



# **Tracing the formation of wheat spikelets through time and space**

**Anna Elisabeth Backhaus**

A thesis submitted to the University of East Anglia for the degree of Doctor  
of Philosophy

John Innes Centre

December 2022

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

This Work Is Dedicated To Heide, Catrin And Eva. The Women That Got Me Here.

*"Forscht, wo ihr was zum Forschen findet.*

*Das Unerforschbare laßt unergründet."*

(Erich Kästner, 1933)

## Abstract

Throughout domestication, humans have significantly altered the structure of the wheat (*Triticum aestivum* L.) spike to increase its yield. Spikelets are the fundamental building blocks of *Poaceae* inflorescences, and their development and branching patterns determine the inflorescence architectures and grain yield of grasses. Across the wheat spike, not all spikelets produce the same amount of grain: central spikelets produce the most grains, while this gradually decreases acro- and basipetally. The overall aim of this thesis was to understand the differences in spikelet development across the spike, and specifically the genetic and developmental causes of rudimentary basal spikelets in wheat. Basal spikelets are initiated first, yet they fall behind in development and size compared to central spikelets shortly after initiation. Using low-input transcriptomics, we were able to observe larger differences in gene expression profiles between the apical, central and basal sections of a single spike than between consecutive timepoints. We found that *SVP* (*SHORT VEGETATIVE PHASE*) MADS-box transcription factors are expressed highest in the basal section and display the opposite expression gradient to flowering genes. Higher expression of *SVP* (*VRT-A2*) in the basal sections was associated with increased numbers of rudimentary basal spikelets. Next, we developed a protocol to cultivate wheat meristems *ex situ*, to image their growth at cellular resolution, and quantify the effect of *SVPs* on basal spikelet establishment directly. Furthermore, experiments in the field suggested that the abortion of florets within a spikelet is linked to their developmental age, supporting the idea that it is the delayed development of basal spikelets that causes their preferential abortion, and thus rudimentary shape. These results highlight the complexity of transitioning from one developmental program to the next and that such changes in genetic networks do not happen instantly but require a gradual transition, during which the basal spikelets are formed.

## **Access Condition and Agreement**

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

## Acknowledgments

My greatest and deepest gratitude to my supervisor Cristobal Uauy. Thank you for all your support, encouragement, freedom to explore and guidance over the last four years. In Germany, the main supervisor is termed “Doktorvater”, and I could not think of a more fitting term for your role during my PhD. It’s been an honour to be part of your group and you have taught me everything. Thank you. I would also like to thank my secondary supervisors Richard Morris and Wilfried Haerty. You agreed to supervise me after just one short rotation, this has opened so many doors for me. Thank you for all your support and for bearing with me, despite my absolute lack of knowledge in your fields of research! Richard, thank you so much for all your encouragement. Long live Bayes! Wilfried, thank you for making complex tasks and collaborations seem easy, also thank you for a great interview, which got me into the program! I would also like to thank everyone who has made it possible for me to be at JIC, especially the rotation program JIF funding and Steph Bornemann, who always knew how to calm down a bunch of nervous PhD students. Special thanks also to my mentor Diane. You were randomly allocated to supervise my first year but then kept supervising me through the WiW program. Thank you for all your ideas and help, but especially for being such an outstanding role model and giving us so much support.

Of course, I also want to thank everyone that was there for the highs and lows of the PhD journey. My PhD siblings, Aura, Jesus, Marina, Josh, Svenja, Giuseppe and Isa. Thanks for so many wonderful, hilarious and tasty moments (in the lab, in Norwich and beyond). Without you my life would be so dull and I can’t wait until the next time when we are all together. A big thanks also to Nikolai, the wise, and Ricardo, the patient. Thanks for all your answers and support. But especially thank you for being great friends. I would also like to thank Jemima and Sophie (H.) for helping me when I first joined the group. Your welcoming gestures made a big difference, and you were inspirations to me for what is possible during a PhD.

Thank you also to all my friends outside of the JIC bubble, thank you for bearing with me during four very hectic and slightly absent years. Ronan, thanks for looking after my mental and physical health daily by providing me with food, sofas, bikes and lifts at night. Also, a big thanks to my true norfolkians and great friends Steve and Monty. My home friends: Lara, Isotta, Maja and Alice. Thank you for everything. We are scattered all over the world and yet not even a PhD was able to break our friendships. Thank you for every visit in the fine city of Norwich, thanks for being only one phone call away. Alice, thanks for always giving me guidance and purpose. A humongous thanks to Miguel and the Rabatis. Thanks for bearing with me over the last four years and for re-living the ups and downs of a PhD with me! Thanks for our daily calls while commuting through no-signal-Norwich and for every journey to see me. I still don’t understand how I was so lucky to have you in my life, but I cannot wait for our next adventures.

Lastly, I want to thank my family. You didn’t only support me during my studies but have been my backbone, my inspirations, and my best friends for 27 years. Thank you Catrin, Heide, Eva, Maike, Anette, Britta and Trude. Thanks for encouraging me to go on, for showing me the world and for being there for me when I need you the most. But especially for all the small things you’ve done for me, for revising Latin vocabulary with me so I pass school, for picking me up from rainy bus stops, for looking after brany and bello, and so much more.

## Table of contents

Abstract .....	ii
Acknowledgments .....	iii
Table of contents .....	iv
Table of figures .....	viii
Table of tables .....	ix
List of abbreviations.....	x
1 General Introduction.....	1
1.1 Increasing wheat productivity in the next decade.....	1
1.1.1 Pressure to increase yields while safeguarding nature .....	1
1.1.2 Role of wheat breeding and fundamental research for future yield gains.....	3
1.2 Architecture and development of the wheat spike .....	5
1.2.1 Evolution of the grass inflorescence .....	5
1.2.2 Architecture of the wheat and cereal crops inflorescence.....	7
1.2.3 Developmental mechanisms underlying the wheat spike architecture .....	11
1.3 What we know about rudimentary basal spikelets .....	15
1.3.1 Lanceolate shape of the wheat inflorescence .....	15
1.3.2 Rudimentary basal spikelets in the literature.....	16
1.4 Current state of techniques for developmental studies of wheat .....	17
1.4.1 RNA-seq for macro and micro dissected tissues.....	17
1.4.2 Imaging of the wheat spike.....	19
1.5 Thesis Aim .....	20
2 High expression of MADS-box gene <i>VRT2</i> increases the number of rudimentary basal spikelets in wheat .....	21
2.1 Chapter Summary .....	21
2.2 Introduction .....	21
2.3 Results.....	24
2.3.1 Low-input sequencing enables spatial analysis of the wheat spike transcriptome..	24
2.3.2 Transcriptome-wide differences are largest between the apical and basal sections of the spike.....	26
2.3.3 The <i>SVP</i> MADS-box transcription factors have opposing expression profiles to flowering E-class genes .....	29
2.3.4 <i>SVP</i> expression is higher in basal and peduncle sections and increased across all sections in <i>T. polonicum VRT-A2b isogenic lines</i> .....	33
2.3.5 Misexpression of <i>VRT-A2b</i> in <i>T. polonicum</i> increases rudimentary basal spikelet numbers 38	
2.4 Discussion.....	42

