice, knockdown lines of *OsSPL14* (orthologous to maize *UB2*/*UB3*) exhibit reduced panicle branching (Wang et al. 2015). Wheat presents a slightly different outcome: simultaneous loss-of-function mutations in orthologs of *TSH4* and *UB2* (*SPL17* and *SPL14*) in a hexa-mutant (*TaSPL14-aabbdd_TaSPL17-aabbdd*) do not directly mirror the panicle branching phenotypes of maize and rice. Instead, hexa-mutant lines show a reduced number of spikelets per spike alongside bract outgrowth at the base of the inflorescence (Chen et al. 2023). This phenotype is consistent with a conserved role in bract suppression; however, the interpretation of reduced spikelet number as it pertains to a 'reduced branching' phenotype should be further explored.

3.2.2.3 *AP1/FUL-like* and *SVP* transcription factors in spikelet meristem transition

Whether examined through the framework of an unbranched spike-type inflorescence in *Triticeae,* or a branching panicle-type inflorescence seen in *Oryza sativa* and *Zea mays,* the transition to SM identity is accompanied by fundamental changes in meristem–bract pairings. In either inflorescence type, the establishment of an SM is ultimately marked by the outgrowth of two bracts (the glumes) with suppressed AMs, followed by one or more large bracts (lemmas) subtending outgrown FMs (See Figure 1.5A; Kellogg 2022). While the timing of the transition to SM identity is variable across grass species (Kyozuka 2014), the regulatory programmes underlying this transition exhibit some level of conservation.

Among the key regulators of SM identity in grasses are *APELATA1 (AP1) /FRUITFULL (FUL)*-like MIKC-type MADS box transcription factors. In wheat, single loss-of-function mutants of the *AP1/FUL*-like genes *VERNALIZATION1* (*VRN1*), *FUL2*, and *FUL3* retain a normal spikelet phenotype. However, in *vrn1 ful2* double mutants, although a typical double ridge forms, the AMs develop vegetative structures with occasional residual floral organs. This phenotype is further exacerbated in *vrn1 ful2 ful3* triple mutants, where AMs give rise to complete vegetative tillers subtended by de-repressed bracts, rather than spikelets. These observations indicate that *AP1/FUL*-like genes are essential for specifying AM and LR fate, and for promoting the transition to SM identity (Li et al. 2019, 2021b). A comparable role for *AP1/FUL*-like genes is observed in rice. In the *osmads14 osmads15* double mutant (orthologous to wheat *VRN1* and *FUL2*), secondary branches and spikelets are entirely absent; instead, AMs on primary branches develop into leaf-like organs (Wu et al. 2017).

A related function is seen for the *SEPALLATA (SEP)* transcription factor *PANICLE PHYTOMER 2* (*PAP2*), where loss-of-function (*pap2-1*) mutants fail to specify SM identity and produce branching structures in place of spikelets (Kobayashi et al. 2010). Notably, when these SM identity genes are disrupted together, the earlier vegetative-to-inflorescence meristem transition is also compromised. Knockdown lines targeting *OsMADS14*, *OsMADS15*, and *OsMADS18* (ortholog to *TaFUL3*) cause a delay in the transition to IM formation; knockdown in a *pap2* loss-of-function background results in the complete loss of transition to an IM, whereby the plants formed multiple shoots in place of the inflorescence (Kobayashi et al. 2010, 2012). Together, these findings suggest that *AP1/FUL*-like and *SEP* transcription factors function cooperatively to specify SM identity and to promote the transition from vegetative to IM identity.

To further clarify the role of *VRN1* and *FUL2* in meristem transitions, transcriptomic profiling of *vrn1 ful2* double mutants identified upregulation of three *SHORT VEGETATIVE PHASE (SVP)* MADS-box transcription factors: *SVP1, VEGETATIVE TO REPRODUCTIVE TRANSITION2 (VRT2),* and *SVP3*. This suggests that *SVP* genes act in an antagonistic manner to SM identity. Supporting this, constitutive expression of *VRT2* transforms basal spikelets into branches with multiple spikelets, further indicating that high SVP activity confers a branching identity in AMs (Li et al. 2019, 2021b). A comparable role for SVP-like genes has also been observed in barley, where ectopic expression of *BARLEY MADS1* (*BM1*, orthologous to *TaSVP3*) or *BM3* (orthologous to *TaSVP1*) results in basal spikelets being replaced by tillers or inflorescence-like branches (Trevaskis et al. 2007).

Further evidence suggests that *SVP* downregulation is required for proper SM identity. In barley, loss-of-function mutations in *MANY NODED DWARF 1* (*MND1*), encoding an acyl-CoA N-acyltransferase, produce phenotypes reminiscent of the wheat *vrn1 ful2 ful3* triple mutants (Walla et al. 2020). The *mnd1.a* allele causes a pronounced delay in the vegetative-to-reproductive transition and a high-tillering phenotype. Even after reproductive initiation, these plants continue to produce vegetative structures: triple spikelet meristems (TSMs) revert to a branch meristem-like identity, and bracts become derepressed. *In situ* hybridisation revealed that *MND1* is expressed in the vasculature of SAMs undergoing the transition to IM but is absent in earlier (SAM) or later (IM) stages. *MND1* expression is also restricted to AMs subtending leaves below the inflorescence, consistent with the aerial node branching observed in *mnd1.a* mutants. Notably, *SVP*-like *HvBM1* was upregulated in *mnd1.a* apices, suggesting that the reversion of SMs to BM identity in these mutants may result from misregulation of *SVPs* (Walla et al. 2020).

A related connection is observed in rice, where the ALOG protein TAWAWA1 (TAW1) regulates *SVP* gene activity. In the dominant gain-of-function mutant *taw1-D*, the transition to SM identity is delayed, leading to excessive branch formation. Conversely, reduced *TAW1* expression accelerates spikelet formation, producing compact inflorescences with fewer branches. During vegetative development, *TAW1* is expressed in the SAM, AMs, and young leaves, while in reproductive stages, it is active in the IM and BMs but excluded from SMs. Consistent with its role in promoting branching, *taw1-D* mutants show upregulation of *OsMADS22* (orthologous to *SVP1*), *OsMADS55* (orthologous to *VRT2*), and *OsMADS47* (orthologous to *SVP3*), strengthening the evidence that *SVP* genes act to delay the transition to SM identity (Yoshida et al. 2013).

3.2.2.4 - Regulation of spikelet meristem identity with *COMPOSTIUM* pathways

The transition to SM identity is also influenced by the COMPOSITUM pathway, which contains two major regulators: *COM1* and *COM2* (AP2-ERF transcription factors). Mutations in these regulators induce spike branching, indicating that *COM1* and *COM2* are key determinants of the direct transition from AM to SM characteristic of spike-type inflorescences. Beyond this role, however, both genes also contribute more broadly to inflorescence architecture in other grasses, including maize and rice (Koppolu et al. 2022).

COMPOSITUM2 has multiple names across the grasses: *COM2* in barley, *FRIZZY PANICLE (FZP)* in wheat and rice, and *BRANCHED SILKLESS1* (*BD1*) in maize. In wheat loss-of-function *fzp* mutants, SMs revert to BM or IM-like identity, leading to the indeterminate production of ectopic spikelets that give rise to branch-like structures. In barley, a comparable phenotype is observed (Dobrovolskaya et al. 2015; Poursarebani et al. 2015). In maize, *bd1* mutants exhibit a related but distinct outcome: IM and SPM development appear normal, yet presumptive SMs fail to initiate glumes and instead produce SPMs indeterminately (Chuck et al., 2002). Similarly, in rice, *fzp* mutants produce higher-order branches in place of spikelets (Bai et al. 2016). Interestingly, overexpression of *FZP* in rice panicle branching was dramatically reduced, with spikelets forming on primary branches, indicating that the overexpression of *FZP* leads to the premature transition of BMs to SMs. These overexpression lines also show induction of B- and E-class MADS-box floral identity genes, consistent with a role for *FZP* in promoting floral organ specification (Bai et al. 2016). *In situ* hybridisation reveals that *BD1* transcripts localise to the glume–meristem junction in wild-type maize inflorescences, with expression persisting at the base of developing florets, a pattern conserved in sorghum, rice, and barley (Chuck et al. 2002; Poursarebani et al. 2015).

A similar phenotype, where spikelets revert to branches, is observed in another key SM identity gene, *COMPOSITUM 1 (TCP24),* a grass-specific TCP transcription factor. In barley, the *com1.a* loss-of-function allele results in basal AMs that fail to form spikelets, instead producing indeterminate multimeric spikelets or branch-like structures, indicating a failure to establish SM identity. *COM1* is expressed within the boundary region adjacent to SMs (Poursarebani et al. 2020), with a similar expression pattern observed in maize (Bai et al. 2012). However, functional divergence is evident: whereas barley *COM1* suppresses branching, maize orthologs promote it. (Bai et al. 2012; Poursarebani et al. 2020). Loss-of-function in the maize gene *WAVY AURICLE IN BLADE1* (*WAB1,* ortholog to *COM1*), for example, disrupts boundary formation between tassel branches and the rachis when mutated, resulting in organ fusion (Bai et al. 2012), a phenotype mirrored in sorghum

(*Sorghum bicolor*) (Poursarebani et al. 2020). Despite this divergence, phenotypic and expression data suggest that *COM1* provides critical positional information to adjacent SMs (Koppolu et al. 2022). Notably, while barley *com1* and *com2* mutants share similar branching phenotypes, double mutants display enhanced and more consistent branching, indicating that these two pathways act independently to regulate SM fate and inflorescence architecture (Poursarebani et al. 2015, 2020).

3.2.2.5 - Floral meristem determinacy & floral organ specification

During inflorescence development, SMs progressively form FMs in an alternating, distichous pattern along their flanks. Each new floral meristem is first signalled by the emergence of a subtending bract, the lemma. In wheat, the FM subsequently gives rise to the floral organ primordia, including the palea as part of the outer perianth, the lodicules as inner perianth organs, and reproductive structures such as stamens and a gynoecium enclosing a single ovule (Kirby & Appleyard, 1987). The FM meristem is referred to as determinant- the meristematic cells are consumed by the formation of floral organs (Tanaka et al. 2013).

Many of these genes that control FM determinacy also play a role in floral organ patterning. A central framework for explaining how these organs acquire their identities is provided by the ABCDE model of floral organ identity. This model, along with the related “floral quartet” concept, proposes that floral identity is determined by specific combinations of MADS-domain transcription factors forming higher-order protein complexes that activate downstream developmental programs specification (Coen and Meyerowitz 1991; Ma and dePamphilis 2000; Theißen et al. 2016). In this framework, the flower is organised into four concentric whorls: sepals (whorl 1) are specified by A- and E-class proteins; petals (whorl 2) by A-, B-, and E-class proteins (Honma and Goto 2001); stamens (whorl 3) by B-, C-, and E-class proteins (Coen and Meyerowitz 1991); and carpels (whorl 4) by C- and E-class proteins. Ovule development within the carpel is regulated primarily by D- and E-class proteins, although it is not considered a separate whorl (Colombo et al. 1995; Becker and Theißen 2003).

In *Arabidopsis thaliana*, extensive genetic characterisation has identified genes with these proposed ABCDE functions: A-class identity is encoded by *APETALA1* (*AP1*) and *APETALA2* (*AP2*), with *AP2* representing the sole non-MADS-box gene in the model (Mandel et al. 1992; Jofuku et al. 1994; Irish 2010). B-class function is controlled by *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) (Jack et al. 1992; Goto and Meyerowitz 1994). C-class identity is specified by

AGAMOUS (*AG*), while D-class roles are carried out by *AGAMOUS*-like genes such as *SEEDSTICK* (*STK*), *SHATTERPROOF1* (*SHP1*), and *SHP2* (Favaro et al. 2003; Pinyopich et al. 2003). E-class function is encoded by several *AGAMOUS*-like genes, including *SEPALLATA1* (*SEP1*), *SEP2*, *SEP3*, and *SEP4* (Pelaz et al. 2000; Ditta et al. 2004).

Although the ABCDE model was initially developed in eudicots, it is also broadly conserved in monocots. Grass flowers follow a similar organisational logic, with floral organ identity also governed by MADS-box gene combinations. In this context, A- and E-class proteins specify the palea (whorl 1), A-, B-, and E-class proteins determine lodicule identity (whorl 2), B-, C-, and E-class specify stamens (whorl 3), and C- and E-class define carpels (whorl 4). D-class genes play central roles in ovule development. Comparative studies in rice, maize, barley, and wheat demonstrate strong conservation of B-, C-, and D-class gene functions, although the contributions of A- and E-class genes show greater divergence across species (Whipple and Schmidt 2006; Thompson and Hake 2009; Murai 2013).

Together, these studies highlight a conserved regulatory framework that coordinates the balance between meristem identity, transition, and floral patterning across the grasses. While much of this knowledge has been gained from genetic and transcriptomic studies in model cereals such as maize, rice, and barley, less is known about how these regulators function in wheat, particularly within their native spatial and developmental context. In the following section, I use spatial transcriptomic data to map the expression of key regulators across wheat inflorescence development and explore the dynamics of gene expression domains that underlie tissue identity and developmental transitions.

3.3 - Results

3.3.1 - Unsupervised clustering of 50,000 cells reveals 18 expression domains

A common approach for analysing spatial transcriptomic data is to group cells into transcriptionally similar clusters, which can reveal distinct tissue types. To explore these spatial and temporal dynamics, we performed clustering analysis across all samples collected in this study, which spanned four developmental stages and two genotypes. Transcript counts were normalised for samples individually, followed by sample integration to group 50,731 total cells across the four developmental stages and two genotypes, and subsequent Leiden clustering (Wolf et al. 2018; Palla et al. 2022). This pipeline yielded 18 distinct expression domains (ED, Figure 3.1A), which were visualised as spatially resolved maps (Fig. 3.1B-I. The total number of clusters per sample increases over developmental time, with WT samples W2.5, W3.25, W4, W5, summarised by 11, 12, 17, and 18 domains, respectively (Table 3.1, removing clusters representing less than 0.5% of cells per sample).

We observed that EDs remained consistent in their patterning across biological replicates, in addition to across time points. For example, expression domain 3 (ED3) maps consistently to the developing rachis in W2.5, W3.25, and W4 in both $P1^{WT}$ and $P1^{POL}$, followed by the relative decrease in ED3 and increase in ED6 in the vasculature of the W4 spike. Similarly, in both replicates, we observe the consistent patterning of the L1 layer of spikelet ridges in stages W2.5 and W3.25 assigned to ED4. As the spikelets gain complexity in stages W4 and W5, this patterning of ED4 to the L1 layer becomes less prominent.

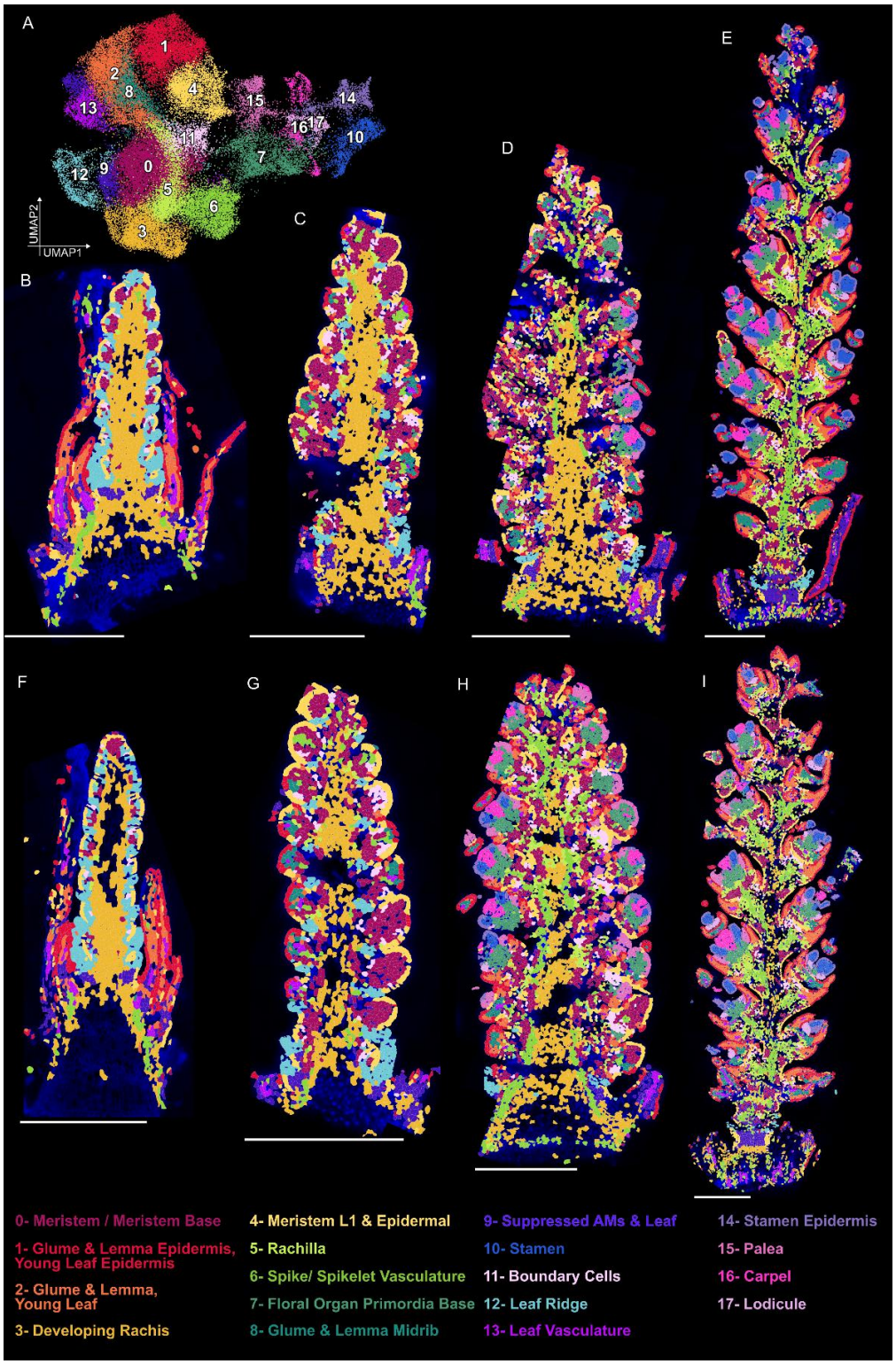


Figure 3.1 - Leiden clustering identifies 18 expression domains over four developmental stages
A) UMAP projection 50,731 cells from eight samples (four developmental stages, two genotypes), and expression domain assignment. (**B-E**) Spatial maps of Leiden clustering in *P1*^{WT} samples across time points W2.5 **B)**, W3.25 **C)**, W4 **D)**, and W5 **E)** using Squidpy (v1.4.1), Scanpy (v 1.10.0), and Scanorama (v1.7.4); Scale bar = 500µm. **F–I)** Corresponding spatial maps in *P1*^{POL} samples at W2.5 (F), W3.25 (G), W4 (H), and W5 (I), showing consistent expression domain patterns; scale bar = 500 µm.

Table 3.1- Number of cells assigned to 18 expression domains

Time Point	Genotype	Cluster Number																	Total Cells per Sample
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	
W2.5	*P1*^{WT}	165	251	166	599	233	23	98	1	6	142	0	41	367	116	0	1	0	0	2209
W2.5	*P1*^{POL}	79	191	138	357	198	6	72	1	5	143	0	21	307	51	0	0	0	0	1569
W3.25	*P1*^{WT}	782	138	62	761	355	90	104	6	58	120	0	151	226	40	0	3	0	0	2896
W3.25	*P1*^{POL}	614	95	21	384	284	54	83	1	59	123	0	99	261	27	0	2	0	0	2107
W4	*P1*^{WT}	1196	385	284	965	659	248	299	233	157	186	40	284	141	82	35	134	76	1	5405
W4	*P1*^{POL}	1264	439	432	667	538	447	558	631	135	78	90	271	57	93	90	383	202	10	6385
W5	*P1*^{WT}	1558	1954	1743	349	803	1550	1260	990	814	598	871	362	105	399	711	393	444	281	15185
W5	*P1*^{POL}	1393	1899	2159	559	969	1513	1248	949	674	309	645	383	56	626	586	397	292	318	14975

3.3.2 - Gene expression analysis defines identity of expression domains

Beyond cell clustering, all generated expression domains require tissue-specific annotations in order to gain biological context. To do so, we annotated the 18 domains by combining knowledge of anatomical features with the identification of domain-enriched genes. In total, we characterised 54 genes from our panel to show domain-specific enrichment. Gene enrichment was determined using rank gene functions in Scanpy, with a logistic regression model (Ntranos et al., 2019). In brief, this analysis generated a ranked list of genes most likely to be enriched gene markers, from which we identified the top enriched genes using the logistic regression score (+2 STD threshold, see Methods).

A majority of these markers (57.4%) were enriched across multiple expression domains, with 23 genes enriched in a single domain. Ten genes were enriched in >3 clusters. These included markers of vegetative tissues and bracts: (*YABBY7 (YAB7))*, outer cell layers (orthologs to rice *ONION1 (ONI1), RICE OUTERMOST CELL-SPECIFIC 7 (ROC7), RICE OUTERMOST CELL 3 (ROC3))*, meristematic and ground tissue (*KNOTTED HOMEOBOX-LIKE 5* (*KNOX5))*, and floral tissues (*APETALA3* (AP3)*, AGAMOUS-LIKE 6 (AGL6) , SEPALLATA3-1 (SEP3-1),* and *SEPALLATA3-2 (SEP3-2*)). Across domains, the total number of significantly enriched marker genes ranged from 3 to 10.

Domains were annotated based on the functional characterisation of marker genes in wheat and related grasses. Orthologs in maize, rice, and barley were identified, and their reported expression patterns were reviewed through comparative literature analysis. In the following section, I will document the annotation of each domain with a tissue type description (Table 3.2)

Table 3.2 - Enriched genes and tissue type annotation in 18 expression domains
For each domain, the top enriched genes are listed alongside their corresponding wheat gene names and orthologous genes in *Oryza sativa* L. ssp. *japonica*.

Domain	Annotation	Gene ID	Gene Name	*Oryza sativa* gene ID	*Oryza sativa* Gene Name
0	Meristem / Meristem Base	TraesCS4A02G256700	*KNOX5*	Os03g0727000	*OSH1*
		TraesCS4B02G064000		Os03g0733600	*OsGIF3*
		TraesCS3D02G284200	*AGL14*	Os01g0726400	*MADS32*
1	Glume & Lemma Epidermis, Young Leaf Epidermis	TraesCS4D02G296400	*FDH*	Os03g0181500	*ONI1*
		TraesCS6D02G220400	*YABBY7*	Os02g0643200	*YAB4*
		TraesCS7A02G308400		Os08g0136100	*Roc7(t)*
		TraesCS1D02G197300		Os10g0575600	*Roc3(t)*
		TraesCS5A02G185600		Os09g0334500	*WRKY74*
		TraesCS4A02G191300		Os11g0124300, Os12g0122000	*OsSCR2*
2	Glume & Lemma, Young Leaf	TraesCS6D02G220400	*YABBY7*	Os02g0643200	*YAB4*
		TraesCS2B02G403100	*YABBY3*	Os04g0536300	*YAB5*
		TraesCS3D02G284200	*AGL14*	Os01g0726400	*MADS32*
		TraesCS1D02G162600	*YABBY1*	Os10g0508300	*YAB3*
		TraesCS5B02G246700		Os09g0470500	*Oshox4*
3	Developing Rachis	TraesCS1B02G042200	*MT2B*	Os05g0111300	*OsMT2b*
		TraesCS4A02G256700	*KNOX5*	Os03g0727000	*OSH1*
		TraesCS2B02G399800		Os04g0516200	*OsG1L4*
		TraesCS7A02G175200	*VRT2*	Os06g0217300	*OsMADS55*
		TraesCS5A02G401800	*SP3*	Os03g0764900	*OsDof15*
		TraesCS7D02G191600		Os06g0232300	*PIN1C*
		TraesCS6A02G335900	*GRF10-6*	Os02g0776900	*GRF1*
		TraesCS7A02G246500	*SPL14*	Os08g0509600	*WFP*
4	Meristem L1 & Epidermal	TraesCS7A02G308400		Os08g0136100	*Roc7(t)*
		TraesCS4D02G296400	*FDH*	Os03g0181500	*ONI1*
		TraesCS1D02G197300		Os10g0575600	*Roc3(t)*
		TraesCS7D02G246100	*CUC3*	Os08g0511200	*OsCUC3*
		TraesCS4B02G064000		Os03g0733600	*OsGIF3*
		TraesCS1D02G343400		Os08g0562500	
		TraesCS3D02G357400	*RIL1*	Os01g0848400	*qSH1, RIL1*
5	Rachilla	TraesCS3D02G284200	*AGL14*	Os01g0726400	*MADS32*
		TraesCS1B02G042200	*MT2B*	Os05g0111300	*OsMT2b*
		TraesCS1D02G075700	*KNOX3*	Os05g0129700	*OsKn2*
		TraesCS3A02G155200		Os01g0231000	*OsIAA3*
6	Spike / Spikelet Vasculature	TraesCS1B02G479300		Os05g0595800	
		TraesCS6A02G176400		Os02g0203700	*SRZ1*
		TraesCS7D02G191600		Os06g0232300	*PIN1C*
		TraesCS1A02G077800		Os10g0147400	*OsLAX4*
		TraesCS2A02G192600	*SHR*	Os07g0586900	*OsSHR1*
		TraesCS6A02G335900	*GRF10-6*	Os02g0776900	*GRF1*
		TraesCS4A02G256700	*KNOX5*	Os03g0727000	*OSH1*
		TraesCS6D02G011600			
7		TraesCS6A02G259000	*AGL6*	Os02g0682200	*MFO1*
		TraesCS4A02G256700	*KNOX5*	Os03g0727000	*OSH1*

	Floral Organ Primordium Base	TraesCS2B02G464200	*LFY*	Os04g0598300	*RFL*
		TraesCS5A02G401800	*SP3*	Os03g0764900	*OsDof15*
		TraesCS7A02G383800	*AP3*	Os06g0712700	*SPW1*
	Glume & Lemma Midrib	TraesCS4D02G245300	*YABBY4*	Os03g0215200	*DL*
		TraesCS2B02G403100	*YABBY3*	Os04g0536300	*YAB5*
8		TraesCS6D02G220400	*YABBY7*	Os02g0643200	*YAB4*
	Suppressed AxMs, Leaf	TraesCS4B02G064000		Os03g0733600	*OsGIF3*
		TraesCS6A02G313800	*SVP-1*	Os02g0761000	*OsMADS22*
		TraesCS6A02G335900	*GRF10-6*	Os02g0776900	*GRF1*
		TraesCS7B02G413900	*MND1*	Os06g0650300	*OsglHAT1*
		TraesCS7A02G175200	*VRT2*	Os06g0217300	*OsMADS55*
		TraesCS5A02G356100		Os09g0555700	*OsIDD6*
		TraesCS2B02G399800		Os04g0516200	*OsG1L4*
9		TraesCS5A02G473800	*AP2-5*	Os03g0818800	*AP2/EREBP #033*
	Stamen	TraesCS7A02G383800	*AP3*	Os06g0712700	*SPW1*
		TraesCS7D02G261600	*SEP3-1*	Os08g0531700	*OsMADS7*
		TraesCS6A02G313800	*SVP-1*	Os02g0761000	*OsMADS22*
		TraesCS1A02G264300	*PI1*	Os05g0423400	*OsMADS4*
		TraesCS4B02G084800		Os03g0416300	
		TraesCS1D02G127700	*AG1*	Os05g0203800	*MADS58*
		TraesCS5A02G230500		Os09g0410700	*OsbHLH039*
		TraesCS3A02G406500	*PI2*	Os01g0883100	*OsMADS2*
10		TraesCS5A02G286800	*SEP3-2*		*OsMADS24*
	Boundary Cells	TraesCS7D02G246100	*CUC3*	Os08g0511200	*OsCUC3*
		TraesCS3D02G284200	*AGL14*	Os01g0726400	*MADS32*
		TraesCS5A02G161000		Os09g0111100	*CycD3*
		TraesCS5B02G246700		Os09g0470500	*Oshox4*
11		TraesCS7D02G339600		Os06g0336200	*OsTIP2*
	Leaf Ridge	TraesCS6A02G287300	*LEC1*	Os02g0725700	*OsLEC1*
		TraesCS1A02G418200	*TSH1*	Os05g0578900	*NL1*
		TraesCS7A02G246500	*SPL14*	Os08g0509600	*WFP*
		TraesCS6D02G245300	*GRF9-6*		
		TraesCS5A02G265900	*SPL17*	Os09g0491532	*OsSPL17*
12		TraesCS1D02G162600	*YABBY1*	Os10g0508300	*YAB3*
	Leaf Vasculature	TraesCS6D02G220400	*YABBY7*	Os02g0643200	*YAB4*
		TraesCS1D02G162600	*YABBY1*	Os10g0508300	*YAB3*
		TraesCS2A02G192600	*SHR*	Os07g0586900	*OsSHR1*
		TraesCS7D02G191600		Os06g0232300	*PIN1C*
		TraesCS1B02G479300		Os05g0595800	
13		TraesCS1A02G077800		Os10g0147400	*OsLAX4*
	Stamen Epidermis	TraesCS4D02G296400	*FDH*	Os03g0181500	*ONI1*
		TraesCS1D02G197300		Os10g0575600	*Roc3(t)*
		TraesCS1A02G264300	*PI1*	Os05g0423400	*OsMADS4*
		TraesCS7A02G383800	*AP3*	Os06g0712700	*SPW1*
		TraesCS7A02G308400		Os08g0136100	*Roc7(t)*
		TraesCS5A02G286800	*SEP3-2*		*OsMADS24*
		TraesCS7D02G261600	*SEP3-1*	Os08g0531700	*OsMADS7*
		TraesCS3A02G406500	*PI2*	Os01g0883100	*OsMADS2*
		TraesCS1D02G127700	*AG1*	Os05g0203800	*MADS58*
14		TraesCS7D02G246100	*CUC3*	Os08g0511200	*OsCUC3*

		TraesCS6A02G259000	*AGL6*	Os02g0682200	*MFO1*
		TraesCS2B02G464200	*LFY*	Os04g0598300	*RFL*
		TraesCS6D02G220400	*YABBY7*	Os02g0643200	*YAB4*
		TraesCS4D02G296400	*FDH*	Os03g0181500	*ONI1*
		TraesCS1D02G197300		Os10g0575600	*Roc3(t)*
15	**Palea**	TraesCS7A02G308400		Os08g0136100	*Roc7(t)*
		TraesCS6A02G259000	*AGL6*	Os02g0682200	*MFO1*
		TraesCS1D02G127700	*AG1*	Os05g0203800	*MADS58*
		TraesCS5A02G286800	*SEP3-2*		*OsMADS24*
		TraesCS7D02G261600	*SEP3-1*	Os08g0531700	*OsMADS7*
		TraesCS4D02G245300	*YABBY4*	Os03g0215200	*DL*
		TraesCS1B02G283900		Os05g0438800	
16	**Carpel**	TraesCS3A02G314300	*AG2*		
		TraesCS6A02G259000	*AGL6*	Os02g0682200	*MFO1*
		TraesCS7A02G383800	*AP3*	Os06g0712700	*SPW1*
		TraesCS5A02G286800	*SEP3-2*		*OsMADS24*
		TraesCS7D02G261600	*SEP3-1*	Os08g0531700	*OsMADS7*
		TraesCS1A02G264300	*PI1*	Os05g0423400	*OsMADS4*
		TraesCS6D02G220400	*YABBY7*	Os02g0643200	*YAB4*
		TraesCS2B02G464200	*LFY*	Os04g0598300	*RFL*
		TraesCS1A02G052000		Os05g0118700	*LSY1*
		TraesCS2D02G256600	*PARG-2D*	Os07g0124700	*OsPLT8*
17	**Lodicules**	TraesCS3A02G406500	*PI2*	Os01g0883100	*OsMADS2*

Bracts & lateral organ domains

ED1/2 are identified in leaves at stage W2.5, and in the bract tissues of spikelets at stages W3.25, W4, and W5. These domains are enriched for *YABBY7*, *YABBY3,* and *YABBY1*, which are orthologs to the rice *TONGARI-BOUSHI* genes *(TOB1/TOB2/TOB3),* which are characterised in rice to be expressed in bracts and leaf primordia (Figure 3.2A-B; Tanaka et al. 2017). This shared expression pattern observed between these two tissue types potentially highlights their transcriptional similarities- consistent with the classification of glumes and lemmas as leaf-like bracts (Patterson et al. 2023). ED8 is positioned adjacent to ED2 in spatial arrangement and shares a similar transcriptional profile in UMAP space. This domain shares the enrichment of *YABBY7* and *YABBY3*, in addition to *YABBY4*, the ortholog to rice *DROOPING LEAF (DL)*, a gene characterised to specify the central region (midrib) of leaf primordia (Figure 3.2C; Ohmori et al. 2011). This suggests a potential role for *YABBY4* in midrib patterning and central domain formation in glumes and lemmas.

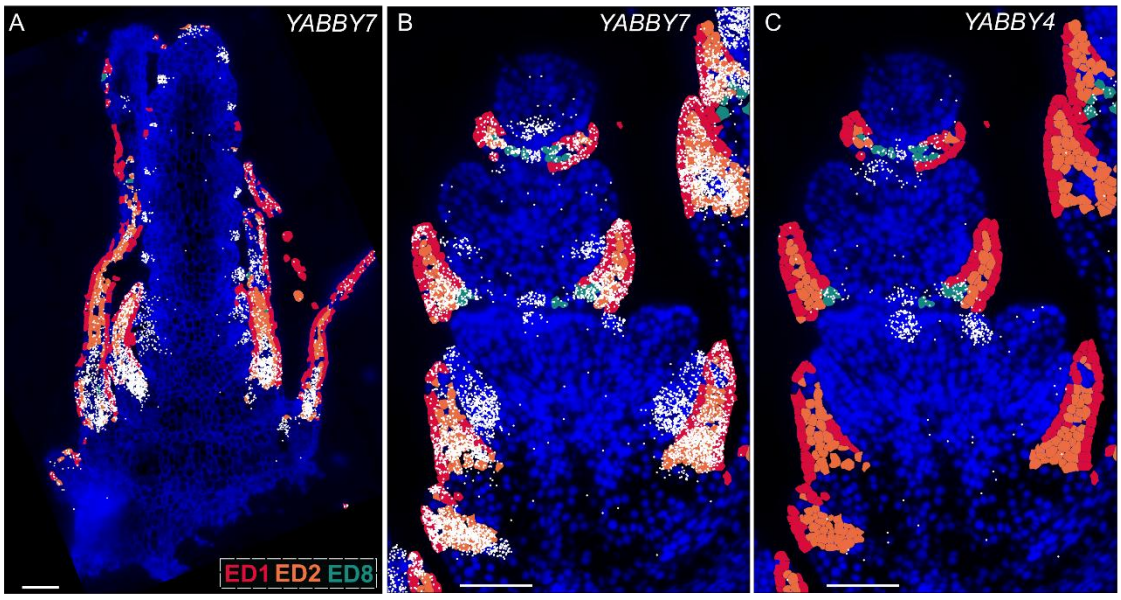


Figure 3.2 - Gene enrichment in bract and lateral organ domains
A) Transcript detection and domain map in $P1^{WT}$ W2.5 spikes. Expression of *YABBY7,* Domains ED1, ED2, ED8. **B-C)** Transcript detection and domain map in $P1^{WT}$ W5 spikes. Domains ED1, ED2, ED8. **B)** Expression of *YABBY7*. **C)** Expression of *YABBY4*. Detected transcript represented by white. Scale bar = 100 µm.

Rachis & Rachilla domains

Two domains, ED3 and ED6, are located within the developing rachis of the inflorescence. Both are enriched for *KNOX5*, whose rice ortholog, OsKN1, is expressed in meristems, ground tissue, and vascular strands (Jackson et al. 1994). ED3 and ED6 are distinguished by their temporal dynamics and gene expression profiles. ED3 is prominent during early developmental stages in the developing ground tissues, comprising 27.1% and 26.2% of total cells in the $P1^{WT}$ sample at W2.5 and W3.25, respectively. By stage W4 in $P1^{WT}$ and $P1^{POL}$ replicates, distinct bands of ED6 formed in patterns indicating localisation to the vasculature. ED6 is enriched for *SHORT ROOT* (*SHR*), whose orthologs have been shown to regulate vascular cell identity and ground tissue proliferation in rice and *Setaria viridis* (Liu et al. 2023a), as well as vascular patterning in maize leaves (Slewinski et al. 2014). A spatial transcriptomic analysis of the maize shoot apex identified a vascular domain marked by *SHR1* and *SHR2* expression (Perico et al. 2024), further supporting the vascular identity of ED6. Based on these data, ED6 was annotated as 'spike and spikelet vasculature,' while ED3 was designated as 'developing rachis.'

ED5 was localised to the base of each spikelet and observed at stages W3.25, W4, and W5. This domain is enriched for the wheat ortholog of *OsMADS32*, a rice gene expressed in floral organ primordia and the rachilla (Sang et al. 2012). ED5 also shows enrichment for *KNOX3*, orthologous to rice *KNOTTED2*, which is expressed in the basal regions of spikelets and in pedicels (Postma-Haarsma et al. 2002). Given its spatial localisation and gene expression profile, ED5 was annotated as ‘rachilla.’

Early inflorescence development domains

Early during wheat inflorescence development, in the double ridge stage (W2.5), the IM initiates pairs of ridges, comprising a lower suppressed bract (leaf ridge, LR) and an upper axillary meristem (AM). The inflorescence can be summarised by five domains at this stage: ED0, ED3, ED4, ED11, and ED12 (Figure 3.3A).

ED12 was annotated as the leaf ridge based on its strong enrichment for *TraesCS1A02G418200*, the wheat ortholog of maize *TASSELSHEATH1* (*TSH1*), which is expressed in the suppressed bract (leaf ridge) of the inflorescence (Figure 3.3B; Whipple et al. 2010). ED0 was assigned as L2/L3 meristematic cells of the spikelet ridge due to the enrichment of *KNOX5*. *KNOX5* is enriched across multiple domains in stage W2.5: ED3 and ED0. We observe *KNOX5* expression in meristematic cells and ground tissue; however, it is excluded from the tunicate/L1 layer, consistent with observations of its maize ortholog *KNOTTED1* (Figure 3.3C; Jackson et al. 1994). ED3 was differentiated from meristematic cells by the unique expression of genes, including *METALLOTHIONEIN 2* (*MT2;* Figure 3.3D). ED4 was annotated as the L1 layer of the spikelet meristem, based on the enrichment of L1-specific markers including *TraesCS7A02G308400* (ortholog to rice *OsROC7t*), *TraesCS4D02G296400* (*ONION1, OsONI1*), and *TraesCS1D02G197300* (*OsROC3t*; Figure 3.3E). Flanking the adaxial boundary of the spikelet meristem are cells assigned to ED11. These genes are enriched in ortholog to rice *OsCUC3*, which is characterised to be expressed between meristem/organ boundaries (Wang et al. 2010). In addition, the expression of *LAX1*, a gene expressed in the cell layers surrounding the meristems (Komatsu et al. 2001), further provided evidence that ED11 cells are located in the 'boundary' region between bracts and meristems (Figure 3.3F). Therefore, this domain is annotated as *boundary cells*.

Below the inflorescence, we observed ED9 marked by strong localised expression of *MANY NODED DWARF 1 (MND1),* a gene whose barley ortholog suppresses axillary meristem outgrowth (Walla et al., 2020). This domain includes multiple tissues: ED9 marks both developing young leaves and suppressed AM below the inflorescence *per se,* with *MANY NODED DWARF 1* (*MND1*) marking suppressed AMs only (Figure 3.3G). We hypothesise that the spike-focused probe panel lacked sufficient genes to distinguish these tissues as distinct ED. This domain was annotated as *suppressed axillary meristems & leaf*.

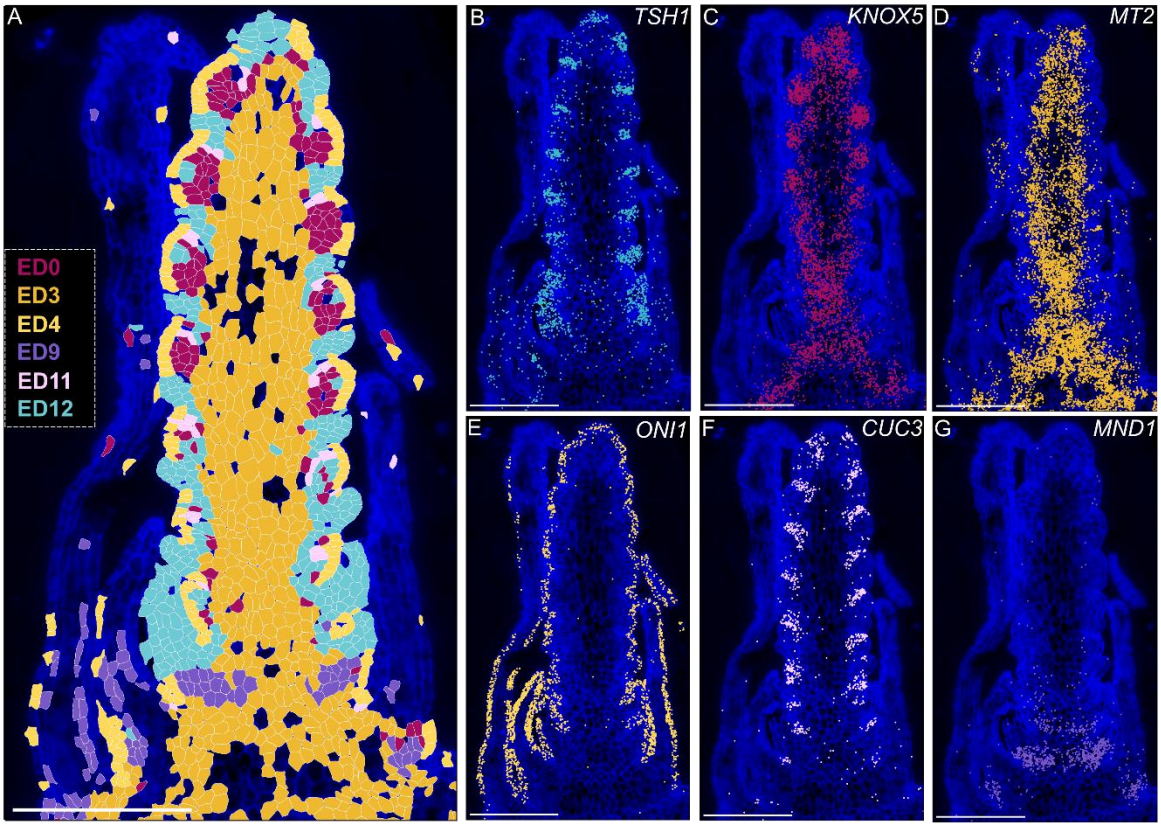


Figure 3.3 - Gene enrichment in early inflorescence development
A) Spatial map of six expression domains in W2.5 spikes, highlighting domains enriched with transcripts **(B-E), B)** *TSH1*, **C)** *KNOX5,* **D)** *MT2B*, **E)** *ONI1,* **F)** *CUC3,* and **G)** *MND1*. Scale bar = 250 µm. Blue stain = DAPI. ED0 = meristem, ED3 = developing rachis, ED4 = meristem L1, ED9 = suppressed AMs, ED11 = boundary, ED12

Floral domains

We identified eight expression domains corresponding to floral organs and their subtending bracts in wheat, including glumes and lemmas (ED1,2), palea (ED15), lodicules (ED17), stamens (ED10 and ED14), and the carpel (ED16; Figure 3.4). Six of these domains (ED16, 14, 7, 15, 10, 17) were primarily detected at developmental stages W4 and W5, consistent with the onset of floral organ specification proceeding from stage W3.5 onwards. We observed that these domains exhibited consistent spatial organisation between *P1*^{WT} and *P1*^{POL} replicates (Figure 3.1). Notably, these domains form a distinct cluster in UMAP space, reflecting their high transcriptional similarity and separation from other expression domains (Figure 3.1A).

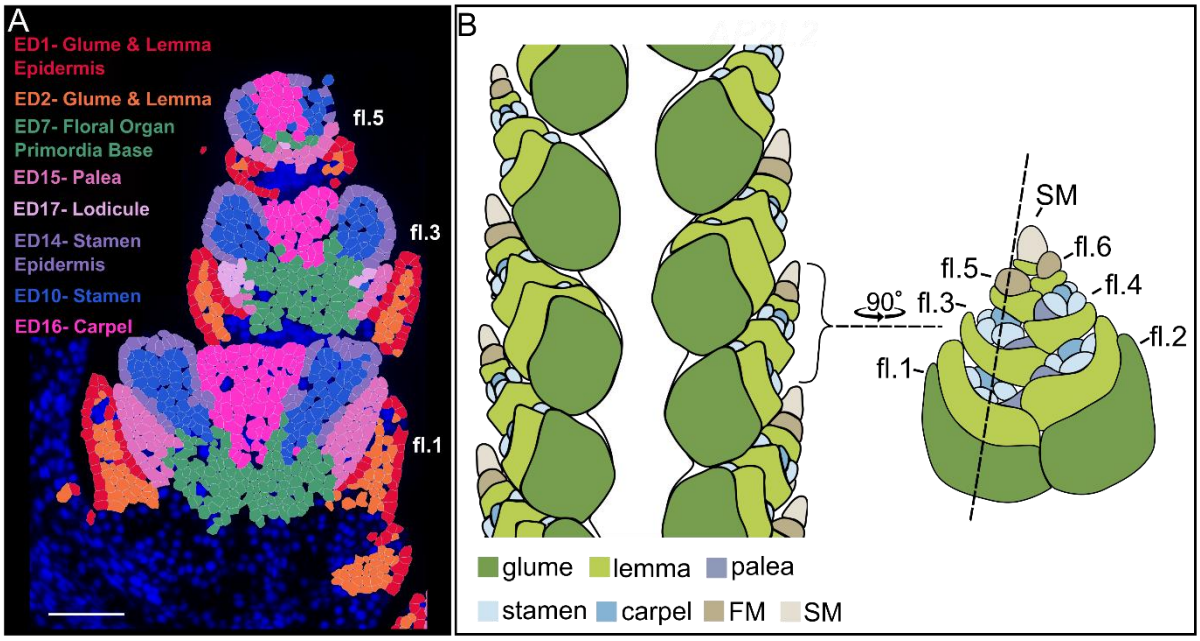


Figure 3.4 - Eight expression domains represent floral organs and subtending bracts
A) Spatial map of floral and bract expression domain (ED1,2 = glume/lemma, ED7 = floral organ primordia base, ED15 = palea, ED17 = lodicule, ED14,10 = stamen, ED16 = carpel) fl = floret. **B)** Diagrammatic representation of wheat floral anatomy and cryosection angle (black dashed line). The sectioning angle through spikelet is representative of the cryosection captured in panel A. fl = floret, SM = Spikelet meristem, FM = floral meristem.

We hypothesised that concatenation of samples across developmental time prior to clustering would reveal floral cells in their earliest stages of emergence. For example, in W3.25 spikes, we identified only three ED15 cells (Figure 3.5A), a predominant domain in more mature spikes (Figure 3.5B) and which later localises to paleae in W5 samples (Figure 3.5B-C). At W3.25, these three ED15 cells were positioned just above the lemma primordia (ED1, 2), a spatial arrangement consistent across stages (Figure 3.5D-F). In all ED15 cells, we observe the expression of *AGAMOUS-LIKE6 (AGL6)*, a key regulator of palea identity in wheat (Kong et al. 2022). This indicates that ED15 at W3.25 most likely represents the first cells with palea identity in the developing spike.

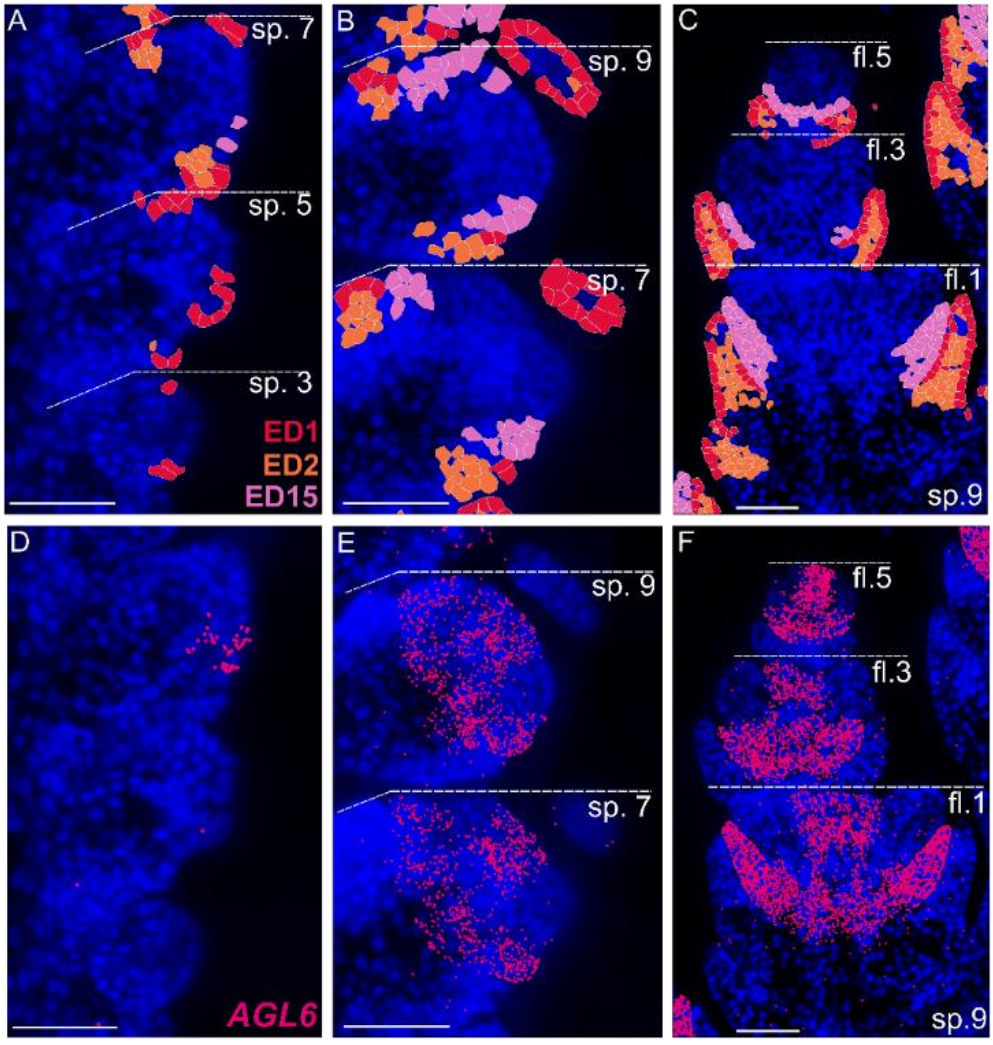


Figure 3.5 - Expression domains and gene enrichment trace floral identity across developmental time

Spatial plots of cell segmentation and assigned expression domains 1, 2, & 15 in **A)** W3.25 spikelets 3,5,7 **B)** W4 spikelets 7,9 **C)** W5 spikelet 9 (florets = fl; 1,3,5). (**D-F**) Expression of *AGL6* in same tissues as F-H, respectively. Blue= DAPI staining; Scale bar = 100 µm.

In total, the floral expression domains are enriched for 29 genes, including 10 MIKC-type MADS-box transcription factors. We hypothesised that the spatial gene expression of these genes would align with the putative functions defined by the ABCDE model.

Consistent with this model, we observe spatially restricted expression of B-, C-, and D-class genes in their expected floral organ domains. B-class genes, including *APETALA3* (*AP3*), *PISTILLATA1* (*PI-1*), and *PISTILLATA2* (*PI-2*), are enriched in lodicules and stamens (Figure 3.F-H). C-class genes *AGAMOUS1* (*AG-1*) and *AGAMOUS2* (*AG-2*) show enrichment in both stamens and carpels (Figure 3.6I-J), while the D-class gene *SEEDSTICK1* (*STK1*) is exclusively expressed in the carpel domain (Figure 3.6K).