2.4.1	High-resolution spatial transcriptomics in crops .....	42
2.4.2	The composite nature of spikes .....	43
2.4.3	Delayed transition of basal spikelets from vegetative to floral developmental programmes.....	44
2.4.4	A model for the regulation of leaf and spikelet ridge outgrowth in the base of the spike	46
2.5	Materials and Methods.....	49
2.5.1	Plant materials .....	49
2.5.2	Low input RNA sequencing .....	49
2.5.3	Bioinformatic analysis .....	49
2.5.4	Reverse transcription quantitative PCR analysis.....	51
2.5.5	Field experiments and phenotyping .....	52
2.5.6	Glasshouse phenotyping.....	53
2.5.7	Phenotyping of transgenic lines.....	53
2.5.8	Modelling .....	53
3	Tracing spikelet primordia growth using MorphoGraphX .....	55
3.1	Summary .....	55
3.2	Introduction .....	55
3.3	Results.....	57
3.3.1	Development of wheat inflorescence meristem <i>ex situ</i> .....	57
3.3.2	<i>TRD1</i> leaf ridge suppressing function is conserved in wheat .....	59
3.3.3	Live imaging of TRD1 mutants confirms enhanced leaf ridge growth.....	62
3.3.4	Proof of concept of live cell growth protocol for wheat inflorescence meristem....	63
3.3.5	Live cell imaging of <i>VRT-A2b</i> basal spikelets.....	64
3.4	Discussion.....	65
3.4.1	Development of live cell imaging in wheat.....	65
3.4.2	GA role in wheat vegetative to floral transition .....	66
3.4.3	MorphoGraphX analysis of live cell growth in wheat.....	68
3.5	Materials and Methods.....	70
3.5.1	Construct assembly.....	70
3.5.2	Wheat transformation .....	71
3.5.3	Live cell imaging.....	72
4	Delayed development of basal spikelets in wheat explains their increased floret abortion and rudimentary nature.....	73
4.1	Summary .....	73
4.2	Introduction .....	73
4.3	Results.....	76
4.3.1	Basal spikelet development ceases two weeks pre-anthesis and is sensitive to resource availability at that time .....	76
4.3.2	Complete floret abortion likely cause of rudimentary basal spikelet formation .....	80

4.3.3	Florets in basal spikelets are less developed than same florets in central spikelets	84
4.3.4	Differences in development pre-abortion can predict reduced grain numbers in basal spikelets	86
4.4	Discussion	89
4.4.1	Basal spikelet abortion is likely the consequence of complete floret abortion	89
4.4.2	Lower resource availability is an unlikely cause for low grain set in basal spikelets	90
4.4.3	Differences in grain set across the spike can be predicted using floret development stages pre-abortion	91
4.4.4	Improvement of basal spikelet fertility through targeting pre-abortion development	94
4.5	Materials and Methods	96
4.5.1	Genetic material and plant growth	96
4.5.2	Shading treatments	97
4.5.3	Phenotyping	97
4.5.4	Sugar measurements	98
4.5.5	Data processing and analysis	99
5	General Discussion	101
5.1	The rise of basal spikelets: Developmental and transcriptional differences in the basal spikelets	101
5.1.1	Establishment of basal spikelets	101
5.1.2	Delayed transition of vegetative to floral signals in base	102
5.1.3	The wheat spike as a cascade of phytomeres	103
5.2	New technologies to investigate primordia development in grasses	104
5.2.1	Spatially resolved sequencing technologies for wheat	104
5.2.2	Novel microscopy to analyse cell growth	106
5.2.3	Importance of field-based physiological experiments	107
5.3	The fall of basal spikelets: Increased floret abortion in basal spikelets	108
5.4	Concluding Statement	110
6	References	111
	Appendices	A1
	Appendix 1	A1
	Appendix 2	A19

## Table of Figures

Figure 1.1: Agnes Denes, Wheatfield – A Confrontation .....	3
Figure 1.2: Phylogenetic distribution of inflorescence phyllotaxis. ....	7
Figure 1.3: Schematic representation to compare inflorescence meristem. ....	9
Figure 1.4: Light Microscopy images of the floral development of the wheat spike .....	13
Figure 1.5: The wheat spikelet structure. ....	14
Figure 2.1: Low-input G&T sequencing of developing wheat spikes.....	24
Figure 2.2: Principal Component analysis (PCA) of RNA-seq libraries .....	25
Figure 2.3: Transcriptome wide differences. ....	27
Figure 2.4: Comparison of Coefficient of variation.....	28
Figure 2.5: Expression patterns of all DEG clusters. ....	29
Figure 2.6: Clustering of differentially expressed genes across the wheat spike .....	30
Figure 2.7: Effect of VRT-A2b allele on VRT-A2 and SEP1-4 expression in the spike.....	34
Figure 2.8: Phenotypic difference between VRT-A2a (blue) and VRT-A2b (orange) on rudimentary basal spikelet numbers (RBS).....	38
Figure 2.9: Dissected Floret 1 and 2 of basal and central spikelets of BC <sub>6</sub> NILs before anthesis .....	39
Figure 2.10: Expression of barley genes in Thiel et al. (2021) .....	47
Figure 2.11: Example of Netlogo simulation outcome with default parameters .....	48
Figure 3.1: Stereomicroscopy images of wheat (A) and barley (B) inflorescence meristem development on wheat meristem media .....	58
Figure 3.2: <i>THIRD OUTER GLUME 1 (TRD1)</i> mutants in wheat. ....	60
Figure 3.3: Alignment of amino acid sequences of <i>TRD1</i> .....	61
Figure 3.4: <i>TRD1</i> phenotypic effect at double ridge stage .....	62
Figure 3.5: Live cell growth analysis of wheat inflorescence.....	63
Figure 3.6: Confocal z-stack images of individual spike meristems visualised in MorphoGraphX ...	65
Figure 3.7: Construct plasmid map of the vector containing guide 2 and guide 3.....	71
Figure 4.1: Pre-anthesis shading treatment phenotypic effects .....	77
Figure 4.2: Spikelet fertility pre-anthesis in control conditions.....	81
Figure 4.3: Waddington development stage of florets across the spikes 20 days pre-anthesis .....	84
Figure 4.4: Prediction of grains per spikelet from floret development pre-abortion.....	87
Figure 4.5: Predictions of grains per spikelet using floret count per spikelet pre-abortion.....	88
Figure 4.6: Re-analysis of Ferrante et al (2020) Figure 9 .....	93
Figure 4.7: Schematic of floret development and sugar sampling in 2021 and 2022 .....	98
Figure 5.1: Summarising figure of wheat spike meristem development from initiation to mature spike, visualising the major findings of this thesis.....	109

## Table of Tables

Table 2.1: Average number of reads aligned to the RefSeqv1.0 genome and expressed genes .....	25
Table 2.2: Summary of raw read number .....	26
Table 2.3: Enrichment of Gene Ontology (GO)-terms in the seven identified clusters of DEGs .....	32
Table 2.4: Enrichment of Transcription factor families and MADS-box transcription factors.....	33
Table 2.5: Relative expression of <i>VRT-A2</i> and <i>SEP1-4</i> measured in Paragon NILs .....	36
Table 2.6: Statistical analysis of qRT-qPCR data from <i>VRT-A2</i> NIL spike sections .....	37
Table 2.7: Field evaluations for rudimentary basal spikelets (RBS) in <i>VRT-A2</i> NILs. ....	40
Table 2.8: Simplified table of rudimentary basal spikelet phenotypic data from <i>VRT-A2</i> transgenic lines.....	40
Table 2.9: Graphical genotype from Adamski et al. (2021) and RBS phenotype of BC6 recombinant inbred lines (RILs) for <i>VRT-A2</i> .....	41
Table 2.10: List of primers used in qRT-qPCR in this chapter .....	52
Table 3.1: Number of CRISPR/Cas9 construct copies detected in T <sub>0</sub> plants from 18 independent events.....	60
Table 3.2: Primer sequences specific to this study and their amplicon size, annealing temperature, and usage .....	72
Table 4.1: Solar radiation (umol/sec/m <sup>2</sup> ) measured in CF 2021 control and shading plots under various light conditions.....	78
Table 4.2: Flowering dates of all genotypes analysed in this study in Church Farm (CF) trials in 2021 and 2022. ....	78
Table 4.3: Statistical analysis of phenotypic data .....	79
Table 4.4: Number of florets pre-abortion (20 DPA) and the number of florets aborted from 20 DPA to 6 DPA abortion.....	82
Table 4.5: Sugar concentrations (µg/mg tissue weight) after shading treatments.....	83
Table 4.6: Statistical analysis of fructose and glucose concentrations (µg/mg tissue weight) .....	83
Table 4.7: Average development stage (Waddington) of floret 1-8 in basal (B) six and central (C) two spikelets .....	85
Table 4.8: T-test comparison of predicted versus actual grain values per spikelet.....	86
Table 4.9: Average development stage (Waddington) of florets <i>VRT-A2</i> NILs.....	95

## List of Abbreviations

AM	Axillary Meristem
CDF	Cumulative Distribution Function
CF	Church Farm
CLSM	Confocal Laser Scanning Microscopy
cv.	Cultivar
DPA	Days Pre Anthesis
DR	Double Ridge
GP	Glume Primordium
KAPS	Kompetitive allele specific PCR
LR	Leaf Ridge
LS	Linsmaier and Skoog
MS	Murashige and Skoog
NIL	Near Isogenic Line
PI	Propidium Iodide
RBS	Rudimentary Basal Spikelets
RNA-seq	RNA sequencing
SAM	Shoot Apical Meristem
SD	Standard Deviation
SE	Standard Error
SEM	Scanning Electron Microscopy
SR	Spikelet Ridge
ssp.	Subspecies
TEM	Transmission Electron Microscopy
W	Waddington
wMM	Wheat Meristem Media
WSC	Water-Soluble Carbohydrates

# 1 General Introduction

## 1.1 Increasing wheat productivity in the next decade

### 1.1.1 Pressure to increase yields while safeguarding nature

In 1982, Agnes Denes planted 285 furrows of wheat (*Triticum aestivum* L.) on a landfill site besides the Hudson River, between the Statue of Liberty and Wall Street (Figure 1.1). This was possibly the first wheat cultivation in Manhattan in over 300 years (<https://www.architecturaldigest.com/story/agnes-denes-prophetic-wheatfield-remains-as-relevant-as-ever>). Agnes Denes initiated this project to raise attention to “*mismanagement, waste, world hunger and ecological concerns*” (<http://www.agnesdenesstudio.com/works7.html>). Denes’s *Wheatfield* is a pioneering piece of environmental land art and reminds us today that the ecological cost of human existence on the planet was already well known over 40 years ago. The work has been praised for its gloomy prediction of the future of our existence, and its’ surreal beauty (<https://www.architecturaldigest.com/story/agnes-denes-prophetic-wheatfield-remains-as-relevant-as-ever>). But why did Danse choose to plant wheat, rather than any other crop that is traded only two blocks away at Wall Street? The global wheat trade is enormous, with 150 million tons of wheat being exported annually. In 2018 an estimated 725 million tonnes of bread wheat were produced globally (FAO, 2018), making it one of the most important staple food crops. It was also one of the earliest domesticated crops and is today cultivated on more land area across the globe than any other (FAO, 2018). Besides its high market share and global economic importance, wheat is a major source of calories and nutrients, as well as protein, in developing countries, where wheat feeds around 1.2 billion people who live on less than US\$ 2 a day (<https://www.cimmyt.org/blogs/wheat-is-not-a-rich-mans-crop/>). With a growing global population and increasing negative effects of climate change on environmental conditions, one of the biggest challenges facing humanity is to feed the growing population while safeguarding nature. Thus, Danse’s choice of placing wheat at the heart of her humanitarian and environmental land art piece seems very fitting.

Today, research and breeding efforts are focused on further improving the yield of wheat to meet the 60% higher demand predicted by 2050 (Godfray et al., 2010). Having more efficient crops, reducing the amount of harvest waste and maximising the output of the energy spent on food production will be three key areas to make future farming more sustainable. More efficient and resilient wheat varieties have been bred over the last century (e.g., Sanchez-Garcia et al. (2013). Major advances in wheat yields, and yield stability, were achieved through increasing genetic resistance to pathogens, resilience to environmental stress, and improving the potential yield of a cultivar. Potential yields can be improved through a variety of targets, at the beginning of the last

century the negative selection of yield reducing traits was a very successful strategy (reduced shattering, height, lodging; Fischer (2022)). Alternatively, the ability of plants to produce energy (source traits such as photosynthetic activity, nitrogen use efficiency) or their ability to translate this produced energy into yield (increased harvest index, increased spike size, tiller numbers) have been targeted in breeding programs (Fischer, 2022). One successful approach to increase yield potential has been the alteration of the wheat plants' architecture, also known as sink strength alterations (Koppolu and Schnurbusch, 2019). Throughout domestication, humans have significantly altered the structure of the wheat spike with respect to the wild progenitors. Modern varieties have an increased number of grains per spike and grains are bigger. Improving sink traits was believed to be most beneficial to high input environments with optimum management, such as the UK, France, and New Zealand (Reynolds et al., 2005). However, increasing the maximum achievable yield in varieties grown mostly in low yield environments can also be desirable. Firstly, increasing sink strength under all conditions can be advantageous and secondly, this will lead to increased yields in the rarer, high-yielding years within these environments.