E-class function encoding genes in monocots are divided into three groups: *SEP3*, *LOFSEP*, and *AGL6*-like (Malcomber and Kellogg 2005; Wu et al. 2018; Dreni and Ferrándiz 2022). In the case of E-class LOFSEP clade, expression of orthologs to *SEPALLATA1* (*SEP1-1, SEP1-2, SEP1- 4, SEP1-5, SEP1-6*) are mostly absent from floral organ primordia (Figure 3.6L-S). However, the E-class genes *SEPALLATA 3-1* (*SEP3-1*), *SEP3-2*, and *AGL6* display overlapping, not fully redundant, enrichment across all floral organs. *AGL6* is expressed in palea, carpels, and lodicules, while *SEP3-1* and *SEP3-2* are expressed in stamens, carpels, and lodicules.

Variable expression patterns are observed within the A-class members of the *AP1/FUL*-like clade include *VRN1, FUL3, and FUL2*. We observe that *VRN1* and *FUL3* are expressed broadly across floral tissues, while *FUL2* displays restricted expression, primarily in the lemmas and glumes (Figure 3.6A-C). The AP2-like genes *AP2-5* and *AP2-2* differ in their distribution: *AP2-5* is broadly expressed, whereas *AP2-2* is expressed in the lemma, palea, lodicules, and stamen primordia (Figure 3.6D-E).

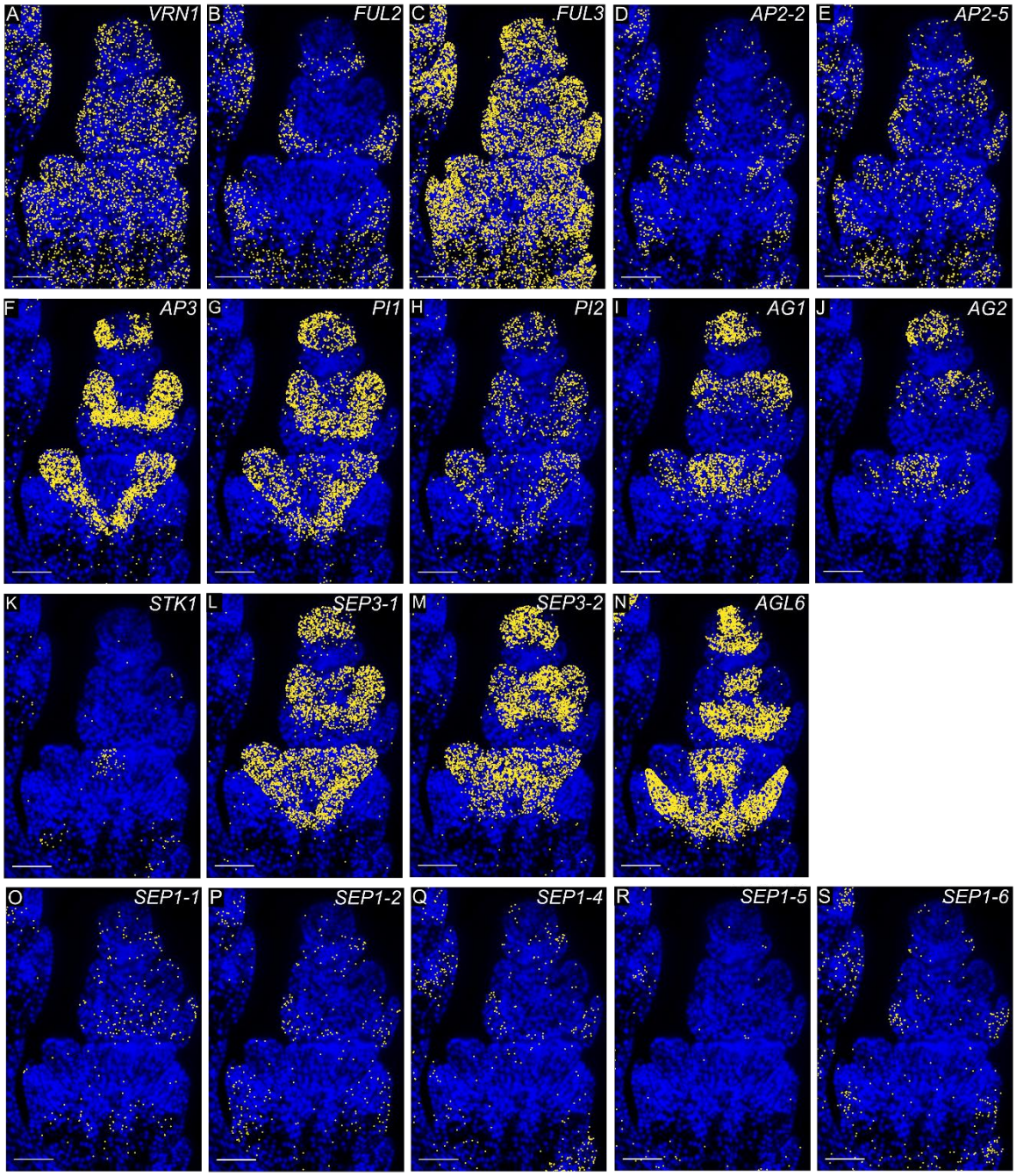


Figure 3.6 - Expression of ABCDE-class genes in floral organ domains
Spatial Transcript Localisation in W5 *P1*^{WT} florets, in **(A-F)** A-class genes **B)** *VRN1,* **B)** *FUL2,* **C)** *FUL3,* **D)** *AP2-2,* **E)** *AP2-5*, **(F-H)** B-class genes **F)** *AP3,* **G)** *PI1,* **H)** *PI2,* **(I-J)** C-class genes **I)** *AG1*, **J)** *AG2,* D-class genes **K)**, and **(L-S)** E-class genes, **L)** *SEP3-1*, **M)** *SEP3-2,* **N)** *AGL6,* **O)** *SEP1-1,* **P)** *SEP1-2,* **Q)** *SEP1-4,* **R)** *SEP1-5,* **S)** *SEP1-6.* Scale bar = 100 µm. Yellow dots represent a detected transcript.

The spatial co-expression of floral identity genes—including those encoding proposed ABCDE-class functions—provides a benchmark for evaluating MERFISH's capacity to resolve the combinatorial gene expression patterns underlying cell identity. We observed that specific combinations of transcription factors differentiate the eight floral and bract domains.

Lodicules, palea, glumes, and lemma all exhibit enrichment of *YABBY7*, a marker of lateral organs in rice (*OsTOB2,* Tanaka et al. 2017). While glumes and lemmas cluster together in ED1+2, the separation of palea and lodicules into ED15 and ED17 is distinguished in part by the enrichment of E-class *AGL6*. This expression pattern is consistent with the classification of palea and lodicules as the first two whorls in floral tissues, in contrast to the bract-like (non-floral) identity in glumes and lemmas.

According to the ABCDE model, B-, C-, and E-class genes specify stamens (whorl 3), while C- and E-class genes determine carpels (whorl 4). In line with this, we observe the enrichment of B-class genes PI-1 and PI-2, as well as AP3, in stamen domains (ED10, ED14), and their absence in carpel domains (ED16). Carpel and stamen domains are also distinguished through the domain-specific expression of E-class genes. Both tissues express the E-class genes *SEP3-1* and *SEP3-2;* however, carpels are uniquely enriched in *AGL6*. Notably, restricted regions within the carpel domain express the *YABBY4* (ortholog to rice *DROOPING LEAF 1),* previously characterised to specify carpel midrib identity (Yamaguchi et al. 2004), suggesting fine-scale tissue patterning is captured within single MERFISH-defined domains.

Of particular interest is ED7, a domain located at the base of floral organ primordia. This domain is enriched in B-class encoding gene *AP3,* E-class encoding *AGL6,* as well as transcription factors outside of the ABCDE model, including *LEAFY (LFY)* and *KNOX5 (KN5)*. The maize *LFY* orthologs (*ZmFL1* and *ZmFL2*) are expressed in floral meristems and within the floral organ primordia as they develop (Bomblies et al. 2003). *KNOX5* orthologs, such as maize KNOTTED1, are floral meristematic markers (Jackson et al. 1994), and the rice AGL6 ortholog (MFO1) is also characterised as being expressed in floral meristems. Both the domain positioning and gene enrichment suggest that these cells are meristematic floral tissues and are therefore given the annotation name: *floral organ primordia base*.

3.3.3 - The spatial restriction of *VRT2* and *SEP1-4* is disrupted in the *VRT-A2b* mutant

We previously identified differences in gene expression between spike sections using microdissection. *VRT2* was most highly expressed in basal sections, with lower expression in the apex, whereas *SEP* MADS-box transcription factors displayed the opposite pattern (Backhaus et al. 2022). *VRT2* is proposed to disrupt SEP–SQUAMOSA complex formation, critical for normal spikelet development, through a protein competition model (Li et al. 2021b). However, the co-expression of *VRT2* and *SEP* genes remains uncharacterised. To quantify these profiles, we computationally dissected $P1^{WT}$ spikes into 30 transverse bins at W4, confirming opposing expression patterns along the apical-basal axis to a finer scale than microdissections. We confirmed the high expression of *VRT2* at the base of the inflorescence. Mean counts per cell of *VRT2* in bins 1-10 (basal) are 3.4x higher than in bins 11-30 (central/apical; Figure 3.7A). Additionally, MERFISH results revealed the spatial restriction of these two genes across the inflorescence. *VRT2* was primarily expressed in ED3 developing rachis cells, with 32.2% of ED3 cells expressing *VRT2*, and showed minimal expression in spikelet tissues such as glumes/lemmas (ED2, 3.3%). In contrast, *SEP1-4* was largely absent from ED3 (1.4%) but enriched in spikelet tissues, including glumes/lemmas (ED2, 31.1%; Fig. 3.7C). Across the $P1^{WT}$ spike, only 0.7% of cells co-expressed *VRT2* and *SEP1-4*.

We next asked if MERFISH could quantify the mis-expression of *VRT2* in $P1^{POL}$, a near-isogenic line carrying the *VRT-A2b* allele. MERFISH revealed ectopic *VRT2* expression and disruption of the heightened *VRT2* signal at the base of the inflorescence, whereas the *SEP1-4* expression pattern remained intact (Figure 3.7B). This ectopic expression pattern led to increased co-localisation of *VRT2* and *SEP1-4*, with 8.2% of cells co-expressing both transcripts along the spike. Co-expression was most pronounced in tissues exhibiting the strongest phenotypic effects in $P1^{POL}$, glumes and lemmas, where 26.3% of ED2 cells co-expressed both genes (compared to 1.1% in $P1^{WT}$; Figure 3.7C-D). These findings establish the low rates of co-expression in *VRT2* and *SEP1-4*, which is disrupted in *VRT-A2b* mutants, and demonstrate the ability of MERFISH to detect tissue-specific changes in gene co-localisation in a developmental mutant.

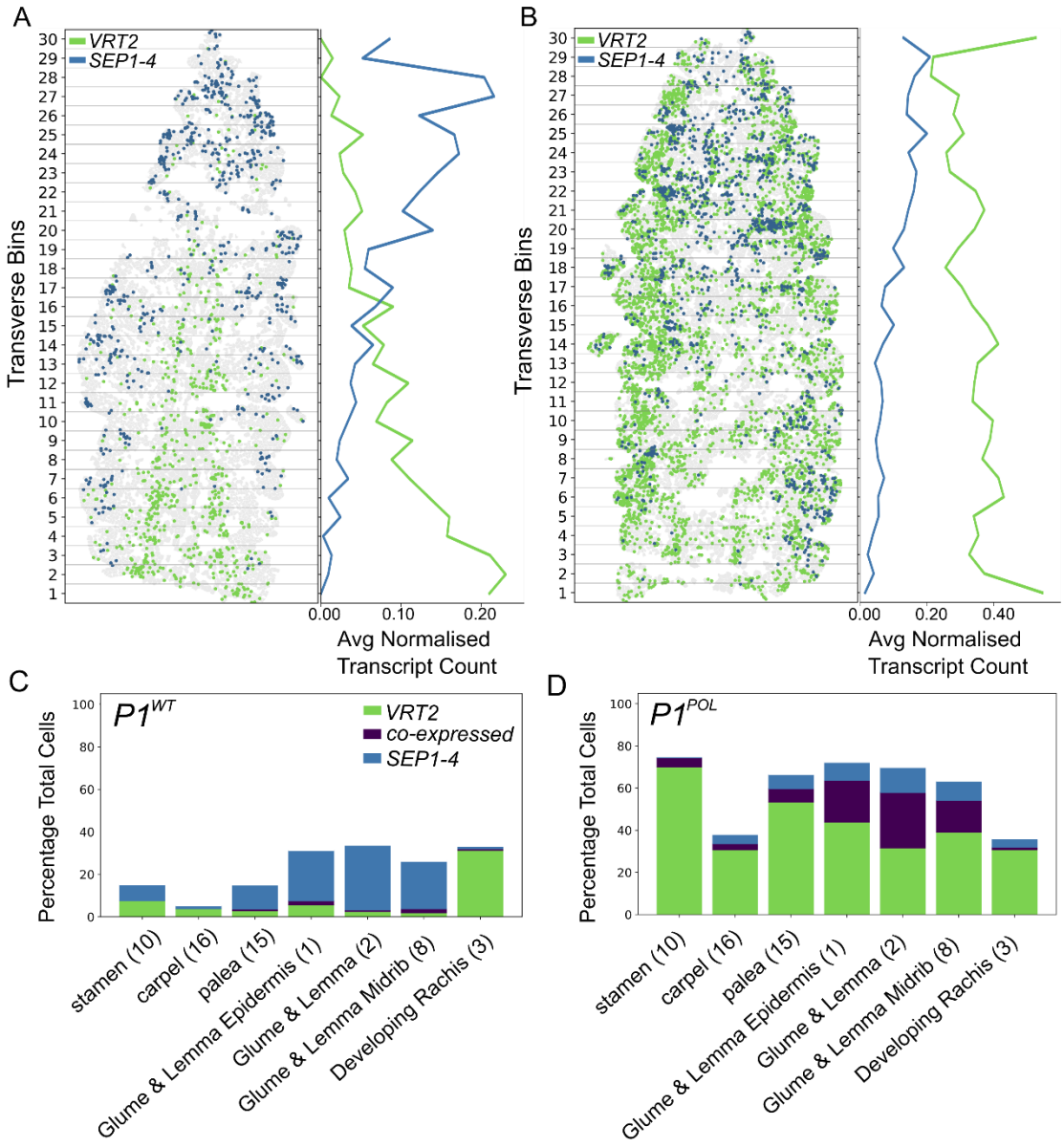


Figure 3.7 - Opposing and spatially restricted gene expression patterns of *VRT2* and *SEP1-4* are disrupted in *P1*^{POL} lines

A, B) Spatial plot of *VRT2* and *SEP1-4* expression in (**A**) *P1*^{WT} and (**B**) *P1*^{POL} W4 spikes, divided into 30 transverse bins along the apical-basal axis, with average normalised expression counts of *VRT2* and *SEP1-4* per transverse bin, normalised with sc.pp.normalize_total and sc.pp.log1p functions (Scanpy v1.10.0**).** Note the difference in scale between (**A**) and (**B**). **C, D)** Proportion of cells expressing either *VRT2, SEP1-4,* or co-expressing both. Calculated as the percentage of total cells per cluster type in (**C**) *P1*^{WT} and (**D**) *P1*^{POL} near-isogenic lines.

3.3.4 - Transcriptional states differentiate spikelet meristems and leaf ridges along the spike

The chronological initiation of AMs does not coincide with their developmental progression. Basal AMs, though first to initiate, lag in development compared to central AMs. By the glume primordium stage (W3), central AMs display visible outgrowth, while basal AMs remain less developed (Bonnett 1966). This is evident from the emergence of glumes, the lateral organs marking SM identity. Due to these morphological differences, we anticipated that central meristems would exhibit a distinct transcriptomic signature, reflecting their transition toward SM identity and the initiation of lateral organ formation. Consistent with this, we observe *FZP* expression adjacent to glume tissues (ED1+2) in central spikelets of $P1^{WT}$ inflorescences at W3 (Figure 3.8). Across the grasses, *in situ* hybridisation shows the *FZP* ortholog is confined to the glume–meristem boundary (Chuck et al. 2002; Poursarebani et al. 2015). Here, we detected a comparable pattern, with *FZP* localised adjacent to ED1+ED2 cells. Given the established role of *FZP* in promoting SM identity across grasses, these results indicate a transcriptional distinction between central and basal meristems by W3.

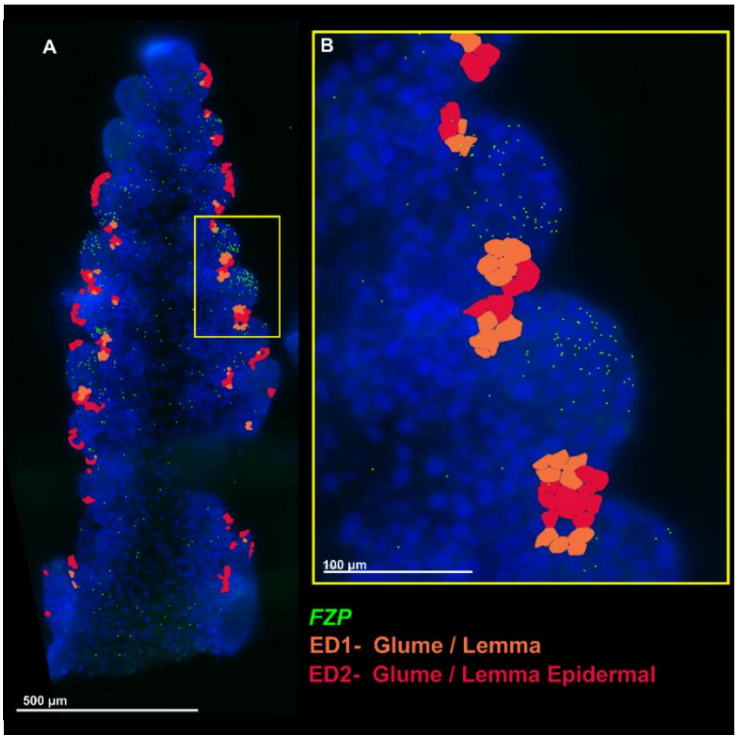


Figure 3.8 - Expression of *FZP* in central spikelets differentiates meristems across the apical-basal axis of the inflorescence
Expression of *FZP* in A) whole inflorescence, and B) central spikelets of $P1^{WT}$ *W3* sample. DAPI = blue. Scale bar = 500 µm.

Therefore, we hypothesised that gene expression patterns may influence these differences before W3, at stage W2.5, when the IM initiates AM + bract pairings. At this stage, the spike has a relatively simple ED composition, with four domains accounting for 94.8% of the inflorescence cells. By contrast, at W3.25, eight domains account for a comparable proportion of cells (94.1%). The W2.5 AM comprises four domains: the L1 layer (ED4), meristematic cells in layers L2/L3 (ED0, ED12), and boundary cells (ED11) marking the adaxial boundary. While all SRs exhibit similar L1 (ED4) and boundary (ED11) patterns, basal AMs lack well-defined ED0 regions (Figure 3.9). Additionally, while all LRs are represented by one domain (ED12), basal LRs are larger, averaging 32.5 ± 15.8 cells per section (LR1-4), compared to 12.5 ± 1.5 cells in central LRs (LR8-11). These findings support the idea that gene expression patterns may differentiate AMs and LRs across the apical-basal axis during or before W2.5.

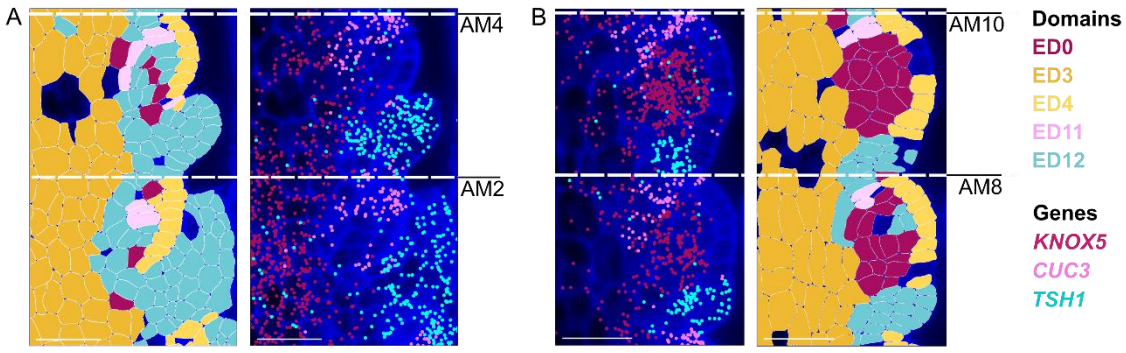


Figure 3.9 - Basal and central axillary meristems differ in expression domain assignment
Domain and gene expression maps of **A)** Basal AMs (2,4) and **B)** central AMs (8,10). *TSH1* and *CUC3* mark the suppressed leaf ridge and the adaxial boundary of the AM, respectively. *KNOX5* in layers L2/L3 highlights transcriptional differences between AMs along the apical-basal axis. Scale bar = 50 µm.

To explore gene expression patterns across the apical-basal axis in W2.5 inflorescences, we first examined the expression of key regulators of the vegetative-to-reproductive transition, particularly those involved in establishing AMs with a reproductive identity (i.e., a transition to SMs without reversion to branch-like or vegetative states). Using expression domains, we defined the onset of the inflorescence at the first appearance of leaf ridges (ED12; Figure 3.10A). A clear transcriptional distinction emerged between vegetative and inflorescence phytomers. *SEP1-6* (ortholog of *PAP2*) is expressed in the inflorescence in a broad expression pattern (Figure 3.10B), and *LAX1* is expressed in inflorescence phytomers only, in a restricted expression pattern adjacent to AMs (Figure 3.10C). In contrast, *TAW1* expression was restricted below the inflorescence in the region of suppressed AMs, closely paralleling the pattern of *MND1* (Figure 3.10D-E). Among the *AP1/FUL-like* genes, *VRN1* and *FUL3* were broadly expressed across vegetative and reproductive tissues, whereas *FUL2* expression was limited to the inflorescence (Figure 3.10F-H). The *SVP*-like transcription factors *VRT2* and *SVP1* were expressed primarily below the inflorescence, but both also exhibited activity in the basal-most inflorescence region and the developing rachis (Figure 3.10I-J). Given this basal restriction of two genes associated with the delay of transition of AM to SM, we anticipated that the most basal meristem-bract pairings may show distinct gene expression patterns beyond *SVP* transcription factors.

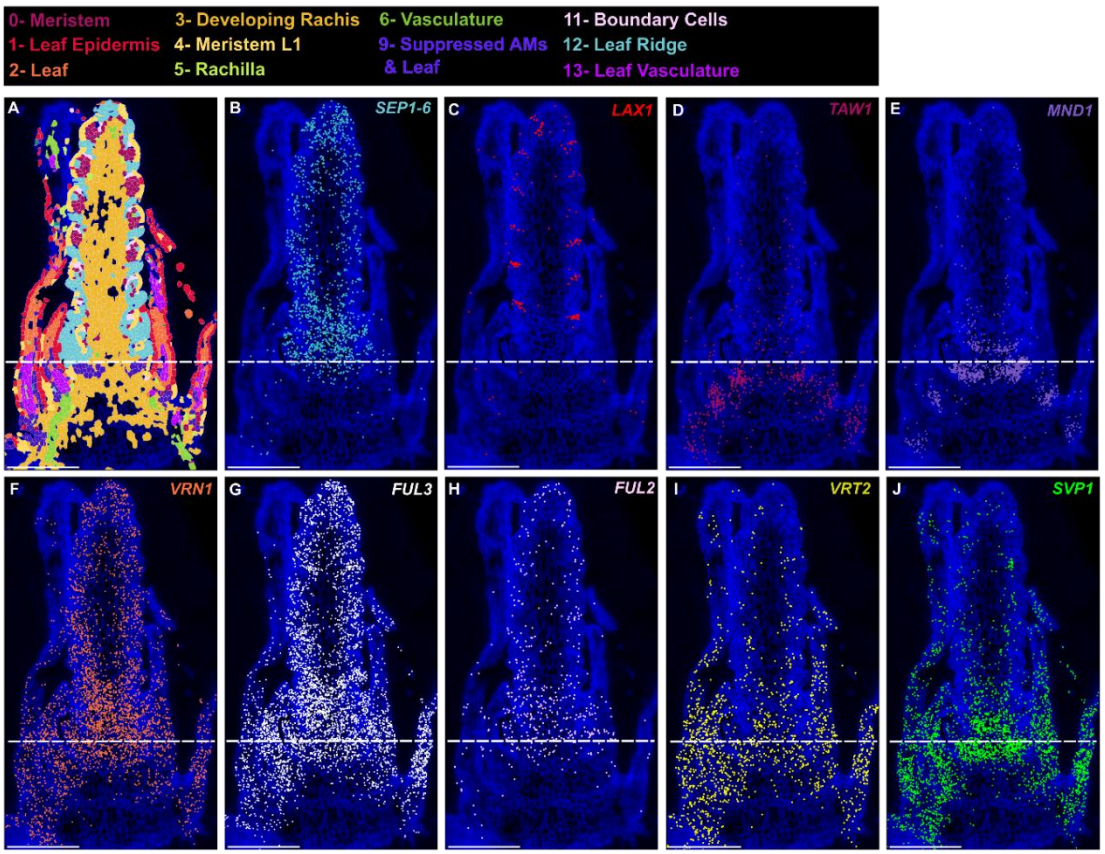


Figure 3.10 - Gene expression signatures differentiate vegetative and reproductive tissues in late double ridge inflorescences

A) Expression domains clarify the start of the inflorescence, which we defined as the boundary between the leaf ridge (ED12) and suppressed axillary meristems in leaf axils (ED9). Transcript localisation in W2.5 $P1^{WT}$ samples of genes **B)** *SEP1-6*, **C)** *LAX1,* **D)** *TAW1*, **E)** *MND1,* **F)** *VRN1,* **G)** *FUL3,* **H)** *FUL2,* **I)** *VRT2,* **J)** *SVP1*. Scale bar = 250 µm. The white dashed line represents the boundary between inflorescence and vegetative tissues, determined by expression domains.

To quantify differential gene expression patterns within the same tissue-type across the bract-meristem pairs across the inflorescence, we grouped cells from each AM, ordered them longitudinally from basal (1) to apical (13; Fig. 3.11A), and performed principal component (PC) analysis using a matrix of average counts per cell in each cell grouping (Fig. 3.11B). PC analyses revealed a strong positive correlation (ρ = 0.68, P = 0.01) between PC1 and location along the apical-basal axis, with lower PC1 scores associated with basal SRs. Given this relationship, we investigated genes with the highest and lowest loading scores on PC1. Genes with the lowest PC1 loadings, such as *AINTEGUMENTA-LIKE6* (*AIL6*, -0.177), *KNOX5* (-0.175), *HOMEOBOX DOMAIN 1* (*HB1*, -0.170), *INCREASED LEAF INCLINATION 3* (*ILI3*, -0.167), are expressed highly in central/apical AMs (Figure 3.11C; Figure 3.12). In the case of *KNOX5,* we observed high levels of expression in the meristem corpus of central SRs, which is absent in basal SRs (Figures 3.9 and 3.12D). Additionally, we observed the specific expression of ortholog to barley *COMPOSITUM 1 (HvCOM1)* to ED11 cells only in SR7,9,10. In contrast, genes with the highest PC1 loadings, including *RACHIS-LIKE1* (*RIL1, 0.186*) and *SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE 14* (*SPL14,* 0.180), showed higher expression in basal AMs (Figure 3.12).

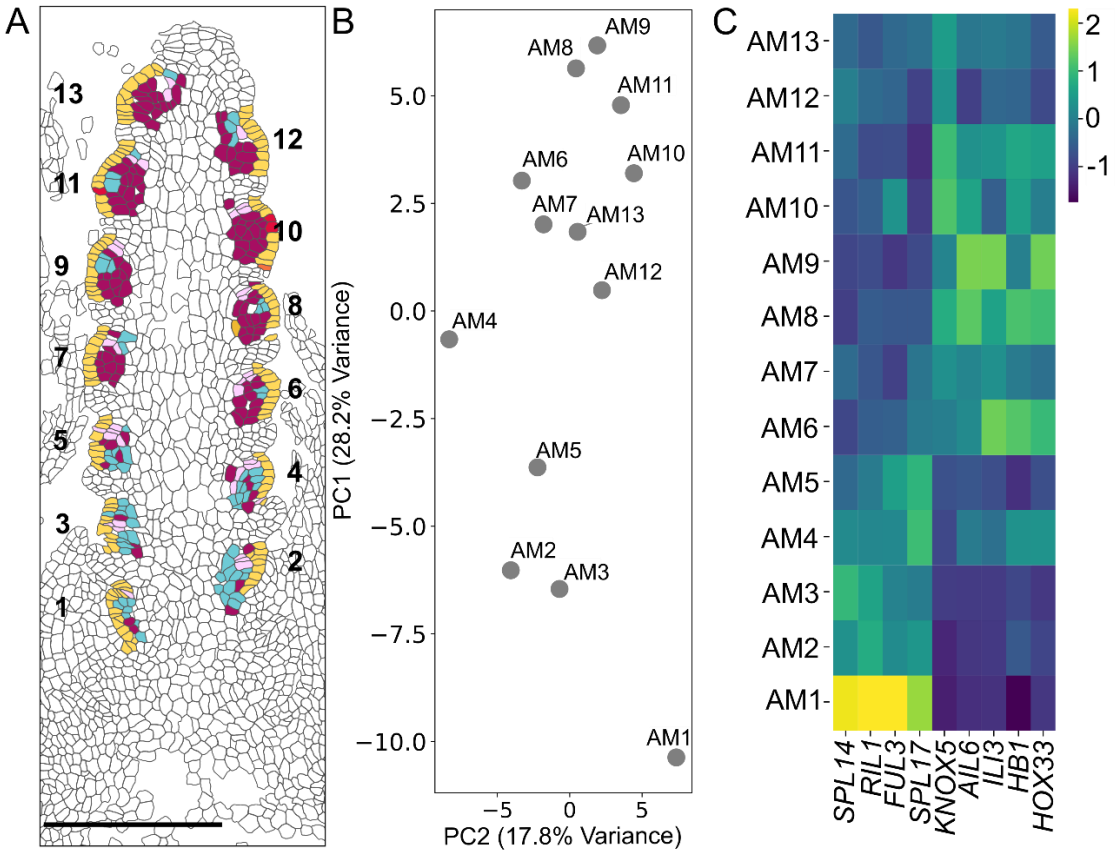


Figure 3.11 - Gene expression patterns define spikelet ridges across the apical-basal axis at the late double ridge stage

A) Expression domains define SRs from 1 (basal) to 13 (apical). **B)** Principal component (PC) analysis of averaged transcripts per SR group. **C)** Normalised gene expression (Z-score) of select genes with the highest/lowest PC1 loading scores from analysis in (B). Scale bar = 250 µm.

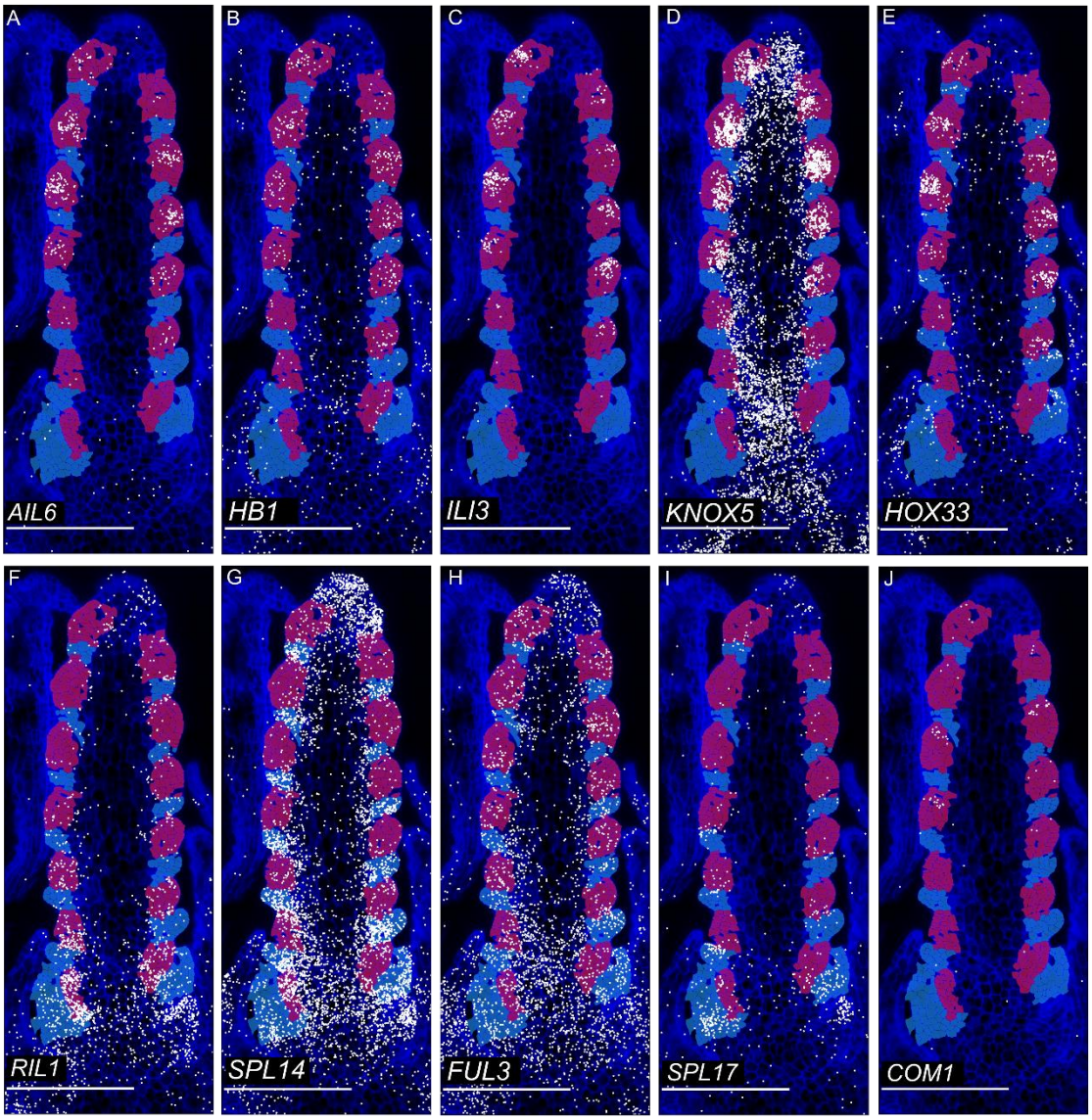


Figure 3.12 - Expression of select genes defining AM across the apical-basal axis at the late double ridge

(A-I) Transcript location of genes **A)** *AIL6*, **B)** *HB1*, **C)** *ILI3,* **D)** *KNOX5*, **E)** *HOX33,* **F)** *RIL1,* **G)** *SPL14,* **H)** *FUL3,* **I)** *SPL17* **J)** *COM1*. Cells in each leaf ridge are highlighted in blue, and axillary meristems are in pink. Scale bar = 250 µm. White dots = transcripts.

The equivalent PC analysis on LRs also revealed a correlation between PC1 and position along the apical-basal axis (ρ = 0.97, P = 2.62e-08; Fig. 3.13A-B). Genes with the highest PC1 loading score, such as *LEAFY COTYLEDON1* (*LEC1*, 0.179), *APETALA2-5* (*AP2-5,* 0.179), *SPL17* (0.171), and *RIL1* (0.159), were more highly expressed in basal LRs, whereas *AGL14* (-0.171) is more highly expressed in central/apical LRs (Figure 3.13C; Figure 3.14). Some genes distinguish between basal and central/apical sections for both LRs and SRs. Ortholog to *SHORT PANICLE 3 (SP3)* is expressed in distinct bands in central/apical regions of both SRs and LRs (Figure 3.14A), while *RIL1* marked both tissues at the base. Notably, *LEC1* was expressed in basal SR1–3 and LR1–3, but from position four onward was restricted to AMs (Figure 3.14D). In contrast, *SPL17* showed LR-specific expression in a position-dependent manner, expressed only in basal LRs (Figure 3.14E). Together, these results provide strong evidence that apical-basal patterning in the developing spike is established before W3 through distinct, tissue-specific gene expression gradients that differentiate SR and LR primordia, and highlight key regulators—such as *RIL1*, *LEC1*, and *SPL17*—that contribute to early axial patterning across multiple cell types.

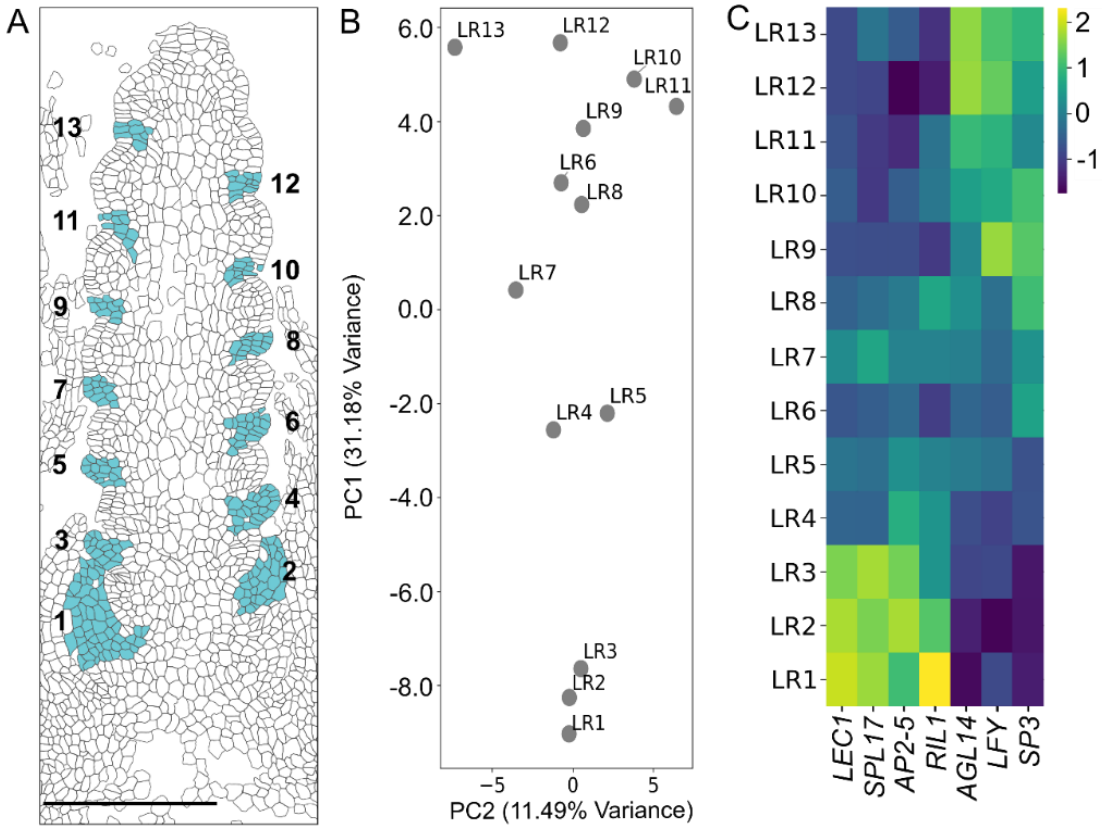


Figure 3.13 - Gene expression patterns define leaf ridges across the apical-basal axis at the late double ridge stage

A) Expression domains define leaf ridges (LRs) from 1 (basal) to 13 (apical). **B)** PC analysis of averaged transcripts per LR group. **C)** Normalised gene expression (Z-score) of select genes with the highest/lowest PC1 loading scores from analysis in (H). Scale bar = 250 µm.

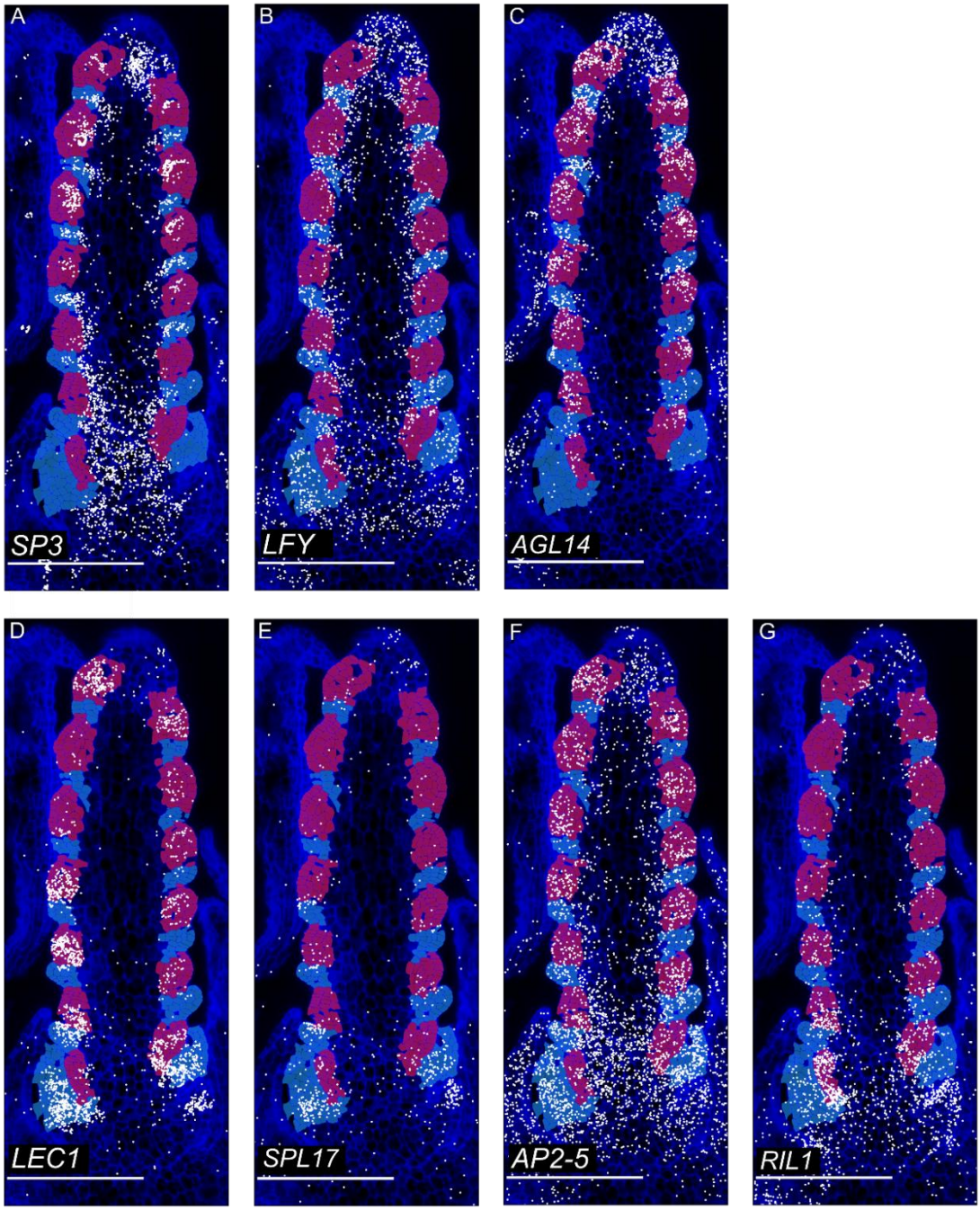


Figure 3.14 - Expression of select genes defining LR across the apical-basal axis at late double ridge

(A-G) Transcript location of genes **A)** *SHORT PANICLE 3* (*SP3*), **B)** *LEAFY 2* (*LFY2*), **C)** *AGAMOUS-LIKE 14 (AGL14),* ***D)*** *LEAFY COTYLEDON1* (*LEC1*), **E)** *SPL17,* **F)** *AP2-LIKE 5* (*AP2-5*), **G)** *RIL1*. Cells in each leaf ridge are highlighted in blue, and axillary meristems are in pink. Scale bar = 250 µm. White dots = transcripts.

3.4 - Discussion

3.4.1 - Expression domains trace phytomer units across developmental time

Using domain assignment and gene enrichment, we quantified the transcriptional programmes distinguishing vegetative and inflorescence phytomers in wheat, expanding upon genetic characterisation across the grasses. In barley, *MND1* is specifically expressed during the transition from the SAM to the IM and otherwise is restricted to AMs below the inflorescence (Walla et al. 2020). This is largely in agreement with the expression patterns observed in wheat. At stage W2.5 in vegetative tissues below the inflorescence, AMs (ED9) located in the axils of leaves (ED1, ED2, ED13) are enriched in *MND1*. Given our limited sample coverage during the transition to an IM (Waddington stages 1.5-2), it is currently unknown whether this pattern is in complete agreement. At W2.5, a similar expression pattern is observed for *TAW1*. In wheat, *TAW1* expression is detected in suppressed axillary meristems below the inflorescence but is absent from the inflorescence meristem (IM) and spikelet ridges. This contrasts with rice, where *TAW1* is expressed in axillary meristems of vegetative tissues as well as in the IMs and BMs, but is excluded from incipient SMs (Yoshida et al. 2013). The restriction of *TAW1* expression to primarily below the inflorescence in wheat may indicate neofunctionalization of *TAW1* to a primary role in the regulation of AMs in vegetative phytomers, however; earlier-stage expression data and further genetic characterisation are needed to clarify its role in the transition between vegetative to reproductive growth. In wheat, *MND1* and *TAW1* expression persists in ED9 cells below the inflorescence at stages W3.25 and W4, perhaps indicating these genes have a continual role in the suppression of meristem outgrowth in vegetative phytomers as inflorescence development continues. This data is consistent with unpublished results from fellow PhD student Isabel Faci in the lab, who has found that between lemma primordia (W3.25) and terminal spikelet (W4), the AMs below the inflorescence are susceptible to environmentally induced differentiation, resulting in the outgrowth of the AMs to pattern leaves, spikes, or tillers. Taken together, these results suggest an opportunity to further define the genetic network conferring a suppressed axillary meristem below the inflorescence.

In contrast, inflorescence phytomers are characterised by the presence of spikelet ridges (SRs; ED0, 4, 11, 12), which alternate with leaf ridges (LRs; ED12). The bract suppression gene *TSH1*, conserved across maize, rice, and barley (Wang et al. 2009; Whipple et al. 2010; Houston et al. 2012), is specifically expressed in LRs, highlighting a shift in lateral

organ fate. This shift coincides with *LAX1* expression in the adaxial boundary of AMs. While *CUC3,* another boundary-related gene, is expressed across both vegetative and inflorescence phytomers, *LAX1* expression is unique to reproductive phytomers. In rice, *LAX1* is required for the initiation and maintenance of AMs in the panicle (Komatsu et al. 2001, 2003), underscoring the unique gene expression signature of reproductive phytomers. At stage W2.5*, SEP1-6*, the wheat orthologue of rice *PAP2*, is broadly expressed across inflorescence phytomers. Additionally, while the *AP2/FUL*-like transcription factors *VRN1* and *FUL3* were broadly expressed, *FUL2* was restricted to the inflorescence. In rice, *PAP2*, together with *AP1/FUL*-like genes, coordinates the vegetative-to-inflorescence meristem transition (Kobayashi et al. 2012). This shift—from suppression of AMs in vegetative phytomers to suppression of LRs and promotion of AM progression in inflorescence phytomers—reflects a fundamental developmental reprogramming event in the wheat shoot apex at the onset of inflorescence formation.

By integrating data across four developmental stages prior to clustering, we could trace how these domain-level programs are deployed along the apical–basal axis of inflorescence phytomers as the spike acquires its lanceolate form. For example, ED12 cells (LRs) were present throughout the inflorescence at W2.5 and W3.25 but became restricted to the basal region by W4 and W5, suggesting *TSH1* mediated bract suppression continues in basal LRs. Similarly, the lanceolate shape can be observed through the transcriptomic signatures marking the transition from SM to FM identity, and the elaboration of floral organs. At stage W3.25 in central spikelet meristems, distinct bands of *COM2/FZP* expression are observed, characterised as a marker of SM identity following formation of the glumes (Poursarebani et al. 2015). In parallel, a small population of cells within a central spikelet expressed *AGL6*, a gene associated with palea identity in wheat (Kong et al. 2022), likely marks the earliest emergence of palea cells (ED15). By W4, central spikelets also contained ED10, ED14, and ED16 domains (associated with stamen and carpel identity), and were enriched for *AG1*, *SEP3-1*, *SEP3-2*, key C- and E-class floral regulators (Cui et al. 2010; Dreni et al. 2011). In contrast, these floral domains were absent from basal spikelets at both stages, reflecting a developmental gradient along the spike and delayed floral progression in the basal region.

3.4.2 - MERFISH quantifies gene co-expression key to development

By detecting hundreds of transcripts simultaneously, MERFISH enabled a tissue-specific analysis of gene co-localisation. We used this to investigate how a cis-regulatory mutation in *VRT2* (*P1^{POL}* allele), previously shown to increase *VRT2* expression (Adamski et al. 2021; Backhaus et al. 2022), alters its spatial expression. MERFISH revealed the *P1^{POL}* allele drives ectopic *VRT2* expression into spikelet and floral tissues, resulting in elevated co-localisation with *SEP1-4*, potentially inhibiting the formation of SEP-SQUAMOSA protein complexes required for spikelet development (Li et al. 2021b). In this context, the utility of MERFISH was not in quantifying relative transcript abundance (which can be captured by bulk RNA-seq), but in resolving co-expression networks as spatial domains are reprogrammed in a regulatory mutant. Additionally, while co-expression of *VRT2* and *SEP1-4* could be inferred from snRNA-seq, MERFISH uniquely revealed both the extent of their ectopic co-expression and their spatial distribution, providing mechanistic insight into how *P1^{POL}* alters spikelet architecture.

Similarly, MERFISH revealed the co-localisation of MADS-MIKC transcription factors in wheat floral tissues, allowing us to connect gene co-expression with floral organ identity. The ABCDE model was an ideal application of a multiplexed spatial technique- this model alongside the subsequent floral quartet model describes the specification of floral organ identity through spatially restricted and combinatorial gene expression (Coen and Meyerowitz 1991; Colombo et al. 1995; Honma and Goto 2001; Mohanty et al. 2022), which lacks thorough characterisation in wheat compared to other model species. Here, we observe the high levels of conservation in B- C- and D- class gene expression across species, indicating their conserved function outlined in the ABCDE model.

These tissue-specific patterns align with phenotypes of homeotic mutants in grasses. For example, B-class genes *AP3*, *PI-1,* and *PI-2* are expressed in lodicules and stamens, consistent with the homoeotic transformation of these organs in rice *ap3* mutants (lodicules to palea-like structures, stamens into carpels (Nagasawa et al. 2003) and in maize *silky1* mutants (Ambrose et al. 2000). Similarly, the expression of C-class *AG1* and *AG2* are expressed in the stamen and carpel, consistent with the rice double mutant (orthologs *osmads3* and *osmads58*), which forms lodicule and carpel-like organs in place of the stamens and carpel (Dreni et al. 2011; Sugiyama et al. 2019). Additionally, *AGL6* distinguishes palea cells (ED15) from glumes and lemmas (ED1+2) in clustering analyses, and is enriched in palea, carpel, and lodicules. In tetraploid wheat, *AGL6* double mutants

produce lemma-like paleas, lack lodicules, develop ectopic organs between the second and third floral whorls, and display loss of carpel determinacy (Kong et al. 2022).

However, MERFISH detected the divergence in gene expression in E- class genes, even among monocots, in direct contrast to their modelled function. Five *SEP1* orthologs- members of the *LOFSEP* gene family in wheat are expressed throughout the spike; however, are mostly absent from floral organ primordia- with most expression observed in glume and lemma primordia tissue. This is consistent with the loss-of-function phenotypes observed in barley and rice- whereby loss of *LOFSEP* members disturbs lemma development and inflorescence branching, while inner floral organs remain unaffected; suggesting further genetic characterisation in wheat may display a similar phenotype (Gao et al. 2010; Wu et al. 2018; Li et al. 2021a; Zhang et al. 2024b; Shen et al. 2025).