Thus, improved genetic material is a fruitful avenue to improve the sustainability of future wheat production. However, this is only the case if the improved material is readily available and introduced to all markets globally. One important focus of future wheat yield improvements must be to close the yield gap between high and low producing areas globally and to strengthen self-sufficiency in production in countries, or sub-regions, rather than relying on global trade, which has been proven to be fragile and create dependencies. Only recently, the invasion of Ukraine by Russian forces essentially stopped all wheat exports from this major exporting country and reminded us of the dangers of relying on imports, especially for developing nations (Bentley, 2022).



Figure 1.1: Agnes Denes, *Wheatfield – A Confrontation: Battery Park Landfill, Downtown Manhattan – The Harvest*, 1982. © Agnes Denes. Courtesy of the artist and Leslie Tonkonow Artworks + Projects, New York. Source: <https://www.artsy.net/article/artsy-editorial-agnes-deness-manhattan-wheatfield-grown-poignant>

#### 1.1.2 Role of wheat breeding and fundamental research for future yield gains

Whilst large improvements in yield were achieved during the Green Revolution, rates of increase in crop yields have slowed in recent years (Ray et al., 2012). Breeding programs have achieved past yield gains through the incremental improvement of the genetic composition of cultivars, relying on the power of crossing existing material and subsequent selection of the new. Through the crosses new combinations of genetic diversity are created, either by introduction of new alleles or the novel recombination of existing alleles in the program. Subsequent selection largely relies on phenotypic screening in various conditions or by targeted screening for certain traits, such as resistance through the application of pests in the field. A wide range of strategies have been developed to increase the efficiency of the selection pipeline, reducing the time and resources it takes to produce new cultivars (Reynolds and Braun, 2022). The challenge for breeders is that they need to improve multiple traits, which have large and complex underlying genetic networks that are rarely controlled by a single gene.

In fundamental research, much effort has been spent on elucidating the effect of single genes on phenotypes, gene expression, metabolism, or cell growth. Therefore, it might be difficult to see how fundamental research can help breeding. Introgression of a single gene to improve one targeted trait has often failed to achieve real-world yield gains. The exception to this rule might have been the introgression of the semi-dwarfing *REDUCED HEIGHT (RHT)* alleles during the Green

Revolution, which achieved astonishing yield gains by efficiently reducing plant heights in all genetic backgrounds (Khush, 2001).

Much of the general knowledge and techniques developed in fundamental research have been applied in modern breeding. For example, genomic selection, in addition to classic phenotypic selection, is today an important part of most breeding programs (Reynolds and Braun, 2022). The use of marker-assisted selection improved the efficiency of backcrossing alleles that are recessive, epistatic, or affecting traits that cannot be easily measured on a single plant basis (Reynolds and Braun, 2022). More recently, the development of cheap genotyping marker arrays, SNP detection technology (e.g., KASP), and functional markers has further helped the adaption of genomic selection in wheat breeding (Adamski et al., 2020). More recently and enabled by the decreasing cost of genome sequencing, breeding programs are in the process of incorporating whole genome sequencing approaches into their programs. The progress in marker/genomic selection in breeding programs was only possible due to the achievements of fundamental research that improved genomic sequencing, provided reference genomes, and developed mathematical models for genomic selections. However, many of these new tools were also achieved by collaborations between fundamental and applied researchers, suggesting that the question of what is more beneficial, applied or fundamental research, should maybe be dismissed altogether. Breeding programs are the driving force of future yield gains, but improvements made to the breeding programs themselves will equally rely on fundamental research.

However, the increasing ability to perform fundamental research in the crop species directly, rather than using a model species, has had multiple advantages. The plant architecture of wheat is difficult to compare to that of model species, such as *Arabidopsis thaliana*, and genetic mechanisms have been lost, changed, or been neo-functionalised since the species last common ancestor. For example, the very important leaf ridge inhibiting effect of the *THIRD OUTER GLUME1 (TRD1)* gene in the grass inflorescence is not shared with its ortholog function in *Arabidopsis* (Whipple et al., 2010). The availability of vital genetic tools, such as a reference genome and robust gene annotations, in crops has been a game-changer for research in crops and has allowed the further applications of tools developed in model species directly in crops. The fluent exchange of tools and knowledge between model organisms, fundamental crop research, and breeding is what drives a fruitful future for wheat genetics, which remains a keystone for the development of more sustainable and efficient agricultural systems.

## 1.2 Architecture and development of the wheat spike

### 1.2.1 Evolution of the grass inflorescence

To improve the grass inflorescence, a substantial understanding of its evolution, architecture, and development has been established. Many morphological distinctions among grass taxa are based on characters of the inflorescence. A unifying feature of the taxa is that the flowers (termed florets in grasses), and thus later the seeds, are produced in spikelets. Spikelets are the flowering units of grasses, composed of glume, lemma, palea, lodicules, stamens, and pistil (Malcomber et al., 2006). Spiral phyllotaxis of spikelet arrangement is widespread in the grasses, and most likely the ancestral arrangement, while two-ranked inflorescences most likely derived independently in *Centrolepidaceae* and the PACMAD clade (Kellogg et al., 2013). Two ranked primary branches initiated at angles of 180° are termed distichous. Overall, grass spikelet architecture is highly diverse and botanical features (such as phyllotaxis or symmetry) are often not consistent with phylogenetic relationships. Phyllotaxis of the spikelets can be the same as the patterning of leaves or it may change (Kellogg et al., 2013). Spikelet determinacy does not always match spike determinacy and even though most grass florets are hermaphroditic, they can also be monoecious (for example maize) or dioecious, albeit the later is very rare among grasses (Connor, 1979).

A robust phylogeny is available for the grass family and provides the basis for the classification of the ca. 10,000 species in 700 genera (Grass Phylogeny Working Group et al., 2001; Sajo et al., 2007). The closest relative to the taxa is the subfamily *Anomochlooideae*, which lacks true spikelets (Clark and Judziewicz, 1996). The inflorescence meristem of members of the *Anomochlooideae* family produces primary branches that are arranged either in a spiral (for example *Streptochaeta*) or distichous (for example *Anomochloa*) pattern (Judziewicz and Soderstrom, 1989; Sajo et al., 2008). The *Streptochaeta* spikelet seems to be morphologically an intermediate between the true grass spikelet and the reproductive units of grass relatives (Sajo et al., 2008). The *Pharoideae* subfamily is the earliest lineage with a true spikelet and is sister to all remaining spikelet-bearing grasses (Kellogg et al., 2013). The spikelets of the species *Pharus* that belongs to this subfamily are directly attached to the rachis and are arranged spirally (Sajo et al., 2007). *Pharus* spikelets are furthermore monoecious and produce only one floret per spikelet (Sajo et al., 2007). It appears that after the divergence of *Pharoideae*, multiflowered spikelets arose (Kellogg et al., 2013). However, there is quite a large diversity in the number of flowers per spikelet, which can be determined and indetermined in grasses. The earliest-diverging lineage of this group is *Puelioideae*, for which no spikelet arrangement data is available (Kellogg et al., 2013). The major radiation of the grasses occurred much later, perhaps 40–50 million years ago, when the BEP clade (*Bambusoideae*, *Ehrhartoideae* (syn *Oryzoideae*), and *Pooideae* (i.e. wheat)) diverged and a lineage leading to the PACMAD clade originated (Kellogg et al., 2013).