Among the E-class genes expressed in floral tissues, *SEP3-2*, and *AGL6* display overlapping, but not fully redundant, expression across all floral organs. In wheat, *AGL6* is expressed in palea, lodicules, and carpels. These spatial patterns align with the phenotypes reported in *mads6* mutants in rice, which exhibit altered palea morphology and homoeotic conversion of lodicules and stamens into glume-like organs (Li et al. 2010). In tetraploid wheat, *AGL6* double mutants produce lemma-like paleas, lack lodicules, develop ectopic organs between the second and third floral whorls, and display loss of carpel determinacy, with some carpels developing into spikelet (Kong et al. 2021). While the functional significance of the partially overlapping expression between E-class genes remains unresolved, a similar expression pattern has been recently observed in barley floral tissues using spatial transcriptomics (Demesa-Arevalo et al., 2025), suggesting conservation of this regulatory arrangement across cereals.

In wheat, the function of A-class genes—postulated to specify sepal and petal identity—remains unclear based on current genetic characterisation. Our MERFISH dataset shows that orthologs in the *AP1/FUL*-like clade, including *VRN1* and *FUL3*, are broadly expressed across all floral whorls, whereas *FUL2* expressed primarily in lemmas and glumes. These spatial patterns are consistent with mutant phenotypes: *vrn1 ful2* double and *vrn1 ful2 ful3* triple mutants exhibit de-repression of the leaf ridge into a true leaf subtending a spikelet ridge that develops into a vegetative tiller, along with floral defects such as leafy paleas and lodicules (Li et al. 2019). These phenotypes implicate *VRN1, FUL2,* and *FUL3*

in a broader role coordinating meristem identity and phase transition in addition to potential A-class functions (Li et al. 2019).

In the case of E-class *AP2* genes, the two wheat orthologs *AP2-2* and *AP2-5* are expressed in the lemma, palea, lodicule, and stamen primordia, or broadly across all floral tissues, respectively. Given its co-expression with E-class genes in the palea and with B- and E-class genes in the lodicules, *AP2-2* may fulfil an A-class function. In *ap2l2*-null mutants, floral organ identity remains largely unaffected. However, *ap2l2 ap2l5* double mutants produce multiple empty bracts before transitioning to florets, and display organ defects including missing paleas and lodicules transformed into carpel-like organs (Debernardi et al. 2020). The upregulation of C-class *AG* genes in these double mutants suggests that the two AP2 genes function analogously to A-class genes, repressing C-class activity in the outer floral whorls (Causier et al. 2010).

3.4.3 - Gene expression patterns differentiate the apical-basal axis

Variation in domain composition and gene expression across SRs prior to visible spikelet initiation may reflect shifts in meristem identity along the apical–basal axis of the developing spike. Given limited functional characterisation in wheat, we infer putative gene function based on orthologous genes characterised in grasses. At W2.5, central SRs exhibit a transcriptional state distinct from basal SRs, marked by expression of genes including *KNOX5*, *AIL6*, and *HB1*. Orthologs of *KNOX5* (*ZmKNOTTED1*/ *OsOSH1*) are expressed in ground tissue and meristems in maize and rice- including in IM, BM, SM, FM- and are excluded from sites of determinate organ initiation (Jackson et al. 1994; Hake et al. 1995; Suzaki et al. 2006). In wheat, *KNOX5* shows a similar pattern yet is notably absent from the meristem corpus of basal SRs. In barley, *HvKN1* is similarly excluded from newly initiated meristems and reactivated upon acquisition of triple spikelet meristem (TSM) identity, suggesting a role in meristem phase transitions (Demesa-Arevalo et al. 2025). *AIL6*, associated with the indeterminate-to-determinate transition and FM identity (Nole-Wilson et al. 2005; Krizek 2009) and *LF1* (orthologous to wheat *HB1)*, a regulator of FM specification (Zhang et al. 2017), further support transcriptomic signatures of meristem transition in central SRs. Additionally, we observe the specific expression of *COM1* to SR7,9,10. In similar spatial transcriptomic studies, the barley ortholog *HvCOM1* is used as a marker of true TSM identity (Demesa-Arevalo et al. 2025). Collectively, these expression

patterns indicate that only central SRs acquire spikelet meristem competency at this stage.

We propose that the specific expression of *KNOX5*, *AIL6*, and *LF1* in central SRs reflects the progressive acquisition of spikelet meristem (SM) identity. In contrast, the identity of basal SRs remains less clear. Basal SRs are enriched in *RIL1* and *SPL14,* genes associated with BM identity in rice. Knockdown of *OsSPL14* reduces panicle branching and spikelet formation (Wang et al. 2015), while a heterozygous *ril1* mutation in a *ri* background- its close paralog- disrupts the spatial and temporal regulation of BM initiation (Ikeda et al. 2019). BM identity is not directly translatable to the unbranched wheat inflorescence, whereby AMs transition directly to SM identity without branching (Koppolu and Schnurbusch 2019). However, the expression of BM-associated genes in basal SRs may reflect a shared indeterminate transcriptional state preceding the transition to the determinate SM fate.

Although the functional consequences of these gene expression patterns have yet to be characterised in wheat, they may help explain phenotypes observed in branching mutants. In spike-type inflorescence such as wheat and barley, loss of function in key regulators such as *FRIZZY PANICLE (FZP)* results in the replacement of spikelets by branch-like structures, most prominent at the base of the spike (Dobrovolskaya et al. 2015; Poursarebani et al. 2015). Similarly, in barley, *mnd1* loss-of-function mutants show a reversion of TSMs to a BM-like identity, accompanied by an outgrowth bracts formed in the most basal phytomer units (Walla et al. 2020). In wheat, *vrn1-null/ful-A2-null Ful-B2* mutants (retaining only one functional copy of FUL2), frequently produce branch-like structures at the base of the inflorescence. In more severe v*rn1ful2ful3-null* mutants, branch-like AMs are often subtended by fully elongated leaves in basal phytomers, whereas central and apical AMs are subtended by bracts (Li et al. 2019). This branching pattern resembles the panicle-type inflorescence of rice, in which primary branches are longest at the base and gradually decrease toward the apex (Bommer and Whipple, 2018). Across species, these observations suggest that the position and timing of meristem–bract pair initiation strongly influence the determinacy and fate of lateral organs.

Genetic studies in grasses suggest that suppressed LRs can function as signalling centres regulating adjacent AM activity (Whipple et al. 2010). At stage W2.5, we detect expression of genes in LRs, whose orthologs involved in both bract suppression and promotion of AM development. For example, we observe expression of *TSH1* and *SPL17* in LRs—genes

whose orthologs in maize (*TSH1* and *TSH4,* respectively), suppress bract/LR outgrowth while regulating BM determinacy (Xiao et al. 2022). Interestingly, *SPL17* expression is restricted to basal LRs and absent from more central and apical LRs (positions 8–13), consistent with the basal specific expression of *RIL1* in LRs and AMs. Both genes have dual roles, in bract suppression and BM initiation (Ikeda et al. 2019; Xiao et al. 2022). These results indicate that the genetic program regulating bract suppression and meristem determinacy shifts along the apical–basal axis of the wheat spike, reinforcing the idea of their coordinated role in spike development.

3.5 - Methods

3.5.1 - Staining, segmentation, and transcript visualisation

We processed the cell segmentation data as GeoDataFrames (Geopandas v0.14.4; Jordahl et al. 2020), and converted the transcript coordinates into a GeoDataFrame from global x and y coordinates. We performed a spatial join operation to assign transcripts to segmented cells, retaining only transcripts located within cell boundaries. We loaded the DAPI staining image as a .tiff file, alongside a transformation matrix enabling conversion between pixel space to physical (micron) space, as generated by the MERSCOPE Instrument Software. The transformation matrix converted to micron space and applied the raw image using Scikit-image (v0.23.3, Walt et al. 2014). After transformation, the image was normalised, and image contrast and brightness was adjusted with Scikit-image. We next rotated segmented cell polygons and transcript coordinates using NumPy (v1.26.3, Harris et al. 2020), and visualised cell geometries as polygons using Matplotlib (v3.8.2, Hunter 2007) with polygon handling and transformations facilitated by Shapely (v2.0.4, Gillies et al. 2022). Transcripts were overlaid as point features. The corresponding image was rotated with Scipy.ngimage (Virtanen et al. 2020; v1.13.0). Full details in implementation scripts, see https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics.

3.5.2 - MERFISH data integration, unsupervised clustering, and gene enrichment analysis

We processed spatial transcriptomic data from eight samples (four timepoints; two NILs) using the Scanorama (Hie et al. 2019, 2024) integration tool, and performed clustering using the Leiden algorithm with a resolution parameter of 1.0. Spatial maps of Leiden cluster assignment were performed as described in 'Staining, Segmentation, and Transcript Visualisation'. We exported the expression domain (ED) assignment of each per cell (see archive folder 'SampleIntegration_Clustering.tar.gz' at doi.org/10.5281/zenodo.14515926). The number of domains representing each sample were summated, with domains representing less than 0.5% of total cells in sample removed. Next, we performed gene enrichment analysis on the integrated AnnData object with metal function sc.tl.rank_genes_groups() using the logistic regression model (Ntranos et al. 2019). This analysis returned a ranked list of genes most probable to be enriched

gene markers, which we displayed alongside the average normalised expressions per ED for each sample. We determined top enriched values (using a +2 standard deviation threshold) and used these to annotate EDs with tissue type identity labels. See archive folder 'GeneEnrichmentAnalysis.tar.gz' at doi.org/10.5281/zenodo.14515926 for domain assignments, gene enrichment analysis output, and domain annotations.

3.5.3 - Transect analysis of *VRT-SEP* gradients

We filtered cells from two samples (W4, *VRT-A2a* and *VRT-A2b NILs*) to include only cells from the inflorescence region, defined as the beginning of ED12 marking leaf ridges. These cells were selected in the MERSCOPE Visualizer tool (MERSCOPE Vizualizer 2023) using the Polygon Lasso Tool, exported as a .csv file, and the segmented cells and transcripts were mapped as previously described. The Y-axis of the spatial plot was divided into 30 transverse bins along the spike. Each cell was assigned to a bin based on its centre Y-coordinate, and we averaged the normalized transcript counts per cell within each bin (Fig. 4, A and B). For both samples, we binarized gene expression data for *VRT2* and *SEP1-4* within each cell, assigning a value of 1 for detected reads and 0 for no detected reads. For each ED, we quantified the number of cells expressing only *VRT2*, only *SEP1-4*, or co-expressing both genes and visualised them as a percentage with Matplotlib (Hunter 2007).

3.5.4 - Gene expression analysis on late double ridge spikes

We selected Late Double Ridge (W2.5) and Lemma Primordia (W3.25) $P1^{WT}$ inflorescence cells using the MERSCOPE Visualizer Polygon Lasso tool and exported cell identity data as a .csv file. We defined the inflorescence boundary by the first suppressed leaf ridge (ED12) and excluded cells outside the inflorescence. Cell counts were summed by ED, and we calculated the cumulative percentage of cells in the most populated ED to assess their contribution to the total cell population. We calculated the top EDs accounting for approximately 94% of the cells in the sample.

Groups of cells comprising the Leaf Ridges (LR) and Spikelet Ridges (SR) (defined by EDs) were annotated as "Custom Cell Groups" in the MERSCOPE Visualizer tool(MERSCOPE Vizualizer 2023). We delineated SR boundaries by ED4 cells along the adaxial and abaxial axes, extending to the start of ED3 cells along the medio-lateral axis. We identified LRs as groups of ED12 cells beginning beneath the end of ED4 cells from adjacent SRs. We labelled SRs and LRs sequentially from 1 (most basal) to 13 (most apical) along the

inflorescence (Fig. 5, C and H). We calculated the total number of cells in basal (LR1-4) and central (LR8-11) leaf ridges and determined the mean cell numbers and summary statistics to compare ridge sizes between these regions.

Normalized gene expression values per cell were averaged by LR or SR group and filtered to include only genes with at least one average expression score above 0.30 across all groups. We standardised the resulting data matrix using StandardScaler and performed PC analysis with scikit-learn (v1.4.2; Pedregosa et al. 2011), to extract the first two principal components (Fig. 5, D and I). We inverted PC1 and PC2 scores to align the axes with the desired biological orientation and calculated the Spearman's Correlation Coefficient between the position along the inflorescence (1-13) and PC1 scoring (Scipy.stats, v1.13.0). We extracted the top genes contributing to PC1 through PC1 loading scores, and calculated average expression (Z-score normalised) in each cell grouping and visualised using Matplotlib (v3.8.2; Hunter 2007) and Seaborn (v0.13.1; Waskom 2021). All scripts and supplementary data are available (https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics/ and doi.org/10.5281/zenodo.14515926).

Chapter 4 – Applying Stereo-seq to the wheat inflorescence

This work was conducted in collaboration with staff members at BGI, Latvia. The STOmics team at BGI performed the permeabilisation testing, full transcriptomic capture experiments, and sequencing at the BGI Facilities in Riga, Latvia. The STOmics bioinformatics team processed raw sequencing data with the SAW pipeline. Thank you to Lili Feng and YenYu Lin for coordinating the project, including overseeing the shipment of samples, organising my visit to the BGI facilities, overseeing data transfer, and providing consultation on results. JIC Horticultural Services facilitated plant growth.

4.1 – Chapter summary

In this chapter, I present the implementation of Stereo-Seq (SpaTial Enhanced REsolution Omics-sequencing) for the spatial transcriptomic profiling of wheat inflorescence tissues, representing the first application of this technique in wheat. I document the preparation of high-quality, fresh-frozen samples, in addition to our results from permeabilisation testing, which optimises the incubation times for a critical step in the Stereo-seq protocol. Following on from these optimisations, we performed a complete transcriptome capture experiment across three sequencing arrays on late double ridge (W2.5) inflorescence tissues. We observed an overall low mapping rate with the Stereo-Seq Analysis Workflow (SAW), with 63.3% of reads remaining unmapped to the wheat reference genome. Of the high-quality reads obtained, we observed a significant lateral diffusion of mRNA outside of tissue boundaries, with 78.5% of detected transcripts falling outside of tissue boundaries across the sequencing array. Due to a combination of these effects, the low number of detected transcripts in key developmental genes limited further analysis of this dataset.

4.2 – Introduction

4.2.1 – Sequencing-based spatial transcriptomics in plant tissues

Spatial transcriptomics techniques encompass a diverse set of methods that aim to measure gene expression while preserving positional context within tissues (Moses and Pachter 2022). Broadly, these methods can be divided into two categories: imaging-based and sequencing-based (Nobori 2025). Imaging-based approaches, derived from single-molecule RNA fluorescence *in situ* hybridisation (smFISH), rely on fluorescently labelled probes and microscopy to localise and quantify RNA molecules within cells. By contrast, sequencing-based approaches capture mRNA transcripts with specialised arrays, followed by RNA sequencing to generate spatially resolved maps of gene expression (Giacomello et al. 2017; Moses and Pachter 2022; Nobori 2025).

Sequencing-based spatial transcriptomic methods have advanced rapidly since their first demonstration in 2016 (Ståhl et al. 2016). In this first approach, tissue cryosections are placed on arrays patterned with spots (each ~100µm diameter) containing millions of capture oligonucleotides. These oligonucleotides are designed with three key features: a spot-specific barcode sequence for spatial localisation (coordinate identity sequence; CID), an oligo(dT) sequence for binding polyadenylated mRNA, and a semi-randomised unique molecular identifier (UMI) for distinguishing individual transcripts. Following their placement on the array, tissue cryosections are permeabilised to allow for the movement of polyadenylated RNA molecules onto the spot array, where they hybridise with the capture probes. Reverse transcription is then performed *in situ,* incorporating the spatial barcode and UMI into the resulting cDNA, which is subsequently processed for sequencing library preparation (Figure 4.1; Ståhl et al. 2016; Giacomello et al. 2017).

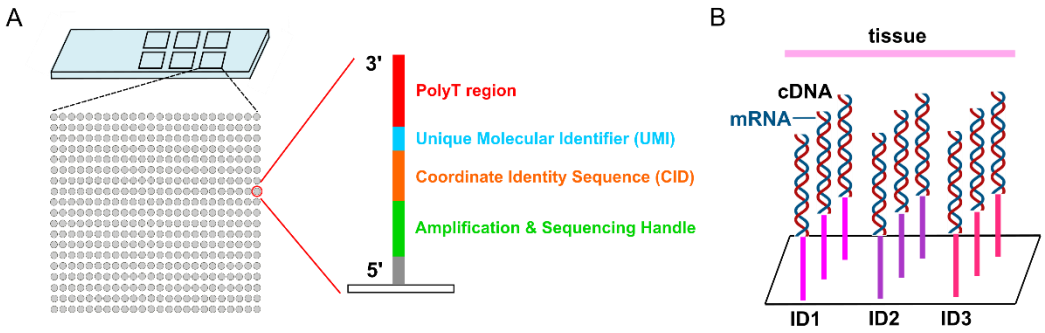


Figure 4.1 - Schematic representation of spatial transcriptomics through an array of barcoded oligos

A) A glass slide is prepared with capture oligonucleotides to generate thousands of capturing spots, each with a unique coordinate identity sequence (CID) assigned to its spatial location. **B)** mRNA molecules diffuse from the tissue onto the glass slide and are captured by oligonucleotide probes. cDNA is synthesised, incorporating the CID and MID sequences. Adapted from Giacomello et al., 2017.

The first application of a sequencing-based spatial method to plants was reported by Giacomello et al. (2017), who adapted the technique pioneered by Ståhl et al. (2016) for *Arabidopsis thaliana* inflorescence meristems, *Populus tremula* leaf buds, and *Picea abies* female cones. This work established key workflows for plants, including adaptations suitable for these tissues, such as changes to fixation, staining, and enzymatic permeabilisation. Importantly, it demonstrated the ability to capture gene expression differences between tissue domains in Arabidopsis; however, resolution was constrained by the 100 µm spot size of the array, limiting analysis to tissue-level rather than single-cell resolution (Giacomello et al. 2017).

The limitations of early low-resolution arrays (Ståhl et al. 2016; Giacomello et al. 2017) were soon addressed by emerging techniques, which improved spatial resolution (from 100 µm to a submicron scale), enhanced capture efficiency, and expanded the field of view. Visium (10x Genomics) reduced spot size to ~55 µm and increased the density of capture probes, thereby improving transcript recovery and enabling analyses at finer spatial scales. Slide-seq (Rodriques et al. 2019) further enhanced spatial resolution by replacing fixed capture spots with DNA-barcoded beads of ~10 µm diameter, positioned at high density and computationally mapped back to spatial coordinates, achieving near single-cell resolution. Stereo-seq (Chen et al. 2022) extended these innovations by using patterned DNA nanoball arrays with diameters measuring 220 nm, offering subcellular resolution while also providing centimetre-scale fields of view. Collectively, these

techniques have been successfully applied to diverse plant species, including *Arabidopsis thaliana* (Xia et al. 2022), *Solanum lycopersicum* (Song et al. 2023), *Glycine max* (Liu et al. 2023b; Zhang et al. 2024a), *Zea mays* (Fu et al. 2023; Wang et al. 2024), *Medicago truncatula* (Serrano et al. 2024) and *Triticum aestivum* (Liu et al. 2025; Qu et al. 2025). Notably, the applications of these techniques are novel to the field of plant science. All of these studies were published during the course of this thesis.

Sequencing-based spatial transcriptomic techniques offer clear advantages: unlike imaging-based techniques that are limited to a predefined gene set, sequencing-based approaches allow for an unbiased capture of the transcriptome (Shi et al. 2023a). The ability to capture a broad range of gene expression, without prior selection or indeed, without known characterisation, has a wide appeal. An additional advantage in plants is that sequencing-based methods require only that tissue sections be applied to a capture array, after which the tissue can be removed following mRNA capture. This circumvents challenges common to imaging-based FISH approaches, where high autofluorescence in plant tissues often necessitates additional optimisation (Donaldson 2020; Giacomello 2021).

4.2.2 – Stereo-seq for plant developmental biology

Building on advances in sequencing-based spatial transcriptomics in plants, I apply Stereo-seq to wheat inflorescence tissues to investigate the transcriptional programs underlying spike development. Stereo-seq (SpaTial Enhanced REsolution Omics-sequencing) is a DNA nanoball (DNB) based technology that enables transcriptome-wide capture of mRNA with both submicron resolution across large tissue areas (Chen et al. 2022). The method relies on a silicon-based patterned array (chip) densely coated with single-stranded DNBs, each ~220 nm in diameter and positioned at 500 nm intervals in a highly regular grid. DNBs are generated by rolling-circle amplification and carry unique coordinate identifiers (CIDs), which serve as spatial barcodes to define the position of each spot on the array (Chen et al. 2022).

To establish the location of each spatial barcode, *in situ* sequencing is first used to read the CID of each DNB (Figure 4.2A; Drmanac et al. 2010; Chen et al. 2022). Subsequently, oligonucleotide capture probes are hybridised to each DNB, which contains a UMI, providing a sequence associated with each captured mRNA molecule, and a poly(T) tail, enabling capture of poly(A)-tailed mRNA molecules (Figure 4.2B). Frozen tissue sections

are mounted directly on the chip, fixed, and permeabilised, allowing endogenous mRNA to hybridise to the poly(T) region of the probes anchored on the DNBs. The captured mRNA is converted into cDNA via reverse transcription, which incorporates both the CID and the UMI (Figure 4.2C). The resulting cDNA is then amplified and used as a template for library preparation, and sequenced (Chen et al. 2022). Computational analysis of the sequencing data determines the mRNA transcript identity and the precise spatial location on the array, yielding high-resolution gene expression maps across tissue sections (Gong et al. 2024).

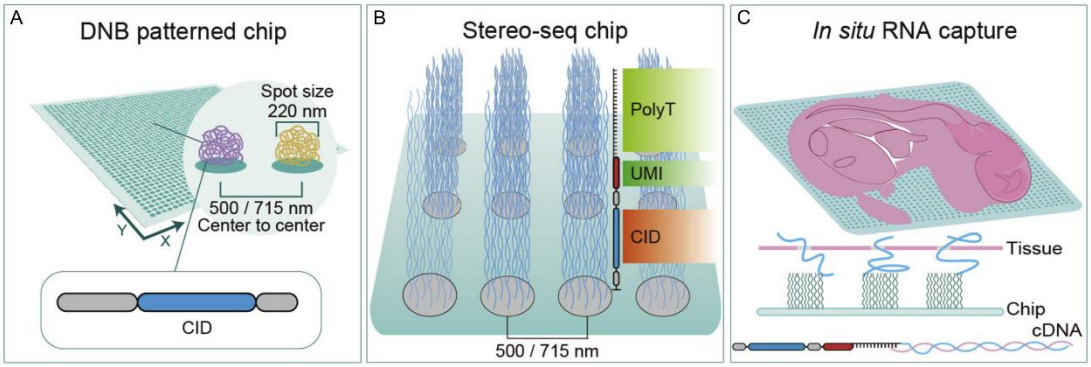


Figure 4.2 - Schematic of Stereo-seq experimental design
A) Patterned arrays contain regularly spaced DNB, each containing a CID determined by *in situ* sequencing. **B)** UMI-polyT containing oligonucleotide probes are ligated to each DNB spot. **C)** *in situ* RNA capture from tissue and cDNA generation. Figure adapted from Chen et al. 2022.

Although first established in mammalian systems (Chen et al. 2022), Stereo-seq was soon adapted to plants. In *Arabidopsis* leaves, Stereo-seq resolved transcriptomic differences between upper and lower epidermal cells and characterised gene expression gradients across the medio-lateral axis of the leaf (Xia et al. 2022). Its application across diverse plant tissues continues to expand, with recent studies demonstrating successful implementation in grass inflorescences, including maize ears (Wang et al. 2024) and wheat spikes and roots (Ke et al. 2025; Liu et al. 2025; Qu et al. 2025). As with other spatial transcriptomic studies, the field is advancing rapidly, with technological improvements and newly released datasets continually expanding our understanding of plant development. In this chapter, I present the first application of Stereo-seq to wheat inflorescence tissues, carried out before the method was widely implemented in plant systems. I describe the limitations encountered when applying the technology to plant samples and compare our results with more recent Stereo-seq datasets as well as with imaging-based MERFISH.

Chapter 4.3 - Results

4.3.1- Sample preparation and optimisation of wheat spike samples for Stereo-seq

The preparation of Stereo-seq requires fresh frozen tissue with high-quality RNA. We optimised a tissue dissection and embedding protocol to maximise RNA integrity and section quality. In brief, this required the dissection of wheat inflorescence tissues (Waddington stage W2.5; Figure 4.3A-B), followed by rapid embedding in O.C.T., ensuring a lack of air bubbles surrounding the tissue (Figure 4.3C). Using this technique, we obtained high-quality tissue cryosections in fresh tissues (Figure 4.3D). To test RNA quality following sample preparation and sectioning, we extracted RNA from sectioned fresh tissue, which yielded an average RIN score of 7.9, exceeding the requirements of RIN ≥ 6 recommended for Stereo-Seq.

Following the optimisation of fresh-frozen samples and cryosectioning techniques, we visited the BGI facilities in Latvia, where we performed the next steps. A further optimisation step involves tissue permeabilisation tests, which use fluorescent labelling and visualisation of captured mRNA to determine the ideal conditions that maximise mRNA captured onto the sequencing chip while preventing excess lateral mRNA diffusion or tissue damage. For this test, we performed tissue mounting onto four sequencing chips, followed by incubation of the tissue in permeabilisation reagents tested at four time points (6, 12, 18, and 24 minutes). Following permeabilisation, we incorporated fluorescently labelled nucleotides into the captured mRNA by reverse transcription, digested the tissue away from the chip, and imaged the resulting signal from the fluorescently labelled nucleotides to determine the localisation and quantity of mRNA released to the chip. A strong signal with minimal lateral diffusion indicates quality permeabilisation conditions. Based on our results (Figure 4.4), we determined an 18-minute permeabilisation to be the optimal time based on the strength of fluorescence and the minimal diffusion outside tissue boundaries, which became more evident in the 24-minute permeabilisation sample.

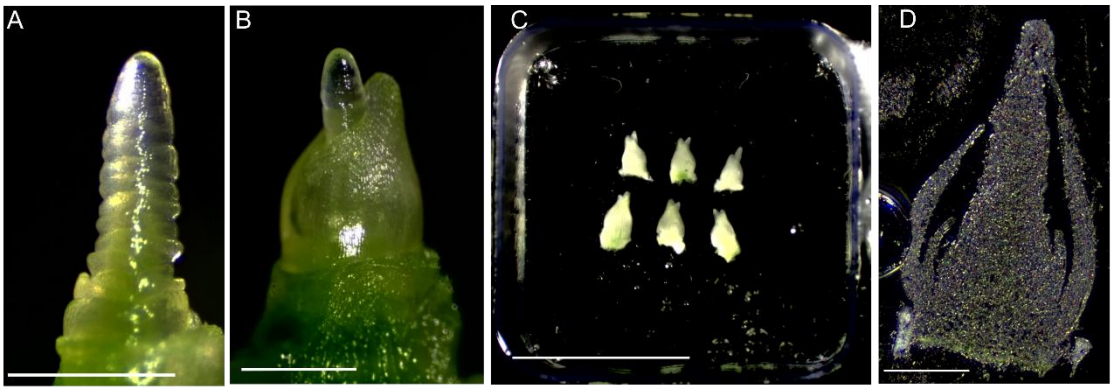


Figure 4.3 -Initial tests of sample preparation and cryosectioning yield high-quality samples of wheat inflorescence tissue

A) Waddington stage 2.5 (double ridge) spikes were dissected for use in Stereo-Seq. Scale bar = 0.5mm. **B)** For ease in sampling and cryosectioning, inflorescence tissues were dissected with surrounding leaf tissue. Scale bar = 0.5 mm. **C)** Orientation of 6 double ridge spikes in a 1 cm x 1 cm mould embedded in OCT before flash freezing. Scale bar = 5mm. **D)** Cryosection of fresh double ridge spike mounted onto a glass slide and imaged on a Leica S9D stereomicroscope. Scale bar = 0.5mm.

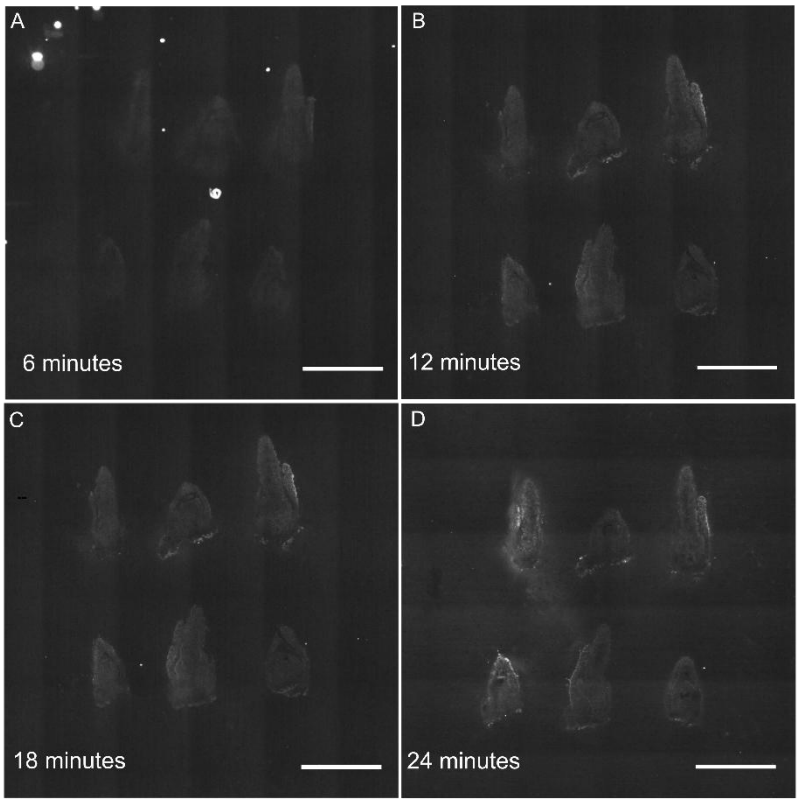


Figure 4.4 - Fluorescence microscopy detects cDNA synthesised on chip after 6, 12, 18, or 24 minutes of permeabilisation

Detection of mRNA capture in four sequential cryosections of wheat inflorescence tissue following incubation of permeabilisation reagent for **A)** 6 minutes, **B)** 12 minutes, **C)** 18 minutes, **D)** 24 minutes. TRITC channel. Scale bar represents 1000µm.

4.3.2- Full transcriptome capture experiment yields low mapping rates and excess lateral mRNA diffusion

After confirming quality sample preparation and optimising permeabilisation time, we proceeded to the transcriptomic capture experiment. We processed three sequencing chips in total, with six inflorescence samples each (*Triticum aestivum* cv. Paragon). One crucial step of the experimental protocol is the capture of quality images of nuclei and cell wall staining, which is used to determine later where tissue boundaries lie on the sequencing chip. Of the three chips processed during the experiment, one passed image quality checks (Table 4.1). The remaining two chips could not be focused during imaging of ssDNA or cell wall staining. Cryosectioning was of variable quality across the sequencing chips. In all three chips, we observed the folding of tissue, indicating tissue detachment during cryosectioning steps (Figure 4.5). In only one chip, C02134B1, we observed optimal section depth through inflorescence tissues and successful adherence of tissue (Figure 4.5A). Due to the high quality of sectioning and pass of imaging quality checks, onward analysis reported on will be from sequencing chip C02134B1, containing six spikes.

Table 4.1- Two samples failed QC during the imaging of ssDNA and cell wall staining

Chip ID	Tissue Type	QC check	Imaging Remarks
C02134B1	Inflorescence	Yes	Clear focus, normal tissue morphology. Some areas of folded tissue
C02134B2	Inflorescence	No	Unfocused; Folded tissue; Sectioning past the centre of the inflorescence
C02134B3	Inflorescence	No	Unfocused; Folded tissue

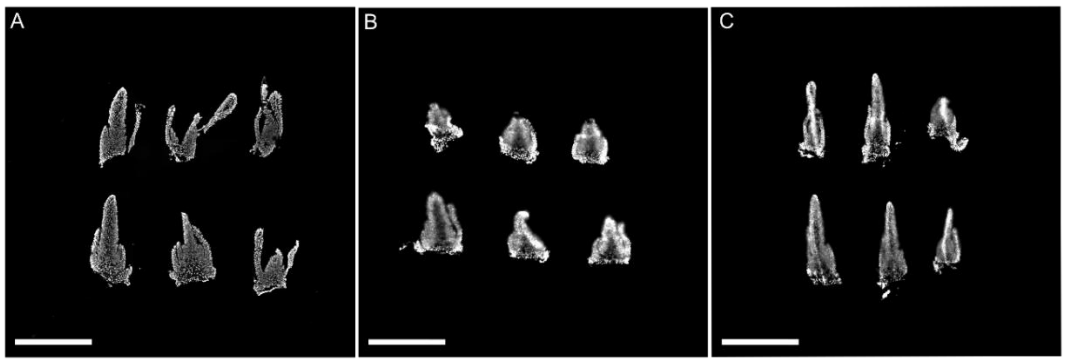


Figure 4.5 - ssDNA staining of six sequencing chips used for transcriptomic experiments **A)** C02134B1, **B)** C02134B2, **C)** C02134B3. Scale bar = 1000µm.

Raw sequencing reads from sequencing chip C02134B1 were aligned to the IWGSC RefSeq v2.1 genome (Zhu et al. 2021), using the Stereo-Seq Analysis Workflow (SAW) standard analysis process pipeline (Gong et al. 2024). From a total of 4.57 billion sequencing reads, 89.8% achieved a Q score ≥30, indicating high base-calling accuracy. Of these, 81.9% contained a valid CID, resulting in 3.40 billion clean reads. However, the standard SAW workflow yielded relatively low alignment rates, with 63.3% of reads unmapped. This is a low alignment rate when compared to other transcriptomic techniques, with an average alignment rate of 85% from the bulk RNA-seq microdissection dataset documented in Chapter 2. Among the mapped reads, 593 million aligned uniquely to the reference genome, while 654 million mapped to multiple genomic loci. Of these, only 305 million reads aligned to annotated exon regions.

Of the sequencing reads mapped to an annotated gene in the reference genome, many represent PCR duplicates derived from the same captured molecule. Collapsing reads by unique combinations of capture ID (CID) and molecular identifier (UMI) yields the number of distinct RNA molecules detected. For this sample, we obtained 42,205,135 unique reads detected across the array, representing 95,398 genes. Notably, we observe the significant expression of 75,962 genes from the microdissection bulk RNA-seq dataset on inflorescence tissues of the equivalent stage (>0.5 TPM, W2.5), indicating that Stereo-seq detected the expression of additional genes. However, these results may reflect the inclusion of vegetative tissues in the samples used for the Stereo-seq array that were not captured in the bulk RNA-seq dataset, including leaf primordia and peduncle tissues. The SAW pipeline summarises spatial resolution at multiple bin sizes; in the standard output at Bin200 (200 × 200 DNA nanoball spots, ~100 × 100 µm area), we observed an average of 2,093 unique reads and 754 genes per bin200 area.

However, because the wheat inflorescence samples are relatively small (~1.2 mm in length), only a minor fraction of the 1 cm × 1 cm chip was covered by tissue. To correct for this, a tissue mask was applied to quantify the number of reads detected under tissue boundaries. Under the tissue mask, we observed an average of 22,532 unique reads detected and 6,716 genes detected per bin 200 area (Table 4.2). Based on their distribution across the chip, we observe a high number of reads detected past tissue boundaries, resulting in a large capture area surrounding the tissue (Figure 4.6). Under the tissue-defined area, we detected 21.5% of unique reads (~9.1 million UMIs), while the remaining 78.5% were located outside tissue boundaries. The high percentage of reads across the sequencing array likely reflects the lateral diffusion of mRNA transcripts during the tissue permeabilisation step, a noted issue with Stereo-seq (You et al. 2024).

Table 4.2 - Summary statistics under captured tissue in bin 20, 50, 100, and 200 regions

Bin Size	Estimated Size	Mean Genes Detected (per bin)	Median Genes Detected (per bin)	Mean Unique Reads Detected (per bin)	Median Unique Reads Detected (per bin)
50	25 x 25 µm	968	961	1,855	1,855
100	50 x 50 µm	2,812	2,971	6,739	7,060
200	100 x 100 µm	6,716	7,255	22,532	23,023

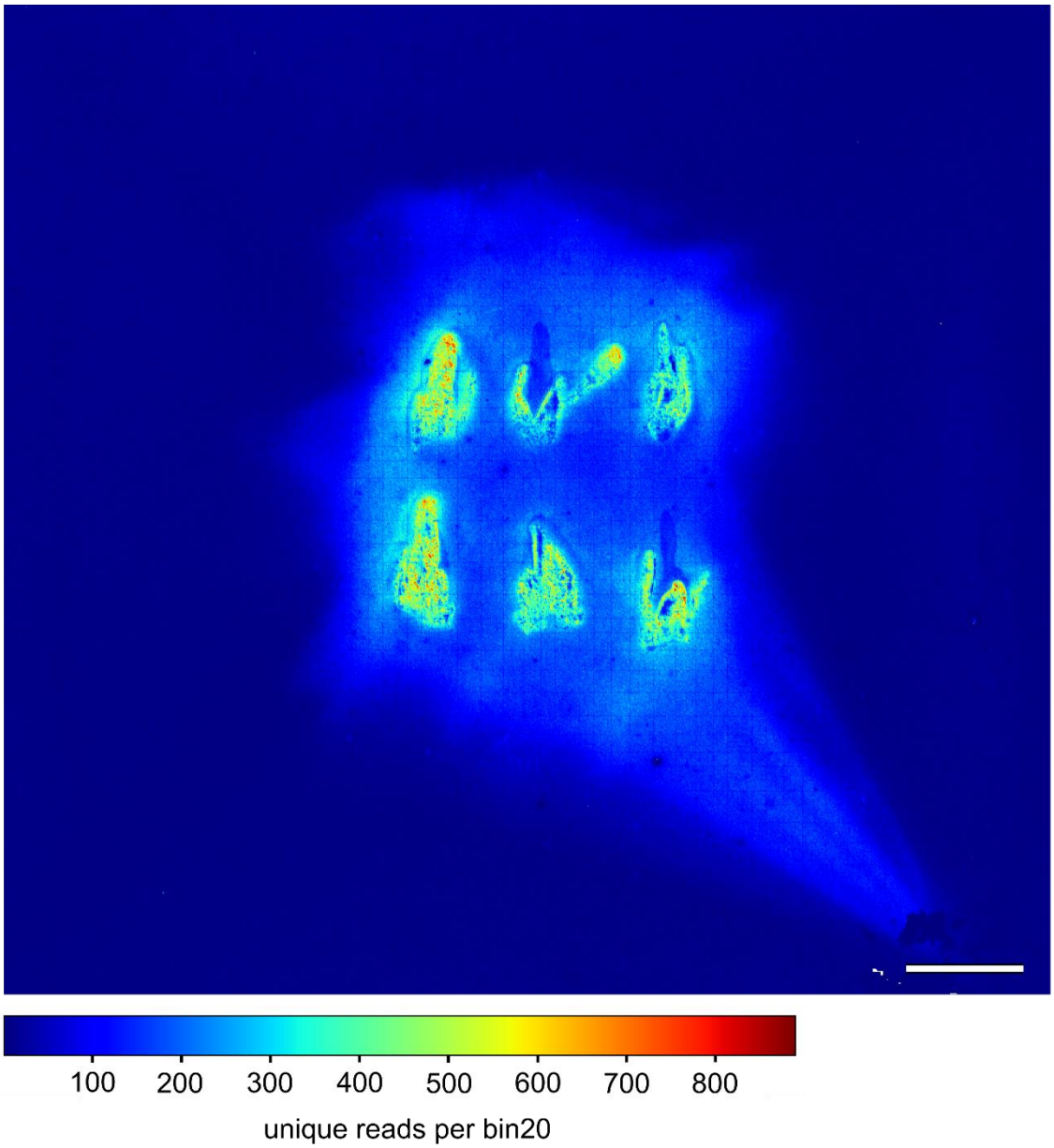


Figure 4.6 - Transcriptome capture with Stereo-seq displays lateral diffusion of mRNA transcripts beyond tissue boundaries.

Heatmap of UMI count (unique reads) per bin20 region (10 µm × 10 µm area). Scale bar = 1000 µm.

4.3.3 - Low capture rates of key developmental genes observed with Stereo-seq

To evaluate the spatial precision of Stereo-seq, given the diffusion of mRNA transcripts across the chip, we analysed the expression of established cell-type and tissue-domain markers. These included *KNOTTED HOMEOBOX-LIKE 5 (KNOX5),* a marker of ground tissues and meristematic cells; *TASSEL SHEATH 1 (TSH1)*, a marker of leaf ridges; *CUP-SHAPED COTYLEDON 3 (CUC3)*, a marker of meristem–leaf ridge boundaries; *ONION1 (ON1)*, an epidermal marker; *MANY NODED DWARF 1 (MND1),* a marker of suppressed axillary meristems below the inflorescence; and *METALLOTHIONEIN 2 (MT2),* a marker of ground tissues (See Chapter 3). For each gene, we assessed gene expression across all homoeologs (see Table 4.3). We found that the majority of transcripts were detected outside their expected tissue boundaries, with an average of 74.1% detected transcripts falling outside annotated regions. Within tissues, transcript detection was generally low, ranging from 40 to 442 counts per gene (excluding the highly expressed *MT2,* Table 4.3). As a result, Stereo-seq spatial patterns were often difficult to resolve.

To evaluate the accuracy of Stereo-seq in capturing spatial gene expression, we compared its expression profiles with those obtained from a comparable stage and section in the MERFISH dataset. In all cases, the number of detected transcripts within the inflorescence was low when compared to MERFISH, further indicating the effects of lateral diffusion limiting biological interpretation (Table 4.3). For *KNOX5*, transcripts were detected in ground tissues but lacked the clear enrichment in meristematic L2/L3 layers observed with MERFISH (Figure 4.7A). *MND1* transcripts were primarily detected in vegetative tissues below the inflorescence, with occasional expression in inflorescence tissues, broadly consistent with MERFISH results (4.7B). For *TSH1*, *CUC3*, and *ONI1*, low transcript counts limited spatial resolution, and patterns were only discernible when compared directly to MERFISH data (Figure 4.7C-E). In contrast, *MT2* showed consistently high expression in ground tissues, closely mirroring patterns observed with MERFISH (Figure 4.7F). This high expression was also supported by the microdissection RNA-seq dataset, in which *MT2* averaged 1570 TPM in basal sections and 1168 TPM in central sections of W2.5 spikes. However, Stereo-seq data exhibited pronounced transcript diffusion, with frequent detection of *MT2* transcripts outside the expected ground tissue boundaries.

Table 4.3 – Total detected transcripts within tissue boundaries of 6 tissue domain marker genes

Gene Name	IWGSC RefSeq v2.1	IWGSC RefSeq v1.1	Detected Transcripts Total	Detected Transcripts Under Tissue	% Transcripts Detected Under Tissue	Detected Transcripts Outside Tissue	% Transcripts Detected Outside Tissue	Detected Transcripts in MERFISH samples (homoeolog non-specific)
ONI1	TraesCS4A03G0012800	TraesCS4A02G007400	1810	442	24.4%	1368	75.6%	4121
	TraesCS4B03G0780900	TraesCS4B02G297500	1185	280	23.6%	905	76.4%	
	TraesCS4D03G0701000	TraesCS4D02G296400	1355	301	22.2%	1054	77.8%	
KNOX5	TraesCS4A03G0671200	TraesCS4A02G256700	989	268	27.1%	721	72.9%	6200
	TraesCS4B03G0126000	TraesCS4B02G057900	578	127	22.0%	451	78.0%	
	TraesCS4D03G0106600	TraesCS4D02G058000	1313	321	24.4%	992	75.6%	
NL1	TraesCS1A03G1020200	TraesCS1A02G418200	781	241	30.9%	540	69.1%	2484
	TraesCS1B03G1202900	TraesCS1B02G448200	249	53	21.3%	196	78.7%	
	TraesCS1D03G0980700	TraesCS1D02G425900	430	122	28.4%	308	71.6%	
CUC3	TraesCS7A03G0569500	TraesCS7A02G247600	257	47	18.3%	210	81.7%	1431
	TraesCS7B03G0391800	TraesCS7B02G143900	482	105	21.8%	377	78.2%	
	TraesCS7D03G0550700	TraesCS7D02G246100	187	40	21.4%	147	78.6%	
MND1	TraesCS7A03G1228100	TraesCS7A02G506400	1061	299	28.2%	762	71.8%	1135
	TraesCS7B03G1114800	TraesCS7B02G413900	516	128	24.8%	388	75.2%	
	TraesCS7D03G1168900	TraesCS7D02G494500	705	204	28.9%	501	71.1%	
MT2	TraesCS1B03G0086800	TraesCS1B02G042200	32323	12021	37.2%	20302	62.8%	10962
	TraesCS1D03G0066100	TraesCS1D02G034800	142330	51244	36.0%	91086	64.0%	

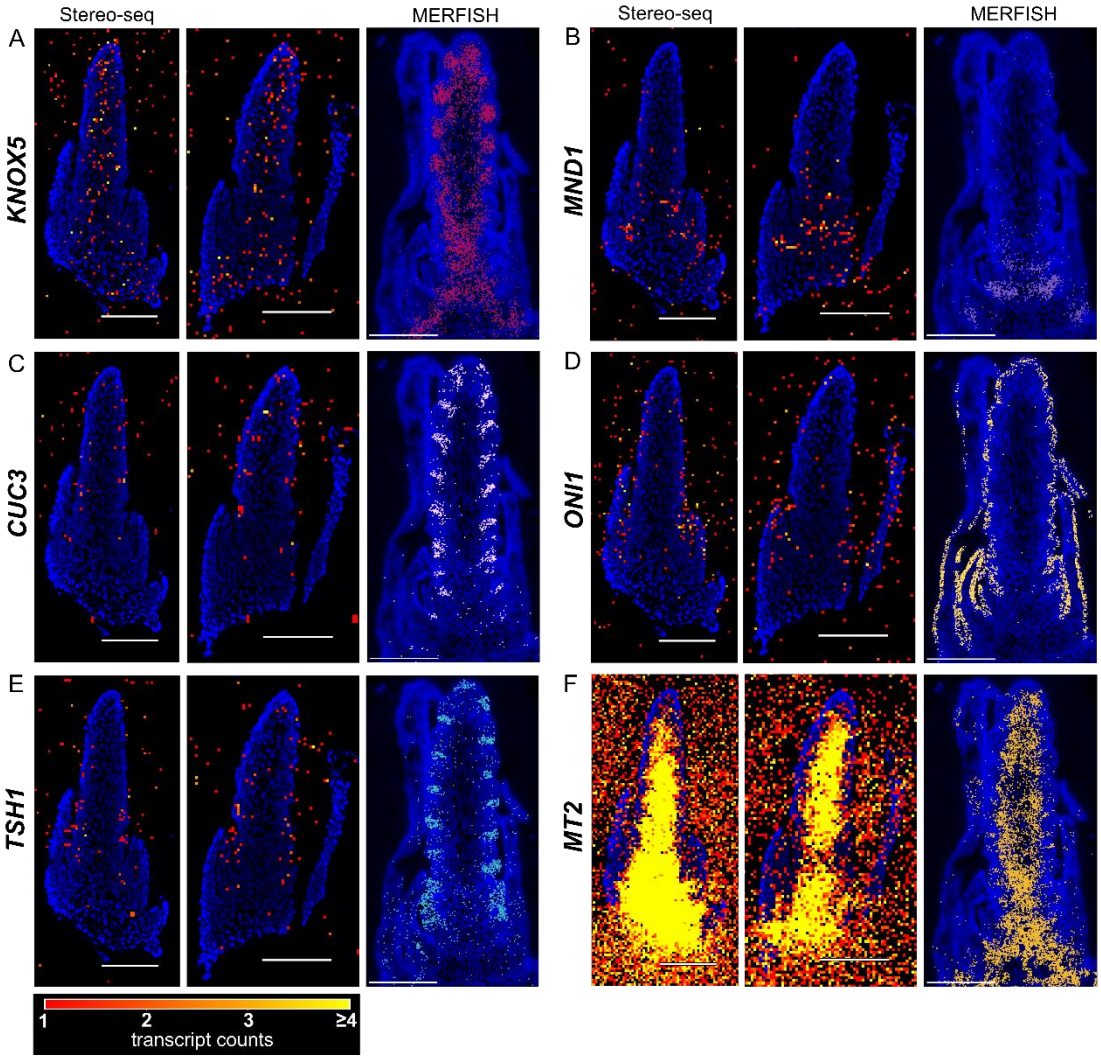


Figure 4.7 - Comparison of MERFISH and Stereo-seq highlights limited transcript capture and poor spatial resolution of Stereo-seq in tissue-domain marker genes

Spatial transcript detection from Stereo-seq (Bin20, 10µm x 10µm area) in two biological replicates of W2.5 inflorescence tissue is compared with a corresponding MERFISH sample. Expression patterns are shown for six developmental marker genes: **A)** *KNOX5*, **B)** *MND1*, **C)** *CUC3*, **D)** *ONI1*, **E)** *TSH1*, and **F)** *MT2*. Scale bar = 250 µm. Blue stain corresponds to ssDNA in Stereo-seq samples and DAPI in MERFISH samples. For each gene name, expression information from all homoeologs is displayed (see Table 4.3).

We next compared our spatial expression data with a published Stereo-seq dataset from late double ridge (W2.5) spikes of *Triticum aestivum* (Qu et al., 2025). From our collection, we selected two high-quality inflorescence sections for comparison with a representative section from Qu et al. In their dataset, a high-quality inflorescence sample yielded on average 1,771 unique reads per bin40 (20 µm × 20 µm). By contrast, our two sections contained 1,389 and 1,479 unique reads per bin40 (replicates 1 and 2, respectively; Figure 4.8), indicating somewhat lower rates of transcript capture.

We then examined four reported marker genes, *KNOX5*, *MND1*, *CUC3*, and *TSH1*. Spatial profiles were broadly consistent across datasets: *KNOX5* localised to ground tissues (Figure 4.8A), *MND1* marked the base of the inflorescence (Figure 4.8B), and *CUC3* and *TSH1* were expressed within the inflorescence (Figure 4.8C–D). However, expression patterns in Qu et al. generally appeared sharper. For example, *TSH1*, a marker of leaf ridges, was strictly confined to those regions in their dataset. In ours, *TSH1* was also detected outside leaf ridges, with transcripts dispersed throughout the inflorescence, a signal absent in both the Qu et al. dataset and our MERFISH dataset. This suggests that lateral diffusion of mRNA may be more pronounced in our Stereo-seq application. A limitation of the Qu et al. dataset is that transcripts outside annotated tissue boundaries were not reported, preventing a direct comparison of lateral mRNA diffusion. Overall, the reduced resolution of marker gene expression in our Stereo-seq sections constrains the biological interpretations that can be drawn from this dataset.