The PACMAD clade includes the remaining subfamilies, which make up about 60% of the grass species (*Panicoideae*, *Arundinoideae*, *Chloridoideae*, *Micrairoideae*, and *Danthonioideae*) (Malcomber et al., 2006). Most species in the PACMAD have spiral phyllotaxis and lack terminal spikelets, which is also the case for maize (*Zea mays*), one of the best studied species of the clade (Kellogg et al., 2013). Within the *Pooideae*, distichous inflorescence branching appeared after the divergence of *Nardus*. Based on current data, the shift in inflorescence phyllotaxis from spiral to distichous appears to have occurred at about the same time as the expansion of genome size and an expansion to cooler, temperate climates that characterises *Triticeae*, *Bromeae*, and *Poeae* (Kellogg et al., 2013). However, causative links between the shift to distichous branching and either event can only be hypothesised, but not proven.

As mentioned in the previous paragraph, many characteristic traits are not consistent with phylogeny in the grasses or within the *Triticeae*, for example, morphology, life cycle, reproductive behaviour, ploidy level, and genomic constitution may vary widely (Barkworth and Bothmer, 2009; Laurie, 2009). The important crops *T. aestivum* (bread wheat), *T. durum* (pasta wheat), *Hordeum vulgare* (barley), *Secale cereale* (rye) and *Triticosecale* (triticale) belong to the *Triticeae* and their inflorescence morphology has been studied in great detail (Love, 1984). The inflorescence of the *Triticeae* is generally termed spike, with spikelets directly attached to the rachis or with only very short branches and no formation of secondary branches. Furthermore, the spikes of the above-mentioned crops are all distichous, however wheat and rye form a determinate spike while barley is indeterminate.

Bread wheat has a complex evolutionary background with three polyploidization events leading to its allohexaploid genome structure (IWGSC et al., 2018). The diploid relative of the species *T. urartu* (AA) and an unknown species related to *Aegilops speltoides* (BB) hybridised and formed the allotetraploid *T. turgidum* ssp. *dicoccoides* (AABB; wild emmer), an ancestor of modern-day pasta wheat. A second hybridization event between *T. turgidum* and the diploid grass species, *Ae. tauschii* (DD), produced the allohexaploid *T. aestivum* (AABBDD) (IWGSC et al., 2018). Even though no detailed studies of the wheat ancestor spike morphology exist, it appears that *Aegilops* and *T. urartu* spikes are already distichous and form terminal spikelets. However, the spikelets of both are much smaller and form less grain.

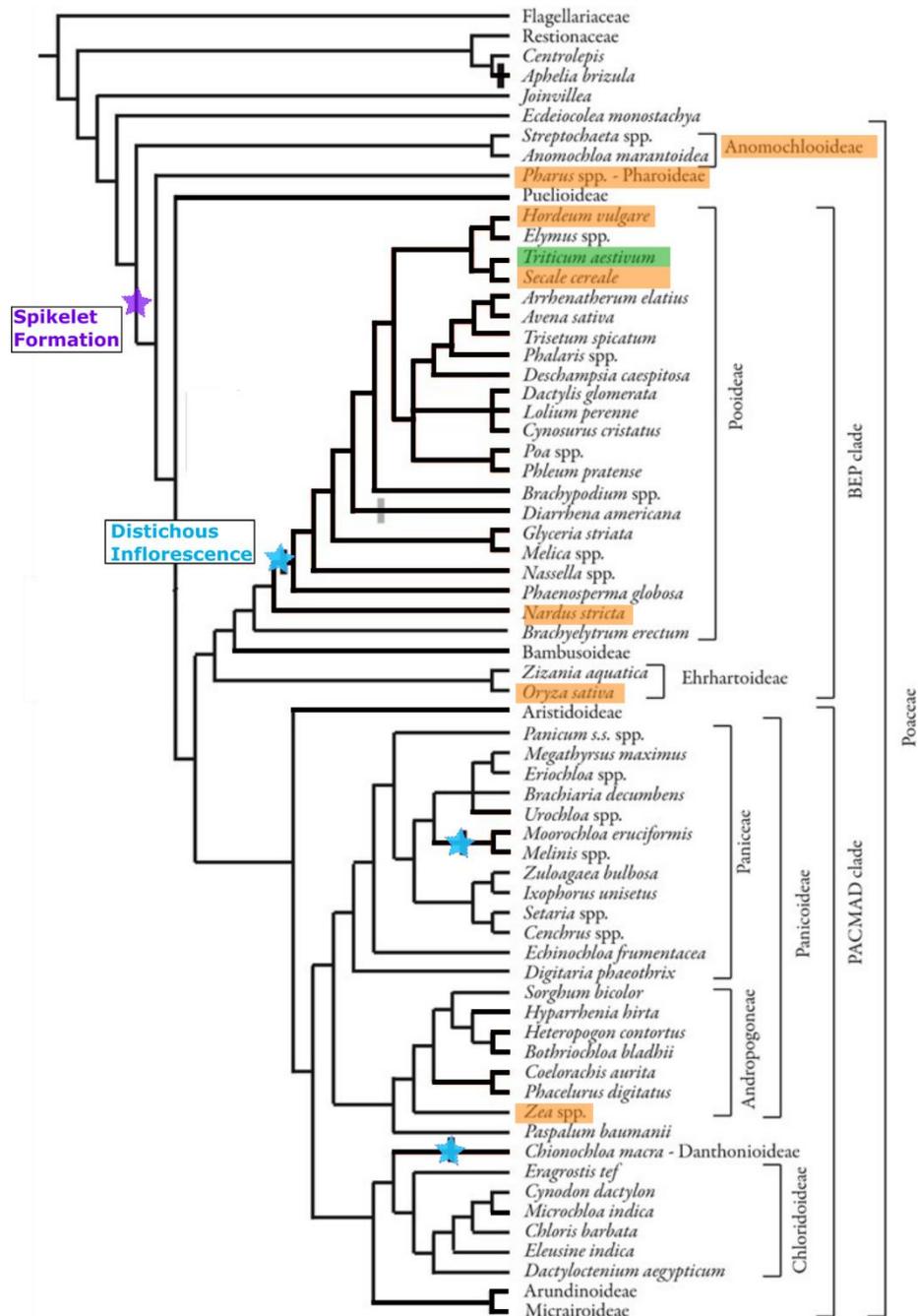


Figure 1.2: Adapted from (Kellogg et al., 2013). Phylogenetic distribution of inflorescence phyllotaxis. Species relationships based on (Grass Phylogeny Working Group et al., 2001). Blue stars indicate the origin of distichous phyllotaxis; vertical grey bar indicates reversion to non-distichous two-ranked phyllotaxis. Purple star indicates origin of true spikelet formation within the grasses. Orange boxes highlight species mentioned within the text specifically, green box indicates wheat (*Triticum aestivum* L.).