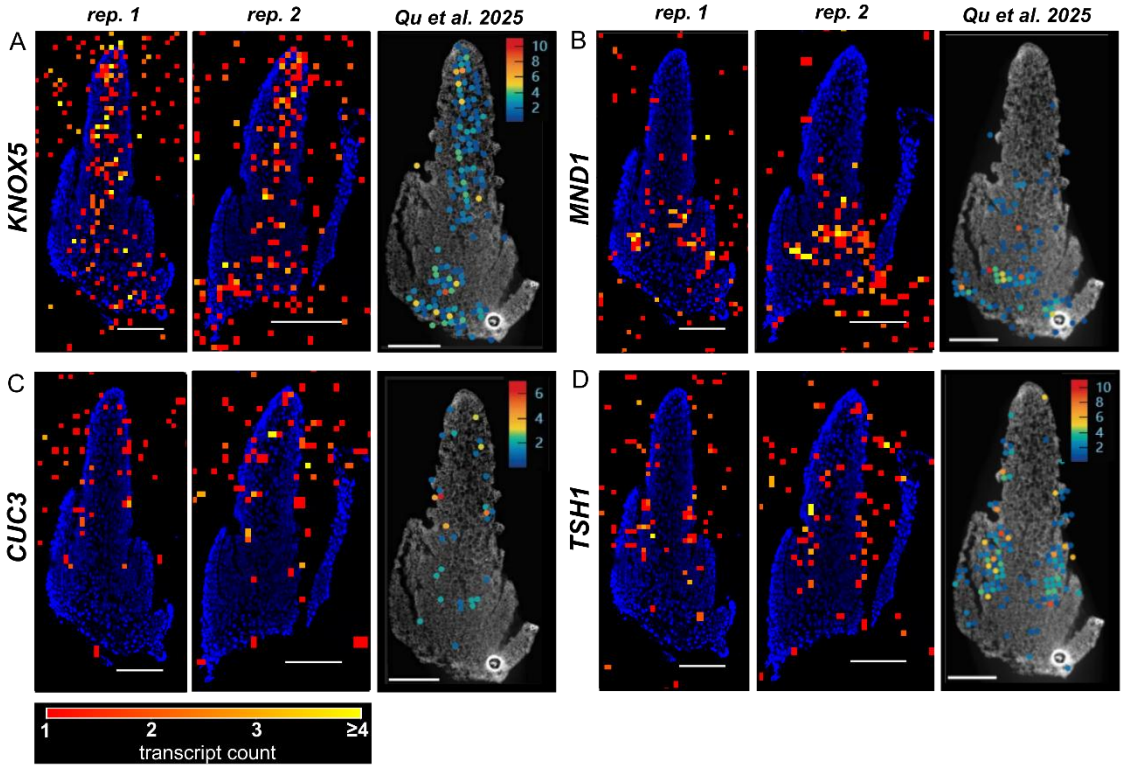


Figure 4.8 - Comparison of Stereo-seq data from our initial trials with a published dataset of late double ridge wheat spikes

Spatial transcript detection at bin40 resolution (20 µm × 20 µm) is shown for two biological replicates of W2.5 inflorescence tissue alongside a corresponding sample from Qu et al. (2025). Expression patterns are presented for four tissue-domain marker genes: **A)** *KNOX5*, **B)** *MND1*, **C)** *CUC3*, and **D)** *TSH1*. Scale bar = 200 µm. Blue staining indicates ssDNA in Stereo-seq samples and DAPI in MERFISH samples. For each marker, expression represents the combined signal from all homoeologs (see Table 4.3). Note: the MND1 panel from Qu et al. does not include a gene expression scale.

Chapter 4.4 – Discussion

4.4.1 - Our application of Stereo-seq to wheat inflorescence has limited utility

Overall, we found that our first implementations of Stereo-seq to wheat inflorescence tissues had limited utility. Due to a low overall mapping rate and lateral transcript diffusion across the sequencing array, the detection of meaningful expression patterns in key developmental genes was limited and difficult to interpret.

The lateral diffusion of transcripts with Stereo-seq is a documented issue reported in mammalian tissues, with this issue being noted in a benchmarking paper applied to mouse (You et al. 2024). The authors commented on the lateral diffusion in Stereo-seq being particularly pronounced when compared to related technologies, Slide-seq (bead-based technique; Rodriques et al. 2019) and PIXEL-seq (gel-based array; Fu et al. 2022), and BMKMANU S1000 (bead-based; BMKGENE). In comparison between these techniques, marker genes selected for their distinct cell-layer specific expression patterns were detected diffusely in the Stereo-seq datasets. However, in the case of one marker gene specific to melanocytes, Stereo-seq showed the best control of lateral diffusion among the techniques benchmarked, indicating that tissue type influences mRNA diffusion (You et al. 2024). Indeed, transcript diffusion is not always a barrier in the application of Stereo-seq to plant tissues, with its implementation in *Arabidopsis* leaves reporting minimal lateral diffusion outside tissue boundaries (Xia et al. 2022). Aside from tissue type, permeabilisation time is a key factor impacting the lateral diffusion of mRNA (You et al. 2024). An obvious first step for improving Stereo-seq in wheat inflorescence tissues is reducing permeabilisation time. Qu et al. (2025) achieved this by implementing a 12-minute permeabilisation step (compared to our 18 minutes), resulting in notably improved results.

Another key limitation of this dataset is the relatively low proportion of reads that align to the wheat reference genome, with 63.3% of reads remaining unmapped, and a small proportion of total reads mapping to an exon (9.0% of total clean reads). Several factors could underlie these reduced mapping rates, including technical issues during library preparation, RNA quality, or limitations in the alignment strategy. In general, I found the reporting of alignment rates found in applications of Stereo-seq to plant tissues to be challenging to find, making a comparative analysis not possible. To address the poor mapping rate of sequencing reads in this dataset, the logical next step is to optimise the alignment settings used in the Stereo-seq Analysis Workflow (SAW) pipeline (Gong et al.

2024)to improve alignment to a reference genome, which may be particularly challenging in a polyploid plant genome.

4.4.2 - Improvements to Stereo-seq in plants further develop our understanding of meristematic development

The first implementation of sequencing-based spatial transcriptomics in plant tissues relied on a technique using barcoded spot arrays, with a resolution of 100 µm (Giacomello et al. 2017). Building on this foundation, subsequent technologies have been applied to plants, which have progressively refined spatial resolution, ranging from ~55 µm with 10x Genomics Visium (Fu et al. 2023; Song et al. 2023; Serrano et al. 2024; Zhang et al. 2025), to submicron scales with Stereo-seq (Xia et al. 2022; Wang et al. 2024; Liu et al. 2025; Qu et al. 2025). These advances now enable transcriptomic data to be resolved at cellular and even subcellular scales, providing powerful opportunities to investigate plant development.

Following the generation of our Stereo-seq dataset, subsequent applications of Stereo-seq to wheat inflorescence tissue have produced higher-quality datasets, enabling more comprehensive spatial transcriptome maps. Liu et al. (2025) applied Stereo-seq to wheat spikes at the lemma primordia (W3.25; Waddington et al. 1983) and terminal spikelet (W4) stages, integrating the spatial data with a complementary single-nuclei RNA-seq (snRNA-seq) dataset. Similarly, Qu et al. (2025) applied Stereo-seq to late double ridge (W2.5) and lemma primordia (W3.25) stages. Compared with our dataset, the Qu et al. study detected the expression of key tissue-domain markers such as *TSH1*, *MND1*, and *MT2* with greater clarity, potentially due to a higher capture of transcripts or a reduction in lateral diffusion. Qu et al. (2025) leveraged this improved resolution to cluster bin40 regions of double ridge spikes into 10 expression domains. Similarly, Liu et al. (2025) utilised the integration of snRNA-seq with Stereo-seq data to resolve 18 transcriptionally defined cell types spatially. These studies demonstrate the capacity of Stereo-seq to achieve finer spatial resolution and more comprehensive transcriptome coverage, facilitating the classification of cell types and tissue domains.

Indeed, the application of Stereo-seq to inflorescence tissues has further advanced our characterisation of meristematic development. For example, in maize ears, meristematic cells from the inflorescence meristem (IM), spikelet meristem (SM), and (floral meristem) FM cannot be resolved by single-cell RNA-seq (scRNA-seq) analysis alone (Xu et al. 2021).

However, the application of Stereo-seq distinguished transcriptional differences between the indeterminate IM and the determinate SM through the expression of the MADS-box transcription factors *ZmMADS8* and *ZmMADS14*. Functional validation confirmed their role, as double mutants disrupted meristem determinacy and produced indeterminate branch-like structures (Wang et al. 2024). In wheat, Stereo-seq clustering resolved two distinct cell populations: SMs, marked by genes such as *RAMOSA2* and *SEPALLATA1–4*, and boundary regions, associated with *SPL14*, *SPL17*, and *TSH1*. Loss-of-function of *RAMOSA2* produced a paired-spikelet phenotype, demonstrating its central role in specifying SM identity and promoting the formation of short lateral branches bearing a single spikelet (Qu et al. 2025). In both cases, these studies highlight the utility of sequencing-based spatial transcriptomics in characterising expression domains in a developing inflorescence tissue, adding to our breadth of knowledge on the key regulators of meristem identity.

Looking forward, the capabilities of sequencing-based spatial transcriptomics continue to expand, offering improved performance in plant systems. New sequencing chemistry released with Stereo-seq V2 is now compatible with formalin-fixed paraffin-embedded samples (Zhao et al. 2025). In brief, updates to the technique include a random-priming-based strategy for *in situ* RNA capture, which is reported to improve on unbiased transcript capture and whole-body gene coverage. As a result of these improvements to the techniques, Stereo-V2 also displays a reduction in lateral diffusion, measured by a higher level of restriction of marker genes (Zhao et al. 2025), noting significant improvement from the results reviewed by You et al. (2024). This updated sequencing chemistry may reduce lateral diffusion in plant tissues while preserving the advantage of high-resolution transcript capture.

4.4.3 – Towards standardised reporting in spatial transcriptomics

Overall, this chapter highlights some of the lessons learnt when first optimising a spatial transcriptomic technique for a plant tissue. Our implementation of Stereo-seq to wheat inflorescence tissue documented some key limitations, including lateral mRNA diffusion, which limited the biological interpretation of our dataset.

Currently, an open framework has been proposed to establish standards for the primary data and metadata released alongside spatial transcriptomic publications (Jackson and Pachter 2023), aligning with the FAIR principles of data management and stewardship (Wilkinson et al. 2016). This is a particular challenge in spatial transcriptomic techniques, where even amongst the sequencing-based methods, data types are variable (Jackson and Pachter 2023). In future, I would encourage the plant science community to adopt such frameworks, while also incorporating additional quality metrics that could inform protocol optimisation and reproducibility. In particular, transparency in overall alignment rates is critical given the complexity of plant genomes, where large size, repetitive content, and polyploidy can strongly influence mapping performance (Claros et al. 2012). Reporting the proportion of reads mapping to exons, together with clear documentation of the alignment parameters used in bioinformatic pipelines, would further support the development of robust computational approaches. Moreover, while excellent work has benchmarked lateral mRNA diffusion in animal tissues (e.g., You et al., 2024), equivalent metrics remain largely absent from plant studies. Systematic reporting of transcripts detected outside tissue boundaries could ultimately help refine wet-lab protocols and enable meaningful comparisons across plant tissue types.

Chapter 4.5 - Methods

4.5.1 - Plant materials and growth conditions

For all samples used in Stereo-seq experiments, we used hexaploid bread wheat, *Triticum aestivum* cv. Paragon. Plants were grown under a 16/8 h light/dark cycle at 20/15 °C, 65% relative humidity and bottom-watering irrigation (Simmonds et al. 2024).

4.5.2 - Tissue dissections and embedding

We cleaned all surfaces and dissection tools with RNABlitz before use. For the dissection of plant tissues for Stereo-seq, we used a published dissection methodology (Faci et al. 2024), but maintained the youngest leaves surrounding meristems (See Figure 4.3B). We dissected inflorescence tissues at Waddington stage W2.5, 'Late Double Ridge' (Waddington et al. 1983). We filled a 1 cm × 1 cm Tissue-Tek mould (Thermo Fisher, AGG4581) with Tissue-Plus O.C.T. compound (Agar Scientific, AGR1180) using clean dissection tools. We also filled a 60 mm Petri dish with O.C.T. compound. Using clean dissection tools, we carefully transferred dissected inflorescence tissues to the OCT-filled Petri dish, where they were mixed with O.C.T. to ensure complete coating. Using a stereomicroscope (Leica S9 with an HXCAM HiChrome HR4 Lite camera and a Photonic Optics light source), we inspected meristems for air bubbles, which were carefully removed with a fine dissection tool. Meristems were then placed into the O.C.T.-filled Tissue-Tek mould and arranged at the bottom of the mould. Each O.C.T. block contained six inflorescences. The O.C.T. blocks were flash-frozen and stored at −70 °C. Frozen blocks were shipped on dry ice to MGI in Riga, Latvia.

4.5.3 – Assessment of RNA integrity

For a block of O.C.T. containing six inflorescence tissues, 10-20 O.C.T. sections of 10 µm section thickness were collected into pre-cooled 1.5ml Eppendorf tubes. Total RNA was extracted with the RNA Clean and Concentrator Kit (ZYMO-Research, R1017). RNA quality was quantified with the Agilent Tapestation system according to the manufacturer's recommendation.

4.5.4 – Tissue cryosectioning, mounting, and fixation

A PCR Thermal Cycler with Stereo-Seq PCR Adaptor plate (Cat. No. 301AUX001) was set to 37°C in advance. Cryostat was set to a chamber temperature of -24°C and -15°C. OCT-embedded tissues were stored at -80°C to equilibrate to the cryostat chamber, and forceps and brushes were placed inside the chamber to pre-cool. Each sample of fresh-frozen wheat tissue was mounted on the tissue block with OCT. We used the Stereo-seq Chip P slide (1cm*1cm; Cat. No. 210CP118). Before sectioning, the chip ID number was recorded, and the chip was equilibrated to room temperature for 1 minute, then rinsed twice with 100 µL 0.01N HCl in nuclease-free and coated in 0.01% Poly-L-Lysine. Excess water was dried with a power dust remover from one side of the chip at a 30-45° angle and wiped around the edges of the chip with a dust-free tissue. 10 µm sections of wheat meristematic tissue in OCT were flattened with a brush, moved to the edge of the cryostat stage, and flipped. A room-temperature chip was picked up with forceps and gently placed down onto the section. Immediately after placement, the chip was placed on the PCR thermal cycler front side up for 5 minutes.

Methanol in a 50ml Corning Tube was precooled for 10-30 minutes at -20°C. Following the drying step, chips were immediately submerged in pre-cooled methanol for 30 minutes at -20°C. Chips were moved to a sterile fume hood, and excess methanol was wiped from the back and sides of the chip with dust-free paper and left for 4-6 minutes to allow for complete evaporation. Then, the chip was transferred to a clean benchtop.

4.5.5 – Stereo-seq tissue permeabilisation testing

Tissue permeabilisation testing was conducted based on the Stereo-Seq Permeabilisation for Chip-on-a-Slide User Manual (Cat. No.: 211SP118, Kit Version: V1.0, Manual Version B). We used the Stereo-seq Permeabilisation Kit (Cat. No 111KP118) and the Stereo-seq Chip P slide (1cm*1cm; Cat. No. 210CP118). In brief, four chips with four consecutive cryosections of inflorescence tissues were prepared in line with section 'Tissue Cryosectioning, Mounting, Fixation'. Following fixation, chips were placed on a PCR thermal cycler with Stereo-Seq PCR Adaptor to 37°C (Cat. No. 301AUX001). 150µL of 1x Permeabilisation Reagent Solution was added onto each chip and incubated at 37°C for either 6 min, 12min, 18min, or 24min. Following permeabilisation, tissues were rinsed with 100 µL of PR Rinse Buffer, and 90 µL of RT QC mix was applied to each sequencing chip. Reverse transcription was carried out at 42°C (on PCR Thermal Cycler) for 3 hours in the dark. RT QC mix was removed, and tissues were washed with 100 µL 0.1x SSC (with 5% RNase inhibitor, Cat. No. 1000028499) solution. Tissue Removal enzyme was added to each chip and incubated for 1 hour at 55°C. Following incubation, chips were washed with 0.1x SSC (with 5% RNase inhibitor, Cat. No. 1000028499) solution, followed by nuclease-free water. Chips were visually inspected to ensure tissue was removed entirely from the chips. Fluorescence imaging was performed in the TRITC channel with a 10x objective lens.

4.5.6 – Stereo-seq transcriptomic capture experiment

Transcriptome capture experiments were conducted based on the Stereo-Seq Transcriptomic Set for Chip-on-a-Slide User Manual (Cat. No.: 211ST114, Kit Version: V1.2, Manual Version A_1). We used the Stereo-seq Transcriptomics T Kit (Cat. No 111KT114) and the Stereo-seq Chip T slide (1cm*1cm; Cat. No. 210CT114). In brief, three chips were mounted with cryosections of inflorescence tissues, which were prepared in line with section 'Tissue Cryosectioning, Mounting, Fixation'. The Stereo-seq chip was placed in a clean 10cm Petri dish, and 100 µL of tissue fluorescent staining solution (5x SSC with calcofluor white, ssDNA stain, and RNAse inhibitor, Cat. No. 1000028499) was added to the chip and incubated for 5 minutes at room temperature in the dark. Staining solution was removed from the corner of the chip with a pipette, and 100 µL of 0.1X SSC (with 5% RNase inhibitor, Cat. No. 1000028499) was added per chip. Chips were transferred to dust-free paper and thoroughly dried with a power dust remover. 5 µL of imaging reagent was added gently to the centre of the tissue. Fluorescence images were taken on the FITC channel and DAPI channel with a 10x objective lens. ssDNA images were checked for quality with STOmics ImageStudio software. To remove the imaging reagent, chips were washed with 0.1xSSC, excess solution was wiped from the sides and back with dust-free paper.

Following imaging, tissues were incubated in 100 µL of permeabilisation reagent solution (PR enzyme 1mg/mL, cat. No. 1000028500 in 0.01N HCL) for 18 minutes at 37°C. Chip was rinsed with 100 µL PR rinse buffer with 5% RNase Inhibitor (Cat. No. 1000033684). Rinse buffer was removed, and 90 µL of reverse transcription mixture (Cat. No. 1000028501) was added per chip, and incubated at 42°C for 3 hrs. Following incubation, the reverse transcription mixture was removed through a slight tilt of the chip, and liquid was pipetted away from the corner without touching the chip surface. 400 µL tissue removal buffer was added to each chip and incubated at 55°C for 10 minutes. cDNA Release Mix (Cat. No. 1000028512) was prepared during the incubation step. Subsequently, the tissue removal buffer was removed, and 400 µL of cDNA Release mix was added per chip, sealed in a Petri dish, and incubated overnight at 55°C. Released cDNA was carefully removed from each chip with a pipette and aliquoted into a 1.5ml tube. 100µL of nuclease-free water was added to each chip, pipetted up and down on the chip surface, and collected into the same 1.5ml tube.

Released cDNA was purified with DNA Cleanup Beads AMPure® XP (Agencourt, Cat. No. A63882) using the manufacturer's recommendation. cDNA was eluted in 44 µL of

nuclease-free water and collected in a 0.2ml PCR tube. cDNA was amplified by PCR with the cDNA Amplification Mix (Cat. No. 1000028514) and cDNA primer (Cat. No. 1000028513). PCR programming was conducted as follows: incubation at 95 °C for 5 minutes, 15 cycles at 98°C for 20 seconds, 58°C for 20 seconds, 72°C for 3 minutes, and a final incubation at 72°C for 5 minutes.

Sequencing library preparation was performed with the Stereo-seq Library Prep Kit (Cat. No. 111KL114).1 µL of cDNA sample was used to measure and record the concentration of purified cDNA with the Qubit dsDNA HS Kit (Thermo, Cat. No. Q32854). A total of 20 ng of DNA was fragmented using transposase-assisted tagmentation. Indexed PCR and library purification were performed according to the manufacturer's recommendations to generate the final sequencing libraries. Stereo-seq libraries were subsequently sequenced with the MGI DNBSEQ-T10x4RS at the MGI Latvia sequencing facility.

4.5.7 – Stereo-seq raw data processing and genome alignment

For onward bioinformatic analysis, we followed the Stereo-seq Analysis Workflow (STOmics, Version A2, 2023). Mapping of raw reads to the IWGSC RefSeq v2.1 genome assembly and annotation (Zhu et al. 2021) was performed with the Stereo-seq Analysis Workflow (SAW, v5.5.3) software suite (Gong et al. 2024). In brief, this bioinformatic pipeline processes the raw sequencing reads from the Stereo-seq sequencing platform, and in combination with chip ID information and microscopy images, generates mRNA spatial position reconstruction, filtering, genome alignment, expression matrix generation, and tissue region expression (Gong et al. 2024). The workflow outputs .gef files (gene expression file) for onward analysis, in addition to a report containing information on genome alignment statistics, and spatial visualisation of gene expression information.

4.5.8 – Stereo-seq data visualisation

For the visualisation of total transcripts density across the sequencing chip, we extracted transcript counts from the bin20 resolution matrix of the .gef file with h5py (v3.11.0). We aggregated per bin to obtain total transcript counts. Resulting counts were visualised as a heatplot with matplotlib (v3.8.2; Hunter 2007).

For visualisation of transcript count for individual genes. A brightfield stain image (.tif) was aligned to the bin20 transcriptomic mask (.tif) using scipy (v1.13.0; Virtanen et al. 2020) and numpy (Harris et al. 2020; v1.24.3). The aligned stain was cropped to a region of interest, and transcript counts for the selected genes were extracted from the .gef file using h5py (v3.11.0) and aggregated per bin with pandas to construct a dense expression grid. The stain and an overlaid heatmap of bin20 counts were visualised with matplotlib (v3.8.2; Hunter 2007)

Chapter 5 – General Discussion

5.1 – Thesis summary

This thesis aimed to investigate spatially restricted gene expression across the wheat inflorescence. To achieve this, I built upon a microdissection-based bulk RNA-sequencing dataset and sought to refine gene expression patterns across the apical–basal axis to a cellular resolution. Specifically, I implemented two spatial transcriptomic approaches in wheat tissues for the first time: multiplexed error-robust fluorescence *in situ* hybridisation (MERFISH) and Stereo-seq, to address the following questions:

1. Can we apply spatial transcriptomic techniques to inflorescence tissues and successfully resolve gene expression patterns at cellular resolution?
2. What are the spatio-temporal dynamics of gene expression underlying spikelet and floret patterning in the wheat inflorescence?
3. Can gene expression signatures distinguish axillary meristems along the apical–basal axis before the establishment of spikelet patterning?

5.2 – Evaluating spatial transcriptomic approaches for plant developmental biology

Spatial transcriptomics has rapidly expanded in scope and application in plant research (Giacomello 2021; Nobori 2025), with a growing variety of methods now available (Moses and Pachter 2022). Broadly, these approaches can be divided into 'targeted' imaging-based techniques and 'untargeted' sequencing-based techniques, each with distinct strengths and limitations. Imaging-based approaches, such as MERFISH, provide subcellular resolution but are constrained by a predefined gene list. In contrast, sequencing-based platforms such as Stereo-seq capture transcriptomes in an unbiased manner, though often at the cost of spatial resolution and cellular precision (Giacomello 2021).

In this study, we applied both imaging-based and sequencing-based approaches for the first time to developing wheat spikes. Our initial attempts with Stereo-seq revealed technical challenges associated with the permeabilisation step required to transfer mRNA molecules onto the sequencing chip. We observed a lateral diffusion effect, in which transcripts were detected beyond their cellular origins, complicating the interpretation of spatial patterns - a noted limitation of stereo-seq (You et al. 2024). This resulted in 78.5% of the total detected transcripts falling outside tissue boundaries, dramatically reducing the overall transcripts available for analysis. This effect, in combination with the low mapping rates (with 63.6% of reads remaining unmapped to the reference), resulted in few developmental genes being detected under tissue boundaries, limiting interpretations. For instance, *TASSEL SHEATH 1 (TSH1),* a marker of leaf ridges (LRs), is expressed at an average of ~56 TPM in late double ridge (W2.5) spikes (central and basal, microdissection RNA-seq dataset). In our MERFISH dataset, we detected 2,484 transcripts at this stage (non-homoeolog specific); whereas Stereo-seq detected only 416 counts within the inflorescence across all three homoeologs.

More recent applications of Stereo-seq in plants have reported improved accuracy of transcript capture within tissue boundaries (Liu et al. 2025; Qu et al. 2025). For example, Qu et al. reported improved expression patterns in wheat inflorescence tissue when compared to our Stereo-seq dataset (2025). In their dataset, expression of *TSH1* was restricted to leaf ridges (LRs), a pattern that appeared more diffuse in our implementation. These improvements may be explained in part by the reduction in permeabilisation time (12 minutes), which prevents excess mRNA diffusion within and outside tissue boundaries. However, even with these improvements in mind, the total number of

detected transcripts observed in key developmental regulators remained sparse in both datasets when compared with imaging-based approaches such as MERFISH and Molecular Cartography (Xu et al. 2025).

A strength of MERFISH in our application to wheat spikes was the ability to segment cells directly from DAPI and PolyT stains, ensuring accurate assignment of transcripts into cellular boundaries. This allowed for the clustering of individual cells into expression domains, which defined key regions patterning the late double ridge spike. For example, in our MERFISH dataset on late double ridge spikes, *KNOTTED-LIKE HOMEOBOX 5 (KNOX5)* is clearly confined to the L2/L3 layers, *ONION1 (ONI1)* to L1 layers, and *CUP SHAPED COTYLEDON 3 (CUC3)* in two to three cell layers formed in the boundary region adjacent to the axillary meristem (AM). Alternating with these domains are LRs, marked by the expression of *TASSEL SHEATH 1 (TSH1)*. These gene expression patterns refined the clustering of cells into expression domain (ED) categories, including meristematic cells (ED0), the L1 layer/epidermis (ED4), boundary cells (ED11), and leaf ridges (ED12).

In contrast, our implementation of Stereo-seq yielded poor-quality stains, as the images of cell wall staining with Calcofluor White often failed quality control, preventing cell segmentation and further downstream analyses. In the implementations of Stereo-seq in wheat spikes (Qu et al. 2025), and maize ears (Wang et al. 2024), cell segmentation was not performed; instead, clustering was conducted on square regions of the chip, termed bin40 (20 µm x 20 µm) and bin50 (25 µm x 25 µm), which serve as an approximation of a single cell. However, such grid-based assignments may obscure aspects of complex tissue organisation observed in meristematic tissues, where gene expression domains define single-cell layers in orientations not aligned with square grids.

In the Qu et al. dataset, clustering distinguished alternating LRs and AMs into two domains. However, the AM cluster was marked by *CUC3,* a marker of boundary cells, and also by the ortholog of *RICE OUTERMOST CELL-SPECIFIC GENE 4 (ROC4),* which in our MERFISH dataset was expressed within AMs at the late double ridge stage. This suggests that the Stereo-seq AM domain encompasses both the boundary cells and AM. At a subsequent developmental stage, the lemma primordia (LP), clustering again resolved two alternating domains: one corresponding to spikelets, and a second termed the 'spikelet boundary.' This boundary domain was marked by *CUC3* together with LR markers *TSH1* and *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 17*, indicating that it clustered together both LRs and boundary cells (Qu et al. 2025). These observations raise the possibility that implementing cellular segmentation could refine Stereo-seq clustering, separating these broader domains into refined categories such as boundary cells and LRs.

Nonetheless, there are examples where square-bin clustering has captured fine-scale meristem organisation. In maize ears, for instance, clustering resolved distinct zones within lateral meristems, including the meristem base, internal region, adaxial periphery, and epidermis (Wang et al. 2024).

Collectively, these comparisons highlight how methodological choices influence the biological conclusions that can be drawn from spatial transcriptomic studies. In our dataset, lateral diffusion of mRNA molecules and low transcript capture limited the ability of Stereo-seq to resolve developmental gene expression. Future improvements, including optimised permeabilisation, cell segmentation, and transcript assignment, may enhance its performance. By contrast, MERFISH consistently provided accurate gene expression patterns to cellular resolution, which aided in the clustering of expression domains. While the gene set of MERFISH was limited to 200 genes, this technique offers a clear advantage for validating or screening cell-type markers identified in bulk or single-cell transcriptomic datasets, where spatial precision is more critical than transcriptome breadth.

Although our findings highlight the advantages of MERFISH, it is equally important to recognise the unique contributions and future promise of sequencing-based methods such as Stereo-seq. Unlike imaging-based approaches, sequencing provides an unbiased survey of the transcriptome, allowing for data-driven exploratory experiments (Giacomello 2021) and enabling the clustering of cells into cell-type categories analogous to those obtained through single-cell RNA-seq (scRNA-seq; as seen in Xia et al. 2022; Wang et al. 2024; Qu et al. 2025). Here, we use MERFISH data to a similar effect. The clustering of cells using expression data from 200 genes provided a valuable framework for tracing developmental trajectories, capturing 18 reproducible ‘expression domains’ that can be mapped across time and tissue contexts. However, a limitation of this approach is that expression domains derived from a targeted 200-gene panel inevitably obscure finer cellular heterogeneity and bias domain assignments toward the selected marker set. For instance, expression domain clustering grouped two distinct tissue types: young leaves surrounding the inflorescence in late double ridge spikes (W2.5), and the glume and lemma tissues in the spikelet (stages W3.25 – W5), suggesting the panel lacked genes distinguishing these tissues. The capabilities of MERFISH continue to expand, with novel chemistry enabling the detection of fragmented and low-quality samples, and panel sizes expanding up to 1000 genes (Vizgen 2025). However, with continued technical advances, sequencing-based methods have the potential not only to replicate this type of spatiotemporal mapping but also to extend it to the whole transcriptome, offering a more comprehensive view of cell-type clustering in complex plant tissues.

A further advantage of sequencing-based platforms in polyploid crops such as wheat is their potential to achieve homoeolog-specific resolution. In our MERFISH dataset, the binding sites of gene probes were designed across the full length of an mRNA transcript of interest to maximise signal, similar to the probe design described for smFISH (Duncan et al. 2016). In hexaploid wheat, there is a high sequence similarity between homoeologs from the three sub-genomes (AABBDD; Ramírez-González et al. 2018). Given the many target sites across a transcript, we confirm this sequence similarity leads to the binding of gene probes across homoeologs in the A, B, and D sub-genomes. This means MERFISH cannot accurately distinguish between homoeolog-specific contributions to gene expression. Yet such differences can be biologically meaningful. Approximately 30% of wheat triads (composed of A, B, and D genome copies) exhibit nonbalanced expression, and in some cases, display high inter-tissue variation (Ramírez-González et al. 2018). To an even finer resolution, single-nuclei RNA-seq (snRNA-seq) on wheat roots shows that the level of non-balanced expression is highly heterogeneous. Within whole roots, bulk RNA-seq determined ~40% of homoeolog triads to be nonbalanced; however, when examining gene expression to a single-cell resolution, nonbalanced expression varied among cell-types, ranging from 31% to 76 % (Zhang et al. 2023). These patterns of tissue-level and cell-level variations in homoeolog balance are undetectable with probe-based imaging methods but could be resolved in future experiments with sequencing-based approaches.

Overall, I believe it is crucial to adopt a nuanced approach, comparing the diverse range of spatial transcriptomics techniques available to plant researchers. This field represents a world that is rapidly expanding, where no single platform provides a complete solution. Continued innovation, combined with a nuanced appreciation of the strengths and limitations of each method, will enable plant researchers to select spatial technologies best suited to their questions.

5.3 – Rapid advancements in spatial and single-cell biology will advance our understanding of meristematic development

Over recent years, single-cell transcriptomics has transformed our ability to interrogate meristematic development at unprecedented resolution. While decades of work have identified networks of individual genes controlling meristem fate, the precise cellular trajectories through developmental transitions remain poorly defined. Single-cell and single-nuclei RNA-seq now provide detailed profiles of individual cells, enabling the reconstruction of developmental trajectories and the detection of rare or transient states (Giacomello 2021). These approaches have already yielded comprehensive cell atlases of floral meristems in *Arabidopsis* (Neumann et al. 2022), and inflorescence tissues in rice and maize (Xu et al. 2021; Zong et al. 2022).

Nevertheless, scRNA-seq alone loses information on the physical location of cells, limiting our ability to connect transcriptional identity to tissue architecture and function. Developmental trajectories inferred from single-cell data can highlight changes in gene expression patterns and shifts in cell fate, but the spatial origin of individual cells can only be inferred indirectly from marker gene expression (Giacomello 2021). The integration of single-cell profiles with spatially resolved datasets overcomes this limitation, enabling the projection of cellular expression patterns onto tissue sections and effectively reconstructing spatially resolved single-cell atlases (implemented in Demesa-Arevalo et al. 2025; Xu et al. 2025)

Such integrative approaches are now advancing our understanding of meristem transitions in grasses. In barley, the combination of scRNA-seq with single-molecule fluorescence *in situ* hybridisation (smFISH) revealed the complexity of spikelet meristem (SM) identity and spikelet maturation. Analysis of *com1a;com2g* mutants—lacking two key regulators that suppress branch-like identity in AMs—showed that central SMs reverted to indeterminate IM-like states. These indeterminate structures produced multiple spikelets from their flanks, and the onset of *RAMOSA2 (RA2)* and *CRABS CLAW (CRC)* expression was reduced and delayed, signifying a dissimilarity from wild-type SMs. In contrast, wild-type SMs progressed smoothly toward FM and floret identities, characterising key regulators in this transition, including *COMPOSITUM1* (*COM1*; Demesa-Arevalo et al. 2025).

A comparable strategy was applied in wheat inflorescence, where a Molecular Cartography dataset comprising 99 genes was used to annotate cell clusters identified by snRNA-seq (Xu et al. 2025). Developmental trajectory analyses revealed that

meristematic cells (IM and SM) diverged along several paths, including toward floral meristems, or toward a distinct "transition zone" and suppressed bracts. This transition zone corresponded closely to the earliest basal phytomers described in our study and was characterised by preferential expression of *TEOSINTE BRANCHED 1 (TB1)*, *TB2*, and *SQUAMOSA PROMOTER BINDING PROTEIN LIKE 17 (SPL17)*. Gene imputation further indicated enrichment of *LEC1*, an observation also supported by our dataset. The authors concluded that these inferred developmental paths reflected the expected relationships among cell types derived from meristematic populations (Xu et al. 2025). Together, these studies highlight the power of scRNA-seq–based trajectory analysis integrated with spatial validation to explore meristem transitions. Looking forward, applying these approaches to finely resolved developmental time courses in wheat will be critical for dissecting how AMs transition to SM identity, providing a higher-resolution view of the dynamics that shape wheat spike architecture.

In parallel, the development of whole-mount spatial transcriptomic approaches offers a unique opportunity to dissect the complex three-dimensional organisation of meristems. Recent advances, such as Phytomap (Plant Hybridisation-based Target Observation of Gene Expression Map), which adapts multiplexed fluorescence in situ hybridisation (FISH) for intact plant tissues (Nobori et al. 2023), demonstrate the feasibility of capturing gene expression patterns without the need for sectioning. Our MERFISH dataset illustrates the value of such approaches: we observed the organisation of distinct gene expression domains associated with SM identity, with *KNOTTED HOMEOBOX LIKE 5 (KNOX5)* restricted to the L2/L3 layers, *COM1* expressed in bands adjacent to each meristem, and *AINTEGUMENTA-LIKE6 (AIL6)* forming discrete domains within single meristems. As development progresses, this spatial complexity increases further, with SMs giving rise to FMs on their flanks and *FRIZZLE PANICLE (FZP)* expressed in bands between glumes and floral meristems. While section-based profiling captures a portion of these spatially organised expression patterns, whole-mount strategies promise a more continuous and integrated view of spatial patterning, enabling a higher-resolution reconstruction of the earliest events in meristem differentiation. Looking forward, this technique presents promise in its applications to grass inflorescence development.

5.4 – The inflorescence is a cascade of phytomers

Grass morphology is organised into repeating units called phytomers, produced sequentially to form a modular architecture (Briske 1991; Moore and Moser 1995). Throughout this thesis, I have described wheat development using a phytomer-based framework, highlighting how the repetition of meristem–leaf units form the foundation of both vegetative and reproductive growth.

The application of spatial transcriptomics to such a repeating system is particularly powerful, as described by Laureyns et al. in their work, adapting in situ sequencing (ISS) to the maize shoot apex (2021). By sampling and sectioning both the shoot apex and recently initiated lateral primordia, they captured a cascade of phytomers at successive developmental stages, providing a spatiotemporal sequence of transcriptional events (Laureyns et al. 2021). We see clear parallels in our application of spatial transcriptomics to wheat, where our approach captured both inflorescence and vegetative phytomers in a gradient of age and developmental stage.

While microdissection-based bulk RNA-seq has allowed the profiling of several phytomer units at the base, centre, or apex of the inflorescence, spatial transcriptomics refines this to a higher resolution. Within the inflorescence, we detected opposing gradients of gene expression signatures distinguishing AMs across the apical-basal axis. During the double ridge (W2.5) and glume primordia (W3.25) stages, central phytomers exhibited gene expression signatures indicating the transition to SM identity, including *COM1* and *COM2/FZP*. By stages W3.25 and W4, these phytomers expressed floral identity markers such as *AGAMOUS LIKE-6 (AGL6)*, *AGAMOUS 1 (AG1)*, *SEPALLATA 3-1 (SEP3-1)*, and *SEP3-2*. The expression of these genes first observed in central phytomers highlights the establishment of the lanceolate shape of the spike. Importantly, these insights depended on maintaining spatial context. In single-cell approaches, the dissociation of tissue prior to sequencing removes positional information (Giacomello 2021), making it difficult (if not impossible) to reconstruct the sequential arrangement of phytomers through computational methods such as UMAP clustering and cell type identification. In contrast, the application of MERFISH connected the transcriptional state of each phytomer to their position across the apical-basal axis.

In addition, our capture of vegetative tissues below the inflorescence also expanded our understanding of meristem transition beyond what was possible with the micro-dissection experiment. In W2.5 samples, we observed that AMs are enriched in TAW1 and MND1 genes implicated in meristem identity transition (Yoshida et al. 2013; Walla et al. 2020),

with *MND1* specifically characterised in barley to suppress the transition of vegetative AMs to an IM identity (Walla et al. 2020). These signatures of meristematic suppression were absent in inflorescence phytomers, which instead expressed genes associated with LR suppression, such as *TSH1* (Wang et al. 2009; Whipple et al. 2010; Houston et al. 2012; Xiao et al. 2022). Thus, by comparing vegetative and reproductive phytomers, we observed distinct regulatory modules that separate vegetative from reproductive phytomer contexts. These findings provide a framework for functional characterisation, as they highlight candidate genes whose altered expression could shift meristem fate. For example, testing whether ectopic expression of *MND1* or *TAW1* in inflorescence AMs alters the timing or outcome of meristem identity transitions would directly evaluate their regulatory roles. In this way, our spatial dataset not only describes transcriptional signatures but also generates hypotheses for dissecting the genetic control of meristem transitions.

5.5 - Making a model of the wheat spike

Here we interpret inflorescence development within a framework of meristem identity transitions, where shifts in meristem fate across developmental time underpin the patterning of branches, spikelets, and florets, which collectively establish the architecture of the inflorescence. One question we can examine is whether the diversity of inflorescence types can be modelled through a common mechanism of meristem identity transitions (Prusinkiewicz et al. 2007; Harder and Prusinkiewicz 2013).

Early models of inflorescence development describe an iterative series of 'switches' that control the timing of transitions between meristematic identities (Kellogg 2000). Building on this concept, Prusinkiewicz et al. developed a computational model in which meristem identity is governed by a hypothetical variable representing the degree of "vegetativeness" (*veg*) of a meristem (2007). In this framework, *veg* declines over developmental time, and a transition to floral meristem (FM) identity occurs once *veg* falls below a critical threshold. When *veg* remains high, meristems adopt an indeterminate fate and produce branching structures, whereas lower *veg* values promote the switch to FM identity. Thus, temporal variation in *veg* provides a dynamic mechanism for regulating meristem fate, offering a unifying principle to explain inflorescence architectures. Indeed, the model can produce a continuum of inflorescence structures, creating racemes, panicles, and cymes, highlighting its capacity to capture the diversity of plant architectures observed in nature. (Prusinkiewicz et al. 2007).

Such a framework aligns with the broader evolutionary concept of heterochrony, in which shifts in the timing of developmental programs generate variation in organ size, shape, and number (Gould 1988; Geuten and Coenen 2013; Buendía-Monreal and Gillmor 2018). In the Prusinkiewicz et al. model, the developmental threshold at which *veg* falls below a critical value provides a direct source of heterochronic variation (2007). If this transition occurs earlier, meristems switch more rapidly to a floral fate, whereas a delayed transition prolongs indeterminate growth and branching. These temporal shifts in meristem fate do not require changes to the underlying developmental program but rather to its timing, thereby illustrating how heterochrony can diversify inflorescence architecture (Bartlett and Thompson 2014; Bommert and Whipple 2018; Koppolu and Schnurbusch 2019).

While this model highlights how heterochrony in meristem identity transitions shapes inflorescence diversity across species, the same principles can also be applied at a finer scale within a single inflorescence. Variation in the timing of meristem identity transition is also a developmental feature observable along the axis of one spike. Early developmental

observations of the wheat spike reveal that basal meristems are delayed in their progression to SM identity compared with more central meristems (Bonnett 1966; Waddington et al. 1983; Kirby and Appleyard 1984). Consistent with this, our MERFISH dataset identifies transcriptionally distinct states between basal and central meristems, highlighting the specific expression of transcription factors *RI-LIKE1 (RIL1), SPL14* to basal AMs, and *KNOX5, AIL6*, and *COM1* to central AMs. Given the role of these regulatory genes in meristem identity and transition (Jackson et al. 1994; Hake et al. 1995; Nole-Wilson et al. 2005; Poursarebani et al. 2015, 2020; Wang et al. 2015; Ikeda et al. 2019), our results indicate the presence of distinct signatures in basal AMs consistent with delayed progression toward spikelet identity.

Why might basal meristems undergo this delay? One explanation lies in their developmental context: the earliest initiated ridges form during a transition from vegetative to reproductive growth. In this light, the basal ridges may originate under a heightened "vegetative" state, slowing their transition toward spikelet fate (Backhaus et al. 2022). *SHORT VEGETATIVE PHASE (SVP)* transcription factors are key in this transition between vegetative and reproductive cues, which provide a molecular correlate for the *veg* factor in the Prusinkiewicz model. *SVP*s are broadly associated with vegetative growth and are downregulated following floral transition across species, including Arabidopsis (Gregis et al. 2013), wheat (Adamski et al. 2021; Li et al. 2021b; Liu et al. 2021), rice (Sentoku et al. 2005; Lee et al. 2008), and barley (Trevaskis et al. 2007). Functionally, they repress the transition of AMs to an IM or SM identity during vegetative phases, with *vrt2 svp1* double mutants producing axillary spikelets or spikes at sub-peduncle nodes (Li et al. 2021b). It is proposed by Li et al. that *SVPs* disrupt SQUAMOSA–SEP protein interactions required for normal spikelet development (Li et al. 2021b), thereby providing a mechanism of action by which *SVP* expression must decrease for reproductive development to proceed. As observed in both the micro-dissection bulk RNA-seq and MERFISH dataset, *SVP* genes are more highly expressed in the basal ridges (first initiated), suggesting a molecular mechanism that holds these meristems back from transitioning to SM identity. This interpretation agrees with observations that *VRT-A2* dosage correlates with the presence of rudimentary basal spikelets (Backhaus et al. 2022), further reinforcing the idea that elevated *SVP* activity delays or suppresses meristem fate transitions at the base of the spike.

Based on these concepts, Backhaus et al. developed a computational model of a growing wheat spike (2022), which is influenced by the role of *SVPs* and *SEPs* as proposed by Li et al. (2021). In this model, *SVP* is a "vegetative" signal, and *SEP* is a "floral signal". The

model assumes that 1) *SVP* suppresses *SEP* expression, 2) *SEP* promotes spikelet outgrowth, and 3) *SVP* decreases over time. This model recreated the opposing gradients of *SVP* and *SEP* transcription factors across the apical-basal axis of the inflorescence, in addition to recapitulating the formation of the lanceolate shape (Backhaus et al. 2022), providing evidence similar to the Prusinkiewicz model, that a decreasing level of *vegetative* signal can control the timing and behaviour of meristem fate, ultimately influencing the shape and form of the inflorescence (Prusinkiewicz et al. 2007).

Our MERFISH dataset provides an opportunity to refine such models of wheat spike development further. First, transcriptional differences between central and basal meristems could define a molecular 'signature' of meristem identity, which can then be tested in existing germplasm such as transgenic *VRT-A2* overexpression lines (Adamski et al. 2021) or *vrt2svp1*-null double mutants (Li et al. 2021b) to assess how *SVPs* contribute to delayed basal meristem transitions. In addition, MERFISH revealed a distinct set of genes restricted to basal LRs, including *SPL17* and *LEAFY COTYLEDON1 (LEC1)*. Given the hypothesis that LRs act as signalling centres regulating adjacent AMs (Whipple et al. 2010), this expression pattern suggests coordinated regulation between LRs and AMs along the apical–basal axis. An important future direction will be to test whether these LR-specific expression programs operate independently of *SVP* activity, and whether the organisation of LRs adds an additional regulatory dimension to models of meristem identity transitions in wheat. As spatial transcriptomic methods become more widely accessible, including the optimisation of approaches such as Phytomap, there is considerable promise for using these techniques to characterise developmental regulators across diverse genetic backgrounds and fine-scale time courses. In doing so, we can further bridge developmental modelling with functional genetics, providing a framework to dissect how spatiotemporal regulation of meristem fate shapes inflorescence architecture.

5.6 – Interpretations of spatial data require functional validation

The distinct expression signatures observed along the apical–basal axis of late double-ridge spikes within both AMs and adjacent LRs suggest a potential contribution to the delayed progression of basal spikelet ridges. Whether these transcriptional states actively drive such delays remains unresolved. This highlights the value of functional studies, where loss-of-function mutations can test the causal roles of these genes in meristem identity transitions during inflorescence development.

While rudimentary basal spikelet formation has been a central focus of this thesis, the broader literature in grass inflorescence development emphasises that a wider spectrum of phenotypes reflects altered meristem identity transitions. Variation in branch number, spikelet number, bract outgrowth, and the ectopic formation of branch-like structures all represent developmental outcomes of shifts in meristem fate. In this study, MERFISH data identify a suite of candidate genes associated with meristem identity in wheat—including *RIL1*, *AIL6*, *KNOX5*, *COM1*, *LEC1*, *MND1*, and *TAW1*. Functional characterisation of these genes will provide valuable insight into their roles in regulating meristem fate and inflorescence architecture.

In addition to their potential regulatory roles, the candidate genes identified here can serve as transcriptional signatures for follow-up work in genetic lines that exhibit altered inflorescence phenotypes. For instance, assessing whether high-RBS lines share similar expression profiles marked by *RIL1*, *SPL14*, or *LEC1* would provide a means to test whether delayed meristem identity transitions at the spike base are consistently associated with specific gene signatures. Comparative analysis of wild-type and developmental mutant lines could therefore refine the functional roles of these candidate regulators and establish diagnostic markers of basal meristem fate.

Recent advances in transgenic methods also open exciting opportunities for functional characterisation of these genes in wheat. Our lab has optimised a technique to mis-express candidate genes with a transgenic approach. In brief, the semi-spatial microdissection dataset was used to identify genes that are strongly expressed at the base of the inflorescence. To test their regulatory potential, these cis-regulatory elements were placed upstream of two candidate regulators of spikelet development, *SEP1-6* and *MOF1*. Preliminary evidence suggests that this targeted overexpression of *SEP1-6* reduced the number of rudimentary basal spikelets (Jones 2025, unpublished).

The integration of this technique with the spatial data provided by MERFISH is clear: gene expression data from 200 genes identifies highly specific regulatory contexts, such as

SPL17 expression restricted to basal LRs, or *RIL1* expression in basal LRs and AMs. Using these regulatory elements to mis-express genes such as *KNOX5*, which is enriched in central meristems, would provide a powerful strategy to test how altering spatial domains of expression reshapes meristem identity transitions. Alternatively, *MND1* expression, known to suppress axillary meristems outgrowth in vegetative tissues (Walla et al. 2020), may contribute to the delayed basal phenotype when ectopically expressed within the inflorescence. Together, these candidate genes and emerging functional approaches highlight the potential of combining high-resolution spatial datasets with targeted genetic manipulation. Such integration will enable not only the validation of putative regulators but also the dissection of their activity in precise spatial contexts, ultimately refining our understanding of how meristem identity transitions are coordinated across developmental time and space.

5.7 – Concluding statement

This thesis represents a first step toward optimising spatial transcriptomic approaches in wheat, demonstrating how these technologies can be applied to uncover the cellular and molecular organisation of the inflorescence. By generating and analysing a dataset of 200 genes across four developmental stages of wheat spike development, I have provided new insights into the genetic regulation of spikelet and floral development, as well as the patterning events that underlie the early formation of the wheat spike. Beyond these specific findings, I hope the release of this dataset offers a valuable resource to the wheat research community, in addition to enabling broader comparisons across related grasses and supporting efforts to dissect conserved developmental programs.

As single-cell and spatial transcriptomic methods continue to advance, they will open new avenues to understand and characterise development in complex plant tissues such as the grass inflorescence. While these approaches provide fundamental insights into gene regulation, they also hold promise as practical tools to support crop improvement, allowing researchers to rapidly characterise gene expression and extend insights to select candidate genes of interest. By offering high-resolution gene expression data in wheat spikes, I hope this work contributes to the toolkit available to researchers, ultimately supporting efforts to develop higher-yielding and more resilient wheat varieties.

Bibliography

Adamski NM, Borrill P, Brinton J, Harrington SA, Marchal C, Bentley AR, Bovill WD, Cattivelli L, Cockram J, Contreras-Moreira B, et al. A roadmap for gene functional characterisation in crops with large genomes: Lessons from polyploid wheat. eLife. 2020:9:e55646. https://doi.org/10.7554/eLife.55646

Adamski NM, Simmonds J, Brinton JF, Backhaus AE, Chen Y, Smedley M, Hayta S, Florio T, Crane P, Scott P, et al. Ectopic expression of Triticum polonicum *VRT-A2* underlies elongated glumes and grains in hexaploid wheat in a dosage-dependent manner. Plant Cell. 2021:33(7):2296–2319. https://doi.org/10.1093/plcell/koab119

Aida M, Ishida T, Fukaki H, Fujisawa H, and Tasaka M. Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. Plant Cell. 1997:9(6):841–857. https://doi.org/10.1105/tpc.9.6.841

Aida M, Ishida T, and Tasaka M. Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. Dev Camb Engl. 1999:126(8):1563–1570. https://doi.org/10.1242/dev.126.8.1563

Alam O and Purugganan MD. Domestication and the evolution of crops: variable syndromes, complex genetic architectures, and ecological entanglements. Plant Cell. 2024:36(5):1227–1241. https://doi.org/10.1093/plcell/koae013

Ambrose BA, Lerner DR, Ciceri P, Padilla CM, Yanofsky MF, and Schmidt RJ. Molecular and Genetic Analyses of the *Silky1* Gene Reveal Conservation in Floral Organ Specification between Eudicots and Monocots. Mol Cell. 2000:5(3):569–579. https://doi.org/10.1016/S1097-2765(00)80450-5

Arranz-Otaegui A, Gonzalez Carretero L, Ramsey MN, Fuller DQ, and Richter T. Archaeobotanical evidence reveals the origins of bread 14,400 years ago in northeastern Jordan. Proc Natl Acad Sci. 2018:115(31):7925–7930. https://doi.org/10.1073/pnas.1801071115

Backhaus AE, Griffiths C, Vergara-Cruces A, Simmonds J, Lee R, Morris RJ, and Uauy C. Delayed development of basal spikelets in wheat explains their increased floret abortion and rudimentary nature. J Exp Bot. 2023:74(17):5088–5103. https://doi.org/10.1093/jxb/erad233

Backhaus AE, Lister A, Tomkins M, Adamski NM, Simmonds J, Macaulay I, Morris RJ, Haerty W, and Uauy C. High expression of the MADS-box gene *VRT2* increases the number of rudimentary basal spikelets in wheat. Plant Physiol. 2022:189(3):1536–1552. https://doi.org/10.1093/plphys/kiac156

Bai F, Reinheimer R, Durantini D, Kellogg EA, and Schmidt RJ. TCP transcription factor, *BRANCH ANGLE DEFECTIVE 1 (BAD1)*, is required for normal tassel branch angle formation in maize. Proc Natl Acad Sci. 2012:109(30):12225–12230. https://doi.org/10.1073/pnas.1202439109

Bai X, Huang Y, Mao D, Wen M, Zhang L, and Xing Y. Regulatory role of FZP in the determination of panicle branching and spikelet formation in rice. Sci Rep. 2016:6:19022. https://doi.org/10.1038/srep19022

Bartlett ME and Thompson B. Meristem identity and phyllotaxis in inflorescence development. Front Plant Sci. 2014:5:508. https://doi.org/10.3389/fpls.2014.00508

Barton MK and Poethig RS. Formation of the shoot apical meristem in Arabidopsis thaliana: an analysis of development in the wild type and in the shoot meristemless mutant. Development. 1993:119(3):823–831. https://doi.org/10.1242/dev.119.3.823

Becker A and Theißen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol. 2003:29(3):464–489. https://doi.org/10.1016/S1055-7903(03)00207-0

Bettinger RL, Barton L, and Morgan C. The origins of food production in north China: A different kind of agricultural revolution. Evol Anthropol Issues News Rev. 2010:19(1):9–21. https://doi.org/10.1002/evan.20236

Boden SA and Østergaard L. How can developmental biology help feed a growing population? Development. 2019:146(3):dev172965. https://doi.org/10.1242/dev.172965

Bolser DM, Kerhornou A, Walts B, and Kersey P. *Triticeae* Resources in Ensembl Plants. Plant Cell Physiol. 2015:56(1):e3. https://doi.org/10.1093/pcp/pcu183

Bomblies K, Wang R-L, Ambrose BA, Schmidt RJ, Meeley RB, and Doebley J. Duplicate *FLORICAULA/LEAFY* homologs *zfl1* and *zfl2* control inflorescence architecture and flower patterning in maize. Development. 2003:130(11):2385–2395. https://doi.org/10.1242/dev.00457

Bommert P, Satoh-Nagasawa N, Jackson D, and Hirano H-Y. Genetics and evolution of inflorescence and flower development in grasses. Plant Cell Physiol. 2005:46(1):69–78. https://doi.org/10.1093/pcp/pci504

Bommert P and Whipple C. Grass inflorescence architecture and meristem determinacy. Semin Cell Dev Biol. 2018:79:37–47. https://doi.org/10.1016/j.semcdb.2017.10.004

Bonnett OT. Inflorescences of maize, wheat, rye, barley, and oats : their initiation and development (University of Illinois (Urbana-Champaign campus) Agricultural Experiment Station: Champaign, IL).