### 1.2.2 Architecture of the wheat and cereal crops inflorescence

For plants, the arrangement and presentation of seeds is of vital importance for their evolutionary success. To maximise numbers of seeds, and seed dispersal, plants have evolved an impressive suite of floral architectures (Doyle and Gaut, 2000). During domestication, we have further altered the architecture of crop species to improve yield and ease of harvest (Doebley et al., 2006). For example in wheat, the number of seeds per spike (Philipp et al., 2018), as well as the threshability of the spike (Debernardi et al., 2020), have been largely improved during domestication.

The *Poaceae* have the distinguishing feature of arranging their flowering structures in spikelets along a vertically elongated inflorescence (Malcomber et al., 2006). Species within the *Triticeae* tribe (including wheat) form sessile spikelets directly attached to the rachis and this complete structure is commonly referred to as a spike. In contrast, the spikelets of rice are formed after the initiation of primary and secondary branches and the result of this increased branching structure is a more complex inflorescence, commonly termed panicle (Bommert et al., 2005). However, both terms are borrowed terminology from dicots and not entirely accurate (Kellogg, 2022). The spikelets are themselves inflorescences, rather than flowers as in dicots, and the grass inflorescence is in fact a compound structure. Technically, the spikelets are spike-like structures and the correct term for wheat spikes is synflorescence (Weberling, 1992; Vegetti and Weberling, 1996). However, in this thesis, we refer to the wheat synflorescence by its more common name, spike. The important difference is that spikelets are not determinate structures and that the glume, lemma and palea are not homologs of sepals or petals. The spike-like identity of wheat spikelets is especially noticeable when considering how readily sham ramification can be observed in wheat mutants. Sham ramification refers to the development of spikes, rather than spikelets, in the lower rachis nodes as it is the case for example in 'Miracle Wheat' (Dobrovolskaya et al., 2015; Wolde et al., 2019).

The floral structure of wheat (and more generally all grasses) is termed a floret. In wheat, an indeterminate number of florets is initiated per spikelet, of which typically only the basal 3-5 become fertile and set grain. Each wheat floret contains one carpel and three stamens but lacks petal and sepal structures (like all grasses) (Ciaffi et al., 2011). All grasses form only one ovule per carpel (unlike *Arabidopsis*) (Shen et al., 2021). The carpel of rice, barley, and wheat are very similar in that their carpel is derived from the early fusion of three primordia, with two partially fused styles and two feathery stigmas covered with papillae cells where pollen is deposited (Kellogg, 2001; Rudall et al., 2005; Dreni et al., 2013; Shen et al., 2021). The reproductive organs are surrounded by two lodicules, these swell during anthesis forcing the lemma and palea apart to facilitate pollination of the stigma from the dehisced anther. There is evidence that the lodicules in grasses are the result of modified petals in dicots (such as *Arabidopsis*), which is supported by the previously reported control of lodicule development by classic petal development genes of *Arabidopsis* (Yoshida, 2012).

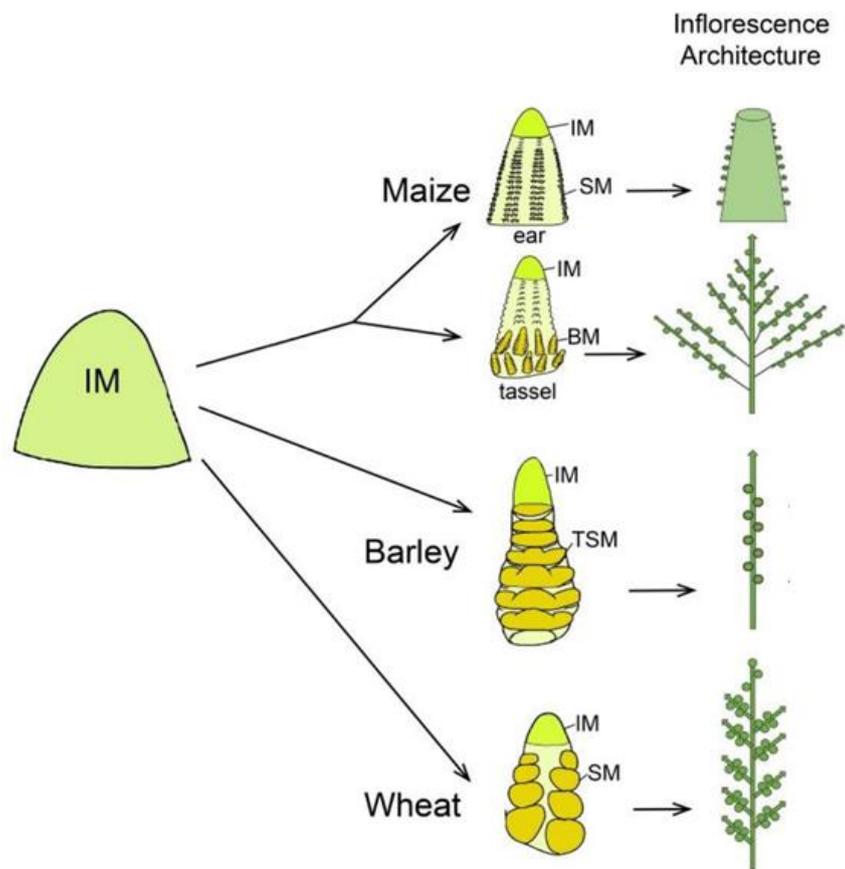


Figure 1.3: Adapted from (Wang et al., 2021). Schematic representation to compare inflorescence meristem (IM) differentiated fate and inflorescence architecture in maize, barley and wheat. In maize, the IM of a tassel is converted from shoot apical meristem, while the axillary meristem converts into an ear. The ear IM initiates a series of determinate axillary meristems (spikelet meristems; SM), giving rise to pairs of spikelets. The tassel produces branches from the branch meristems (BM), which then form pairs of spikelets. Each spikelet of ear and tassel further initiates two floret meristems. In barley and wheat, the axillary meristems formed by the IM directly differentiate into spikelet meristems. Barley has a triple spikelet meristem (TSM) structure composed of a central spikelet and two lateral spikelets, whose development is either suppressed to form a two-rowed type or promoted to form a six-rowed type. Conversely, in wheat, the inflorescence is composed of single spikelet that produce multiple floral meristems.

Elaborate work in *Arabidopsis* and *Antirrhinum* mutants led to the establishment of the “ABCDE” flowering model in eudicots (Coen and Meyerowitz, 1991; Theissen and Saedler, 2001). The model illustrates how the combined functions of five classes of homeotic genes determine the organ identities of the four floral whorls (Ciaffi et al., 2011). Class A genes specify sepal identity in the first whorl while the combined action of class A and B genes control the petal identity in the second whorl. Class B and C activities control the stamen identity and the class C activity alone the formation of carpels in the fourth whorl (Ciaffi et al., 2011). The later discovered D genes also control ovule development. The E class homeotic gene functions are required for sepal, petal, stamen, and carpel development and are provided by the *SEPALLATA* genes, a set of four similar and redundant MADS-box genes (Ciaffi et al., 2011). Analysis of mutants of the homeotic flowering genes in grasses (mainly maize and rice) has been used to elucidate the conservation of the ABCDE model in monocots. Overall, the same principles apply and the function of multiple of the core flowering genes seems to be conserved in grasses. However, multiple genome and gene

multiplications, followed by mutation and diversification of the sequences, has led to a wealth of novel flowering genes in the grasses, allowing for neo- and sub-functionalization. Furthermore, the interpretation of functional conservation is complicated by the differences in inflorescence and floral structures between grasses and *Arabidopsis*, making it difficult to identify equivalent tissues between them. For example, the reproductive organs of grasses are surrounded by lemma and palea, whose correspondence to eudicot organs remains controversial (Ciaffi et al., 2011) and will be discussed in more detail in the following section.

In wheat (and *Brachypodium*), multiple florets are initiated along the spikelet rachilla, whose tip is indeterminate and, like in all grasses, each spikelet is subtended by a pair of glumes. Glumes are the bract tissue in which axil florets are formed and they can be large and encompass much of the florets, as in wheat, or they can be severely reduced structures at the base of the floret, as in rice (Ciaffi et al., 2011). The number and development of florets within spikelets is highly diverse between grass species. Unlike wheat, barley forms determinate spikelets with strictly one floret per spikelet (of which either one or three are formed per node, depending on row-type) (Sakuma and Schnurbusch, 2020). Maize forms two florets per spikelet in both the tassel and ear, however the second floret is aborted in the latter (Weatherwax, 1923). Rice also forms only one floret per spikelet, and as it only develops reduced glumes, it could be hypothesised that the lemma and palea partially fulfil the functional role of the glume pair.

Whether the inflorescence terminates in a spikelet or is indeterminate also varies among grass species. The inflorescence ends as an undifferentiated dome in maize, rice, barley, signal grass (*Brachiaria decumbens*), and finger millet (*Eleusine coracana*), whereas it ultimately becomes a spikelet in wheat, ryegrass (*Lolium* spp.), sorghum (*Sorghum bicolor*), and oats (*Avena sativa*; (Kellogg, 2022)). Terminal spikelet formation is conserved within species and is not altered by environmental conditions (Kellogg, 2022). The formation of the spikelets along the inflorescence, florets within the spikelets, and the development of florets into grains is a highly controlled and genetically conserved process that largely controls the yield potential of a plant. The inflorescence architecture among the *Triticeae* bears many similarities, even though key features of the spike may vary and impact how yields per spike can be increased. In wheat, the increase in number of grains per spikelet played a major role during wheat domestication and modern breeding (Philipp et al., 2018). In barley it was mainly the increase in spikelet number and lateral spikelet fertility that led to increased yields per spike rather than the number of grains per spikelet as the later has little plasticity in barley. This suggests that breeding has in both (barley and wheat) mainly achieved to increase the survival rate of already initiated primordia, rather than finding ways to form new primordia from, or next to, original meristems.

### 1.2.3 Developmental mechanisms underlying the wheat spike architecture

The architecture of plant inflorescences has fascinated botanists for many centuries and is a distinguishing feature for the taxonomic classification of species. Over the last century, a new fascination arose: What are the developmental and genetic mechanisms that lead to the final shape of the inflorescence?

In flowering plants, all above-ground organs are formed from the shoot apical meristem. After germination, the apical meristem (also termed apex) continuously forms new organs, termed primordia, on its flanks while maintaining a zone of undifferentiated, pluripotent stem cells in the central zone (Figure 1.3). Shoot development occurs in repeating modules called phytomers, consisting of a leaf and axillary meristem (AM) primordia and an internode. Plant architecture is largely dictated by the activity of the meristem and the determinacy of the initiated primordia (Pautler et al., 2013).

In the wheat seed the shoot apex has already initiated three to four leaf primordia, and the tiller buds of the coleoptile and first leaf can be seen. After imbibition, the coleoptile grows until it is just above the soil surface after which the first leaf emerges from it (Kirby and Appleyard, 1987). Upon germination this apex resumes activity and further phytomers are initiated. During the first phase, the wheat apex is vegetative and has a broad-based conical shape (Kirby and Appleyard, 1987). The vegetative apex is located very close to the base of the plant, is mostly below ground and covered by the emerging leaves. During the early vegetative growth, the phytomers are forming fully developed leaves and the AMs are of shoot identity, forming the tillers of the plant.

Leaf primordia appear as ridges of tissue on the flank of the shoot apex and at a constant rate that is temperature and photoperiod dependent. The primordia appear first as a ridge and then extend around the apex, until eventually its margins overlap. The leaf primordium is then growing upwards, forming a collar-like structure, enclosing the apex and the later-initiated primordia. The number of leaves on the main shoot generally varies from about 10 to 14 leaves, which is however temperature and photoperiod dependent (Kirby and Appleyard, 1987).

An AM of tiller identity is initiated in the axils of all leaves as they are formed. Primordia are usually positioned adjacent to the overlapping margin of the subtended leaf. As the tiller meristems grows, the prophyll is initiated on its flanks and subsequent tiller development resembles that of the main shoot (Kirby and Appleyard, 1987). The further development of the tillers depends upon the position of the tiller along the plant and the environment. If unfavourable, growth quickly slows and stops, and the bud does not grow to a length of more than 2-3 mm. Generally, more tiller primordia are initiated than the plant can support and only the first 1-4 tillers are fully developed in the mature plant of modern wheat. Tiller survival is strongly controlled by environmental factors, such as resource availability and planting density (Shang et al., 2021). During this phase, the internodes between the phytomers are very short and only marginally growing (Kirby and

Appleyard, 1987). Thus, in the field wheat plants are very short and of grassy appearance for most of the season.

After the vegetative phase, the apex transitions into the reproductive phase, which can be first recognised by the elongation of the apex and its more cylindrical shape (Kirby and Appleyard, 1987). Primordia pairs (termed double ridges) are initiated on the flanks of the apex in a distichous pattern and with slightly longer internodes separating them (Bonnett, 1966). A comprehensive review of the known genetic gradients that are underlying this gradual transition in grasses has been recently published by Kellogg (2022). In terms of phytomers, the transitioning apex is still initiating a leaf ridge, an AM and an internode, which is slightly longer than during the vegetative phase (the spike internodes are collectively termed the rachis). However, the lower primordia ridge, the leaf, is now suppressed while the AMs are activated, and their identity has switched from shoot/tiller to spikelet primordia (Figure 1.3). The suppression of the lower/leaf ridge has been shown to be mainly facilitated by the expression of a GATA domain zinc-finger transcription factor across all grasses (*TASSELSHEATH1* (*TSH1*) and *TASSELSHEATH4* (*TSH4*) in maize (Whipple et al., 2010), *TRD1* in barley (Houston et al., 2012) and *NECKLEAF1* (*NL1*) in rice (Wang et al., 2009)). We will further discuss these paralogous genes in Chapter 3. The apex at the tip of the developing spike continues to initiate new phytomers until molecular signals initiate the formation of the terminal spikelet, thus determining the potential number of spikelets early on in development. Li et al. (2019) recently showed that the major vernalisation and flowering time controlling genes *VERNALIZATION1* (*VRN1*) and *FRUITFULL2* (*FUL2*) affect terminal spikelet formation. In the single mutants (*vrn1* or *ful2*) the transition to terminal spikelet formation was delayed, whereas in the *vrn1ful2* double mutant the inflorescence remained indeterminate and failed to produce a terminal spikelet. *VRN1* and *FUL2* belong to some of the best studied flowering genes in wheat. They belong to the *SQUAMOSA*-clade of MADS box transcription factors and their sequence is most similar to the A-class flowering genes *APETALA1* (*AP1*)/*FRUITFULL* (*FUL*) of *Arabidopsis* (Murai et al., 2003; Trevaskis et al., 2003). In complete loss-of-function mutant in *Arabidopsis*, no flowers are produced, and the meristem reiterates the development of leafy shoots (Ferrándiz et al., 2000).