Borrill P, Adamski N, and Uauy C. Genomics as the key to unlocking the polyploid potential of wheat. New Phytol. 2015:208(4):1008–1022. https://doi.org/10.1111/nph.13533

Borrill P, Harrington SA, and Uauy C. Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat. Plant J. 2019:97(1):56–72. https://doi.org/10.1111/tpj.14150

Borrill P, Ramirez-Gonzalez R, and Uauy C. expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform. Plant Physiol. 2016:170(4):2172–2186. https://doi.org/10.1104/pp.15.01667

Bray NL, Pimentel H, Melsted P, and Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016:34(5):525–527. https://doi.org/10.1038/nbt.3519

Brinton J and Uauy C. A reductionist approach to dissecting grain weight and yield in wheat. J Integr Plant Biol. 2019:61(3):337–358. https://doi.org/10.1111/jipb.12741

Briske DD. Developmental morphology and physiology of grasses. . In. Grazing Management: An Ecological Perspective. (Timber Press: Portland), pp. 85–108.

Buendía-Monreal M and Gillmor CS. The Times They Are A-Changin': Heterochrony in Plant Development and Evolution. Front Plant Sci. 2018:9:1349. https://doi.org/10.3389/fpls.2018.01349

Calderini DF and Ortiz-Monasterio I. Grain Position Affects Grain Macronutrient and Micronutrient Concentrations in Wheat. Crop Sci. 2003:43(1):141–151. https://doi.org/10.2135/cropsci2003.1410

Carbajal-Friedrich AAJ and Burgess AJ. The role of the ideotype in future agricultural production. Front Plant Physiol. 2024:2. https://doi.org/10.3389/fphgy.2024.1341617

Causier B, Schwarz-Sommer Z, and Davies B. Floral organ identity: 20 years of ABCs. Semin Cell Dev Biol. 2010:21(1):73–79. https://doi.org/10.1016/j.semcdb.2009.10.005

Cervantes-Pérez SA, Zogli P, Amini S, Thibivilliers S, Tennant S, Hossain MS, Xu H, Meyer I, Nooka A, Ma P, et al. Single-cell transcriptome atlases of soybean root and mature nodule reveal new regulatory programs that control the nodulation process. Plant Commun. 2024:5(8):100984. https://doi.org/10.1016/j.xplc.2024.100984

Chang Z, Xu R, Xun Q, Liu J, Zhong T, Ding Y, and Ding C. OsmiR164-targeted *OsNAM*, a boundary gene, plays important roles in rice leaf and panicle development. Plant J Cell Mol Biol. 2021:106(1):41–55. https://doi.org/10.1111/tpj.15143

Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, Qiu X, Yang J, Xu J, Hao S, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell. 2022:185(10):1777-1792.e21. https://doi.org/10.1016/j.cell.2022.04.003

Chen KH, Boettiger AN, Moffitt JR, Wang S, and Zhuang X. Spatially resolved, highly multiplexed RNA profiling in single cells. Science. 2015. https://doi.org/10.1126/science.aaa6090

Choi J, Li J, Ferdous S, Liang Q, Moffitt JR, and Chen R. Spatial organization of the mouse retina at single cell resolution by MERFISH. Nat Commun. 2023:14(1):4929. https://doi.org/10.1038/s41467-023-40674-3

Chuck G and Bortiri E. The unique relationship between *tsh4* and *ra2* in patterning floral phytomers. Plant Signal Behav. 2010:5(8):979–981. https://doi.org/10.4161/psb.5.8.12220

Chuck G, Muszynski M, Kellogg E, Hake S, and Schmidt RJ. The Control of Spikelet Meristem Identity by the *branched silkless1* Gene in Maize. Science. 2002:298(5596):1238–1241. https://doi.org/10.1126/science.1076920

Chuck G, Whipple C, Jackson D, and Hake S. The maize SBP-box transcription factor encoded by *tasselsheath4* regulates bract development and the establishment of

meristem boundaries. Development. 2010:137(8):1243–1250. https://doi.org/10.1242/dev.048348

Chuck GS, Brown PJ, Meeley R, and Hake S. Maize *SBP*-box transcription factors *unbranched2* and *unbranched3* affect yield traits by regulating the rate of lateral primordia initiation. Proc Natl Acad Sci U S A. 2014:111(52):18775–18780. https://doi.org/10.1073/pnas.1407401112

Cisar C, Keener N, Ruffalo M, and Paten B. A unified pipeline for FISH spatial transcriptomics. Cell Genomics. 2023:3(9):100384. https://doi.org/10.1016/j.xgen.2023.100384

Clark SE, Jacobsen SE, Levin JZ, and Meyerowitz EM. The *CLAVATA* and *SHOOT MERISTEMLESS* loci competitively regulate meristem activity in Arabidopsis. Dev Camb Engl. 1996:122(5):1567–1575. https://doi.org/10.1242/dev.122.5.1567

Claros MG, Bautista R, Guerrero-Fernández D, Benzerki H, Seoane P, and Fernández-Pozo N. Why Assembling Plant Genome Sequences Is So Challenging. Biology. 2012:1(2):439–459. https://doi.org/10.3390/biology1020439

Clayton, W. D. The spikelet. . In. Reproductive versatility in the grasses. (Cambridge University Press: Cambridge, Massachusetts, USA), pp. 32–51.

Clement CR, De Cristo-Araújo M, Coppens D'Eeckenbrugge G, Alves Pereira A, and Picanço-Rodrigues D. Origin and Domestication of Native Amazonian Crops. Diversity. 2010:2(1):72–106. https://doi.org/10.3390/d2010072

Clifford TH. Spikelet and floral morphology. . In. Grass systematics and evolution. (Smithsonian Institution Press: Washington DC), pp. 21–30.

Coen ES and Meyerowitz EM. The war of the whorls: genetic interactions controlling flower development. Nature. 1991:353(6339):31–37. https://doi.org/10.1038/353031a0

Colombo L, Franken J, Koetje E, van Went J, Dons HJ, Angenent GC, and van Tunen AJ. The petunia MADS box gene *FBP11* determines ovule identity. Plant Cell. 1995:7(11):1859–1868. https://doi.org/10.1105/tpc.7.11.1859

Covert I, Gala R, Wang T, Svoboda K, Sümbül U, and Lee S-I. Predictive and robust gene selection for spatial transcriptomics. Nat Commun. 2023:14:2091. https://doi.org/10.1038/s41467-023-37392-1

Cui R, Han J, Zhao S, Su K, Wu F, Du X, Xu Q, Chong K, Theissen G, and Meng Z. Functional conservation and diversification of class E floral homeotic genes in rice (*Oryza sativa*). Plant J Cell Mol Biol. 2010:61(5):767–781. https://doi.org/10.1111/j.1365-313X.2009.04101.x

Debernardi JM, Greenwood JR, Jean Finnegan E, Jernstedt J, and Dubcovsky J. APETALA 2-like genes *AP2L2* and *Q* specify lemma identity and axillary floral meristem development in wheat. Plant J. 2020:101(1):171–187. https://doi.org/10.1111/tpj.14528

Debernardi JM, Lin H, Chuck G, Faris JD, and Dubcovsky J. microRNA172 plays a crucial role in wheat spike morphogenesis and grain threshability. Dev Camb Engl. 2017:144(11):1966–1975. https://doi.org/10.1242/dev.146399

Demesa-Arevalo E, Dörpholz H, Vardanega I, Maika JE, Pineda-Valentino I, Eggels S, Lautwein T, Köhrer K, Schnurbusch T, Korff M von, et al. Imputation integrates single-cell and spatial gene expression data to resolve transcriptional networks in barley shoot meristem development. 2025:2025.05.09.653223. https://doi.org/10.1101/2025.05.09.653223

Denyer T, Wu P-J, Colt K, Abramson BW, Pang Z, Solansky P, Mamerto A, Nobori T, Ecker JR, Lam E, et al. Streamlined spatial and environmental expression signatures characterize the minimalist duckweed Wolffia australiana. Genome Res. 2024:34(7):1106–1120. https://doi.org/10.1101/gr.279091.124

Dillehay TD, Rossen J, Andres TC, and Williams DE. Preceramic adoption of peanut, squash, and cotton in northern Peru. Science. 2007:316(5833):1890–1893. https://doi.org/10.1126/science.1141395

Ditta G, Pinyopich A, Robles P, Pelaz S, and Yanofsky MF. The *SEP4* Gene of *Arabidopsis thaliana* Functions in Floral Organ and Meristem Identity. Curr Biol. 2004:14(21):1935–1940. https://doi.org/10.1016/j.cub.2004.10.028

Dobrovolskaya O, Pont C, Sibout R, Martinek P, Badaeva E, Murat F, Chosson A, Watanabe N, Prat E, Gautier N, et al. *FRIZZY PANICLE* drives supernumerary spikelets in bread wheat. Plant Physiol. 2015:167(1):189–199. https://doi.org/10.1104/pp.114.250043

Donald CM. The breeding of crop ideotypes. Euphytica. 1968:17(3):385–403. https://doi.org/10.1007/BF00056241

Donaldson L. Autofluorescence in Plants. Molecules. 2020:25(10):2393. https://doi.org/10.3390/molecules25102393

Dreni L and Ferrándiz C. Tracing the Evolution of the *SEPALLATA* Subfamily across Angiosperms Associated with Neo- and Sub-Functionalization for Reproductive and Agronomically Relevant Traits. Plants. 2022:11(21):2934. https://doi.org/10.3390/plants11212934

Dreni L, Pilatone A, Yun D, Erreni S, Pajoro A, Caporali E, Zhang D, and Kater MM. Functional Analysis of All AGAMOUS Subfamily Members in Rice Reveals Their Roles in Reproductive Organ Identity Determination and Meristem Determinacy. Plant Cell. 2011:23(8):2850–2863. https://doi.org/10.1105/tpc.111.087007

Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, Nazarenko I, Nilsen GB, Yeung G, et al. Human Genome Sequencing Using Unchained Base Reads on Self-Assembling DNA Nanoarrays. Science. 2010:327(5961):78–81. https://doi.org/10.1126/science.1181498

Dubcovsky J and Dvorak J. Genome plasticity a key factor in the success of polyploid wheat under domestication. Science. 2007:316(5833):1862–1866. https://doi.org/10.1126/science.1143986

Duncan S, Johansson HE, and Ding Y. Reference genes for quantitative Arabidopsis single molecule RNA fluorescence in situ hybridization. J Exp Bot. 2022:74(7):2405–2415. https://doi.org/10.1093/jxb/erac521

Duncan S, Olsson TSG, Hartley M, Dean C, and Rosa S. A method for detecting single mRNA molecules in Arabidopsis thaliana. Plant Methods. 2016:12(1):13. https://doi.org/10.1186/s13007-016-0114-x

Edgar RC. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun. 2022:13(1):6968. https://doi.org/10.1038/s41467-022-34630-w

Ellis EC, Kaplan JO, Fuller DQ, Vavrus S, Klein Goldewijk K, and Verburg PH. Used planet: A global history. Proc Natl Acad Sci. 2013:110(20):7978–7985. https://doi.org/10.1073/pnas.1217241110

Endress PK. Disentangling confusions in inflorescence morphology: Patterns and diversity of reproductive shoot ramification in angiosperms. J Syst Evol. 2010:48(4):225–239. https://doi.org/10.1111/j.1759-6831.2010.00087.x

Evans CEB, Arunkumar R, and Borrill P. Transcription factor retention through multiple polyploidization steps in wheat. G3 Bethesda Md. 2022:12(8):jkac147. https://doi.org/10.1093/g3journal/jkac147

Evans MW and Grover F. Developmental morphology of the growing point of the shoot and the inflorescence in grasses. J Agric Res. 1940.

Evers T and Millar S. Cereal Grain Structure and Development: Some Implications for Quality. J Cereal Sci. 2002:36(3):261–284. https://doi.org/10.1006/jcrs.2002.0435

Faci I, Backhaus AE, and Uauy C. Wheat spike meristem micro-dissection. 2024. https://doi.org/dx.doi.org/10.17504/protocols.io.3byl49r2zgo5/v1

FAO. The future of food and agriculture – Alternative pathways to 2050 (Rome).

FAO. World Food and Agriculture – Statistical Yearbook 2024 (FAO: Rome).

Favaro R, Pinyopich A, Battaglia R, Kooiker M, Borghi L, Ditta G, Yanofsky MF, Kater MM, and Colombo L. MADS-box protein complexes control carpel and ovule development in *Arabidopsis*. Plant Cell. 2003:15(11):2603–2611. https://doi.org/10.1105/tpc.015123

Femino AM, Fay FS, Fogarty K, and Singer RH. Visualization of Single RNA Transcripts in Situ. Science. 1998:280(5363):585–590. https://doi.org/10.1126/science.280.5363.585

Fischer DS, Theis FJ, and Yosef N. Impulse model-based differential expression analysis of time course sequencing data. Nucleic Acids Res. 2018:46(20):e119. https://doi.org/10.1093/nar/gky675

Fischer T. History of Wheat Breeding: A Personal View. . In. Wheat Improvement: Food Security in a Changing Climate, MP Reynolds and H-J Braun, eds. (Springer International Publishing: Cham), pp. 17–30. https://doi.org/10.1007/978-3-030-90673-3_2

Forster BP, Franckowiak JD, Lundqvist U, Lyon J, Pitkethly I, and Thomas WTB. The Barley Phytomer. Ann Bot. 2007:100(4):725–733. https://doi.org/10.1093/aob/mcm183

Fraser J and Kokko EG. Panicle, spikelet, and floret development in orchardgrass *(Dactylis glomerata)*. Can J Bot. 1993:71(4):523–532. https://doi.org/10.1139/b93-058

Friedman J and Harder LD. Inflorescence architecture and wind pollination in six grass species. Funct Ecol. 2004:18(6):851–860. https://doi.org/10.1111/j.0269-8463.2004.00921.x

Fu X, Sun L, Dong R, Chen JY, Silakit R, Condon LF, Lin Y, Lin S, Palmiter RD, and Gu L. Polony gels enable amplifiable DNA stamping and spatial transcriptomics of chronic pain. Cell. 2022:185(24):4621-4633.e17. https://doi.org/10.1016/j.cell.2022.10.021

Fu Y, Xiao W, Tian L, Guo L, Ma G, Ji C, Huang Y, Wang H, Wu X, Yang T, et al. Spatial transcriptomics uncover sucrose post-phloem transport during maize kernel development. Nat Commun. 2023:14(1):7191. https://doi.org/10.1038/s41467-023-43006-7

Fuller DQ. An Emerging Paradigm Shift in the Origins of Agriculture. Gen Anthropol. 2010:17(2):1–12. https://doi.org/10.1111/j.1939-3466.2010.00010.x

Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, and Sun G-P. The Domestication Process and Domestication Rate in Rice: Spikelet Bases from the Lower Yangtze. Science. 2009:323(5921):1607–1610. https://doi.org/10.1126/science.1166605

Fuller DQ, Willcox G, and Allaby RG. Cultivation and domestication had multiple origins: arguments against the core area hypothesis for the origins of agriculture in the Near East. World Archaeol. 2011:43(4):628–652. https://doi.org/10.1080/00438243.2011.624747

Gall JG and Pardue ML. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci U S A. 1969:63(2):378–383. https://doi.org/10.1073/pnas.63.2.378

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Pè ME, and Schmidt RJ. The role of *barren stalk1* in the architecture of maize. Nature. 2004:432(7017):630–635. https://doi.org/10.1038/nature03148

Gao X, Liang W, Yin C, Ji S, Wang H, Su X, Guo C, Kong H, Xue H, and Zhang D. The *SEPALLATA*-Like Gene *OsMADS34* Is Required for Rice Inflorescence and Spikelet Development. Plant Physiol. 2010:153(2):728–740. https://doi.org/10.1104/pp.110.156711

Geuten K and Coenen H. Heterochronic genes in plant evolution and development. Front Plant Sci. 2013:4:381. https://doi.org/10.3389/fpls.2013.00381

Giacomello S. A new era for plant science: spatial single-cell transcriptomics. Curr Opin Plant Biol. 2021:60:102041. https://doi.org/10.1016/j.pbi.2021.102041

Giacomello S, Salmén F, Terebieniec BK, Vickovic S, Navarro JF, Alexeyenko A, Reimegård J, McKee LS, Mannapperuma C, Bulone V, et al. Spatially resolved transcriptome profiling in model plant species. Nat Plants. 2017:3:17061. https://doi.org/10.1038/nplants.2017.61

Gillies S, van der Wel C, Van den Bossche J, Taves MW, Arnott J, and Ward B. Shapely. 2022. https://doi.org/10.5281/zenodo.7428463

Glémin S and Bataillon T. A comparative view of the evolution of grasses under domestication. New Phytol. 2009:183(2):273–290. https://doi.org/10.1111/j.1469-8137.2009.02884.x

Glémin S, Scornavacca C, Dainat J, Burgarella C, Viader V, Ardisson M, Sarah G, Santoni S, David J, and Ranwez V. Pervasive hybridizations in the history of wheat relatives. Sci Adv. 2019:5(5):eaav9188. https://doi.org/10.1126/sciadv.aav9188

Gohlke C. cgohlke/tifffile: v2024.5.10. 2024. https://doi.org/10.5281/zenodo.11170947

Gong C, Li S, Wang L, Zhao F, Fang S, Yuan D, Zhao Z, He Q, Li M, Liu W, et al. SAW: an efficient and accurate data analysis workflow for Stereo-seq spatial transcriptomics. GigaByte Hong Kong China. 2024:2024:gigabyte111. https://doi.org/10.46471/gigabyte.111

Goto K and Meyerowitz EM. Function and regulation of the Arabidopsis floral homeotic gene *PISTILLATA*. Genes Dev. 1994:8(13):1548–1560. https://doi.org/10.1101/gad.8.13.1548

Gould SJ. The Uses of Heterochrony. . In. Heterochrony in Evolution: A Multidisciplinary Approach, ML McKinney, ed. (Springer US: Boston, MA), pp. 1–13. https://doi.org/10.1007/978-1-4899-0795-0_1

Gregis V, Andrés F, Sessa A, Guerra RF, Simonini S, Mateos JL, Torti S, Zambelli F, Prazzoli GM, Bjerkan KN, et al. Identification of pathways directly regulated by *SHORT VEGETATIVE PHASE* during vegetative and reproductive development in Arabidopsis. Genome Biol. 2013:14(6):R56. https://doi.org/10.1186/gb-2013-14-6-r56

Groiss S, Pabst D, Faber C, Meier A, Bogdoll A, Unger C, Nilges B, Strauss S, Föderl-Höbenreich E, Hardt M, et al. Highly resolved spatial transcriptomics for detection of rare events in cells. 2021:2021.10.11.463936. https://doi.org/10.1101/2021.10.11.463936

Guillotin B, Rahni R, Passalacqua M, Mohammed MA, Xu X, Raju SK, Ramírez CO, Jackson D, Groen SC, Gillis J, et al. A pan-grass transcriptome reveals patterns of cellular divergence in crops. Nature. 2023:617(7962):785–791. https://doi.org/10.1038/s41586-023-06053-0

Hake S, Char BR, Chuck G, Foster T, Long J, and Jackson D. Homeobox genes in the functioning of plant meristems. Philos Trans R Soc Lond B Biol Sci. 1995:350(1331):45–51. https://doi.org/10.1098/rstb.1995.0136

Hammer K. Das Domestikationssyndrom. Kult. 1984:32(1):11–34. https://doi.org/10.1007/BF02098682

Harder LD and Prusinkiewicz P. The interplay between inflorescence development and function as the crucible of architectural diversity. Ann Bot. 2013:112(8):1477–1493. https://doi.org/10.1093/aob/mcs252

Harris CR, Millman KJ, Van Der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, et al. Array programming with NumPy. Nature. 2020:585(7825):357–362. https://doi.org/10.1038/s41586-020-2649-2

Harrop TWR, Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, and Kater MM. Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. Plant J Cell Mol Biol. 2016:86(1):75–88. https://doi.org/10.1111/tpj.13147

Hawkesford MJ, Araus J-L, Park R, Calderini D, Miralles D, Shen T, Zhang J, and Parry MAJ. Prospects of doubling global wheat yields. Food Energy Secur. 2013:2(1):34–48. https://doi.org/10.1002/fes3.15

Hayta S, Smedley MA, Demir SU, Blundell R, Hinchliffe A, Atkinson N, and Harwood WA. An efficient and reproducible Agrobacterium-mediated transformation method for hexaploid wheat (Triticum aestivum L.). Plant Methods. 2019:15(1):121. https://doi.org/10.1186/s13007-019-0503-z

He G, Zhang Y, Liu P, Jing Y, Zhang L, Zhu Y, Kong X, Zhao H, Zhou Y, and Sun J. The transcription factor *TaLAX1* interacts with *Q* to antagonistically regulate grain threshability and spike morphogenesis in bread wheat. New Phytol. 2021:230(3):988–1002. https://doi.org/10.1111/nph.17235

Hebelstrup KH, Azariadis A, Cordes A, Henriksen PS, and Brinch-Pedersen H. Prehistoric Plant Exploitation and Domestication: An Inspiration for the Science of De Novo Domestication in Present Times. Plants. 2023:12(12):2310. https://doi.org/10.3390/plants12122310

Henry AG, Brooks AS, and Piperno DR. Plant foods and the dietary ecology of Neanderthals and early modern humans. J Hum Evol. 2014:69:44–54. https://doi.org/10.1016/j.jhevol.2013.12.014

Hie B, Bryson B, and Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nat Biotechnol. 2019:37(6):685–691. https://doi.org/10.1038/s41587-019-0113-3

Hie BL, Kim S, Rando TA, Bryson B, and Berger B. Scanorama: integrating large and diverse single-cell transcriptomic datasets. Nat Protoc. 2024:19(8):2283–2297. https://doi.org/10.1038/s41596-024-00991-3

Honma T and Goto K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature. 2001:409(6819):525–529. https://doi.org/10.1038/35054083

Houston K, Druka A, Bonar N, Macaulay M, Lundqvist U, Franckowiak J, Morgante M, Stein N, and Waugh R. Analysis of the barley bract suppression gene *Trd1*. Theor Appl Genet. 2012:125(1):33–45. https://doi.org/10.1007/s00122-012-1814-x

Howe KL, Contreras-Moreira B, De Silva N, Maslen G, Akanni W, Allen J, Alvarez-Jarreta J, Barba M, Bolser DM, Cambell L, et al. Ensembl Genomes 2020-enabling non-vertebrate genomic research. Nucleic Acids Res. 2020:48(D1):D689–D695. https://doi.org/10.1093/nar/gkz890

Howells RM, Craze M, Bowden S, and Wallington EJ. Efficient generation of stable, heritable gene edits in wheat using CRISPR/Cas9. BMC Plant Biol. 2018:18(1):215. https://doi.org/10.1186/s12870-018-1433-z

Hunter JD. Matplotlib: A 2D Graphics Environment. Comput Sci Eng. 2007:9(3):90–95. https://doi.org/10.1109/MCSE.2007.55

Ikeda K, Sunohara H, and Nagato Y. Developmental Course of Inflorescence and Spikelet in Rice. Breed Sci. 2004:54(2):147–156. https://doi.org/10.1270/jsbbs.54.147

Ikeda T, Tanaka W, Toriba T, Suzuki C, Maeno A, Tsuda K, Shiroishi T, Kurata T, Sakamoto T, Murai M, et al. BELL1-like homeobox genes regulate inflorescence architecture and meristem maintenance in rice. Plant J Cell Mol Biol. 2019:98(3):465–478. https://doi.org/10.1111/tpj.14230

Imbert A, Ouyang W, Safieddine A, Coleno E, Zimmer C, Bertrand E, Walter T, and Mueller F. FISH-quant v2: a scalable and modular tool for smFISH image analysis. RNA. 2022:28(6):786–795. https://doi.org/10.1261/rna.079073.121

Irish VF. The flowering of Arabidopsis flower development. Plant J Cell Mol Biol. 2010:61(6):1014–1028. https://doi.org/10.1111/j.1365-313X.2009.04065.x

IWGSC. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science. 2014:345(6194). https://doi.org/10.1126/science.1251788

Jack T, Brockman LL, and Meyerowitz EM. The homeotic gene *APETALA3* of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell. 1992:68(4):683–697. https://doi.org/10.1016/0092-8674(92)90144-2

Jackson D, Veit B, and Hake S. Expression of maize *KNOTTED1*related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. Development. 1994:120(2):405–413. https://doi.org/10.1242/dev.120.2.405

Jackson KC and Pachter L. A standard for sharing spatial transcriptomics data. Cell Genomics. 2023:3(8):100374. https://doi.org/10.1016/j.xgen.2023.100374

Jofuku KD, den Boer BG, Van Montagu M, and Okamuro JK. Control of Arabidopsis flower and seed development by the homeotic gene *APETALA2*. Plant Cell. 1994:6(9):1211–1225. https://doi.org/10.1105/tpc.6.9.1211

John HA, Birnstiel ML, and Jones KW. RNA-DNA hybrids at the cytological level. Nature. 1969:223(5206):582–587. https://doi.org/10.1038/223582a0

Jordahl K, Bossche JV den, Fleischmann M, Wasserman J, McBride J, Gerard J, Tratner J, Perry M, Badaracco AG, Farmer C, et al. geopandas/geopandas: v0.8.1. 2020. https://doi.org/10.5281/zenodo.3946761

Judziewicz and Soderstrom. Morphological, anatomical, and taxonomic studies in *Anomochloa* and *Streptochaeta* (*Poaceae, Bambusoideae*) (Smithsonian Institution Press: Washington, D.C).

Ke R, Mignardi M, Pacureanu A, Svedlund J, Botling J, Wählby C, and Nilsson M. In situ sequencing for RNA analysis in preserved tissue and cells. Nat Methods. 2013:10(9):857–860. https://doi.org/10.1038/nmeth.2563

Ke Y, Pujol V, Staut J, Pollaris L, Seurinck R, Eekhout T, Grones C, Saura-Sanchez M, Van Bel M, Vuylsteke M, et al. A single-cell and spatial wheat root atlas with cross-species

annotations delineates conserved tissue-specific marker genes and regulators. Cell Rep. 2025:44(2):115240. https://doi.org/10.1016/j.celrep.2025.115240

Kellogg E. A model of inflorescence development. . In. Monocots: Systematics and Evolution, KL Wilson and DA Morrison, eds. (CSIRO: Melbourne), pp. 84–88.

Kellogg EA. Floral displays: genetic control of grass inflorescences. Curr Opin Plant Biol. 2007:10(1):26–31. https://doi.org/10.1016/j.pbi.2006.11.009

Kellogg EA. Genetic control of branching patterns in grass inflorescences. Plant Cell. 2022:34(7):2518–2533. https://doi.org/10.1093/plcell/koac080

Kellogg EA, Camara PEAS, Rudall PJ, Ladd P, Malcomber ST, Whipple CJ, and Doust AN. Early inflorescence development in the grasses (*Poaceae*). Front Plant Sci. 2013:4:250. https://doi.org/10.3389/fpls.2013.00250

Kerstetter RA, Laudencia-Chingcuanco D, Smith LG, and Hake S. Loss-of-function mutations in the maize homeobox gene, *knotted1*, are defective in shoot meristem maintenance. Development. 1997:124(16):3045–3054. https://doi.org/10.1242/dev.124.16.3045

Khush GS. Green revolution: the way forward. Nat Rev Genet. 2001:2(10):815–822. https://doi.org/10.1038/35093585

Kirby EJM and Appleyard M. Cereal development guide. 2nd Edition. (Arable Unit, National Agricultural Centre: Kenilworth).

Kislev ME. *Triticum parvicoccum sp. Nov.*, the oldest naked wheat. Isr J Bot. 1979:28(2):95–107. https://doi.org/10.1080/0021213X.1979.10676861

Kobayashi K, Maekawa M, Miyao A, Hirochika H, and Kyozuka J. *PANICLE PHYTOMER2 (PAP2)*, encoding a *SEPALLATA* subfamily MADS-box protein, positively controls spikelet meristem identity in rice. Plant Cell Physiol. 2010:51(1):47–57. https://doi.org/10.1093/pcp/pcp166

Kobayashi K, Yasuno N, Sato Y, Yoda M, Yamazaki R, Kimizu M, Yoshida H, Nagamura Y, and Kyozuka J. Inflorescence meristem identity in rice is specified by overlapping functions of three *AP1/FU*L-like MADS box genes and *PAP2*, a *SEPALLATA* MADS box gene. Plant Cell. 2012:24(5):1848–1859. https://doi.org/10.1105/tpc.112.097105

Kolde R. pheatmap: Pretty Heatmaps. 2019. https://doi.org/10.32614/CRAN.package.pheatmap

Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H, Shimamoto K, and Kyozuka J. *LAX* and *SPA*: Major regulators of shoot branching in rice. Proc Natl Acad Sci. 2003:100(20):11765–11770. https://doi.org/10.1073/pnas.1932414100

Komatsu M, Maekawa M, Shimamoto K, and Kyozuka J. The *LAX1* and *FRIZZY PANICLE 2* Genes Determine the Inflorescence Architecture of Rice by Controlling Rachis-Branch and Spikelet Development. Dev Biol. 2001:231(2):364–373. https://doi.org/10.1006/dbio.2000.9988

Kong X, Wang F, Geng S, Guan J, Tao S, Jia M, Sun G, Wang Z, Wang K, Ye X, et al. The wheat *AGL6*-like MADS-box gene is a master regulator for floral organ identity and a target

for spikelet meristem development manipulation. Plant Biotechnol J. 2022:20(1):75–88. https://doi.org/10.1111/pbi.13696

Koppolu R, Chen S, and Schnurbusch T. Evolution of inflorescence branch modifications in cereal crops. Curr Opin Plant Biol. 2022:65:102168. https://doi.org/10.1016/j.pbi.2021.102168

Koppolu R and Schnurbusch T. Developmental pathways for shaping spike inflorescence architecture in barley and wheat. J Integr Plant Biol. 2019:61(3):278–295. https://doi.org/10.1111/jipb.12771

Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, Simmonds J, Ramirez-Gonzalez RH, Wang X, Borrill P, et al. Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci. 2017:114(6):E913–E921. https://doi.org/10.1073/pnas.1619268114

Krausmann F, Erb K-H, Gingrich S, Haberl H, Bondeau A, Gaube V, Lauk C, Plutzar C, and Searchinger TD. Global human appropriation of net primary production doubled in the 20th century. Proc Natl Acad Sci. 2013:110(25):10324–10329. https://doi.org/10.1073/pnas.1211349110

Krizek B. AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. Plant Physiol. 2009:150(4):1916–1929. https://doi.org/10.1104/pp.109.141119

Kyozuka J. Chapter Seven - Grass Inflorescence: Basic Structure and Diversity. . In. Advances in Botanical Research, F Fornara, ed, The Molecular Genetics of Floral Transition and Flower Development. (Academic Press), pp. 191–219. https://doi.org/10.1016/B978-0-12-417162-6.00007-9

Latting J. Differentiation in the grass inflorescence. . In. The Biology and Utilization of Grasses. (Academic Press: New York, NY), pp. 365–399.

Laureyns R, Joossens J, Herwegh D, Pevernagie J, Pavie B, Demuynck K, Debray K, Coussens G, Pauwels L, Van Hautegem T, et al. An in situ sequencing approach maps *PLASTOCHRON1* at the boundary between indeterminate and determinate cells. Plant Physiol. 2021:188(2):782–794. https://doi.org/10.1093/plphys/kiab533

Lee S, Choi SC, and An G. Rice SVP-group MADS-box proteins, *OsMADS22* and *OsMADS55*, are negative regulators of brassinosteroid responses. Plant J Cell Mol Biol. 2008:54(1):93–105. https://doi.org/10.1111/j.1365-313X.2008.03406.x

Lee TA, Illouz-Eliaz N, Nobori T, Xu J, Jow B, Nery JR, and Ecker JR. A single-cell, spatial transcriptomic atlas of the *Arabidopsis* life cycle. Nat Plants. 2025:11(9):1960–1975. https://doi.org/10.1038/s41477-025-02072-z

Levy AA and Feldman M. Evolution and origin of bread wheat. Plant Cell. 2022:34(7):2549–2567. https://doi.org/10.1093/plcell/koac130

Li C, Lin H, Chen A, Lau M, Jernstedt J, and Dubcovsky J. Wheat *VRN1, FUL2* and *FUL3* play critical and redundant roles in spikelet development and spike determinacy. Development. 2019:146(14):dev175398. https://doi.org/10.1242/dev.175398

Li G, Kuijer HNJ, Yang X, Liu H, Shen C, Shi J, Betts N, Tucker MR, Liang W, Waugh R, et al. *MADS1* maintains barley spike morphology at high ambient temperatures. Nat Plants. 2021a:7(8):1093–1107. https://doi.org/10.1038/s41477-021-00957-3

Li H, Liang W, Jia R, Yin C, Zong J, Kong H, and Zhang D. The *AGL6*-like gene *OsMADS6* regulates floral organ and meristem identities in rice. Cell Res. 2010:20(3):299–313. https://doi.org/10.1038/cr.2009.143

Li K, Debernardi JM, Li C, Lin H, Zhang C, Jernstedt J, Korff M von, Zhong J, and Dubcovsky J. Interactions between SQUAMOSA and SHORT VEGETATIVE PHASE MADS-box proteins regulate meristem transitions during wheat spike development. Plant Cell. 2021b:33(12):3621–3644. https://doi.org/10.1093/plcell/koab243

Li L-F, Zhang Z-B, Wang Z-H, Li N, Sha Y, Wang X-F, Ding N, Li Y, Zhao J, Wu Y, et al. Genome sequences of five Sitopsis species of *Aegilops* and the origin of polyploid wheat B subgenome. Mol Plant. 2022:15(3):488–503. https://doi.org/10.1016/j.molp.2021.12.019

Li S, Zhang C, Li J, Yan L, Wang N, and Xia L. Present and future prospects for wheat improvement through genome editing and advanced technologies. Plant Commun. 2021c:2(4):100211. https://doi.org/10.1016/j.xplc.2021.100211

Li T, Horsfall D, Basurto-Lozada D, Roberts K, Prete M, Lawrence JEG, He P, Tuck E, Moore J, Yoldas AK, et al. WebAtlas pipeline for integrated single-cell and spatial transcriptomic data. Nat Methods. 2024. https://doi.org/10.1038/s41592-024-02371-x

Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X, Teng S, Hiroshi F, et al. Control of tillering in rice. Nature. 2003:422(6932):618–621. https://doi.org/10.1038/nature01518

Li X, Xie C, Cheng L, Tong H, Bock R, Qian Q, and Zhou W. The next Green Revolution: integrating crop architectype and physiotype. Trends Biotechnol. 2025. https://doi.org/10.1016/j.tibtech.2025.04.002

Liu J, Chen Z, Wang Z, Zhang Z, Xie X, Wang Z, Chai L, Song L, Cheng X, Feng M, et al. Ectopic expression of *VRT-A2* underlies the origin of Triticum polonicum and Triticum petropavlovskyi with long outer glumes and grains. Mol Plant. 2021:14(9):1472–1488. https://doi.org/10.1016/j.molp.2021.05.021

Liu J, Tran V, Vemuri VNP, Byrne A, Borja M, Kim YJ, Agarwal S, Wang R, Awayan K, Murti A, et al. Concordance of MERFISH spatial transcriptomics with bulk and single-cell RNA sequencing. Life Sci Alliance. 2022:6(1):e202201701. https://doi.org/10.26508/lsa.202201701

Liu Q, Teng S, Deng C, Wu S, Li H, Wang Y, Wu J, Cui X, Zhang Z, Quick WP, et al. *SHORT ROOT* and *INDETERMINATE DOMAIN* family members govern *PIN-FORMED* expression to regulate minor vein differentiation in rice. Plant Cell. 2023a:35(8):2848–2870. https://doi.org/10.1093/plcell/koad125

Liu X, Lin X, Kang J, Long KA, Yue J, Chen C, Wang D, Lister A, Macaulay IC, Liu X, et al. Decoding cellular transcriptional regulatory networks governing wheat inflorescence development. 2025:2025.01.18.633750. https://doi.org/10.1101/2025.01.18.633750

Liu Z, Kong X, Long Y, Liu S, Zhang H, Jia J, Cui W, Zhang Z, Song X, Qiu L, et al. Integrated single-nucleus and spatial transcriptomics captures transitional states in soybean nodule maturation. Nat Plants. 2023b:9(4):515–524. https://doi.org/10.1038/s41477-023-01387-z

Liu ZH, Wang HY, Wang XE, Zhang GP, Chen PD, and Liu DJ. Genotypic and spike positional difference in grain phytase activity, phytate, inorganic phosphorus, iron, and zinc contents in wheat (*Triticum aestivum* L.). J Cereal Sci. 2006:44(2):212–219. https://doi.org/10.1016/j.jcs.2006.06.001

Long JA, Moan EI, Medford JI, and Barton MK. A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of Arabidopsis. Nature. 1996:379(6560):66–69. https://doi.org/10.1038/379066a0

Long KA, Lister A, and Uauy C. Embedding and Sectioning of Wheat Inflorescence Tissue. 2025. https://dx.doi.org/10.17504/protocols.io.rm7vzqwb4vx1/v1

Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014:15(12):550. https://doi.org/10.1186/s13059-014-0550-8

Ma H and dePamphilis C. The ABCs of Floral Evolution. Cell. 2000:101(1):5–8. https://doi.org/10.1016/S0092-8674(00)80618-2

Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner SG, et al. Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet. 2019:51(5):885–895. https://doi.org/10.1038/s41588-019-0381-3

Maher LA, Richter T, and Stock JT. The Pre-Natufian Epipaleolithic: Long-term Behavioral Trends in the Levant. Evol Anthropol Issues News Rev. 2012:21(2):69–81. https://doi.org/10.1002/evan.21307

Malcomber ST and Kellogg EA. *SEPALLATA* gene diversification: brave new whorls. Trends Plant Sci. 2005:10(9):427–435. https://doi.org/10.1016/j.tplants.2005.07.008

Mandel MA, Gustafson-Brown C, Savidge B, and Yanofsky MF. Molecular characterization of the Arabidopsis floral homeotic gene *APETALA1*. Nature. 1992:360(6401):273–277. https://doi.org/10.1038/360273a0

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, International Wheat Genome Sequencing Consortium, Jakobsen KS, Wulff BBH, Steuernagel B, Mayer KFX, et al. Ancient hybridizations among the ancestral genomes of bread wheat. Science. 2014:345(6194):1250092. https://doi.org/10.1126/science.1250092

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011:17(1):10–12. https://doi.org/10.14806/ej.17.1.200

McKim SM. Moving on up – controlling internode growth. New Phytol. 2020:226(3):672–678. https://doi.org/10.1111/nph.16439

McSteen P, Laudencia-Chingcuanco D, and Colasanti J. A floret by any other name: control of meristem identity in maize. Trends Plant Sci. 2000:5(2):61–66. https://doi.org/10.1016/S1360-1385(99)01541-1

MERSCOPE Vizualizer. 2023.

Moffitt JR, Hao J, Bambah-Mukku D, Lu T, Dulac C, and Zhuang X. High-performance multiplexed fluorescence *in situ* hybridization in culture and tissue with matrix imprinting and clearing. Proc Natl Acad Sci. 2016:113(50):14456–14461. https://doi.org/10.1073/pnas.1617699113

Mohanty JN, Sahoo S, and Mishra P. A genetic approach to comprehend the complex and dynamic event of floral development: a review. Genomics Inform. 2022:20(4):e40. https://doi.org/10.5808/gi.21075

Moore KJ and Moser LE. Quantifying Developmental Morphology of Perennial Grasses. Crop Sci. 1995:35(1):37–43. https://doi.org/10.2135/cropsci1995.0011183X003500010007x

Moses L and Pachter L. Museum of spatial transcriptomics. Nat Methods. 2022:19(5):534–546. https://doi.org/10.1038/s41592-022-01409-2

Murai K. Homeotic Genes and the ABCDE Model for Floral Organ Formation in Wheat. Plants. 2013:2(3):379–395. https://doi.org/10.3390/plants2030379

Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H, and Nagato Y. *SUPERWOMAN1* and *DROOPING LEAF* genes control floral organ identity in rice. Dev Camb Engl. 2003:130(4):705–718. https://doi.org/10.1242/dev.00294

Nesbitt, M. Wheat evolution: integrating archaeological and biological evidence. . In. Wheat Taxonomy: The Legacy of John Percival. (Linnean Society: London), pp. 37–59.

Neumann M, Xu X, Smaczniak C, Schumacher J, Yan W, Blüthgen N, Greb T, Jönsson H, Traas J, Kaufmann K, et al. A 3D gene expression atlas of the floral meristem based on spatial reconstruction of single nucleus RNA sequencing data. Nat Commun. 2022:13(1):2838. https://doi.org/10.1038/s41467-022-30177-y

Nobori T. Exploring the untapped potential of single-cell and spatial omics in plant biology. New Phytol. 2025. https://doi.org/10.1111/nph.70220

Nobori T, Monell A, Lee TA, Sakata Y, Shirahama S, Zhou J, Nery JR, Mine A, and Ecker JR. A rare PRIMER cell state in plant immunity. Nature. 2025:638(8049):197–205. https://doi.org/10.1038/s41586-024-08383-z

Nobori T, Oliva M, Lister R, and Ecker JR. Multiplexed single-cell 3D spatial gene expression analysis in plant tissue using PHYTOMap. Nat Plants. 2023:9(7):1026–1033. https://doi.org/10.1038/s41477-023-01439-4

Nole-Wilson S, Tranby TL, and Krizek BA. *AINTEGUMENTA-like (AIL)* genes are expressed in young tissues and may specify meristematic or division-competent states. Plant Mol Biol. 2005:57(5):613–628. https://doi.org/10.1007/s11103-005-0955-6

Ntranos V, Yi L, Melsted P, and Pachter L. A discriminative learning approach to differential expression analysis for single-cell RNA-seq. Nat Methods. 2019:16(2):163–166. https://doi.org/10.1038/s41592-018-0303-9

Ohmori Y, Toriba T, Nakamura H, Ichikawa H, and Hirano H-Y. Temporal and spatial regulation of *DROOPING LEAF* gene expression that promotes midrib formation in

rice. Plant J Cell Mol Biol. 2011:65(1):77–86. https://doi.org/10.1111/j.1365-313X.2010.04404.x

Pachitariu M and Stringer C. Cellpose 2.0: how to train your own model. Nat Methods. 2022:19(12):1634–1641. https://doi.org/10.1038/s41592-022-01663-4

Palla G, Spitzer H, Klein M, Fischer D, Schaar AC, Kuemmerle LB, Rybakov S, Ibarra IL, Holmberg O, Virshup I, et al. Squidpy: a scalable framework for spatial omics analysis. Nat Methods. 2022:19(2):171–178. https://doi.org/10.1038/s41592-021-01358-2

Pankin A and von Korff M. Co-evolution of methods and thoughts in cereal domestication studies: a tale of barley (*Hordeum vulgare*). Curr Opin Plant Biol. 2017:36:15–21. https://doi.org/10.1016/j.pbi.2016.12.001

Patil V, McDermott HI, McAllister T, Cummins M, Silva JC, Mollison E, Meikle R, Morris J, Hedley PE, Waugh R, et al. *APETALA2* control of barley internode elongation. Development. 2019:146(11):dev170373. https://doi.org/10.1242/dev.170373

Patterson EL, Richardson A, and Bartlett M. Pushing the boundaries of organ identity: Homology of the grass lemma. Am J Bot. 2023:110(4):e16161. https://doi.org/10.1002/ajb2.16161

Pautler M, Tanaka W, Hirano H-Y, and Jackson D. Grass Meristems I: Shoot Apical Meristem Maintenance, Axillary Meristem Determinacy and the Floral Transition. Plant Cell Physiol. 2013:54(3):302–312. https://doi.org/10.1093/pcp/pct025

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011:12(null):2825–2830.

Pelaz S, Ditta GS, Baumann E, Wisman E, and Yanofsky MF. B and C floral organ identity functions require *SEPALLATA* MADS-box genes. Nature. 2000:405(6783):200–203. https://doi.org/10.1038/35012103

Perico C, Zaidem M, Sedelnikova O, Bhattacharya S, Korfhage C, and Langdale JA. Multiplexed in situ hybridization reveals distinct lineage identities for major and minor vein initiation during maize leaf development. Proc Natl Acad Sci. 2024:121(28):e2402514121. https://doi.org/10.1073/pnas.2402514121

Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, and Yanofsky MF. Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature. 2003:424(6944):85–88. https://doi.org/10.1038/nature01741

Piperno DR, Ranere AJ, Holst I, Iriarte J, and Dickau R. Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. Proc Natl Acad Sci. 2009:106(13):5019–5024. https://doi.org/10.1073/pnas.0812525106

Postma-Haarsma AD, Rueb S, Scarpella E, den Besten W, Hoge JHC, and Meijer AH. Developmental regulation and downstream effects of the knox class homeobox genes *Oskn2* and *Oskn3* from rice. Plant Mol Biol. 2002:48(4):423–441. https://doi.org/10.1023/a:1014047917226

Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawroński P, Bini F, Govind G, Rutten T, Sakuma S, Tagiri A, et al. The Genetic Basis of Composite Spike Form in Barley and 'Miracle-Wheat.' Genetics. 2015:201(1):155–165. https://doi.org/10.1534/genetics.115.176628

Poursarebani N, Trautewig C, Melzer M, Nussbaumer T, Lundqvist U, Rutten T, Schmutzer T, Brandt R, Himmelbach A, Altschmied L, et al. *COMPOSITUM 1* contributes to the architectural simplification of barley inflorescence via meristem identity signals. Nat Commun. 2020:11(1):5138. https://doi.org/10.1038/s41467-020-18890-y

Preece C, Livarda A, Christin P-A, Wallace M, Martin G, Charles M, Jones G, Rees M, and Osborne CP. How did the domestication of Fertile Crescent grain crops increase their yields? Funct Ecol. 2017:31(2):387–397. https://doi.org/10.1111/1365-2435.12760

Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, and Coen E. Evolution and development of inflorescence architectures. Science. 2007:316(5830):1452–1456. https://doi.org/10.1126/science.1140429

Qu Y, Tan C, Yang L, Pasquariello M, Alabdullah AK, Sun S, Iqbal M, Salamon J, and Boden S. Spatial transcriptomics identifies distinct domains regulating yield-related traits of the wheat ear. 2025:2025.08.12.670006. https://doi.org/10.1101/2025.08.12.670006

R Core Team. R: A Language and Environment for Statistical Computing. 2018.

Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, and Tyagi S. Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods. 2008:5(10):877–879. https://doi.org/10.1038/nmeth.1253

Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, van Ex F, Pasha A, et al. The transcriptional landscape of polyploid wheat. Science. 2018:361(6403):eaar6089. https://doi.org/10.1126/science.aar6089

Ranere AJ, Piperno DR, Holst I, Dickau R, and Iriarte J. The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. Proc Natl Acad Sci. 2009:106(13):5014–5018. https://doi.org/10.1073/pnas.0812590106

Reynolds M, Bonnett D, Chapman SC, Furbank RT, Manès Y, Mather DE, and Parry MAJ. Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. J Exp Bot. 2011:62(2):439–452. https://doi.org/10.1093/jxb/erq311

Reynolds MP, Slafer GA, Foulkes JM, Griffiths S, Murchie EH, Carmo-Silva E, Asseng S, Chapman SC, Sawkins M, Gwyn J, et al. A wiring diagram to integrate physiological traits of wheat yield potential. Nat Food. 2022:3(5):318–324. https://doi.org/10.1038/s43016-022-00512-z

Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, Welch J, Chen LM, Chen F, and Macosko EZ. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. Science. 2019:363(6434):1463–1467. https://doi.org/10.1126/science.aaw1219

Rosyara U, Kishii M, Payne T, Sansaloni CP, Singh RP, Braun H-J, and Dreisigacker S. Genetic Contribution of Synthetic Hexaploid Wheat to CIMMYT's Spring Bread Wheat Breeding Germplasm. Sci Rep. 2019:9:12355. https://doi.org/10.1038/s41598-019-47936-5

Running SW. A Measurable Planetary Boundary for the Biosphere. Science. 2012:337(6101):1458–1459. https://doi.org/10.1126/science.1227620

Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, and Martin W. Genetics and geography of wild cereal domestication in the near east. Nat Rev Genet. 2002:3(6):429–441. https://doi.org/10.1038/nrg817

Sang X, Li Y, Luo Z, Ren D, Fang L, Wang N, Zhao F, Ling Y, Yang Z, Liu Y, et al. CHIMERIC FLORAL ORGANS1, Encoding a Monocot-Specific MADS Box Protein, Regulates Floral Organ Identity in Rice. Plant Physiol. 2012:160(2):788–807. https://doi.org/10.1104/pp.112.200980

Sax K. Chromosome Relationships in Wheat. Science. 1921:54(1400):413–415. https://doi.org/10.1126/science.54.1400.413

Sax K. STERILITY IN WHEAT HYBRIDS. II. CHROMOSOME BEHAVIOR IN PARTIALLY STERILE HYBRIDS. Genetics. 1922:7(6):513–552. https://doi.org/10.1093/genetics/7.6.513

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012:9(7):676–682. https://doi.org/10.1038/nmeth.2019

Schneider CA, Rasband WS, and Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012:9(7):671–675. https://doi.org/10.1038/nmeth.2089

Schultze-Motel J. *Triticum parvicoccum* Kislev in Transcaucasia. Genet Resour Crop Evol. 2019:66(7):1363–1366. https://doi.org/10.1007/s10722-019-00795-5

Scofield S, Dewitte W, and Murray JA. *STM* sustains stem cell function in the Arabidopsis shoot apical meristem and controls *KNOX* gene expression independently of the transcriptional repressor *AS1*. Plant Signal Behav. 2014:9:e28934. https://doi.org/10.4161/psb.28934

Scofield S, Murison A, Jones A, Fozard J, Aida M, Band LR, Bennett M, and Murray JAH. Coordination of meristem and boundary functions by transcription factors in the *SHOOT MERISTEMLESS* regulatory network. Dev Camb Engl. 2018:145(9):dev157081. https://doi.org/10.1242/dev.157081

Sentoku N, Kato H, Kitano H, and Imai R. *OsMADS22*, an *STMADS11*-like MADS-box gene of rice, is expressed in non-vegetative tissues and its ectopic expression induces spikelet meristem indeterminacy. Mol Genet Genomics MGG. 2005:273(1):1–9. https://doi.org/10.1007/s00438-004-1093-6

Sentoku N, Sato Y, Kurata N, Ito Y, Kitano H, and Matsuoka M. Regional expression of the rice *KN1*-type homeobox gene family during embryo, shoot, and flower development. Plant Cell. 1999:11(9):1651–1664. https://doi.org/10.1105/tpc.11.9.1651

Serrano K, Bezrutczyk M, Goudeau D, Dao T, O'Malley R, Malmstrom RR, Visel A, Scheller HV, and Cole B. Spatial co-transcriptomics reveals discrete stages of the arbuscular mycorrhizal symbiosis. Nat Plants. 2024:10(4):673–688. https://doi.org/10.1038/s41477-024-01666-3

Shahan R, Hsu C-W, Nolan TM, Cole BJ, Taylor IW, Greenstreet L, Zhang S, Afanassiev A, Vlot AHC, Schiebinger G, et al. A single-cell *Arabidopsis* root atlas reveals developmental trajectories in wild-type and cell identity mutants. Dev Cell. 2022:57(4):543-560.e9. https://doi.org/10.1016/j.devcel.2022.01.008

Shen C, Yang X, Wang D, Li G, and Tucker MR. Functional retrogression of *LOFSEPs* in specifying floral organs in barley. aBIOTECH. 2025:6(1):1–11. https://doi.org/10.1007/s42994-024-00182-4

Shi J, Pan Y, Liu X, Cao W, Mu Y, and Zhu Q. Spatial Omics Sequencing Based on Microfluidic Array Chips. Biosensors. 2023a:13(7):712. https://doi.org/10.3390/bios13070712

Shi X, Yang Y, Ma X, Zhou Y, Guo Z, Wang C, and Liu J. Probabilistic cell/domain-type assignment of spatial transcriptomics data with SpatialAnno. Nucleic Acids Res. 2023b:51(22):e115. https://doi.org/10.1093/nar/gkad1023

Simmonds J, Faci I, Buffagni V, and Uauy C. How to grow wheat in controlled and semi-controlled environment. 2024. https://doi.org/dx.doi.org/10.17504/protocols.io.6qpvr37jbvmk/v1

Slewinski TL, Anderson AA, Price S, Withee JR, Gallagher K, and Turgeon R. *Short-root1* plays a role in the development of vascular tissue and kranz anatomy in maize leaves. Mol Plant. 2014:7(8):1388–1392. https://doi.org/10.1093/mp/ssu036

Smedley MA, Hayta S, Clarke M, and Harwood WA. CRISPR-Cas9 Based Genome Editing in Wheat. Curr Protoc. 2021:1(3):e65. https://doi.org/10.1002/cpz1.65

Song X, Guo P, Xia K, Wang M, Liu Y, Chen L, Zhang J, Xu M, Liu N, Yue Z, et al. Spatial transcriptomics reveals light-induced chlorenchyma cells involved in promoting shoot regeneration in tomato callus. Proc Natl Acad Sci U S A. 2023:120(38):e2310163120. https://doi.org/10.1073/pnas.2310163120

Souer E, van Houwelingen A, Kloos D, Mol J, and Koes R. *The No Apical Meristem* Gene of Petunia Is Required for Pattern Formation in Embryos and Flowers and Is Expressed at Meristem and Primordia Boundaries. Cell. 1996:85(2):159–170. https://doi.org/10.1016/S0092-8674(00)81093-4

Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, Giacomello S, Asp M, Westholm JO, Huss M, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science. 2016:353(6294):78–82. https://doi.org/10.1126/science.aaf2403

Sugiyama S-H, Yasui Y, Ohmori S, Tanaka W, and Hirano H-Y. Rice Flower Development Revisited: Regulation of Carpel Specification and Flower Meristem Determinacy. Plant Cell Physiol. 2019:60(6):1284–1295. https://doi.org/10.1093/pcp/pcz020

Suzaki T, Toriba T, Fujimoto M, Tsutsumi N, Kitano H, and Hirano H-Y. Conservation and Diversification of Meristem Maintenance Mechanism in Oryza sativa : Function of

the FLORAL ORGAN NUMBER2 Gene. Plant Cell Physiol. 2006:47(12):1591–1602. https://doi.org/10.1093/pcp/pcl025

Tamagno S, Carrera CS, Marchese SI, Savin R, and Slafer GA. Sterility of basal spikelets in wheat: predetermined fate or a matter of resources? J Exp Bot. 2024:75(22):7160–7173. https://doi.org/10.1093/jxb/erae373

Tanaka W, Pautler M, Jackson D, and Hirano H-Y. Grass Meristems II: Inflorescence Architecture, Flower Development and Meristem Fate. Plant Cell Physiol. 2013:54(3):313–324. https://doi.org/10.1093/pcp/pct016

Tanaka W, Toriba T, and Hirano H-Y. Three *TOB1*-related *YABBY* genes are required to maintain proper function of the spikelet and branch meristems in rice. New Phytol. 2017:215(2):825–839. https://doi.org/10.1111/nph.14617

Tao Y, Zhou X, Sun L, Lin D, Cai H, Chen X, Zhou W, Yang B, Hu Z, Yu J, et al. Highly efficient and robust π-FISH rainbow for multiplexed in situ detection of diverse biomolecules. Nat Commun. 2023:14(1):443. https://doi.org/10.1038/s41467-023-36137-4

The International Wheat Genome Sequencing Consortium (IWGSC), Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, Pozniak CJ, Choulet F, Distelfeld A, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science. 2018:361(6403):eaar7191. https://doi.org/10.1126/science.aar7191

Theißen G, Melzer R, and Rümpler F. MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. Dev Camb Engl. 2016:143(18):3259–3271. https://doi.org/10.1242/dev.134080

Thiel J, Koppolu R, Trautewig C, Hertig C, Kale SM, Erbe S, Mascher M, Himmelbach A, Rutten T, Esteban E, et al. Transcriptional landscapes of floral meristems in barley. Sci Adv. 2021:7(18):eabf0832. https://doi.org/10.1126/sciadv.abf0832

Thompson BE and Hake S. Translational Biology: From *Arabidopsis* Flowers to Grass Inflorescence Architecture. Plant Physiol. 2009:149(1):38–45. https://doi.org/10.1104/pp.108.129619

Torre E, Dueck H, Shaffer S, Gospocic J, Gupte R, Bonasio R, Kim J, Murray J, and Raj A. Rare Cell Detection by Single-Cell RNA Sequencing as Guided by Single-Molecule RNA FISH. Cell Syst. 2018:6(2):171-179.e5. https://doi.org/10.1016/j.cels.2018.01.014

Trevaskis B, Tadege M, Hemming MN, Peacock WJ, Dennis ES, and Sheldon C. *Short Vegetative Phase*-Like MADS-Box Genes Inhibit Floral Meristem Identity in Barley. Plant Physiol. 2007:143(1):225–235. https://doi.org/10.1104/pp.106.090860

Uauy C, Nelissen H, Chan RL, Napier JA, Seung D, Liu L, and McKim SM. Challenges of translating *Arabidopsis* insights into crops. Plant Cell. 2025:37(5):koaf059. https://doi.org/10.1093/plcell/koaf059

Vegetti AC and Weberling F. The structure of the paracladial zone in Poaceae. TAXON. 1996:45(3):453–460. https://doi.org/10.2307/1224137

Virshup I, Rybakov S, Theis FJ, Angerer P, and Wolf FA. anndata: Annotated data. 2021:2021.12.16.473007. https://doi.org/10.1101/2021.12.16.473007

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020:17(3):261–272. https://doi.org/10.1038/s41592-019-0686-2

Waddington SR, Cartwright PM, and Wall PC. A Quantitative Scale of Spike Initial and Pistil Development in Barley and Wheat. Ann Bot. 1983:51(1):119–130. https://doi.org/10.1093/oxfordjournals.aob.a086434

Walkowiak S, Gao L, Monat C, Haberer G, Kassa MT, Brinton J, Ramirez-Gonzalez RH, Kolodziej MC, Delorean E, Thambugala D, et al. Multiple wheat genomes reveal global variation in modern breeding. Nature. 2020:588(7837):277–283. https://doi.org/10.1038/s41586-020-2961-x

Walla A, Wilma van Esse G, Kirschner GK, Guo G, Brünje A, Finkemeier I, Simon R, and von Korff M. An Acyl-CoA N-Acyltransferase Regulates Meristem Phase Change and Plant Architecture in Barley1. Plant Physiol. 2020:183(3):1088–1109. https://doi.org/10.1104/pp.20.00087

Walt S van der, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, and Yu T. scikit-image: image processing in Python. PeerJ. 2014:2:e453. https://doi.org/10.7717/peerj.453

Wang G, Moffitt JR, and Zhuang X. Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. Sci Rep. 2018:8:4847. https://doi.org/10.1038/s41598-018-22297-7

Wang J, Bao J, Zhou B, Li M, Li X, and Jin J. The osa-miR164 target *OsCUC1* functions redundantly with *OsCUC3* in controlling rice meristem/organ boundary specification. New Phytol. 2021:229(3):1566–1581. https://doi.org/10.1111/nph.16939

Wang L, Sun S, Jin J, Fu D, Yang X, Weng X, Xu C, Li X, Xiao J, and Zhang Q. Coordinated regulation of vegetative and reproductive branching in rice. Proc Natl Acad Sci. 2015:112(50):15504–15509. https://doi.org/10.1073/pnas.1521949112

Wang L, Yin H, Qian Q, Yang J, Huang C, Hu X, and Luo D. *NECK LEAF 1*, a *GATA* type transcription factor, modulates organogenesis by regulating the expression of multiple regulatory genes during reproductive development in rice. Cell Res. 2009:19(5):598–611. https://doi.org/10.1038/cr.2009.36

Wang Y, Luo Y, Guo X, Li Y, Yan J, Shao W, Wei W, Wei X, Yang T, Chen J, et al. A spatial transcriptome map of the developing maize ear. Nat Plants. 2024:10(5):815–827. https://doi.org/10.1038/s41477-024-01683-2

Waskom ML. seaborn: statistical data visualization. J Open Source Softw. 2021:6(60):3021. https://doi.org/10.21105/joss.03021

Wayne A Kelly. Describing the grass inflorescence. Journal of Range Management. 1982:35(5):672–675.

Weberling. Morphology of Flowers and Inflorescences (Cambridge University Press: Cambridge).

Weide A. On the Identification of Domesticated Emmer Wheat, *Triticum turgidum subsp. dicoccum (Poaceae)*, in the Aceramic Neolithic of the Fertile Crescent. Archäol Informationen. 2015:38:381–424. https://doi.org/10.11588/ai.2015.1.26205

Weir I, Lu J, Cook H, Causier B, Schwarz-Sommer Z, and Davies B. *CUPULIFORMIS* establishes lateral organ boundaries in Antirrhinum. Dev Camb Engl. 2004:131(4):915–922. https://doi.org/10.1242/dev.00993

Whipple CJ. Grass inflorescence architecture and evolution: the origin of novel signaling centers. New Phytol. 2017:216(2):367–372. https://doi.org/10.1111/nph.14538

Whipple CJ, Hall DH, DeBlasio S, Taguchi-Shiobara F, Schmidt RJ, and Jackson DP. A Conserved Mechanism of Bract Suppression in the Grass Family. Plant Cell. 2010:22(3):565–578. https://doi.org/10.1105/tpc.109.073536

Whipple CJ and Schmidt RJ. Genetics of Grass Flower Development. . In. Advances in Botanical Research, Developmental Genetics of the Flower. (Academic Press), pp. 385–424. https://doi.org/10.1016/S0065-2296(06)44010-6

Wiggin T and Yu C. VizGen Post Processing Tool. 2024.

Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016:3(1):160018. https://doi.org/10.1038/sdata.2016.18

Wolf FA, Angerer P, and Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018:19(1):15. https://doi.org/10.1186/s13059-017-1382-0

Wu D, Liang W, Zhu W, Chen M, Ferrándiz C, Burton RA, Dreni L, and Zhang D. Loss of *LOFSEP* Transcription Factor Function Converts Spikelet to Leaf-Like Structures in Rice. Plant Physiol. 2018:176(2):1646–1664. https://doi.org/10.1104/pp.17.00704

Wu F, Shi X, Lin X, Liu Y, Chong K, Theißen G, and Meng Z. The ABCs of flower development: mutational analysis of *AP1/FUL*-like genes in rice provides evidence for a homeotic (A)-function in grasses. Plant J. 2017:89(2):310–324. https://doi.org/10.1111/tpj.13386

Wyatt R. Inflorescence Architecture: How Flower Number, Arrangement, and Phenology Affect Pollination and Fruit-Set. Am J Bot. 1982:69(4):585–594. https://doi.org/10.1002/j.1537-2197.1982.tb13295.x

Xia K, Sun H-X, Li J, Li J, Zhao Y, Chen L, Qin C, Chen R, Chen Z, Liu G, et al. The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in *Arabidopsis* leaves. Dev Cell. 2022:57(10):1299-1310.e4. https://doi.org/10.1016/j.devcel.2022.04.011

Xiao Y, Guo J, Dong Z, Richardson A, Patterson E, Mangrum S, Bybee S, Bertolini E, Bartlett M, Chuck G, et al. Boundary domain genes were recruited to suppress bract growth and promote branching in maize. Sci Adv. 2022:8(24):eabm6835. https://doi.org/10.1126/sciadv.abm6835

Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu Z, Wang L, Fox N, et al. Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. Dev Cell. 2021:56(4):557-568.e6. https://doi.org/10.1016/j.devcel.2020.12.015

Xu X, Lin H, Zhang J, Burguener G, Paraiso F, Tumelty C, Li C, Liu Y, and Dubcovsky J. Spatial and single-cell expression analyses reveal complex expression domains in early wheat spike development. 2025:2025.02.15.638402. https://doi.org/10.1101/2025.02.15.638402

Xu X, Passalacqua M, Rice B, Demesa-Arevalo E, Kojima M, Takebayashi Y, Harris B, Sakakibara H, Gallavotti A, Gillis J, et al. Large-scale single-cell profiling of stem cells uncovers redundant regulators of shoot development and yield trait variation. 2024:2024.03.04.583414. https://doi.org/10.1101/2024.03.04.583414

Yamaguchi T, Nagasawa N, Kawasaki S, Matsuoka M, Nagato Y, and Hirano H-Y. The YABBY Gene *DROOPING LEAF* Regulates Carpel Specification and Midrib Development in Oryza sativa. Plant Cell. 2004:16(2):500–509. https://doi.org/10.1105/tpc.018044

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, Chen R, Yamazaki R, Tokunaga H, Kitaguchi Y, Sato Y, et al. *TAWAWA1*, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. Proc Natl Acad Sci U S A. 2013:110(2):767–772. https://doi.org/10.1073/pnas.1216151110

Yoshida H. Is the lodicule a petal: Molecular evidence? Plant Sci. 2012:184:121–128. https://doi.org/10.1016/j.plantsci.2011.12.016

You Y, Fu Y, Li L, Zhang Z, Jia S, Lu S, Ren W, Liu Y, Xu Y, Liu X, et al. Systematic comparison of sequencing-based spatial transcriptomic methods. Nat Methods. 2024:21(9):1743–1754. https://doi.org/10.1038/s41592-024-02325-3

Zeder MA. The Origins of Agriculture in the Near East. Curr Anthropol. 2011:52(S4):S221–S235. https://doi.org/10.1086/659307

Zhang L, He C, Lai Y, Wang Y, Kang L, Liu A, Lan C, Su H, Gao Y, Li Z, et al. Asymmetric gene expression and cell-type-specific regulatory networks in the root of bread wheat revealed by single-cell multiomics analysis. Genome Biol. 2023:24(1):65. https://doi.org/10.1186/s13059-023-02908-x

Zhang MJ, Ntranos V, and Tse D. Determining sequencing depth in a single-cell RNA-seq experiment. Nat Commun. 2020:11(1):774. https://doi.org/10.1038/s41467-020-14482-y

Zhang T, Li Y, Ma L, Sang X, Ling Y, Wang Y, Yu P, Zhuang H, Huang J, Wang N, et al. *LATERAL FLORET 1* induced the three-florets spikelet in rice. Proc Natl Acad Sci U S A. 2017:114(37):9984–9989. https://doi.org/10.1073/pnas.1700504114

Zhang X, Fonseca A, Kutashev K, Sicard A, Duncan S, and Rosa S. Quantitative RNA spatial profiling using single-molecule RNA FISH on plant tissue cryosections. 2024a:2024.04.09.588031. https://doi.org/10.1101/2024.04.09.588031

Zhang X, Luo Z, Marand AP, Yan H, Jang H, Bang S, Mendieta JP, Minow MAA, and Schmitz RJ. A spatially resolved multi-omic single-cell atlas of soybean development. Cell. 2025:188(2):550-567.e19. https://doi.org/10.1016/j.cell.2024.10.050

Zhang Y, Shen C, Li G, Shi J, Yuan Y, Ye L, Song Q, Shi J, and Zhang D. *MADS1*-regulated lemma and awn development benefits barley yield. Nat Commun. 2024b:15(1):301. https://doi.org/10.1038/s41467-023-44457-8

Zhao Y, Li Y, He Y, Wu J, Liu Y, Li X, Li Z, Yuan Q, Li J, Zhang X, et al. Stereo-seq V2: Spatial mapping of total RNA on FFPE sections with high resolution. Cell. 2025:0(0). https://doi.org/10.1016/j.cell.2025.08.008

Zheng Y, Crawford GW, Jiang L, and Chen X. Rice Domestication Revealed by Reduced Shattering of Archaeological rice from the Lower Yangtze valley. Sci Rep. 2016:6(1):28136. https://doi.org/10.1038/srep28136

Zhong W, Dong L, Zhang G, Li Z, Wang Q, Jackson D, and Yang F. *CUC* genes control boundary domains for shoot meristem maintenance and organ initiation in maize. 2025. https://doi.org/10.1111/nph.70441

Zhu M, Hsu C-W, Peralta Ogorek LL, Taylor IW, La Cavera S, Oliveira DM, Verma L, Mehra P, Mijar M, Sadanandom A, et al. Single-cell transcriptomics reveal how root tissues adapt to soil stress. Nature. 2025:642(8068):721–729. https://doi.org/10.1038/s41586-025-08941-z

Zhu T, Wang L, Rimbert H, Rodriguez JC, Deal KR, De Oliveira R, Choulet F, Keeble-Gagnère G, Tibbits J, Rogers J, et al. Optical maps refine the bread wheat *Triticum aestivum cv. Chinese Spring* genome assembly. Plant J Cell Mol Biol. 2021:107(1):303–314. https://doi.org/10.1111/tpj.15289

Zimmermann R and Werr W. Pattern formation in the monocot embryo as revealed by *NAM* and *CUC3* orthologues from *Zea mays L*. Plant Mol Biol. 2005:58(5):669–685. https://doi.org/10.1007/s11103-005-7702-x

Zohary and Hopf. Domestication of Plants in the Old World - The Origin and Spread of Domesticated Plants in South-west Asia, Europe, and the Mediterranean Basin (Oxford University Press: New York).

Zong J, Wang L, Zhu L, Bian L, Zhang B, Chen X, Huang G, Zhang X, Fan J, Cao L, et al. A rice single cell transcriptomic atlas defines the developmental trajectories of rice floret and inflorescence meristems. New Phytol. 2022:234(2):494–512. https://doi.org/10.1111/nph.18008

Appendices

Appendix 1

Spatial Transcriptomics Reveals Expression Gradients in Developing Wheat Inflorescences at Cellular Resolution

Katie A. Long[1 †], Ashleigh Lister[2 †], Maximillian R. W. Jones[1], Nikolai M. Adamski[1], Rob E. Ellis[1], Carole Chedid[3], Sophie J. Carpenter[1], Xuemei Liu[4], Anna E. Backhaus[1,5], Andrew Goldson[2], Vanda Knitlhoffer[2], Yuanrong Pei[1], Martin Vickers[1], Burkhard Steuernagel[1], Gemy G. Kaithakottil[2], Jun Xiao[4], Wilfried Haerty[2], Iain C. Macaulay[2], Cristobal Uauy[1*]

https://doi.org/10.1101/2024.12.19.629411

[1]John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom

[2]Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ, United Kingdom

[3]Vizgen Inc, 61 Moulton Street, Cambridge, MA 02138, USA

[4]Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

[5]ICARDA, Biodiversity and Integrated Gene Management, P.O. Box 6299, Rabat Institutes, Rabat Morocco

[†]These authors contributed equally: K. A. Long and A. Lister

Spatial Transcriptomics Reveals Expression Gradients in Developing Wheat Inflorescences at Cellular Resolution

Katie A. Long[1†], Ashleigh Lister[2†], Maximillian R. W. Jones[1], Nikolai M. Adamski[1], Rob E. Ellis[1], Carole Chedid[3], Sophie J. Carpenter[1], Xuemei Liu[4], Anna E. Backhaus[1,5], Andrew Goldson[2], Vanda Knitlhoffer[2], Yuanrong Pei[1], Martin Vickers[1], Burkhard Steuernagel[1], Gemy G. Kaithakottil[2], Jun Xiao[4], Wilfried Haerty[2], Iain C. Macaulay[2], Cristobal Uauy[1*]

[1]John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom
[2]Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ, United Kingdom
[3]Vizgen Inc, 61 Moulton Street, Cambridge, MA 02138, USA
[4]Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[5]ICARDA, Biodiversity and Integrated Gene Management, P.O. Box 6299, Rabat Institutes, Rabat Morocco

[†]These authors contributed equally: K. A. Long and A. Lister

* Author for correspondence: cristobal.uauy@jic.ac.uk

Abstract

The diversity of plant inflorescence architectures is specified by gene expression patterns. In wheat (*Triticum aestivum*), the lanceolate-shaped inflorescence (spike) is defined by rudimentary spikelets at the base which initiate first but subsequently lag in development compared with central spikelets. While previous studies identified gene expression differences between central and basal inflorescence sections, the spatio-temporal dynamics and gradients along the apical-basal axis remain poorly resolved due to bulk tissue-level techniques. Here, using spatial transcriptomics, we profiled 200 genes across four stages of wheat inflorescence development to cellular resolution. Cell segmentation and unsupervised clustering identified 18 expression domains and their enriched genes, revealing dynamic spatio-temporal organisation along the apical-basal axis of the inflorescence. Along this axis, we uncovered distinct and spatially coordinated gene expression gradients patterning meristems prior to the visible delay in basal spikelet development. This study demonstrates the potential for spatial transcriptomics time-series to advance plant developmental biology.

Introduction

Nature provides a stunning diversity of vegetative and flowering structures contributing to plant fitness. Understanding the process of tissue patterning requires characterizing regulatory genes and their spatial expression, as cell fate depends on positional cues within developing tissues[1–4]. Grass morphology is patterned through phytomers, a basic unit consisting of a node, internode, leaf, and axillary meristem (AM)[5,6]. During vegetative growth, the shoot apical meristem (SAM) initiates phytomers, with lateral leaf primordia outgrowing while AMs remain dormant[7]. This developmental trajectory shifts as the SAM transitions to the inflorescence meristem. Phytomer initiation continues, however leaf (bract) outgrowth is suppressed, and AMs transition to spikelet meristems (SM) to pattern reproductive growth[8,9].

Bread wheat (*Triticum aestivum*) forms an unbranched spike-type inflorescence where pairs of AMs and suppressed bracts are initiated sequentially forming a gradient of meristem ages along the inflorescence, with basal meristems being the oldest [10,11]. The timing of transition of each AM (spikelet ridges (SRs) in wheat) into SMs, however, does not align with their developmental progression. Central SRs along the apical-basal axis are the first to initiate spikelet development, while basal SRs, despite having more time to develop, lag behind[11]. This delay persists throughout inflorescence patterning, resulting in basal spikelets forming immature floral structures that fail to produce grains[12,13]. This highlights the composite nature of the developing wheat spike, where meristems of different ages and developmental stages coexist within an ~1000 µm inflorescence, ultimately influencing its final structure.

Low-input RNA-seq of apical, central and basal sections of micro-dissected wheat spikes revealed large differences in gene expression profiles among them[13]. For example, the MADS-box transcription factor *VEGETATIVE TO REPRODUCTIVE TRANSITION 2 (VRT2)* exhibited its highest expression in basal sections, with decreasing levels toward the apex, in a proposed gradient along the inflorescence. Increased expression of *VRT2* led to a subtle delay in basal spikelet development and elongated organs within the spikelet (glumes and lemmas)[13,14]. While informative, the semi-spatial resolution of microdissections leaves precise expression patterns and putative gradients undefined. Given the high levels of differential expression observed across the spike in this experiment, we hypothesized that other genetic factors contribute to apical-basal axis patterning that warrant further investigation.

Here, we employ an adapted Multiplexed Error Robust Fluorescence *In Situ* Hybridization (MERFISH)[15,16] protocol to map the spatial expression of 200 genes along the apical-basal axis of the wheat spike to cellular resolution. Cell segmentation and unsupervised clustering across four timepoints identified 18 expression domains and their enriched gene markers, offering detailed insights into gene expression at tissue and cellular levels[17–20]. We uncovered detailed spatial and temporal organization patterns in the developing wheat spike, including coordinated transcriptional gradients that distinguish and define leaf and spikelet ridges along the apical-basal axis, prior to central meristem outgrowth. To support the broader research community, we developed an open-access

including transcripts localized to a single cell layer of the epidermis (Fig. 1f) and clear separation of floral tissues by the profiles of floral homeotic regulators (e.g. *AGAMOUS-LIKE6* (*AGL6*), *PISTILLATA1* (*PI1*); Fig. 1g). These observations are consistent with MERFISH providing spatial expression data at cellular resolution.

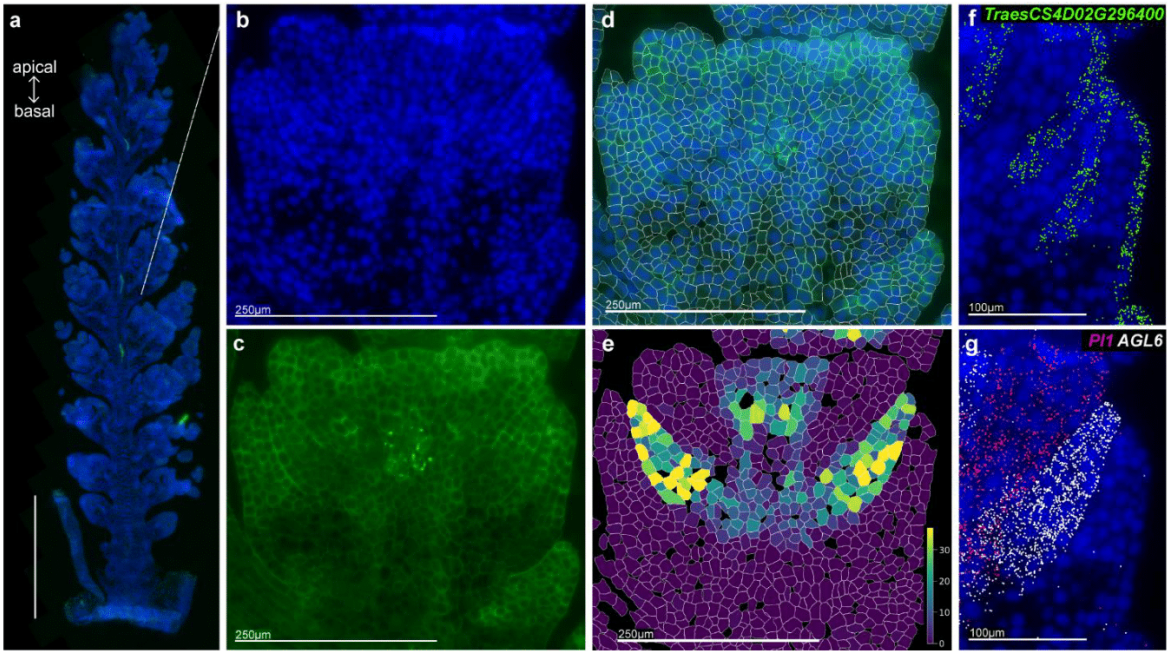


Figure 1. Development and implementation of a 200-gene MERFISH panel in wheat inflorescence tissue

a, Cryosection of W5 spike showing DAPI stain (blue) and polyT stain (green). Inset: higher magnification of floral tissues with individual staining of DAPI (**b**) and PolyT (**c**). **d,** Cellular segmentation with cellpose2, filtered for high quality cells. **e**, Heatmap displaying 'transcript counts per cell' for *AGL6*, with assignment conducted using the Vizgen Post-Processing Tool. **f**, *TRAESCS4D02G296400* transcripts (green) localized in the first cell layer of floral tissues. **g**, Tissue-specific expression patterns of *AGL6* (white) and *PI1* (pink) with distinct, non-overlapping spatial localization towards tips of palea and stamen, respectively.

MERFISH yields low off-target rates and consistent expression patterns with *in situ* hybridisations

We next implemented quality control measures to evaluate the performance of MERFISH in plants. We detected minimal off-target hybridisation based on 15 blank controls which represented <0.3% of total counts per cell (range: 0.24%–0.35%; Supplementary Table 6). As expected, probes were non-homoeolog specific, an important consideration for polyploids (Supplementary Note 1; Extended Data Fig. 2). To further validate our results, we generated *in silico* sections equivalent to physical microdissections used in RNA-seq, yielding an average Spearman's correlation coefficient of 0.66 (range 0.60-0.73; Supplementary Fig. 2c, Supplementary Table 7), supporting consistency between approaches. Additionally, MERFISH data was consistent with published *in situ* hybridizations in cereals (Supplementary Table 8, Supplementary Fig. 3). These results demonstrate that our MERFISH data exhibits

WebAtlas interface[21] (www.wheat-spatial.com), enabling visualization of all measured genes and expression domains. This work highlights the power of spatial transcriptomics time-series to investigate gene expression patterns to single cell resolution while retaining the tissue morphology and spatial context of each cell *in planta*.

Results

MERFISH of wheat inflorescence resolves gene expression to cellular resolution across four developmental timepoints

To study transcriptional gradients in wheat spikes over time, we selected genes of interest for the MERSCOPE spatial transcriptomics platform. To define these genes we first examined a micro-dissection RNA-seq dataset[13], however, given its limited developmental range and high variability, we conducted a more extensive analysis across spike development. We generated RNA-seq from central and basal spike sections across five development stages (Extended Data Fig. 1a,b; Supplementary Table 1)[22]; Early and Late double ridge (EDR, LDR; Waddington stage W2, W2.5; respectively), Lemma Primordia (LP; W3.25), Terminal Spikelet (TS; W4), and Carpel Extension (CE; W5). Individual samples expressed, on average, 49,387 high-confidence genes, with 55,346 unique genes expressed across all samples. We identified 12,384 genes with significant differential expression between central and basal sections over time ($p_{adj} < 0.001$; Extended Data Fig. 1c; Supplementary Table 2), consistent with distinct developmental pathways along the apical-basal axis.

We selected 116 genes from our differential expression dataset, supplemented with 73 additional genes including known inflorescence development genes. The final panel included 189 inflorescence related genes, 100 grain development genes for a separate project, and eleven house-keeping genes (Supplementary Table 3). After designing and synthesising the 300-probe set, we performed the Vizgen MERSCOPE workflow (see Methods). Developing wheat spikes were dissected at four stages (W2.5, W3.25, W4, W5), embedded in optimal cutting temperature (OCT) compound, flash frozen, and cryo-sectioned (Extended Data Fig. 1d). Each OCT block included 5–36 spikes (depending on stage) from two near-isogenic lines: one carrying the wildtype *VRT-A2a* allele ($P1^{WT}$) and the other the misexpression *VRT-A2b* allele from *T. turgidum* ssp. *polonicum*[14] ($P1^{POL}$; Extended Data Fig. 1e; Supplementary Fig. 1). We hybridised the probes to the samples and imaged across five experimental runs.

For downstream analysis, we selected two representative samples at each timepoint (one per genotype) for a total of eight samples. Cell segmentation[17] and transcript assignment[23] (Fig. 1a-e) yielded a cell-by-gene matrix detailing transcript counts per cell for each gene (Supplementary Table 4). After filtering for low-quality cells (see Methods), we assessed sample quality using 'total transcript counts' and 'gene counts per cell' metrics. Across all samples, the total transcript counts per cell averaged between 77.9 to 152.4 counts, while average gene counts per cell ranged from 29.6 to 42.1 (Supplementary Fig. 2a,b, Supplementary Table 5). We observed gene expression patterns at cellular resolution,

minimal off-target activity, is non-homoeolog specific, and is consistent with established gene expression patterns, confirming the technique's robustness in plant tissues.

Unsupervised clustering defines 18 expression domains traced through developmental time

We integrated eight samples and applied unsupervised clustering[19,18] which identified 18 distinct expression domains (Fig. 2a). These were visualized as spatially resolved maps across developmental stages and genotypes (Fig. 2b-e; Extended Data Fig. 3). By concatenating all time points before clustering, we traced the spatio-temporal dynamics of the domains. Notably, six domains were primarily detected at stages W4 and W5 (Fig. 2f; Supplementary Table 9), which we identified as floral tissues (Fig. 2a-e) based on their spatial distribution and organization. We hypothesised that concatenating samples across time would reveal cells in their earliest stages of differentiation. For example, in W3.25 spikes, we identified only three expression domain 15 (ED15) cells (Fig. 2g), a predominant domain in mature spikes and which later localises to paleae in W5 samples (Fig. 2h). At W3.25, these three ED15 cells were positioned just above the lemma primordia (ED1, ED2), a spatial arrangement consistent across stages. ED15 cells expressed *AGL6* (Extended Data Fig. 4), suggesting an early role for *AGL6* in palea identity consistent with wheat *agl6* mutants where paleae transform into lemma-like organs[24]. This indicates that ED15 at W3.25 represents the earliest palea progenitor cells. The presence of ED15 cells on the seventh spikelet ridge again highlights developmental differences between central and basal spikelets and the composite nature of meristems along the spike.

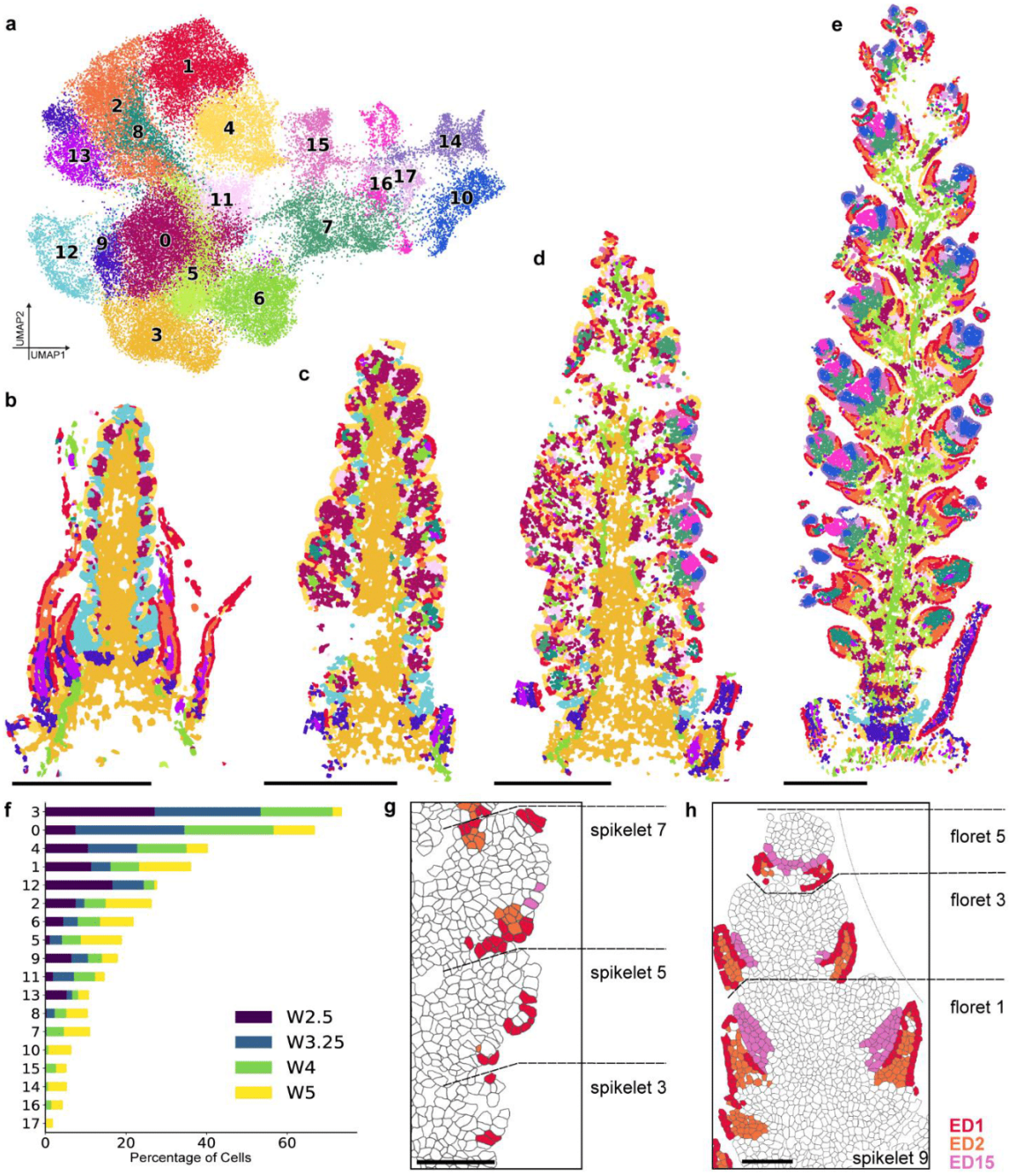


Figure 2. Unsupervised clustering identifies 18 expression domains mapped over four developmental stages.

a, UMAP projection of cells from eight samples, and expression domain assignment. **b-e**, Spatial maps of Leiden clustering across time points W2.5 (**b**), W3.25 (**c**), W4 (**d**), and W5 (**e**) using Squidpy (v1.4.1), Scanpy (v 1.10.0), and Scanorama (v1.7.4) scale bar = 500µm. f, Proportion of cells in each expression domain per sample, calculated as the percentage of total cells assigned to each cluster. **g-h**, Spatial plots of cell segmentation and assigned expression domains 1, 2, & 15 in (**g**) W3.25 spikelets and (**h**) W5 florets highlight the tracing of cell groups across developmental time, with two ED15 cells in W2.5 found above glumes & lemmas (ED1&2) tracing to palea in W5 florets. Scale bar = 100µm**.**

Gene Expression Analysis defines identity of Expression Domains

Expression domains require biological context for tissue-specific annotation. We annotated the 18 domains by combining anatomical features, such as ED15's association with palea, with domain-enriched genes (Methods; Supplementary Table 10-11). These annotations remained consistent across time and genotypes. For instance, at W2.5, ED12 is enriched for *TraesCS1A02G418200* (Fig. 3a-b), the wheat ortholog of maize *TASSEL SHEATH1* (*TSH1*) which suppresses bract outgrowth in cereal inflorescences[8,25–28]. Furthermore, at W2.5 ED12 cells and *TSH1* expression form a repeating pattern along the spike, corresponding to the suppressed leaf ridge (LR). By W3.25, LRs are no longer visible in the central and apical spike sections, yet *TSH1* persists in the few ED12 cells at the base of each spikelet ridge (SR), indicating active LR suppression during early floral organ differentiation. Furthermore, we validated the ABCDE model of floral development in wheat (Supplementary Note 2; Extended Data Fig. 5; Supplementary Table 12).

Some genes uniquely mark single domains, such as *METALLOTHIONEIN 2* (*MT2*), which identifies early vasculature in ED3 (Fig. 3a,c). Others are enriched across multiple domains, such as *KNOTTED1-LIKE HOMEOBOX 5* (*KNOX5*) in ED3 (developing rachis) and ED0 (meristem cells; Fig. 3a,d). *KNOX5* exclusion from the L1 layer is consistent with observations of its maize ortholog *KNOTTED1*[29]. Certain domains encompass multiple tissues: ED9 marks both developing young leaves and suppressed AM below the inflorescence *per se,* with *MANY NODED DWARF 1* (*MND1*) marking suppressed AMs only (Fig. 3a,e). We hypothesize that the spike-focused probe panel lacked sufficient genes to distinguish these tissues as distinct ED. Despite the limitations of the 200-gene panel, these findings demonstrate how expression domains capture complex gene expression profiles, revealing distinct developmental identities within the spike.

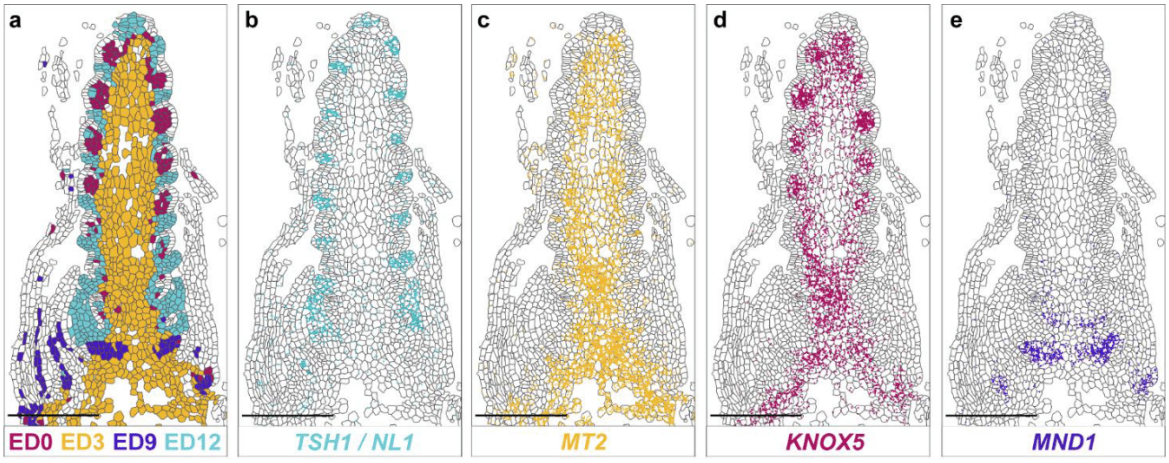


Figure 3. Gene expression analysis defines identity and enriched genes in Expression Domains

a, Spatial map of four Expression Domains in W2.5 spikes, highlighting domains enriched with transcripts of (**b**) *TASSEL SHEATH1 (TSH1)/ NECK LEAF 1 (NL1)*; (**c**) *METALLOTHIONEIN 2 (MT2)*; (**d**) *KNOTTED1-LIKE HOMEOBOX 5 (KNOX5)*, and (**e**) *MANY NODED DWARF 1 (MND1)*. Scale bar = 250µm.

Spatial analysis of *VRT2* and *SEP1-4* gradients

We previously identified differences in gene expression between spike sections using microdissection. *VRT2* was most highly expressed in basal sections, with lower expression in the apex, whereas *SEPALLATA* (*SEP*) MADS-box transcription factors displayed the opposite pattern[13]. To quantify these profiles at W4, we computationally dissected spikes into 25 transverse bins, revealing opposing expression gradients along the apical-basal axis (Fig. 4a,b). Expression domain analysis revealed spatial segregation of *VRT2* and *SEP1-4*. *VRT2* was primarily expressed in ED3 developing rachis cells with 32.2% of ED3 cells expressing *VRT2* and showed minimal expression in spikelet tissues such as glumes/lemmas (ED2, 3.3%). In contrast, *SEP1-4* was largely absent from ED3 (1.4%) but enriched in spikelet tissues including glumes/lemmas (ED2, 31.1%; Fig. 4c; Supplementary Table 13). Across the spike, only 0.7% of cells co-expressed *VRT2* and *SEP1-4*.

Next, we analysed the spatial profiles of *VRT2* and *SEP1-4* in $P1^{POL}$. MERFISH revealed ectopic *VRT2* expression and disruption of its gradient, whereas the *SEP1-4* gradient remained intact (Fig. 4b). This led to increased co-localisation of *VRT2* and *SEP1-4*, with 8.2% of cells co-expressing both transcripts along the spike. Co-expression was most pronounced in tissues exhibiting the strongest phenotypic effects in $P1^{POL}$, glumes and lemmas, where 26.3% of ED2 cells co-expressed both genes (compared to 1.1% in $P1^{WT}$; Fig. 4d; Supplementary Table 13). These results support the 'protein competition' model proposed by Li et al[30], where VRT2 interferes with SEP-FRUITFULL2 protein interactions essential for normal spikelet development. Overall, these findings demonstrate the precision of MERFISH in detecting gene expression gradients and uncovering tissue-specific co-expression patterns in developmental mutants.

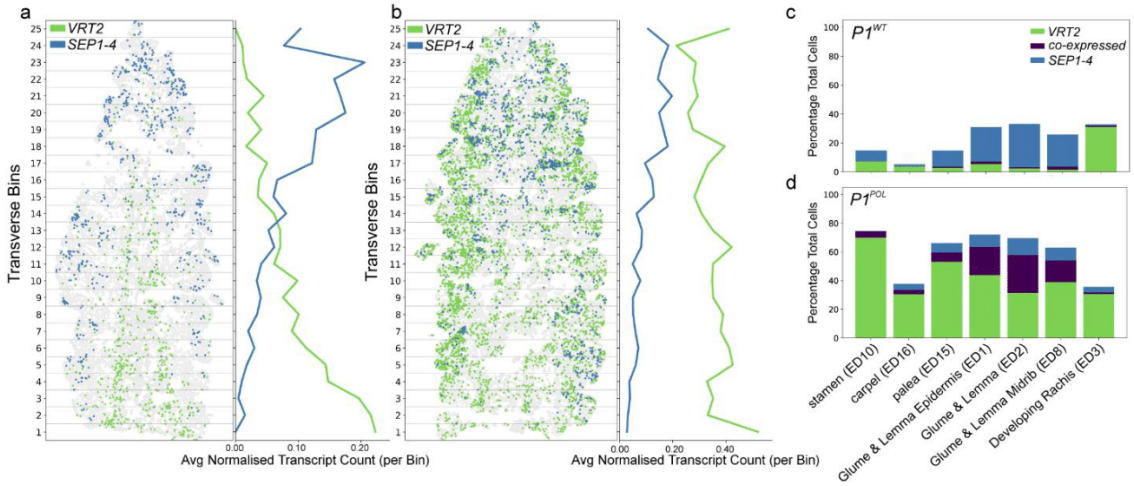


Figure 4. Opposing spatial gradients of *VRT2* and *SEP1-4* along the apical-basal axis disrupted by ectopic expression in $P1^{POL}$.
a, Spatial plot of *VRT2* and *SEP1-4* expression in $P1^{WT}$ W4 spikes, divided into 25 transverse bins along the apical-basal axis, with average normalized expression counts of *VRT2* and *SEP1-4* per transverse bin, normalised with sc.pp.normalize_total and sc.pp.log1p functions (Scanpy v1.10.0). **b**, Spatial plot of *VRT2* and *SEP1-4* expression in $P1^{POL}$ W4 spikes, divided into 25 transverse bins along the apical-basal axis. with average normalized expression

counts of *VRT2* and *SEP1-4* per transverse bin, normalised with sc.pp.normalize_total and sc.pp.log1p functions (Scanpy v1.10.0). Note difference in scale with (**a**). **c-d**, Proportion of cells expressing either *VRT2, SEP1-4,* or co-expressing both. Calculated as the percentage of total cells per cluster type in (**c**) $P1^{WT}$ and (**d**) $P1^{POL}$ isogenic lines.

Transcriptional states differentiate spikelet meristems and suppressed leaf ridges along the spike

The chronological initiation of SRs does not coincide with their developmental progression. Basal SRs, though first to initiate, lag behind in development compared to central SRs. By the Glume Primordium stage (W3), central SRs display visible outgrowth, while basal SRs remain less developed[11]. We hypothesized that additional gene expression gradients, beyond *VRT2* and *SEP1-4*, may influence these differences before W3. To explore this, we analysed transcriptional and ED patterns in W2.5 spikes. At this stage, the spike has a relatively simple ED composition, with four domains accounting for 94.8% of the inflorescence. By contrast, by W3.25, eight domains account for a comparable proportion (94.1%). The W2.5 SR comprises four domains: the L1 layer (ED4), meristematic cells in layers L2/L3 (ED0, ED12), and boundary cells (ED11) marking the adaxial boundary (Supplementary Table 14). While all SRs exhibit similar L1 (ED4) and boundary (ED11) patterns, basal SRs lack distinct ED0 regions and the expression of *KNOX5* typical of central SRs (Fig. 5a,b). Additionally, basal LR are larger, averaging 32.5 ± 15.8 cells per section (LR1-4), compared to 12.5 ± 1.5 cells in central LRs (LR8-11; Supplementary Table 14). These findings suggest developmental gradients between basal and central SR during or before W2.5.

To quantify these gradients, we grouped cells from each SR and LR, ordered them longitudinally from basal (1) to apical (13; Fig. 5c,h), and performed principal component (PC) analysis using average normalised counts per cell in each group (Fig. 5d,I; Supplementary Table 15). PC analyses revealed a consistent developmental progression for SRs and LRs along PC1 (Fig. 5d). Genes positively correlated with PC1, such as *KNOX5* (0.867), *AINTEGUMENTA-LIKE6* (*AIL6,* 0.877), and *INCREASED LEAF INCLINATION 3 (ILI3,* 0.828), have elevated expression in more central/apical SRs (Fig. 5f,g; Supplementary Table 16). Negatively correlated genes, including *SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE 14* (*SPL14,* -0.890), *FRUITFULL3* (-0.820), and *RACHIS-LIKE1* (*RIL1,* -0.923), were more prominent in basal SRs (Fig. 5f; Supplementary Table 16). Leaf ridges showed a similar separation along PC1; *APETALA2-5* (-0.932), *SPL17* (-0.889), and *LEAFY COTYLEDON 1 (LEC1,* -0.934*)* showed elevated expression in basal LRs, contrasting to *YABBY1* (0.944) and *MADS32* (0.894) with higher expression in more central/apical LRs (Fig. 5k,l; Supplementary Table 16). Thus, based on the highly correlated genes across SRs and LRs, we observed two opposing gradients along the apical-basal axis, which comprise different sets of genes, but which cross at a mirrored position along the inflorescence (SR5 & LR6; Fig. 5e,j). This suggests that adjacent spikelet and leaf ridges coordinate gene expression gradients, and collectively shift their specific expression profiles at the position where SM first start to actively differentiate.

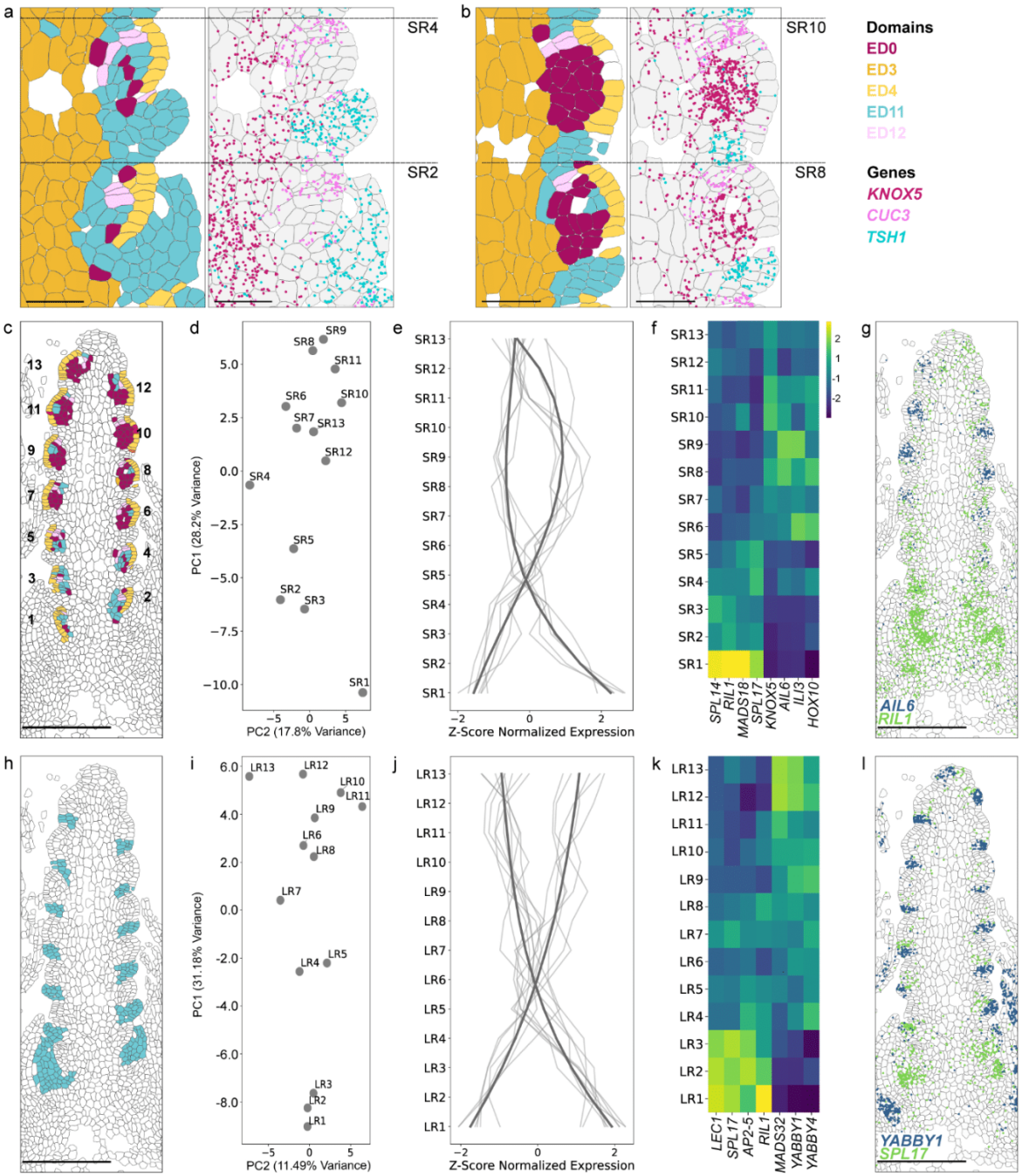


Figure 5. Coordinated and distinct transcriptional gradients define leaf ridges and spikelet ridges across the apical-basal axis at double ridge stage

a-b, Basal (2,4) and central (8,10) spikelet ridges (SRs) differ in expression domain assignment and marker gene expression. *TSH1* and *CUC3* mark the suppressed leaf ridge and the adaxial boundary of the spikelet meristem, respectively. *KNOX5* in layers L2/L3 highlights transcriptional differences between SRs along the apical-basal axis. **c**, Expression domains annotate and group SRs from 1 (basal) to 13 (apical). **d**, Principal component (PC) analysis of averaged transcripts per SR group. **e**, Normalized gene expression (Z-score) of genes highly correlated with PC1 (correlation > 0.80 or < −0.80), with polynomial regression trend lines for positively and negatively correlated groups. **f**, Normalized gene expression (Z-score) of select highly correlated genes. **g**, Expression of *AIL6* and *RIL1* in W2.5 spikes. **h**, Expression domains

annotate and group leaf ridges (LRs) from 1 (basal) to 13 (apical). **i**, PC analysis of averaged transcripts per LR group. **j**, Normalized gene expression (Z-score) of highly correlated genes to PC1 (> 0.80 and < −0.80), with polynomial regression trend lines shown separately for positively and negatively correlated groups. **k**, Normalized gene expression (Z-score) of select highly correlated genes. **l**, Expression of *YABBY1* and *SPL17* expressed in W2.5 spike.

Discussion

A systematic and unbiased approach is critical to uncover biologically meaningful insights from spatial transcriptomics data. To achieve this, we clustered cells across all developmental stages into EDs and identified enriched gene markers. By concatenating samples, we traced EDs across time, revealing novel insights into tissue identity and differentiation (Fig. 2). For example, a small population of three ED15 cells identified at W3.25 marked the earliest palea progenitors, a discovery that would have been missed through single-sample clustering. EDs also revealed similarities between tissue types. Both developing leaves at W2.5 and lemmas at W5 consist of ED1 and ED2, supporting the classification of lemmas as leaf-like organs, rather than floral-tissues[31]. Concatenating samples before clustering provided novel biological insights, therefore, we propose this strategy to analyse spatial transcriptomics data across developmental time-courses.

By utilising expression domains, we characterised each phytomer unit along the apical-basal axis, capturing the transcriptional programmes patterning vegetative and inflorescence tissues. At W2.5, ED9 cells marked by *MND1*, were identified in axillary meristems below the inflorescence (Fig. 3). These cells, forming in the axils of leaves (ED1,2,13,9), are specific to the transcriptional program driving leaf outgrowth and axillary meristem suppression. In contrast, the inflorescence is patterned by LRs (ED12) and SRs (Fig. 5), with SR adaxial boundaries consistently marked by *CUC3* and *BARREN STALK1* (ED11)[32–35]. However, SR patterning also varied along the apical-basal axis. Basal SRs (SR1–5) show increased ED12-type cells and higher expression of genes involved in AM suppression like *MND1*[36], while central and apical SRs (SR6–13) expressed meristem-specific genes like *KNOX5*[29] and floral identity-promoting factors such as *AIL6*[37,38]. Additionally, the cellular resolution of MERFISH revealed not only differences between meristems but also within them, uncovering spatially restricted A*IL6* and *ILI3* expression along the adaxial-abaxial axis (Extended Data Fig. 6).

Gene expression patterns distinguishing SRs form opposing gradients that define their developmental progression, as revealed by PC analysis (Fig. 5). A similar apical-basal gradient emerged in LRs. While all LRs clustered into ED12, basal LRs specifically expressed additional bract suppression genes (*SPL17/TSH4* and *RIL1*)[8,39], suggesting a more complex LR suppression network at the base of the spike involving factors beyond *TSH1*. Models of inflorescence development propose that the suppressed LR serve as a signalling centre orchestrating neighbouring SR meristem activity[40]. The fined-tuned transcriptional patterns of basal LRs and SRs, which are unique compared to central/apical regions, support their coordinated activity. Importantly, this insight into their spatial coordination is only achievable

with spatially resolved techniques, which overcome the limitations of bulk tissue or single-nuclei RNA sequencing.

We hypothesize that the opposing gradients observed along the inflorescence may define meristem phase transition[41]. Prusinkiewicz et al. proposed that meristem transitions are governed by the degree of 'vegetativeness', which must drop below a threshold to enable flowering[42]. Backhaus et al. put forward a wheat-specific model to meristem transition in which basal phytomers, initiated shortly after the SAM transitions to an inflorescence meristem, encounter heightened 'vegetative' signals, which may account for reduced SR development and enhanced LR growth[13]. We suggest that the distinct yet spatially coordinated transcriptional states of basal SRs and LRs adds to this understanding, potentially highlighting novel factors involved in meristem transition.

In summary, applying the cellular resolution of MERFISH to plant tissues allowed us to map the expression of a curated 200-gene panel. While the panel is limited in the breadth of genes assayed, our systematic approach combining sample integration, expression domain clustering, and gene enrichment analysis, provided novel candidate factors contributing to apical-basal patterning in wheat spikes. To support further research, all raw data is made available in addition to analysed data accessible via a WebAtlas[21] interface (www.wheat-spatial.com). We aim to empower the research community to leverage spatial transcriptomics analyses as a valuable tool in plant biology.

Author Contributions

K.A.L.: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation; **A.L.**: Investigation, Data Curation, Methodology, Platform management, Project management, Data management, Validation, Visualization; **M.R.W.J.**: Data Curation, Formal Analysis, Investigation, Resources, Writing – Original Draft Preparation; **N.M.A.**: Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **R.E.E.**: Data Curation, Resources, Software, Visualization; **C.C.**: Methodology, Resources; **S.J.C.**: Formal Analysis, Validation, Writing – Review & Editing; **X.M.L.**: Resources; **A.E.B.**: Resources, Writing – Review & Editing; **A.G.**: Resources, Writing – Review & Editing; **V.K.**: Methodology, Resources; **Y.P.**: Resources; **M.V.**: Resources, Software, Supervision; **B.S.**: Resources, Software, Supervision; **G.G.K.**: Data Curation, Software; **J.X.**: Resources; **W.H.**: Conceptualization, Funding Acquisition; **I.C.M.**: Funding Acquisition, Methodology, Project Administration, Supervision; **C.U.**: Conceptualization, Formal Analysis, Funding Acquisition, Methodology, Project Administration, Supervision, Visualization, Writing – Original Draft Preparation.

Acknowledgments

The authors would like to thank Susan Duncan (JIC) for help with training in sample preparation, David Wright (EI) and David Swarbreck (EI) for supporting data management discussions, and Philippa Borrill (JIC) and Chen Ji (JIC) for contributing grain-related genes for panel design. This work was supported the European Research Council (ERC-2019-COG-866328) and the UK Biotechnology and Biological Sciences Research Council (BBSRC) through the Delivering Sustainable Wheat (BB/X011003/1), Building Robustness in Crops (BB/X01102X/1), and Cellular Genomics (BB/X011070/1) Institute Strategic Programmes. This work was delivered via Transformative Genomics, the BBSRC funded National Bioscience Research Infrastructure (BBS/E/23NB0006) at EI by members of the Single-Cell and Spatial Analysis Group and Technical Genomics Group. K.A. L. was supported by the Gatsby Charitable Foundation; S.J.C was supported by the UKRI Biotechnology and Biological Sciences Research Council Norwich Research Park Biosciences Doctoral Training Partnership (BB/T008717/1); and J.X was supported by the Beijing Natural Science Foundation Outstanding Youth Project (JQ23026).

Competing interests

C.C. is employed by Vizgen. The remaining authors declare no competing interests.

Code and Data Availability

Code associated with this project is available at Github including all implementation scripts (https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics). All files, images, and segmentation outputs used in the analyses and described in GitHub are available on an individual sample basis (https://zenodo.org/records/14515927). The RNA-sequencing data generated in this study were submitted to NCBI under BioProject number XXXXX. Visualisation of transcript counts per cell for *P1*^{WT} samples are available at www.wheat-spatial.com.

Methods

Plant materials

We used two BC_6 near isogenic lines (NIL) differing for *VRT-A2* alleles in a hexaploid wheat (cv Paragon) background. One NIL carried the wildtype Paragon *VRT-A2a* allele, here named *P1*^{WT}, whereas the second NIL carried the *VRT-A2b* allele from *Triticum turgidum* ssp. *polonicum* (named *P1*^{POL})[14]. Plants were grown under a 16/8 h light/dark cycle at 20/15 °C, 65% relative humidity and bottom-watering irrigation[43].

Dissections and Sample Preparation

The *VRT-A2a* NIL was used for semi-spatial RNA-seq, whereas both NILs were used for MERFISH. For semi-spatial RNA-seq we used a published dissection methodology[44] to produce basal and central/apical sections. At the Early Double Ridge stage (EDR, W2), spikes were bisected, whereas for the Late Double Ridge (LDR, W2.5), Lemma Primordia (LP, W3.25), Terminal Spikelet (TS, W4) and Carpel Extension (CE, W5) stages, the basal section consisted of the most basal four spikelets from each spike. Two spikelets were skipped, then the subsequent four spikelets were harvested to comprise the central section (Extended Data Fig. 1; Supplementary Table 1). Samples were stored at -70 °C until RNA extractions which were conducted from the pooled microdissected spikes using Qiagen RNeasy Plant Mini and Zymo Direct-zol RNA Microprep kits as described in the manufacturer's manual. Total RNA (1 ug) was sent to Novogene UK for PCR-free library preparation and Illumina sequencing (PE150; 50M reads per sample).

For MERFISH, we used a similar dissection protocol[44], but maintained the youngest leaves surrounding meristems (Supplementary Fig. 1). After dissection, meristems were transferred using an RNase-free pipette tip into 4% paraformaldehyde (PFA) in 1× PBS (prepared from 6% formaldehyde [w/v], methanol-free; Pierce 28906) in 2 mL RNase-free Eppendorf tubes. Samples were vacuum infiltrated for 10 minutes or until tissue sank and incubated overnight at 4 ºC. The PFA solution was removed, and the samples were washed three times with 1× PBS. Tissue was then immersed in 15% sucrose in 1× PBS at 4 °C for 6 hours, followed by immersion in 30% sucrose in 1× PBS at 4 °C overnight.

Analysis of semi-spatial RNA-seq data

We trimmed raw reads with cutadapt (v1.9.1)[45] and generated read counts and transcripts per million (TPM) values using Kallisto pseudo-alignment (v0.44.0)[46] for all genes in the IWGSC RefSeq v1.1 annotation[47] (Supplementary Table 1). We conducted subsequent analyses for high confidence gene models with non-zero counts in at least one sample. We transformed read counts (rlog function; DESeq2 (v1.34.0))[48], and performed principal component (PC) analysis with prcomp[49]; we identified no outliers (Extended Data Fig. 1b). We calculated differential expression ($p < 0.001$; Benjamini-Hochberg corrected) between central and basal sections across time using ImpulseDE2 (v 3.6.1)[50], on genes with average > 0.5 TPM for at least one stage-section combination (Supplementary Table 2). We clustered the 12,384 differentially expressed genes using k-means (k1:10) and displayed with pheatmap (v1.0.12)[51].

Gene Panel Selection and Design for MERFISH

We designed a 300-gene panel for MERFISH, comprising 200 genes associated with spike development (116 from differential expression dataset, 73 additional genes, and 11 housekeeping genes), and 100 genes from a separate wheat grain project which are not described here. We removed genes that could not accommodate at least 25 specific probes,

10400012) with all instruments (including hybridisation box (Brabantia, 203480)) cleaned using both 70% ethanol and then RNAseZAP (Invitrogen, AM9782) or Blitz RNase Spray (Severn Biotech Ltd, 40-1735-05). We performed checks for autofluorescence at 10X using an EVOS FL2 microscope under a DAPI light cube, recording light intensity levels to decide on reduction of autofluorescence before and after photobleaching (performed for between 3 and 8 hours in EtOH 70%, Vizgen 10100003). A 300-gene probe set (Vizgen product number 20300007) was applied and hybridised for 48 hours. On the days of a run, we re-checked autofluorescence levels and topped up using the photobleacher for 3 hours if necessary. After DAPI staining, we also made checks for efficiency of staining. Clearing times varied depending on run slots. A standard clearing at 47 °C for 1 day with clearing buffer (Vizgen) containing proteinase K (NEB, P81070S) addition, was always performed, whereas additional days (1 to 4 days) of clearing at 37 °C without proteinase K in the buffer was performed. Tissue never fully cleared by eye nor when using a light microscope before MERSCOPE runs.

Upon imaging, care was taken to minimise smears and lint on the slide by cleaning with 80% ethanol and lens cleaning tissue (2105-841, Whatman). We outlined regions of interest around individual spikes on the slide overview using DAPI staining (Supplementary Figure 1). Following 60x imaging, we decoded transcripts using the panel specific MERSCOPE Codebook. We processed raw data with the MERSCOPE Instrument Software to generate and output file structures as described in MERSCOPE instrument User Guide.

Cell segmentation and processing

We performed cell segmentation on stitched images of DAPI and PolyT staining. Prior to segmentation and to minimize error, we lightened seam lines in the stitched images using FIJI (version 1.54f)[53]. Dark stitching lines were processed using the following steps: ‘Process>Filters>Maximum’ (radius = 2 pixels, applied twice) and ‘Process>Filters>Median’ (radius = 2 pixels, applied twice). Then applied three times across the entire image: ‘Process>Filters>Gaussian Blur’ (radius = 4 pixels). See Supplementary Fig. 4 for image edits and segmentation results.

We performed cell segmentation and transcript assignment using the Vizgen Post-Processing Tools (VPT, version 1.2.2)[23], within a Python virtual environment on Ubuntu 20.04. We used Cellpose2 cyto2 model[17] with DAPI (blue channel) as the nuclear marker and PolyT (green channel) as the cytoplasmic marker. For segmentation parameters, see logs (https://zenodo.org/records/14515927). We exported segmentation results as polygon geometry in both mosaic and micron space, assigned transcripts to cell boundaries using the partition-transcripts function in VPT, and generated cell metadata with the derive-entity-metadata function. Finally, we integrated cell boundaries into existing .vzg files for visualization in the Vizgen MERSCOPE Visualizer Tool[54] using the update-vzg function. All implementation scripts are available (https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics) with example commands (https://zenodo.org/records/14515927).

based on Vizgen's probe design software, except for three genes targeted by >20 probes. MERFISH probes were designed and synthesized by Vizgen (Supplementary Table 3).

Meristem Embedding and Sectioning

We cleaned all surfaces and dissection tools with RNABlitz before use. We marked a 1 cm × 1 cm area on the back of a Tissue-Tek mold (25 × 20 × 5 mm; Thermo Fisher, AGG4580) and filled with Tissue-Plus OCT compound (Agar Scientific, AGR1180). We also filled a 60 mm Petri dish with OCT. We removed individual meristems from the 30% sucrose solution using clean dissection tools, aided by a drop of OCT on the tool to adhere the meristems during collection. We transferred meristems to the OCT-filled Petri dish, where they were mixed with OCT to remove residual sucrose and ensure complete coating. Using a stereomicroscope (Leica S9 with an HXCAM HiChrome HR4 Lite camera and a Photonic Optics light source), we inspected meristems for air bubbles, which were carefully removed with a fine dissection tool. We trimmed excess vegetative tissue as needed. Meristems were then placed into the OCT-filled Tissue-Tek mold, arranged within the marked 1 cm^2 region according to genotype and developmental stage. Each OCT block contained 5–36 meristems, depending on the developmental stage (Supplementary Fig. 1), and we imaged them using GX Capture-T[52]. The OCT blocks were flash-frozen and stored at −70 °C.

We performed sectioning using a Leica CryoStar NX70. All inside surfaces and tools were cleaned with Blitz RNase Spray (Severn Biotech Ltd, 40-1735-05), and a fresh blade (MX35 Ultra™ Microtome Blade, 3053835) was used. We set the chuck temperature to −20 °C, and the blade temperature to −18 °C. We pre-warmed samples at the back of the cryostat for 30 minutes before sectioning, and brought the MERSCOPE Slides (Vizgen, 20400001) to room temperature. We trimmed OCT blocks to remove excess OCT, mounted to the chuck, and further trimmed until tissue was exposed; 10 µm sections were cut to inspect tissue regions on glass slides. Once we identified the region of interest at the optimal depth and angle, 10 µm sections were flattened with paintbrushes, flipped, and mounted onto room-temperature MERSCOPE slides, following the placement and technique outlined in the MERSCOPE user guide. After mounting, we placed the slides in 60 mm Petri dishes and incubated at the back of the cryostat for 30 minutes. We then fixed the slides in 4% methanol-free PFA in 1× PBS for 10 minutes. We washed the slides three times with 1× PBS, incubating for 5 minutes per wash. We aspirated residual PBS and air-dried the slides for 1 hour in a cell culture hood with the Petri dish lid closed. We then incubated the slides with 5 mL of 70% ethanol prepared in RNase-free water. Petri dishes were sealed with parafilm and stored at 4 °C, either overnight or for up to 7 days.

MERSCOPE Workflow

We performed slide preparation following guidelines for non-resistant fixed frozen tissue clearing (91600002_MERSCOPE Fresh and Fixed Frozen Tissue Sample Preparation User Guide_Rev E (Vizgen)). We prepared slides using Vizgen sample preparation kit (Vizgen,

Quality Checks and Filtering

We loaded the spatial transcriptomic data from the eight samples into AnnData objects (anndata v0.10.7)[55] and processed using Squidpy (v1.4.1)[19] and Scanpy (v 1.10.0[18]. We filtered expression data to include the 200 spike development and control genes, selected cells from a single inflorescence within the imaged area, and excluded low-quality cells based on volume (>500 pixels) and transcript count (>25 counts)(Supplementary Table 4). For each sample we calculated quality control (QC) metrics (total counts per cell, number of genes detected per cell, percentage of counts from top-expressed genes) and summary statistics (total counts per cell, detected genes per cell; Supplementary Table 5; Supplementary Fig. 2). We assessed off-target hybridization using blank probes (Supplementary Table 6), and across wheat homoeologs (Supplementary Note 1; Extended Data Fig. 2). Finally, we normalised expression data for all samples using scanpy functions sc.pp.normalize_total() and sc.pp.log1p() (v 1.10.0)[18].

As a further quality control metric, we performed *in silico* spike dissections equivalent to those captured by physical microdissection for 16 high-quality MERFISH sections using the MERSCOPE Visualiser 'draw ROI polygon' tool[54]. We exported frequency tables of MERFISH transcripts within basal and central/apical regions and calculated Spearman's rank correlation coefficients between these values and the mean TPM values from the relevant genotype-section-stage combination of the semi-spatial RNA-seq data (Supplementary Table 7; Supplementary Fig. 2).

Additionally, we identified *in situ* hybridisation results in wheat, barley, rice, and maize from equivalent tissues and time points as those used for MERFISH (Supplementary Table 8) and visualised them in side-by-side comparisons (Supplementary Fig. 3). We also compared expression profiles of genes involved in the ABCDE model of floral development with mutant phenotypes and knowledge from orthologs in rice, maize, barley, and wheat (Supplementary Note 2; Supplementary Table 12), and visualised using www.wheat-spatial.com (Extended Data Fig. 5).

Cell Segmentation and Transcript Visualisation

We processed the cell segmentation data as GeoDataFrames using the GeoPandas (v0.14.4)[56], and converted the transcript coordinates into a GeoDataFrame from global x and y coordinates. We performed a spatial join operation to assign transcripts to segmented cells, retaining only transcripts located within cell boundaries. We next rotated segmented cell polygons and transcript coordinates using NumPy (v1.26.3)[57], and visualised cell geometries as polygons using Matplotlib (v 2.0.4)[58] with polygon handling and transformations facilitated by Shapely (v2.0.4)[59]. Transcripts were overlaid as point features (Fig. 2b-e,g-h, 3a-e, 4a-b, 5a-c,g-h,l). Full details in implementation scripts (https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics).

MERFISH Data Integration, Unsupervised Clustering, and Gene Enrichment Analysis

We processed spatial transcriptomic data from eight samples (four timepoints; two NILs) using the Scanorama (v 1.7.4)[60,61] integration tool, and performed clustering using the Leiden algorithm with a resolution parameter of 1.0 (Fig. 2a). Spatial maps of Leiden cluster assignment were performed as described in 'Cell Segmentation and Transcript Visualisation' (Fig. 2b-e; Extended Data Fig. 3). We exported the expression domain (ED) assignment of each per cell (Supplementary Table 9), and visualised as percentage of cells in each ED, per sample (Fig. 2f). Next, we performed gene enrichment analysis on the integrated AnnData object with Scanpy function sc.tl.rank_genes_groups() using the logistic regression model[20]. This analysis returned a ranked list of genes most probable to be enriched gene markers, which we displayed alongside the average normalised expressions per ED for each sample (Supplementary Table 10). We determined top enriched values (using a +2 standard deviation threshold) and used these to annotate EDs with tissue type identity labels (Supplementary Table 11).

Transect analysis of *VRT-SEP* gradients

We filtered cells from two samples (W4, *VRT-A2a* and *VRT-A2b NILs*) to include only cells from the inflorescence region, defined as the beginning of ED12 marking leaf ridges. These cells were selected in the MERSCOPE Visualizer tool[54] using the Polygon Lasso Tool, exported as a .csv file, and the segmented cells and transcripts were mapped as previously described. The Y-axis of the spatial plot was divided into 25 transverse bins along the spike. Each cell was assigned to a bin based on its centre Y-coordinate, and we averaged the normalized transcript counts per cell within each bin (Fig. 4a-b). For both samples, we binarized gene expression data for *VRT2* and *SEP1-4* within each cell, assigning a value of 1 for detected reads and 0 for no detected reads. For each ED, we quantified the number of cells expressing only *VRT2*, only *SEP1-4*, or co-expressing both genes and visualised them as a percentage with Matplotlib (v 3.8.2[58]; Supplementary Table 13; Fig. 4c,d).

Gene expression analysis on Late Double Ridge Spikes

We selected Late Double Ridge (W2.5) and Lemma Primordia (W3.25) *P1*^{WT} inflorescence cells using the MERSCOPE Visualizer Polygon Lasso tool and exported cell identity data as a .csv file. We defined the inflorescence boundary by the first suppressed leaf ridge (ED12) and excluded cells outside the inflorescence (Supplementary Table 14). Cell counts were summed by ED, and we calculated the cumulative percentage of cells in the most populated ED to assess their contribution to the total cell population. We calculated the top EDs accounting for approximately 94% of the cells in the sample.

Groups of cells comprising the Leaf Ridges (LR) and Spikelet Ridges (SR) (defined by EDs) were annotated as "Custom Cell Groups" in the MERSCOPE Visualizer tool[54]. We delineated

SR boundaries by ED4 cells along the adaxial and abaxial axes, extending to the start of ED3 cells along the medio-lateral axis. We identified LRs as groups of ED12 cells beginning beneath the end of ED4 cells from adjacent SRs. We labelled SRs and LRs sequentially from 1 (most basal) to 13 (most apical) along the inflorescence (Fig. 5c,h; Supplementary Table 14). We calculated the total number of cells in basal (LR1-4) and central (LR8-11) leaf ridges and determined the mean cell numbers and summary statistics to compare ridge sizes between these regions.

Normalized gene expression values per cell were averaged by LR or SR group and filtered to include only genes with at least one average expression score above 0.30 across all groups (Supplementary Table 15). We standardised the resulting data matrix using StandardScaler and performed PC analysis with scikit-learn (v 1.4.2)[62], to extract the first two principal components (Fig. 5d,i). We inverted PC1 and PC2 scores to align the axes with the desired biological orientation, and calculated correlations between individual genes and PC1 (Supplementary Table 16).

We selected genes with strong positive or negative correlations to PC1 ($> |0.80|$). We calculated average expression (Z-score normalised) trends for these groups and smoothed values using a Savitzky-Golay filter (SciPy v1.13.0)[63]. Individual gene trends were also smoothed for visualisation. The smoothed average trends were fitted with cubic polynomial curves (NumPy v 1.26.3)[57] to highlight overall expression gradients (Fig. 5e,j). Plots were generated using Matplotlib (v3.8.2)[58] and Seaborn (v0.13.1)[64], and all scripts are available (https://github.com/katielong3768/Wheat-Inflorescence-Spatial-Transcriptomics).

Next, we developed maps of gene expression inside spikelet meristems focused on SR8–11. We selected individual spikelet meristems using predefined cell groups and rotated the cells to achieve a consistent orientation across the four SRs. We filtered transcript coordinates to include only those falling within selected cells and defined a 25 × 25 grid over the SR cells. We calculated transcript counts for each gene within each grid bin using NumPy digitize function (v1.26.3)[57], resulting in frequency matrices for genes of interest . We then averaged frequencies across SR8-11 to generate composite frequency matrices for each gene. To enhance spatial patterns, we applied smoothing to the composite matrices using a Gaussian filter with SciPy (v1.13.0)[63]. To suppress low-frequency noise, we also applied thresholding by setting the bottom 20% of values as minimum the threshold. Finally, we generated contour plots to visualize smoothed composite matrices (Matplotlib, v3.8.2)[58].

Extended Data Figure Legends

Extended Data Figure 1. Micro-dissection of wheat inflorescence and pooled RNA-sequencing distinguishes samples by section and stage.
a, Location of central and basal sections for each developmental time point used in pooled tissue RNA-sequencing. **b**, Principal component (PC) analysis separate samples by timepoint (PC1) and spatial section (PC2). EDR, Early Double Ridge; LDR, Late Double Ridge; LP, Lemma Primordium TS, Terminal Spikelet; CE, Carpel Extension. **c,** 12,384 genes differentially expressed between central and basal microdissections of wheat inflorescences across five developmental time points, genes selected for panel annotated with black bars. **d**, Representative cryosections of four developmental stages: Late Double Ridge (LDR, W2.5), Lemma Primordia (LP, W3.25), Terminal Spikelet (TS, W4), and Carpel Extension (CE, W5); DAPI staining in blue, PolyT in green, visualised with MERSCOPE Visualizer Tool. **e**, OCT-embedded block of 32 dissected meristems from stages EDR and LP, representing $P1^{WT}$ and $P1^{POL}$ genotypes.

Extended Data Figure 2. Hybridisation patterns indicate non-homoeolog specificity and low off-target rates for MERFISH probes.
a, Average count per cell metrics of 15 blank probes (grey) and two gene encoding probes, *TraesCS1B02G448400* and *TraesCS2A02G314100*, designed to test off-target activity, as described in Supplementary Note 1. **b,c**, Expression patterns in W5 spikes of gene encoding probes (**b**) *TraesCS1B02G448400* and (**c**) *TraesCS2A02G314100* visualised in the MERSCOPE Visualizer Tool. DAPI stain in blue and detected transcripts in white.

Extended Data Figure 3. Spatial maps of 18 expression domains mapped over four developmental stages in $P1^{POL}$ NIL replicates.
a, UMAP projection of cells from eight samples, and expression domain assignment as in Fig. 2a. **b-e,** Spatial maps of Leiden clustering across time points (**b**) W2.5, (**c**) W3.25, (**d**) W4, and (**e**) W5 using Squidpy (v1.4.1), Scanpy (v 1.10.0), and Scanorama (v1.7.4). Scale bar = 500 µm.

Extended Data Figure 4. Spatial expression patterns of *AGL6* marks palea and palea progenitor cells.
a,d, Expression patterns of *AGL6* in wildtype spikes at (**a**) LDR/W2.5, (**b**) LP/W3.25, (**c**) TS/W4, and (**d**) CE/W5 developmental stages. DAPI channel in blue, *AGL6* transcripts in pink, visualised with the MERSCOPE Visualizer Tool. LDR, Late Double Ridge; LP, Lemma Primordium TS, Terminal Spikelet; CE, Carpel Extension. **e-g,** Higher magnification of *AGL6* expression at (**e**) LP, (**f**) TS, (**g**) CE stages.

Extended Data Figure 5. Spatial expression patterns of putative floral organ identity genes.
To test the assumptions of the ABCDE model in wheat, we use MERFISH to investigate spatial expression patterns of (**a**) B-class genes *AP3*, *PI1*, and *PI2*, (**b**) C-class genes *AG1* and *AG2*, (**c**) D-class gene *STK1*, and (**d**) E-class genes *SEP3-1*, *SEP3-2*, and *AGL6* in W5 florets. The genes

in **a**-**d** are expressed in a highly spatially defined manner, suggesting their role as floral organ identity genes as postulated by the ABCDE model. In contrast, genes belonging to (**e**) the *LOFSEP* clade or (**f**) encoding putative A-class function (*AP1/FUL*-like or *AP2*-like) show either little or unrestricted expression in developing florets, suggesting that they do not contribute to defining floral organ identity. **g**, Expression domains (EDs) obtained through cell segmentation and unsupervised clustering accurately predict floral organ identity. This image corresponds to the left-hand spikelet of each panel in **a**-**f**. Note that the heatmaps in **a**-**f** are not scaled uniformly, but instead to maximize visibility of each gene's expression domain. Hence, transcript levels in these images are not comparable. For further details see www.wheat-spatial.com.

Extended Data Figure 6. Composite maps of gene expression within spikelet ridges reveals phased gene expression patterns across meristems.
a, Individual SR8 meristem spatial map with overlaid expression of *KNOX5, CUC3, ILI3,* and *AIL6*. **b-d**, Equivalent spatial maps for SR9-11, respectively. **e-h,** Composite maps of average expression from SR8-11 summarised for genes (**e**) *KNOX5*, (**f**) *CUC3*, **(g**) *AIL6*, and (**h**) *ILI3*.

Supplementary Figure Legends

Supplementary Figure 1. OCT block layout and 'Region of Interest' selections for five MERSCOPE experimental runs.
Layout of wheat inflorescence in OCT blocks (**a,c,e,g,i**), annotated with genotype and developmental time point annotations. Asterisk annotations denote spikes selected for final analysis. Images of OCT blocks taken on Leica Dissection Microscope. DAPI stain overview and experimental region selections from MERSCOPE Instrument output (**b,d,f,h,j**).

Supplementary Figure 2. Quality control checks of eight samples and high correlations between MERFISH data and bulk RNA-seq data indicate high sample quality.
a,b, Quality control metrics of (**a**) Total Transcript Counts per cell and (**b**) Total Gene Counts per cell in high quality cells only, in eight samples, across four time points and two genotypes used for onward analysis**. c,** Example of *in silico* microdissection used to select transcript counts in $P1^{WT}$ LDR (W2.5) inflorescence correlated to bulk RNA-seq data (Supplementary Table 7).

Supplementary Figure 3. *In situ* hybridisation results in cereals from equivalent tissues and time points as those used for MERFISH.
a-l, Transcript of wheat genes in MERFISH $P1^{WT}$ samples (left) compared to *in situ* hybridisation of wheat gene or cereal ortholog at equivalent inflorescence stage in published studies (right). Details of growth stages, gene IDs and publication are found in Supplementary Table 8.

Supplementary Figure 4. Editing seam lines in staining images improves segmentation along cell boundaries.

a, Segmentation outputs without staining image edits, visualised in Vizgen MERSCOPE Visualiser Tool. DAPI stain in blue, PolyT in green, and cell segmentation boundaries in white. White arrow denotes seam line detected through segmentation resulting in false cell boundaries. **b**, Raw DAPI stain, visualised in ImageJ. White arrow denotes seam line detected through segmentation resulting in false cell boundaries. **c**, Image J filters (Maximum Filter, x3 and Median Filter, x3) applied to stitching edge and Gaussian blur applied to full DAPI stain image. **d**, Cell segmentation of edited staining images, after filtering of high-quality cells.

References

1. Costa, S. Cell identity: a matter of lineage and neighbours. *New Phytol.* **210**, 1155–1158 (2016).
2. Xu, X., Smaczniak, C., Muino, J. M. & Kaufmann, K. Cell identity specification in plants: lessons from flower development. *J. Exp. Bot.* **72**, 4202–4217 (2021).
3. Costa, S. & Shaw, P. Chromatin organization and cell fate switch respond to positional information in Arabidopsis. *Nature* **439**, 493–496 (2006).
4. Kidner, C., Sundaresan, V., Roberts, K. & Dolan, L. Clonal analysis of the Arabidopsis root confirms that position, not lineage, determines cell fate. *Planta* **211**, 191–199 (2000).
5. Briske, D. D. Developmental morphology and physiology of grasses. in *Grazing Management: An Ecological Perspective* 85–108 (Timber Press, Portland).
6. Moore, K. J. & Moser, L. E. Quantifying Developmental Morphology of Perennial Grasses. *Crop Sci.* **35**, 37–43 (1995).
7. Pautler, M., Tanaka, W., Hirano, H.-Y. & Jackson, D. Grass Meristems I: Shoot Apical Meristem Maintenance, Axillary Meristem Determinacy and the Floral Transition. *Plant Cell Physiol.* **54**, 302–312 (2013).
8. Whipple, C. J. *et al.* A Conserved Mechanism of Bract Suppression in the Grass Family. *Plant Cell* **22**, 565–578 (2010).
9. Kellogg, E. A. Genetic control of branching patterns in grass inflorescences. *Plant Cell* **34**, 2518–2533 (2022).
10. Waddington, S. R., Cartwright, P. M. & Wall, P. C. A Quantitative Scale of Spike Initial and Pistil Development in Barley and Wheat. *Ann. Bot.* **51**, 119–130 (1983).
11. Bonnett, O. T. *Inflorescences of Maize, Wheat, Rye, Barley, and Oats : Their Initiation and Development*. p 721 (1966).
12. Backhaus, A. E. *et al.* Delayed development of basal spikelets in wheat explains their increased floret abortion and rudimentary nature. *J. Exp. Bot.* **74**, 5088–5103 (2023).
13. Backhaus, A. E. *et al.* High expression of the MADS-box gene VRT2 increases the number of rudimentary basal spikelets in wheat. *Plant Physiol.* **189**, 1536–1552 (2022).
14. Adamski, N. M. Ectopic expression of Triticum polonicum VRT-A2 underlies elongated glumes and grains in hexaploid wheat in a dosage-dependent manner. *Plant Cell* **33**, 2296–2319 (2021).
15. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* (2015) doi:10.1126/science.aaa6090.

16. Moffitt, J. R. *et al.* High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proc. Natl. Acad. Sci.* **113**, 14456–14461 (2016).
17. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).
18. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
19. Palla, G. *et al.* Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022).
20. Ntranos, V., Yi, L., Melsted, P. & Pachter, L. A discriminative learning approach to differential expression analysis for single-cell RNA-seq. *Nat. Methods* **16**, 163–166 (2019).
21. Li, T. *et al.* WebAtlas pipeline for integrated single-cell and spatial transcriptomic data. *Nat. Methods* (2024) doi:10.1038/s41592-024-02371-x.
22. Kirby, E. J. M. & Appleyard, M. *Cereal Development Guide. 2nd Edition.* (Arable Unit, National Agricultural Centre, Kenilworth, 1984).
23. Wiggin, T. & Yu, C. VizGen Post Processing Tool. Vizgen (2024).
24. Kong, X. *et al.* The wheat AGL6-like MADS-box gene is a master regulator for floral organ identity and a target for spikelet meristem development manipulation. *Plant Biotechnol J* **20**, 75–88 (2022).
25. Bommert, P. & Whipple, C. Grass inflorescence architecture and meristem determinacy. *Semin. Cell Dev. Biol.* **79**, 37–47 (2018).
26. Wang, L. *et al.* NECK LEAF 1, a GATA type transcription factor, modulates organogenesis by regulating the expression of multiple regulatory genes during reproductive development in rice. *Cell Res.* **19**, 598–611 (2009).
27. Houston, K. *et al.* Analysis of the barley bract suppression gene Trd1. *Theor. Appl. Genet.* **125**, 33–45 (2012).
28. Xiao, Y. *et al.* Boundary domain genes were recruited to suppress bract growth and promote branching in maize. *Sci. Adv.* (2022) doi:10.1126/sciadv.abm6835.
29. Jackson, D., Veit, B. & Hake, S. Expression of maize KNOTTED1 related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. *Development* **120**, 405–413 (1994).
30. Li, K. *et al.* Interactions between SQUAMOSA and SHORT VEGETATIVE PHASE MADS-box proteins regulate meristem transitions during wheat spike development. *Plant Cell* **33**, 3621–3644 (2021).
31. Patterson, E. L., Richardson, A. & Bartlett, M. Pushing the boundaries of organ identity: Homology of the grass lemma. *Am. J. Bot.* **110**, e16161 (2023).
32. Raman, S. *et al.* Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in Arabidopsis thaliana. *Plant J.* **55**, 65–76 (2008).
33. Tanaka, W., Pautler, M., Jackson, D. & Hirano, H.-Y. Grass Meristems II: Inflorescence Architecture, Flower Development and Meristem Fate. *Plant Cell Physiol.* **54**, 313–324 (2013).
34. Komatsu, K. *et al. LAX* and *SPA*: Major regulators of shoot branching in rice. *Proc. Natl. Acad. Sci.* **100**, 11765–11770 (2003).

35. Gallavotti, A. *et al.* The role of barren stalk1 in the architecture of maize. *Nature* **432**, 630–635 (2004).
36. Walla, A. *et al.* An Acyl-CoA N-Acyltransferase Regulates Meristem Phase Change and Plant Architecture in Barley1. *Plant Physiol.* **183**, 1088–1109 (2020).
37. Krizek, B. A. *et al.* AINTEGUMENTA and AINTEGUMENTA-LIKE6 directly regulate floral homeotic, growth, and vascular development genes in young Arabidopsis flowers. *J. Exp. Bot.* **72**, 5478–5493 (2021).
38. Krizek, BethA. AINTEGUMENTA and AINTEGUMENTA-LIKE6 Act Redundantly to Regulate Arabidopsis Floral Growth and Patterning. *Plant Physiol.* **150**, 1916–1929 (2009).
39. Ikeda, T. *et al.* BELL1-like homeobox genes regulate inflorescence architecture and meristem maintenance in rice. *Plant J. Cell Mol. Biol.* **98**, 465–478 (2019).
40. Whipple, C. J. Grass inflorescence architecture and evolution: the origin of novel signaling centers. *New Phytol.* **216**, 367–372 (2017).
41. Bartlett, M. E. & Thompson, B. Meristem identity and phyllotaxis in inflorescence development. *Front. Plant Sci.* **5**, 508 (2014).
42. Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, L. D. & Coen, E. Evolution and development of inflorescence architectures. *Science* **316**, 1452–1456 (2007).
43. Simmonds, J., Faci, I., Buffagni, V. & Uauy, C. How to grow wheat in controlled and semi-controlled environment. Preprint at https://doi.org/dx.doi.org/10.17504/protocols.io.6qpvr37jbvmk/v1.
44. Faci, I., Backhaus, A. E. & Uauy, C. Wheat spike meristem micro-dissection. Preprint at https://doi.org/dx.doi.org/10.17504/protocols.io.3byl49r2zgo5/v1 (2024).
45. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
46. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
47. chromosome-based draftsequence of the hexaploid breadwheat (Triticum aestivum) genom. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* **345**,.
48. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
49. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2017).
50. Fischer, D. S., Theis, F. J. & Yosef, N. Impulse model-based differential expression analysis of time course sequencing data. *Nucleic Acids Res.* **46**, e119 (2018).
51. Kolde, R. pheatmap: Pretty Heatmaps.
52. GX Capture-T. GT Vision (2022).
53. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
54. MERSCOPE Vizualiser. Vizgen (2023).
55. Virshup, I., Rybakov, S., Theis, F. J., Angerer, P. & Wolf, F. A. anndata: Annotated data. 2021.12.16.473007 Preprint at https://doi.org/10.1101/2021.12.16.473007 (2021).
56. Jordahl, K. *et al.* geopandas/geopandas: v0.8.1. Zenodo https://doi.org/10.5281/zenodo.3946761 (2020).
57. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

58. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
59. Gillies, S. *et al.* Shapely. (2022).
60. Hie, B., Bryson, B. & Berger, B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat. Biotechnol.* **37**, 685–691 (2019).
61. Hie, B. L., Kim, S., Rando, T. A., Bryson, B. & Berger, B. Scanorama: integrating large and diverse single-cell transcriptomic datasets. *Nat. Protoc.* **19**, 2283–2297 (2024).
62. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J Mach Learn Res* **12**, 2825–2830 (2011).
63. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
64. Waskom, M. L. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).

Appendix 2 - MERFISH 200 gene panel

RefSeq v1.1 Gene ID	Gene Name	Homoeologous genes Refseq v1.1 gene id	*Hordeum vulgare* gene ID	*Hordeum vulgare* Gene Name	*Oryza sativa Japonica* gene ID	*Oryza sativa Japonica* Gene Name	*Zea mays* gene ID	*Zea mays* Gene Name	Transcription Factor Family (Evans et al. 2022)
TraesCS1B02G448400		TraesCS1A02G418400, TraesCS1D02G426100			Os05g0579100		Zm00001eb296790, Zm00001eb344580		
TraesCS2A02G314100		TraesCS2B02G332600, TraesCS2D02G312300			Os05g0468400				
TraesCS2A02G323500		TraesCS2D02G339400, TraesCS2B02G358600	HORVU.MOREX.r3.2HG0176090		Os11g0157100	*CycT1*	Zm00001eb001290, Zm00001eb428220		
TraesCS2B02G274200	*ATG8g*	TraesCS2D02G229900, TraesCS2A02G224000	HORVU.MOREX.r3.2HG0141620		Os07g0512200	*Atg8*	Zm00001eb105850		
TraesCS7A02G336600		TraesCS7D02G344300, TraesCS7B02G248200	HORVU.MOREX.r3.7HG0712610		Os03g0565500	*mEF-G*	Zm00001eb064310		
TraesCS7B02G413800		TraesCS7D02G494400, TraesCS7A02G506300	HORVU.MOREX.r3.7HG0742760		Os06g0650100	*OsUBQ5*	Zm00001eb275020, Zm00001eb349470		
TraesCS7D02G276300		TraesCS7A02G276400			Os01g0600900, Os01g0720500, Os09g0346500		Zm00001eb296090		
TraesCS1B02G283900		TraesCS1A02G274400, TraesCS1D02G274400	HORVU.MOREX.r3.1HG0075220		Os05g0438800		Zm00001eb348450	*ACT1_5*	
TraesCS6A02G213700		TraesCS7B02G234400, TraesCS6B02G243700, TraesCS6D02G196300	HORVU.MOREX.r3.6HG0592050		Os02g0601300	*OsGapC1*	Zm00001eb184000, Zm00001eb246370	*GAPC3_1, GAPC3_2*	
TraesCS6B02G144000		TraesCS6A02G116200, TraesCS6D02G105000	HORVU.MOREX.r3.6HG0558540		Os06g0701100	*eIF-4a*	Zm00001eb226920, Zm00001eb272920		
TraesCS7A02G313100		TraesCS7B02G213300, TraesCS7D02G309500	HORVU.MOREX.r3.7HG0703580		Os08g0126300	*OsGapC3*	Zm00001eb173410, Zm00001eb261430	*GAPC2*	

TraesCS2D02G256600	*PARG-2D*	TraesCS2A02G267600, TraesCS2B02G268100	HORVU.MOREX.r3.2HG0154350		Os07g0124700	*OsPLT8*	Zm00001eb117830, Zm00001eb299300		AP2/ERF-AP2
TraesCS2D02G515800	*AP2-2*	TraesCS2B02G542400, TraesCS2A02G514200	HORVU.MOREX.r3.2HG0204770		Os04g0649100	*SHAT1*	Zm00001eb068520, Zm00001eb432100		AP2/ERF-AP2
TraesCS4A02G011600		TraesCS4D02G291300, TraesCS4B02G292900	HORVU.MOREX.r3.4HG0404560		Os03g0176300	*OsPLT10*	Zm00001eb005740		AP2/ERF-AP2
TraesCS4A02G123800	*AIL6*	TraesCS4B02G180600, TraesCS4D02G182100	HORVU.MOREX.r3.4HG0381400		Os03g0313100	*AP2/EREBP#086*	Zm00001eb015020	*EREB26*	AP2/ERF-AP2
TraesCS5A02G473800	*AP2-5*	TraesCS5B02G486900, TraesCS5D02G486600	HORVU.MOREX.r3.5HG0525620		Os03g0818800	*AP2/EREBP#033*	Zm00001eb062460	*EREB11*	AP2/ERF-AP2
TraesCS3B02G293000		TraesCS3A02G259900, TraesCS3D02G260100	HORVU.MOREX.r3.3HG0281730		Os01g0693400	*AP2/EREBP#127*	Zm00001eb156040, Zm00001eb360750		AP2/ERF-RAV
TraesCS2B02G168900		TraesCS2D02G148000, TraesCS2A02G144100	HORVU.MOREX.r3.2HG0119460		Os01g0234100				B3
TraesCS4D02G299700		TraesCS4A02G003900, TraesCS4B02G301400	HORVU.MOREX.r3.4HG0405940			*OsUBQ8*			B3
TraesCS5B02G073400		TraesCS5D02G077800	HORVU.MOREX.r3.5HG0438190				Zm00001eb030510		B3
TraesCS3A02G246000		TraesCS3B02G273400, TraesCS3D02G245400			Os01g0670800	*OsARF2*	Zm00001eb157270		B3-ARF
TraesCS1A02G418200	*TSH1*	TraesCS1D02G425900, TraesCS1B02G448200	HORVU.MOREX.r3.1HG0091840	*TRD*	Os05g0578900	*NL1*	Zm00001eb296770, Zm00001eb344540		C2C2-GATA
TraesCS1D02G162600	*YABBY1*	TraesCS1B02G203800, TraesCS1A02G176300	HORVU.MOREX.r3.1HG0049220		Os10g0508300	*YAB3*	Zm00001eb043940, Zm00001eb220660	*YAB9*	C2C2-YABBY
TraesCS2B02G403100	*YABBY3*	TraesCS2D02G382700, TraesCS2A02G386200	HORVU.MOREX.r3.2HG0184460		Os04g0536300	*YAB5*	Zm00001eb075520, Zm00001eb427470		C2C2-YABBY
TraesCS4D02G245300	*YABBY4*	TraesCS4A02G058800, TraesCS4B02G245900	HORVU.MOREX.r3.4HG0396510		Os03g0215200	*DL*	Zm00001eb008680, Zm00001eb400130	*ZmDL2*	C2C2-YABBY
TraesCS5A02G371500	*YABBY6*	TraesCS5D02G380900, TraesCS5B02G373600	HORVU.MOREX.r3.5HG0507700		Os03g0650000	*OsSh1*	Zm00001eb218500	*YAB6*	C2C2-YABBY

TraesCS6D02G220400	*YABBY7*	TraesCS6B02G266200, TraesCS6A02G237700	HORVU.MOREX.r3.6HG0598850		Os02g0643200	*YAB4*	Zm00001eb248640		C2C2-YABBY
TraesCS2D02G310500		TraesCS2A02G312200, TraesCS2B02G329000	HORVU.MOREX.r3.2HG0170820		Os04g0444100	*NSG1*	Zm00001eb081940, Zm00001eb423170		C2H2
TraesCS5A02G356100		TraesCS5D02G364900, TraesCS5B02G358600	HORVU.MOREX.r3.5HG0504500		Os09g0555700	*OsIDD6*	Zm00001eb104490, Zm00001eb320600		C2H2
TraesCS5D02G133600		TraesCS5B02G125100, TraesCS5A02G125900	HORVU.MOREX.r3.5HG0458170		Os12g0158800		Zm00001eb265700, Zm00001eb292470, Zm00001eb389000, Zm00001eb407120	*E2FA_1*	E2F-DP
TraesCS2B02G420900	*MOF1*	TraesCS2D02G400100, TraesCS2A02G402700	HORVU.MOREX.r3.2HG0187690		Os04g0566600	*MOF1/MFS2*	Zm00001eb073790, Zm00001eb428890		GARP-G2-like
TraesCS6D02G248300		TraesCS6A02G266400, TraesCS6B02G293700	HORVU.MOREX.r3.6HG0606010		Os02g0696900	*OsKANADI1*	Zm00001eb187880, Zm00001eb251390	*KAN2_1*	GARP-G2-like
TraesCS7A02G298700		TraesCS7B02G200700, TraesCS7D02G294400	HORVU.MOREX.r3.7HG0698540						GARP-G2-like
TraesCS2A02G192600	*SHR*	TraesCS2D02G194800, TraesCS2B02G214600	HORVU.MOREX.r3.2HG0130150		Os07g0586900	*OsSHR1*	Zm00001eb108090, Zm00001eb326020	*GRAS85*	GRAS
TraesCS4A02G191300		TraesCS4D02G122000, TraesCS4B02G124000	HORVU.MOREX.r3.4HG0353780		Os11g0124300, Os12g0122000	*OsSCR2*	Zm00001eb093670, Zm00001eb195650	*SCR*	GRAS
TraesCS4A02G430600			HORVU.MOREX.r3.7HG0642830		Os06g0127800	*DLT*	Zm00001eb136140		GRAS
TraesCS2A02G435100	*GRF3*	TraesCS2D02G435200, TraesCS2B02G458400	HORVU.MOREX.r3.2HG0193490		Os04g0600900	*GRF3*			GRF
TraesCS6D02G245300	*GRF9-6*	TraesCS6A02G269600, TraesCS6B02G296900	HORVU.MOREX.r3.6HG0606810				Zm00001eb071830, Zm00001eb430300		GRF
TraesCS3D02G357400	*RIL1*	TraesCS3B02G396100, TraesCS3A02G363900	HORVU.MOREX.r3.3HG0302040	*BEL2*	Os01g0848400	*qSH1, RIL1*	Zm00001eb147970	*HB48*	HB-BELL
TraesCS4D02G022500		TraesCS4A02G289500, TraesCS4B02G025400	HORVU.MOREX.r3.4HG0334350		Os03g0680700		Zm00001eb051910		HB-BELL

TraesCS1D02G197300		TraesCS1A02G193400, TraesCS1B02G208400	HORVU.MOREX.r3.1HG0053250	Os10g0575600	*Roc3(t)*	Zm00001eb024680	*ROC3_1*	HB-HD-ZIP
TraesCS2B02G419200		TraesCS2D02G398600, TraesCS2A02G401200	HORVU.MOREX.r3.2HG0187460	Os04g0569100	*ROC4*	Zm00001eb428740		HB-HD-ZIP
TraesCS5A02G549700	*HB-1*	TraesCS4D02G359600, TraesCS4B02G385200	HORVU.MOREX.r3.4HG0417970	Os03g0109400	*HOX10, LF1*	Zm00001eb404260		HB-HD-ZIP
TraesCS5B02G047200		TraesCS5A02G043400, TraesCS5D02G052300	HORVU.MOREX.r3.5HG0429100	Os12g0612700	*HOX33*	Zm00001eb031670		HB-HD-ZIP
TraesCS5D02G385300		TraesCS5B02G378000, TraesCS5A02G375800	HORVU.MOREX.r3.5HG0508470	Os03g0640800	*OsHB4*	Zm00001eb050660, Zm00001eb218730		HB-HD-ZIP
TraesCS7A02G167900		TraesCS7D02G168700, TraesCS7B02G072700	HORVU.MOREX.r3.7HG0663320	Os06g0208100	*Roc8(t)*	Zm00001eb278870	*HDZIV15_OCL15*	HB-HD-ZIP
TraesCS1D02G075700	*KNOX3*	TraesCS1A02G072800, TraesCS1B02G091700	HORVU.MOREX.r3.1HG0016370	Os05g0129700	*OsKn2*	Zm00001eb264910, Zm00001eb354880	*lg4a, lg4b*	HB-KNOX
TraesCS2B02G268200		TraesCS2D02G256400, TraesCS2A02G267400	HORVU.MOREX.r3.2HG0154270	Os07g0129700	*OsKN3, OSH15*	Zm00001eb117820, Zm00001eb299420	*gn1*	HB-KNOX
TraesCS4A02G256700	*KNOX5*	TraesCS4D02G058000, TraesCS4B02G057900	HORVU.MOREX.r3.4HG0339120	Os03g0727000	*OSH1*	Zm00001eb055920	*KN-1*	HB-KNOX
TraesCS5A02G405900		TraesCS5B02G410600, TraesCS5D02G415900	HORVU.MOREX.r3.5HG0513530	Os03g0771500, Os03g0772100	*OsH43*	Zm00001eb058930		HB-KNOX
TraesCS5D02G415900		TraesCS5B02G410600, TraesCS5A02G405900	HORVU.MOREX.r3.5HG0513530	Os03g0771500, Os03g0772100	*OsH43*	Zm00001eb058930		HB-KNOX
TraesCS1A02G052000		TraesCS1D02G054000, TraesCS1B02G069000	HORVU.MOREX.r3.1HG0010970	Os05g0118700	*LSY1*	Zm00001eb265710, Zm00001eb355310	*wox3A*	HB-WOX
TraesCS1B02G427400		TraesCS1A02G399400, TraesCS1D02G406900	HORVU.MOREX.r3.1HG0088440	Os05g0564500	*OsWOX9C*	Zm00001eb295920	*wox9C*	HB-WOX
TraesCS2A02G491900			HORVU.MOREX.r3.2HG0201080	Os04g0663600	*WOX1*	Zm00001eb067310, Zm00001eb280440, Zm00001eb433010	*wus2*	HB-WOX

TraesCS3D02G244300		TraesCS3A02G247200, TraesCS3B02G272200	HORVU.MOREX.r3.3HG0278560		Os01g0667400	*DWT1*	Zm00001eb157360, Zm00001eb359810		HB-WOX
TraesCS5D02G162600		TraesCS5A02G157300, TraesCS5B02G156400	HORVU.MOREX.r3.5HG0467090		Os11g0102100, Os12g0101600	*WOX3*	Zm00001eb197430		HB-WOX
TraesCS2B02G464200	*LFY*	TraesCS2A02G443100, TraesCS2D02G442200	HORVU.MOREX.r3.2HG0194240		Os04g0598300	*RFL*	Zm00001eb071990, Zm00001eb430240	*FL_0*	LFY
TraesCS3A02G093200	*RA2*	TraesCS3B02G108500, TraesCS3D02G093500	HORVU.MOREX.r3.3HG0233930	*RA2*		*ra2*	Zm00001eb123060	*ra2*	LOB
TraesCS3B02G435700		TraesCS3A02G402300, TraesCS3D02G397200	HORVU.MOREX.r3.3HG0306540		Os01g0889400	*CRL1L3*	Zm00001eb145150	*LBD6_1*	LOB
TraesCS4A02G236200		TraesCS4B02G078800, TraesCS4D02G077600	HORVU.MOREX.r3.4HG0342720		Os03g0609500	*OsLBD38*	Zm00001eb051620	*LBD37_3*	LOB
TraesCS3A02G434900		TraesCS3D02G427900			Os01g0922800	*MADS51*			MADS-M-type
TraesCS1A02G199600		TraesCS1B02G214500, TraesCS1D02G203300	HORVU.MOREX.r3.1HG0054220		Os10g0536100	*OsMADS56*	Zm00001eb393670	*MADS76*	MADS-MIKC
TraesCS2A02G174300	*FUL3*	TraesCS2B02G200800, TraesCS2D02G181400	HORVU.MOREX.r3.2HG0127410		Os07g0605200	*MADS18*	Zm00001eb327040	*m28*	MADS-MIKC
TraesCS2B02G281000	*FUL2*	TraesCS2D02G262700, TraesCS2A02G261200	HORVU.MOREX.r3.2HG0156870		Os07g0108900	*DEP / OsMADS15*	Zm00001eb298680		MADS-MIKC
TraesCS3A02G314300	*AG2*	TraesCS3B02G157500, TraesCS3D02G140200	HORVU.MOREX.r3.3HG0243770						MADS-MIKC
TraesCS3D02G284200	*AGL14*	TraesCS3B02G318300, TraesCS3A02G284400	HORVU.MOREX.r3.3HG0286170		Os01g0726400	*MADS32*	Zm00001eb154380		MADS-MIKC
TraesCS4B02G245800	*SEP1-1*	TraesCS4D02G245200, TraesCS4A02G078700	HORVU.MOREX.r3.4HG0396440			*OsMADS1*	Zm00001eb008690, Zm00001eb400120	*MADS14*	MADS-MIKC
TraesCS5A02G515500		TraesCS4B02G346700, TraesCS4D02G341700	HORVU.MOREX.r3.4HG0412460		Os03g0122600	*OsMADS50*	Zm00001eb001670, Zm00001eb403750	*mads1*	MADS-MIKC
TraesCS5B02G396600	*VRN1*	TraesCS5D02G401500, TraesCS5A02G391700	HORVU.MOREX.r3.1HG0050060, HORVU.MOREX.r3.5HG0511210		Os03g0752800	*MADS14*	Zm00001eb057540, Zm00001eb118120, Zm00001eb214750		MADS-MIKC

TraesCS5B02G396700	*SEP1-6*	TraesCS5A02G391800, TraesCS5D02G401700	HORVU.MOREX.r3.5HG0511250		*OsMADS34, PAP2*	Zm00001eb057560, Zm00001eb214740		MADS-MIKC
TraesCS5D02G118200		TraesCS5A02G117500, TraesCS5B02G115100	HORVU.MOREX.r3.1HG0024860	Os12g0207000	*OsMADS13*	Zm00001eb138380, Zm00001eb411130	*MADS13_1, MADS2*	MADS-MIKC
TraesCS6A02G313800	*SVP-1*	TraesCS6B02G343900, TraesCS6D02G293200	HORVU.MOREX.r3.6HG0616500	Os02g0761000	*OsMADS22*	Zm00001eb193790, Zm00001eb255670	*MADS73*	MADS-MIKC
TraesCS7A02G175200	*VRT2*	TraesCS7D02G176700, TraesCS7B02G080300	HORVU.MOREX.r3.7HG0664320	Os06g0217300	*OsMADS55*			MADS-MIKC
TraesCS7A02G383800	*AP3*	TraesCS7D02G380300, TraesCS7B02G286600	HORVU.MOREX.r3.7HG0721170	Os06g0712700	*SPW1*	Zm00001eb272490	*silky1*	MADS-MIKC
TraesCS7D02G120500	*SEP1-4*			Os06g0162800	*OsMADS5*	Zm00001eb271400, Zm00001eb375830	*AGL9, MADS5_0*	MADS-MIKC
TraesCS7D02G120600	*SEP1-5*	TraesCS7B02G021000, TraesCS7A02G122100	HORVU.MOREX.r3.7HG0654930	Os06g0162800	*OsMADS5*	Zm00001eb271400, Zm00001eb375830	*AGL9, MADS5_0*	MADS-MIKC
TraesCS7D02G246100	*CUC3*	TraesCS7B02G143900, TraesCS7A02G247600	HORVU.MOREX.r3.7HG0680310	Os08g0511200	*OsCUC3*	Zm00001eb035180		NAC
TraesCS7D02G342300		TraesCS7A02G334800, TraesCS7B02G246300	HORVU.MOREX.r3.7HG0712060	Os06g0344900	*OMTN5*	Zm00001eb264380	*NAC098_0*	NAC
TraesCS6A02G287300	*LEC1*	TraesCS6B02G316600, TraesCS6D02G268000	HORVU.MOREX.r3.6HG0611100	Os02g0725700	*OsLEC1*	Zm00001eb189490, Zm00001eb253260	*Lec1*	NF-YB
TraesCS6D02G286800		TraesCS6A02G307500, TraesCS6B02G336100	HORVU.MOREX.r3.6HG0615300			Zm00001eb372360		RWP-RK
TraesCS5A02G265900	*SPL17*	TraesCS5D02G273900, TraesCS5B02G265600	HORVU.MOREX.r3.5HG0490900	Os09g0491532	*OsSPL17*	Zm00001eb316740	*SPL17, TSH4*	SBP
TraesCS7A02G246500	*SPL14*	TraesCS7B02G144900, TraesCS7D02G245200	HORVU.MOREX.r3.7HG0679980	Os08g0509600	*WFP*	Zm00001eb035030	*ub2*	SBP
TraesCS7A02G260500		TraesCS7D02G261500, TraesCS7B02G158500	HORVU.MOREX.r3.7HG0684000	Os08g0531600	*GW8*	Zm00001eb175150, Zm00001eb175190	*Not1, tga1*	SBP

TraesCS4B02G042700	*TB1*	TraesCS4D02G040100, TraesCS4A02G271300	HORVU.MOREX.r3.4HG0336690, HORVU.MOREX.r3.4HG0336700, HORVU.MOREX.r3.4HG0336720	Os03g0706500	*FC1*	Zm00001eb054440, Zm00001eb216630	*TCP1*	TCP
TraesCS5A02G185600		TraesCS5B02G183700, TraesCS5D02G190700		Os09g0334500	*WRKY74*	Zm00001eb098330, Zm00001eb310260, Zm00001eb310270	*WRKY46_1*	WRKY
TraesCS5D02G232900		TraesCS5A02G225600, TraesCS5B02G224100	HORVU.MOREX.r3.5HG0484070	Os09g0417600	*WRKY76*	Zm00001eb312870		WRKY
TraesCS3A02G000600		TraesCS3D02G009200, TraesCS3B02G003500, TraesCS3B02G002300	HORVU.MOREX.r3.3HG0218140					bHLH
TraesCS3A02G350600	*BA1*	TraesCS3D02G344600, TraesCS3B02G383000		Os01g0831000	*LAX1*	Zm00001eb148990	*BA1*	bHLH
TraesCS4A02G016000	*ILI3*	TraesCS4D02G286800, TraesCS4B02G288000	HORVU.MOREX.r3.4HG0403770	Os03g0171700	*ILI3*	Zm00001eb005540, Zm00001eb402170		bHLH
TraesCS5A02G230500		TraesCS5D02G237300, TraesCS5B02G229000	HORVU.MOREX.r3.5HG0484930	Os09g0410700	*OsbHLH039*	Zm00001eb099390	*BHLH30_1*	bHLH
TraesCS5A02G515800		TraesCS4B02G347100, TraesCS4D02G342000	HORVU.MOREX.r3.4HG0412560	Os03g0122100	*OsbHLH052*	Zm00001eb001630, Zm00001eb403770		bHLH
TraesCS5D02G449200	*BHLH007*	TraesCS5B02G445900, TraesCS5A02G441900	HORVU.MOREX.r3.5HG0519660	Os03g0797600	*OsbHLH098*	Zm00001eb212930		bHLH
TraesCS6A02G373500		TraesCS6D02G357700, TraesCS6B02G411300	HORVU.MOREX.r3.6HG0627040	Os02g0805250	*OsbHLH107*	Zm00001eb258240	*bHLH168*	bHLH
TraesCS2A02G467100		TraesCS2D02G467100, TraesCS2B02G489900	HORVU.MOREX.r3.2HG0194820	Os02g0833600	*OsbZIP24*	Zm00001eb259650		bZIP
TraesCS2A02G495400		TraesCS2D02G495700, TraesCS2B02G523600	HORVU.MOREX.r3.2HG0201700					bZIP
TraesCS3A02G372400		TraesCS3D02G365200, TraesCS3B02G404800	HORVU.MOREX.r3.3HG0300720	Os01g0859500	*OsLG2*	Zm00001eb147220, Zm00001eb366880	*LG2*	bZIP
TraesCS5A02G508600		TraesCS4D02G335000, TraesCS4B02G338700	HORVU.MOREX.r3.4HG0411520		*OsUBQ6*			bZIP

TraesCS7A02G207100		TraesCS7B02G114300, TraesCS7D02G209800	HORVU.MOREX.r3.7HG0670600	Os01g0882200, Os06g0265400, Os10g0566200	*OsbZIP47*	Zm00001eb280500	*FEA4*, *FEA4*, *FEA4*	bZIP
TraesCS4D02G086300		TraesCS4B02G089700, TraesCS4A02G226800	HORVU.MOREX.r3.4HG0344270	Os11g0243300	*OsZHD4*	Zm00001eb168120		zf-HD
TraesCS1A02G154900	*TAW1*	TraesCS1B02G172100, TraesCS1D02G153700	HORVU.MOREX.r3.1HG0043170	Os10g0478000	*G1L5*	Zm00001eb045800, Zm00001eb221810	*G1L5_0*, *G1L5_1*	
TraesCS1D02G373800		TraesCS1B02G386600, TraesCS1A02G388600, TraesCS1A02G367900, TraesCS1B02G386500	HORVU.MOREX.r3.1HG0079250, HORVU.MOREX.r3.1HG0079280, HORVU.MOREX.r3.1HG0080550			Zm00001eb357340	*H2B2_1*, *H2B2_1*, *H2B2_1*	
TraesCS2A02G168900		TraesCS2D02G176500, TraesCS2B02G195200	HORVU.MOREX.r3.2HG0125150	Os07g0188000	*OsAGO14*	Zm00001eb093840, Zm00001eb097490, Zm00001eb097500, Zm00001eb116180, Zm00001eb116620, Zm00001eb236870, Zm00001eb239630, Zm00001eb258420, Zm00001eb276660, Zm00001eb340540, Zm00001eb424950, Zm00001eb429320		
TraesCS2B02G260800		TraesCS2A02G235900, TraesCS2D02G241800	HORVU.MOREX.r3.2HG0151710	Os07g0182900	*MET1B*	Zm00001eb301620, Zm00001eb301630		
TraesCS2B02G399800		TraesCS2D02G378900, TraesCS2A02G382500	HORVU.MOREX.r3.2HG0181760	Os04g0516200	*OsG1L4*	Zm00001eb426660	*G1L3_3*	
TraesCS2D02G079600	*Ppd1*	TraesCS2A02G081900	HORVU.MOREX.r3.2HG0107710	Os07g0695100	*Hd2*	Zm00001eb112080, Zm00001eb331630		
TraesCS3A02G043300		TraesCS3D02G037400, TraesCS3B02G040200	HORVU.MOREX.r3.3HG0222050	Os01g0129200	*SL1*	Zm00001eb125440, Zm00001eb335100		

TraesCS3A02G143100	*FT2*	TraesCS3B02G162000, TraesCS3D02G144500	HORVU.MOREX.r3.3HG0244930	*FT2*	Os01g0218500	*OsFTL1*	Zm00001eb338650	*ZCN14*
TraesCS3A02G311100		TraesCS3B02G161000, TraesCS3D02G143300, TraesCS3D02G143500	HORVU.MOREX.r3.3HG0244570	*CKX2.2*	Os01g0197700	*Gn1a*	Zm00001eb337910, Zm00001eb337920	
TraesCS3A02G441700		TraesCS3D02G434500, TraesCS3B02G475600	HORVU.MOREX.r3.3HG0312290		Os01g0927500		Zm00001eb142560	*CRN*
TraesCS3D02G124500	*GA3ox-D2*	TraesCS3B02G141800, TraesCS3A02G122600	HORVU.MOREX.r3.3HG0240140	*GA3ox2*	Os01g0177400, Os04g0517600	*D18*	Zm00001eb122500	*Dwarf1*
TraesCS4A02G409200	*RCN4-1*	TraesCS4B02G307600, TraesCS4D02G305800	HORVU.MOREX.r3.4HG0407080		Os04g0411400	*Rcn4*	Zm00001eb084030	*ZCN4*
TraesCS4B02G050200		TraesCS4D02G050400, TraesCS4A02G264800	HORVU.MOREX.r3.4HG0337770					
TraesCS4B02G064000		TraesCS4D02G062900, TraesCS4A02G250600	HORVU.MOREX.r3.4HG0339920		Os03g0733600	*OsGIF3*	Zm00001eb056300	*GIF1*
TraesCS4B02G084800		TraesCS4A02G231200, TraesCS4D02G082700	HORVU.MOREX.r3.4HG0343610		Os03g0416300			
TraesCS4D02G076900		TraesCS4B02G078300, TraesCS4A02G236700	HORVU.MOREX.r3.4HG0342640		Os03g0607600	*CycA3*	Zm00001eb051770, Zm00001eb219220	*CYCA3-1_1*
TraesCS4D02G296400	*ONI1*	TraesCS4A02G007400, TraesCS4B02G297500	HORVU.MOREX.r3.4HG0405320	*FDH*	Os03g0181500	*ONI1*	Zm00001eb006020	
TraesCS5A02G098300	*HTA3*	TraesCS5B02G103600, TraesCS5D02G110600	HORVU.MOREX.r3.5HG0446810		Os12g0530000		Zm00001eb410730	
TraesCS5A02G165400		TraesCS5B02G162600, TraesCS5D02G169900	HORVU.MOREX.r3.5HG0468970, HORVU.MOREX.r3.5HG0469020	*RBCS*	Os12g0274700, Os12g0291100, Os12g0291200, Os12g0291400, Os12g0292400	*RBCS*	Zm00001eb092540, Zm00001eb197410	*RBCS*

TraesCS5B02G509600		TraesCS5D02G509700, TraesCS4A02G363800	HORVU.MOREX.r3.5HG0529120, HORVU.MOREX.r3.5HG0529130		Os02g0519900, Os04g0118400		Zm00001eb063810, Zm00001eb063840, Zm00001eb063860, Zm00001eb311640, Zm00001eb371200	*LOS1_2, LOS1_2*	
TraesCS5B02G560300		TraesCS4A02G319100, TraesCS5D02G566200	HORVU.MOREX.r3.5HG0536610	*GA20ox1*	Os03g0856700	*GA20OX1*	Zm00001eb064970		
TraesCS5D02G136300	*TFL1*	TraesCS5A02G128600, TraesCS5B02G127600	HORVU.MOREX.r3.5HG0460080		Os11g0152500	*RCN1*	Zm00001eb164120, Zm00001eb197000, Zm00001eb406740	*ZCN1, ZCN3*	
TraesCS6B02G251600		TraesCS6D02G204800, TraesCS6A02G230100	HORVU.MOREX.r3.6HG0594230		Os02g0613900		Zm00001eb184610, Zm00001eb247050		
TraesCS7A02G076500		TraesCS7D02G072200	HORVU.MOREX.r3.7HG0647300		Os06g0136900	*DP1*	Zm00001eb270060		
TraesCS7B02G364900		TraesCS7A02G464400, TraesCS7D02G452000	HORVU.MOREX.r3.7HG0735700		Os06g0677000	*DEP3*	Zm00001eb224740		
TraesCS7B02G384000	*WAPO*	TraesCS7A02G481600, TraesCS7D02G468700	HORVU.MOREX.r3.7HG0738310		Os06g0665400	*APO1*			
TraesCS7B02G413900	*MND1*	TraesCS7A02G506400, TraesCS7D02G494500	HORVU.MOREX.r3.7HG0742750	*MND1*	Os06g0650300	*OsglHAT1*	Zm00001eb275010		
TraesCS7D02G521200		TraesCS7A02G533800	HORVU.MOREX.r3.7HG0747230		Os06g0717200	*FON1*	Zm00001eb228140	*TD1*	
TraesCSU02G093200			HORVU.MOREX.r3.7HG0654410						
TraesCS7A02G292900		TraesCS7D02G292500, TraesCS7B02G186900					Zm00001eb415010		AP2/ERF-AP2
TraesCS2B02G136100	*FZP*	TraesCS2A02G116900, TraesCS2D02G118200	HORVU.MOREX.r3.2HG0114260		Os07g0669500	*FZP*	Zm00001eb111080, Zm00001eb330200	*ERF086_0, EREB183*	AP2/ERF-ERF
TraesCS4D02G080200		TraesCS4A02G233900, TraesCS4B02G081600	HORVU.MOREX.r3.4HG0343160		Os03g0619600, Os03g0619800		Zm00001eb051360, Zm00001eb051370, Zm00001eb219020		B3
TraesCS5B02G514600		TraesCS4A02G358000, TraesCS5D02G514900	HORVU.MOREX.r3.4HG0366420, HORVU.MOREX.r3.5HG0518220						B3

TraesCS7A02G066900		TraesCS7D02G061300	HORVU.MOREX.r3.7HG0644500	Os06g0130600		Zm00001eb378270		BBR-BPC
TraesCS5A02G401800	*SP3*	TraesCS5B02G406500, TraesCS5D02G412000	HORVU.MOREX.r3.5HG0512830	Os03g0764900	*OsDof15*	Zm00001eb058480, Zm00001eb214170	*DOF24, DOF25*	C2C2-Dof
TraesCS3D02G233200		TraesCS3A02G219300	HORVU.MOREX.r3.3HG0271180					CPP
TraesCS5A02G203200		TraesCS5B02G201900, TraesCS5D02G209600	HORVU.MOREX.r3.5HG0478780	Os09g0395300	*RL9*	Zm00001eb311960		GARP-G2-like
TraesCS6D02G011600		TraesCS6A02G007300, TraesCS6B02G012800	HORVU.MOREX.r3.6HG0540370					GARP-G2-like
TraesCS7D02G379200	*MOC1*	TraesCS7A02G382800, TraesCS7B02G285500	HORVU.MOREX.r3.7HG0720900	Os06g0610350	*MOC1*	Zm00001eb388460		GRAS
TraesCS4A02G434900	*GRF5*	TraesCS7D02G044200	HORVU.MOREX.r3.7HG0641080	Os06g0116200	*OsGRF5*	Zm00001eb269150, Zm00001eb378820		GRF
TraesCS6A02G335900	*GRF10-6*	TraesCS6D02G315700, TraesCS6B02G366700	HORVU.MOREX.r3.6HG0620090	Os02g0776900	*GRF1*	Zm00001eb193180, Zm00001eb256730		GRF
TraesCS7A02G165600		TraesCS7D02G166400, TraesCS7B02G070200	HORVU.MOREX.r3.7HG0662690	Os06g0204800	*OsGRF2*	Zm00001eb278670, Zm00001eb373670	*GRF11*	GRF
TraesCS5B02G246700		TraesCS5A02G249000, TraesCS5D02G256200	HORVU.MOREX.r3.5HG0488170	Os09g0470500	*Oshox4*	Zm00001eb101280, Zm00001eb315750	*hdz10*	HB-HD-ZIP
TraesCS7A02G308400		TraesCS7B02G208600, TraesCS7D02G305200	HORVU.MOREX.r3.7HG0702200	Os08g0136100	*Roc7(t)*	Zm00001eb416980		HB-HD-ZIP
TraesCS3B02G608600		TraesCS3D02G540700, TraesCS3A02G535200	HORVU.MOREX.r3.3HG0330190			Zm00001eb296470, Zm00001eb344240	*AGL61_0, AGL62_0*	MADS-M-type
TraesCS7D02G388600		TraesCS7B02G295200, TraesCS7B02G295100, TraesCS7A02G393100	HORVU.MOREX.r3.7HG0723010	Os01g0340100	*OsMADS93*	Zm00001eb107600, Zm00001eb340810	*MADS46, MADS17*	MADS-M-type
TraesCS1A02G264300	*PI1*	TraesCS1D02G264500, TraesCS1B02G275000	HORVU.MOREX.r3.1HG0065060	Os05g0423400	*OsMADS4*	Zm00001eb349060, Zm00001eb349070	*m18, MADS29*	MADS-MIKC
TraesCS1D02G127700	*AG1*	TraesCS1A02G125800, TraesCS1B02G144800	HORVU.MOREX.r3.1HG0031260	Os05g0203800	*MADS58*	Zm00001eb284010	*ZAG1*	MADS-MIKC

TraesCS3A02G406500	*PI2*	TraesCS3D02G401700, TraesCS3B02G440200	HORVU.MOREX.r3.3HG0307160		Os01g0883100	*OsMADS2*	Zm00001eb145660	*m16*	MADS-MIKC
TraesCS4D02G243700	*SEP1-2*	TraesCS4B02G245700	HORVU.MOREX.r3.4HG0396400			*OsMADS1*	Zm00001eb008690, Zm00001eb400120	*MADS14, nan*	MADS-MIKC
TraesCS4D02G301100	*SVP3*	TraesCS4B02G302600, TraesCS4A02G002600	HORVU.MOREX.r3.4HG0406150		Os03g0186600	*OsMADS47*	Zm00001eb006480		MADS-MIKC
TraesCS5A02G286800	*SEP3-2*	TraesCS5D02G294500, TraesCS5B02G286100	HORVU.MOREX.r3.5HG0494190			*OsMADS24*			MADS-MIKC
TraesCS6A02G259000	*AGL6*	TraesCS6D02G240200, TraesCS6B02G286400	HORVU.MOREX.r3.6HG0604360		Os02g0682200	*MFO1*	Zm00001eb187330, Zm00001eb250710	*ZAG3*	MADS-MIKC
TraesCS7D02G261600	*SEP3-1*	TraesCS7A02G260600, TraesCS7B02G158600	HORVU.MOREX.r3.7HG0684020		Os08g0531700	*OsMADS7*	Zm00001eb036590	*m6*	MADS-MIKC
TraesCS6D02G253300		TraesCS6A02G273200, TraesCS6B02G300600	HORVU.MOREX.r3.6HG0607680		Os02g0706400		Zm00001eb252160		MYB-related
TraesCS7D02G233300		TraesCS7A02G233300	HORVU.MOREX.r3.7HG0676910		Os12g0522516				MYB-related
TraesCS2A02G306800		TraesCS2B02G323500, TraesCS2D02G305300	HORVU.MOREX.r3.2HG0169570	*NAC026*	Os04g0437000	*ONAC079*	Zm00001eb082430		NAC
TraesCS3D02G401200	*nac6D*	TraesCS3B02G439600, TraesCS3A02G406000	HORVU.MOREX.r3.3HG0307040		Os01g0884300	*OsNAC6*	Zm00001eb145580		NAC
TraesCS7A02G209100		TraesCS7D02G210900, TraesCS7B02G116300, TraesCS7D02G211300	HORVU.MOREX.r3.7HG0668720, HORVU.MOREX.r3.7HG0670950, HORVU.MOREX.r3.7HG0670970, HORVU.MOREX.r3.7HG0671010, HORVU.MOREX.r3.7HG0671160, HORVU.MOREX.r3.7HG0671170, HORVU.MOREX.r3.7HG0671220, HORVU.MOREX.r3.7HG0671300, HORVU.MOREX.r3.7HG0671420, HORVU.MOREX.r3.7HG0678520			*OsUBQ11*			NAC
TraesCS7D02G008500		TraesCS4A02G486200	HORVU.MOREX.r3.7HG0635150		Os06g0104200	*OsSWN7*	Zm00001eb379460		NAC

TraesCS6A02G110100	*SPL3*	TraesCS6B02G138400, TraesCS6D02G098500	HORVU.MOREX.r3.6HG0557190		Os02g0139400	*OsSPL3*	Zm00001eb208150, Zm00001eb232050	*SBP20, sbp27*	SBP
TraesCS1A02G241400		TraesCS1D02G241300, TraesCS1B02G253200	HORVU.MOREX.r3.1HG0058900		Os05g0386201	*OsUBQ7*	Zm00001eb287060, Zm00001eb347090		SRS
TraesCS2D02G019300	*TCP22*	TraesCS2B02G025700, TraesCS2A02G018000	HORVU.MOREX.r3.2HG0097880		Os08g0432300	*OsTCP22*	Zm00001eb041640, Zm00001eb113670, Zm00001eb164370, Zm00001eb179560	*TCP11*	TCP
TraesCS5A02G207300		TraesCS5B02G205600, TraesCS5D02G213400	HORVU.MOREX.r3.5HG0479720	*COM1*	Os09g0410500	*REP1*	Zm00001eb099370, Zm00001eb312380		TCP
TraesCS5D02G516300	*PCF2*	TraesCS4A02G355900, TraesCS5B02G516200	HORVU.MOREX.r3.5HG0530230		Os08g0544800	*PCF2*	Zm00001eb037480		TCP
TraesCS7B02G014500		TraesCS7D02G113100, TraesCS7A02G117100	HORVU.MOREX.r3.7HG0654120		Os01g0293100	*OsbHLH142*	Zm00001eb332170	*bHLH16*	bHLH
TraesCS1A02G306300	*FDL2*	TraesCS1B02G317100, TraesCS1D02G306000	HORVU.MOREX.r3.1HG0070040		Os05g0489700	*HBF1*	Zm00001eb352240	*bZIP96*	bZIP
TraesCS3B02G368300		TraesCS3A02G337200, TraesCS3D02G330300	HORVU.MOREX.r3.3HG0296180		Os01g0813100	*HBF2*	Zm00001eb150010, Zm00001eb369160		bZIP
TraesCS1A02G077800		TraesCS1B02G095900, TraesCS1D02G079900	HORVU.MOREX.r3.1HG0017310		Os10g0147400	*OsLAX4*	Zm00001eb026490		
TraesCS1A02G156100		TraesCS1D02G154700, TraesCS1B02G173200	HORVU.MOREX.r3.1HG0043530		Os10g0479500	*LOGL10*	Zm00001eb045690, Zm00001eb221650	*LOGL10_0*	
TraesCS1A02G392600		TraesCS1D02G400800, TraesCS1B02G420900	HORVU.MOREX.r3.1HG0087000						
TraesCS1B02G042200	*MT2B*	TraesCS1D02G034800	HORVU.MOREX.r3.1HG0005600		Os05g0111300	*OsMT2b*	Zm00001eb355770		
TraesCS1B02G274200		TraesCS1D02G263700, TraesCS1A02G263600	HORVU.MOREX.r3.1HG0065220		Os05g0421900	*GA20OX4*			
TraesCS1B02G479300		TraesCS1D02G452700, TraesCS1A02G444500	HORVU.MOREX.r3.1HG0095240		Os05g0595800		Zm00001eb297830		

TraesCS1D02G343400		TraesCS1A02G341300, TraesCS1B02G354000	HORVU.MOREX.r3.1HG0076820	Os08g0562500		Zm00001eb038890	
TraesCS2A02G376400		TraesCS2B02G393900, TraesCS2D02G372800	HORVU.MOREX.r3.2HG0182990	Os04g0525100, Os04g0525200		Zm00001eb076850	
TraesCS2B02G170400		TraesCS2D02G149200					
TraesCS2B02G318100		TraesCS2A02G302300, TraesCS2D02G301000		Os04g0429600	*OsCslH3*		
TraesCS3A02G155200		TraesCS3D02G162700, TraesCS3B02G181500	HORVU.MOREX.r3.3HG0249060	Os01g0231000	*OsIAA3*		
TraesCS3A02G251500		TraesCS3D02G251900, TraesCS3B02G281000	HORVU.MOREX.r3.3HG0284030	Os01g0708500	*LOGL1*	Zm00001eb155240	*LOGL1*
TraesCS3B02G276500		TraesCS3D02G247700, TraesCS3A02G243700	HORVU.MOREX.r3.3HG0284930	Os01g0715600	*OsPIN8*	Zm00001eb154930	
TraesCS4A02G115400		TraesCS4D02G190100, TraesCS4B02G188800	HORVU.MOREX.r3.4HG0383650	Os03g0299200		Zm00001eb013890, Zm00001eb396660	
TraesCS4D02G017800	*RGB*	TraesCS4A02G294000, TraesCS4B02G019900	HORVU.MOREX.r3.4HG0333760	Os03g0669200	*OsWD40-80*	Zm00001eb052290	*GB1*
TraesCS5A02G161000		TraesCS5D02G165900, TraesCS5B02G158400	HORVU.MOREX.r3.5HG0467900	Os09g0111100	*CycD3*	Zm00001eb095620, Zm00001eb307250	*CYCD3-2_0*
TraesCS5B02G353200		TraesCS5A02G350600, TraesCS5D02G357600		Os01g0600900, Os01g0720500, Os09g0346500	*CAB2R*		
TraesCS5B02G377300		TraesCS5A02G375200	HORVU.MOREX.r3.5HG0508300	Os03g0642300	*OsUBQ9*		
TraesCS5B02G507300		TraesCS5D02G507300, TraesCS4A02G366000	HORVU.MOREX.r3.5HG0528540	Os03g0829200		Zm00001eb211290, Zm00001eb211300	
TraesCS5B02G513800		TraesCS4A02G357800, TraesCS5D02G514300		Os03g0789900			
TraesCS5D02G216900	*DEP1*	TraesCS5A02G215100	HORVU.MOREX.r3.5HG0480200		*OsDEP1*	Zm00001eb100230, Zm00001eb314160	

TraesCS6A02G171800	TraesCS6D02G161200, TraesCS6B02G199700	HORVU.MOREX.r3.6HG0574990		Os02g0189800, Os02g0190000, Os02g0190300	*OsABCB9*	Zm00001eb206350	
TraesCS6A02G176400	TraesCS6D02G167500, TraesCS6B02G209100	HORVU.MOREX.r3.6HG0577760		Os02g0203700	*SRZ1*	Zm00001eb205550, Zm00001eb235450	
TraesCS6A02G377300	TraesCS6B02G414700, TraesCS6D02G361900	HORVU.MOREX.r3.6HG0627710		Os02g0811000	*G1L6*	Zm00001eb191830, Zm00001eb258390	*G1L6_1, G1L6_3*
TraesCS7A02G071700	TraesCS7D02G067300	HORVU.MOREX.r3.7HG0645770		Os08g0167000	*OsABCG18*	Zm00001eb415250	
TraesCS7A02G189800	TraesCS7D02G190900, TraesCS7B02G094700	HORVU.MOREX.r3.7HG0666770		Os03g0436600		Zm00001eb020600	
TraesCS7A02G206400	TraesCS7D02G209200, TraesCS7B02G113600	HORVU.MOREX.r3.7HG0670510		Os06g0264500		Zm00001eb380020	
TraesCS7A02G262900	TraesCS7D02G263600, TraesCS7B02G160700			Os08g0535600	*OsUBQ11*	Zm00001eb036900, Zm00001eb175490	*TIM8_0*
TraesCS7A02G341800	TraesCS7D02G337800, TraesCS7B02G241100	HORVU.MOREX.r3.7HG0710810		Os06g0503400		Zm00001eb281500	
TraesCS7A02G372700							
TraesCS7A02G426200							
TraesCS7D02G191600	TraesCS7B02G095500, TraesCS7A02G190600	HORVU.MOREX.r3.7HG0666880		Os06g0232300	*PIN1C*	Zm00001eb372180	*PIN1C*
TraesCS7D02G339600	TraesCS7A02G331300, TraesCS7B02G243000	HORVU.MOREX.r3.7HG0711310	*HvTIP2*	Os06g0336200	*OsTIP2*		

Appendix 3

Embedding and Sectioning of Wheat Inflorescence Tissue

Katie Long, Ashleigh Lister, Cristobal Uauy

Protocols.io, 2025

dx.doi.org/10.17504/protocols.io.rm7vzqwb4vx1/v1

Embedding and Sectioning of Wheat Inflorescence Tissue

Katie Long[1], Ashleigh Ashleigh Lister[2], Cristobal Uauy[1]

[1]John Innes Centre; [2]Earlham Institute

Katie Long
John Innes Centre

Protocol Info: Katie Long, Ashleigh Ashleigh Lister, Cristobal Uauy . Embedding and Sectioning of Wheat Inflorescence Tissue.
protocols.io **https://protocols.io/view/embedding-and-sectioning-of-wheat-inflorescence-ti-g3tnbynmf**

Manuscript citation:
Spatial Transcriptomics Reveals Expression Gradients in Developing Wheat Inflorescences at Cellular Resolution
Katie A. Long, Ashleigh Lister, Maximillian R. W. Jones, Nikolai M. Adamski, Rob E. Ellis, Carole Chedid, Sophie J. Carpenter, Xuemei Liu, Anna E. Backhaus, Andrew Goldson, Vanda Knitlhoffer, Yuanrong Pei, Martin Vickers, Burkhard Steuernagel, Gemy G. Kaithakottil, Jun Xiao, Wilfried Haerty, Iain C. Macaulay, Cristobal Uauy
bioRxiv 2024.12.19.629411; doi: https://doi.org/10.1101/2024.12.19.629411

Created: June 23, 2025

Last Modified: July 04, 2025

Protocol Integer ID: 220750

Keywords: wheat, development, nuclei, DNA, cryosection

Funders Acknowledgements:
European Research Council
Grant ID: ERC-2019-COG-866328
UK Biotechnology and Biological Sciences Research Council (BBSRC)
Grant ID: BB/X011003/1
UK Biotechnology and Biological Sciences Research Council (BBSRC)
Grant ID: BB/X01102X/1
UK Biotechnology and Biological Sciences Research Council (BBSRC)
Grant ID: BB/X011070/1
UK Biotechnology and Biological Sciences Research Council (BBSRC)
Grant ID: BBS/E/23NB0006
UKRI Biotechnology and Biological Sciences Research Council Norwich Research Park Biosciences Doctoral Training Partnership
Grant ID: BB/T008717/1

Abstract

This protocol documents the steps for tissue fixation, embedding, and sectioning of wheat inflorescence tissue in preparation for a Multiplexed Error Robust Fluorescence In Situ Hybridisation (MERFISH) protocol; however, can be applied to other spatial transcriptomic or histology applications. Accompanying information can be found at Long & Lister et al., 2024 doi: https://doi.org/10.1101/2024.12.19.629411

Guidelines

Safety warnings

- WARNING: Formaldehyde is a carcinogen. Use in a chemical fume hood.
- WARNING: Take care when using microtome blades. The cutting edge is extremely sharp. It is recommended to wear cut-resistant safety gloves. Before manipulating or moving a specimen, or taking a break, always lock the handwheel and cover the cutting edge with the knife guard.
- WARNING: Liquid nitrogen rapidly vaporizes to gas. Displacing air in a confined space such as a service lift may kill by asphyxiation. Cold and frostbite burns can occur from direct contact or contact with items cooled by Liquid Nitrogen.

Materials

Materials/Reagents

- 16% formaldehyde [w/v], methanol-free (Pierce, 28906M)
- Tissue-Tek mold, 25 × 20 × 5 mm (Thermo Fisher, AGG4580M)
- Blitz RNase Spray (Severn Biotech Ltd, 40173505M)
- MX35 Ultra Microtome Blade (Epredia, 3053835M)
- Tissue-Plus OCT compound (Agar Scientific, AGR1180M)
- Polysine Adhesion Slides (Epredia, J2800AMNZM)
- RNase-free 2ml Eppendorf Tubes
- Sucrose
- Plastic Petri Dishes
- 1x PBS in nuclease-free water
- Liquid Nitrogen

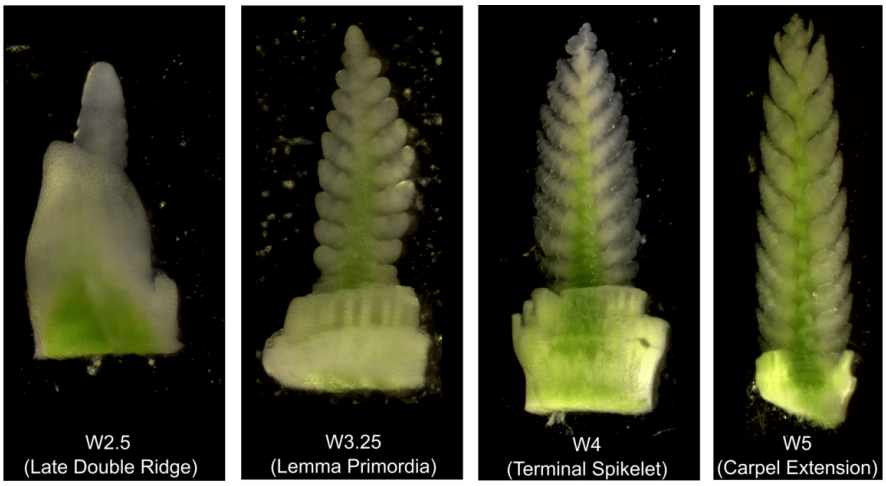
Equipment/Tools

- Liquid Nitrogen Carrier, 4 Litre (Agar Scientific, AGB7477M)
- Dissection Tools (GeeEdge ophthalmic slit knife model CJY01-2.2, GeeEdge MVR CJY04-20g)
- Stereo microscope with lighting & camera (Leica S9, HXCAM HiChrome HR4 Lite camera and a Photonic Optics light source)
- Cryostat (Leica CryoStar NX70M)
- Paintbrushes

Tissue Dissection & Fixation

1 Prepare 4% Paraformaldehyde solution in 1x PBS. Aliquot 1 mL 4% PFA solution into 2ml RNase-free Eppendorf Tube, store on ice.

2 Follow steps for tissue dissection documented in 'Wheat spike meristem microdissection' (dx.doi.org/10.17504/protocols.io.3byl49r2zgo5/v2). We modified this protocol by including the youngest leaves wrapped around the spike meristem.



Dissected spike meristem prepared for MERFISH protocol, in four developmental stages (Waddington stage W2.5, W3.25, W4, and W5). Taken on Leica Stereomicroscope.

3 Carefully transfer dissected spike meristems to 4% PFA solution with RNase-free pipette tip.

4 Vacuum infiltrate tissue submerged in 4% PFA solution for 00:10:00, or until tissue sinks. 10m

5 Incubate tissues in 4% PFA solution overnight at 4 °C

Tissue Cryoprotection

6 Prepare 15% [w/v] sucrose solution with 1.5 g sucrose in 10 mL of 1x PBS.

7 Prepare 30% [w/v] sucrose solution with 3.0 g sucrose in 10 mL of 1x PBS.

8 Carefully aspirate 4% PFA solution from tissue.

9 Wash tissues with 1x PBS solution, repeat step 3x.

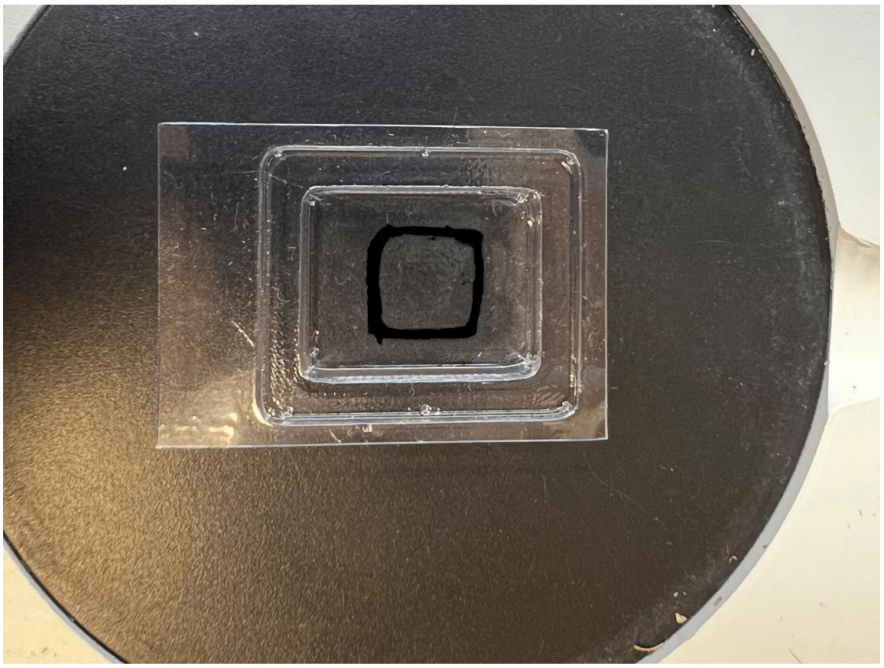
10 Immerse tissue in 15% sucrose solution for 06:00:00 at 4°C. 6h

11 Carefully aspirate 15% sucrose solution from tissues.

12 Immerse tissue in 30% sucrose solution, Overnight at 4 °C. 6h

13 Prior to embedding, clean all work surfaces and dissection tools with Blitz RNase Spray.

14 If protocol requires tissues be embedded within a limited area, proceed with the following step: With a lab marker, draw a 1 cm × 1 cm square on the back of a mold, centering the drawn square in the middle of the mold. Label the front of the mold with sample information.



Tissue-Tek mold (25 × 20 × 5 mm) with a 1cm2 square drawn on the back of the mold with
lab marker

Tissue Embedding

15 Prepare a 60 mm Petri dish filled with O.C.T. compound, adding enough to entirely cover dissected tissues.

16 Collect a drop of O.C.T. compound on a dissection tool, and use to carefully pick up tissue from 30% sucrose solution.

17 With a dissection tool, carefully transfer tissue to the Petri dish, and submerge tissue into O.C.T. compound.

18 Examine tissue in O.C.T compound under a stereo microscope. With a dissection tool, gently move tissue within OCT for 1 minute, ensuring tissue is fully coated in O.C.T. compound. Remove any visible air bubbles from tissue, either by moving tissue gently around in OCT solution, or gently moving bubble with clean dissection tool.

19 Keeping tissue emerged in O.C.T. compound, trim excess tissue with dissection tools as required.

20 Using a dissection tool, move O.C.T. coated tissue into Tissue Tek mold. Push to the bottom of the mold and align tissue to desired orientation for sectioning.

Wheat Inflorescence embedded in OCT, placed within 1cm2 area of TissueTek mold

Freezing & Storage

21 Fill liquid nitrogen carrier with liquid nitrogen.

22 Carefully place a Petri dish lid upside down, floating on liquid nitrogen solution.

23 Place Tissue-Tek mold on top of Petri dish lid. Close liquid nitrogen carrier lid.

24 Once O.C.T is frozen (completely white & opaque), remove Tissue Tek mold.

25 Store Tissue Tek mold in sealed bag, parafilm, or aluminum foil at -70 °C

Cryosectioning

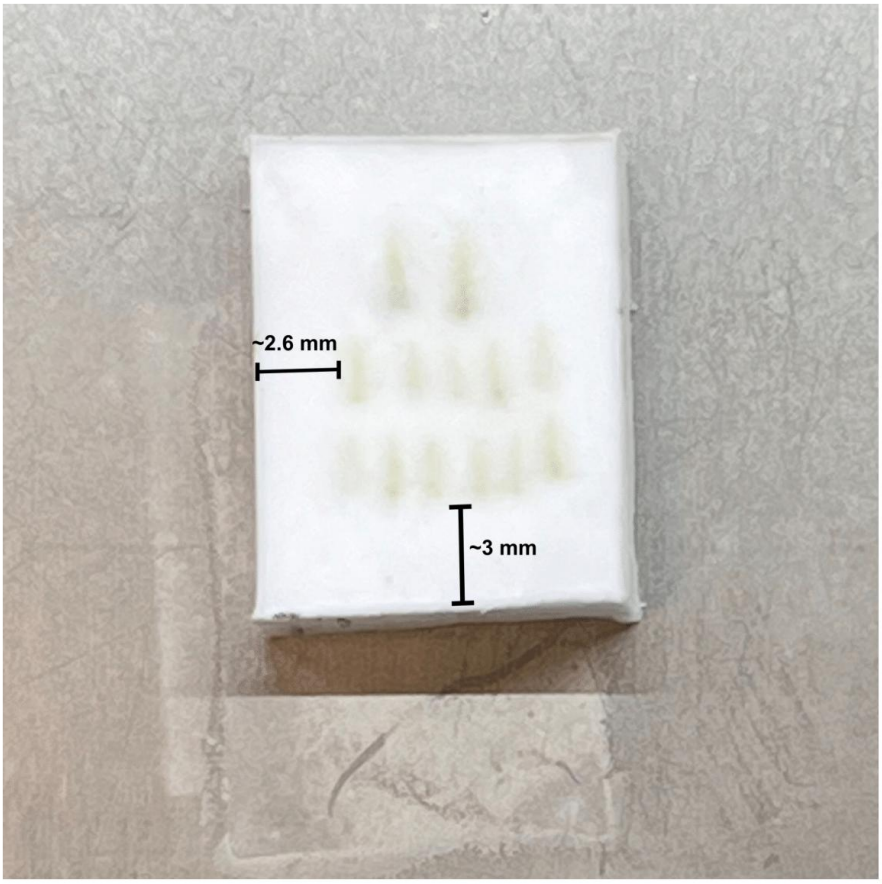
26 Clean the cryostat surface, forceps, and paintbrushes with Blitz RNase Spray before sectioning.

27 Set chamber temperature to -20 °C and blade temperature to -18 °C

28 Position a fresh MX35 Ultra Microtome Blade. Allow to cool completely before locking into place.

29 Warm cryomolds inside the cryostat for 00:30:00 at -20 °C . 30m

30 Trim O.C.T. blocks to remove excess O.C.T., ensuring a margin of O.C.T. left around tissues.



Frozen O.C.T block, trimmed to leave ~3 mm area around tissues

31 Mount O.C.T. block to chuck with room temperature O.C.T. Allow O.C.T. to solidify in cryostat.

32 Mount chuck and align tissues to blade.

33 Take 10 µm sections to trim down block. Throughout trimming, mount section to room temperature glass slides, and inspect sections to assess cutting angle and depth of

tissue.

34 Once the region of interest is at the optimal depth and angle, take 10 µm sections, slowly and flatten with paintbrush as section is cut.

35 Detach section from blade gently, and continue to flatten edges of the section with paintbrush. To prevent further corners peeling or rolling up, flip the section with

paintbrush.

36 Mount section to room temperature slides by quickly and precisely lowering slide into the cryostat chamber and pressing gently to section.

37 Move slide out of the chamber to room temperature. Visually inspect section has completely adhered to glass slide.

Example of wheat inflorescence cryosections adhered to glass slide- taken on Leica Stereomicroscope.

38 Place the slides in 60 mm Petri dishes (or another Petri dish suited to the size of glass slide) and incubate in the cryostat chamber for 00:30:00 30m

39 Wash slides 3x with 5 mL 1x PBS, incubate 00:05:00 each wash.

40 Aspirate 1xPBS. Add 5 mL 70% ethanol or until slide is completely submerged. Seal Petri dish with parafilm. Store at 4 °C for up to 1 month.

41 Next steps will vary depending further experimental design. For MERFISH experiments with the Vizgen MERSCOPE protocol, see MERSCOPE User Guide (Fresh and Fixed Frozen Tissue Sample Preparation, 91600002, Rev F).