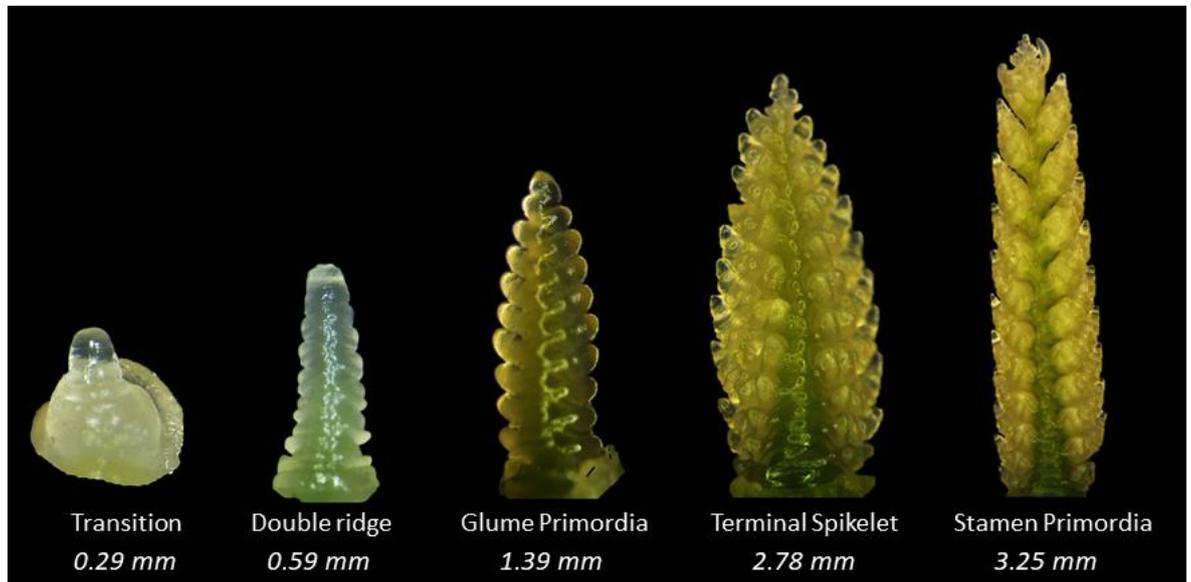


Figure 1.4: Light Microscopy images of the floral development of the wheat spike. Developmental stage according to Waddington et al. (1983). Meristems are not to size, actual size of the meristem from tip to collar ridge is indicated below each spike in millimetres.

Each initiated spikelet continues its developmental trajectory, forming next the glume, followed by lemma, and floret primordia. Each initiated floret primordia first develops palea and lodicules, followed by three stamen primordia, and shortly after the carpel. During the formation of primordia, the spike is only marginally increasing in overall length. Shortly after, however, the spike as well as the internodes below the spike are entering a phase of rapid growth during which excessive floret and spikelet primordia are aborted to adapt final spike size to resource availability. This growth phase of the plant takes place about 10-20 days pre-anthesis and is often termed the ‘critical phase’ of yield potential (Fischer, 1985; Siddique et al., 1989; Savin and Slafer, 1991). We will discuss this phenomenon in greater detail in Chapter 4.

Spikelet differentiation starts in the central spikelets, thus creating inhomogeneous development across the spike. The wheat spikelet is indeterminate and new lemma and floret primordia are initiated continuously at its tip. The most basal florets within each spikelet will begin to differentiate immediately after formation and therefore within each spikelet a gradient of development exists from the base to the tip. It appears that most primordia are developing at similar rates within the spike and therefore the later initiated primordia (at the tip of the spike and spikelet) will lag in development and cannot catch up with their central counterparts. Consistent with this observation, flowering also begins in the centre of the spike. As mentioned in the previous section, the identity of some of these organs is disputed and what marks the beginning of floral development in the wheat spike is not entirely clear. Considering the wheat spike development and mutant phenotypes in terms of phytomers favours the concept that floret primordia formation marks the beginning of true floral development in wheat and not the formation of double ridges. Spikelet primordia are subtended by two glumes, which have suppressed axillary meristems in their axils. In all grasses studied, these AMs are suppressed by *BRANCHED SILKLESS1* (Kellogg, 2022). Thus, the glumes are

also fitting the typical phytomer structure of leaf (modified to glume) plus AM. Next, the spikelet apical meristem produces lemma primordia, in which axils the AMs identity is truly altered to be determinate, floral meristems. Hence making the lemma a modified leaf structure and the palea, lodicules, stamens and a pistil the wheat floral organs developed in the four floral whorls, respectively (Yoshida, 2012).

Examining a large body of data from mutants of homeotic genes mainly in rice, Lombardo & Yoshida (2015) concluded that the lemma and palea of grasses are likely to be sepal equivalents. Alternative interpretations have found the lodicules to be the reduced petals, as the expression of conserved B-class function genes in maize is restricted to the stamen and perianth (Whipple et al., 2004), and lemma to be bracts and palea as modified sepals. In this thesis, we follow the latter interpretation, based on our observations of MADS-box flowering gene mutant effects on the spike, as well as the interpretation of *PIN1* expression patterns in *Brachypodium* spikelets.

In all plants, local auxin maxima precede the initiation of new branching points (Reinhardt et al., 2003). O'Connor et al. (2014) traced the expression patterns of *PIN-FORMED1* (*PIN1a*, *PIN1b*), *SISTER OF PIN1* (*SoPIN*) and *DR5* in the *Brachypodium* spikelet using fluorescence reporters. The micrographs reveal how the lemma is subtending each floret primordia and how *PIN1a* and *PIN1b* expression differs between the two primordia. *PIN1a* is exclusively expressed in the lemma, whereas *PIN1b* is expressed both in the lemma and in the older floret primordia. This leads to *DR5* signals being mostly/strongest in the lemma and rachis (O'Connor et al., 2014). The results of the study show that the lemma is initiated before and outside of the floret primordium.

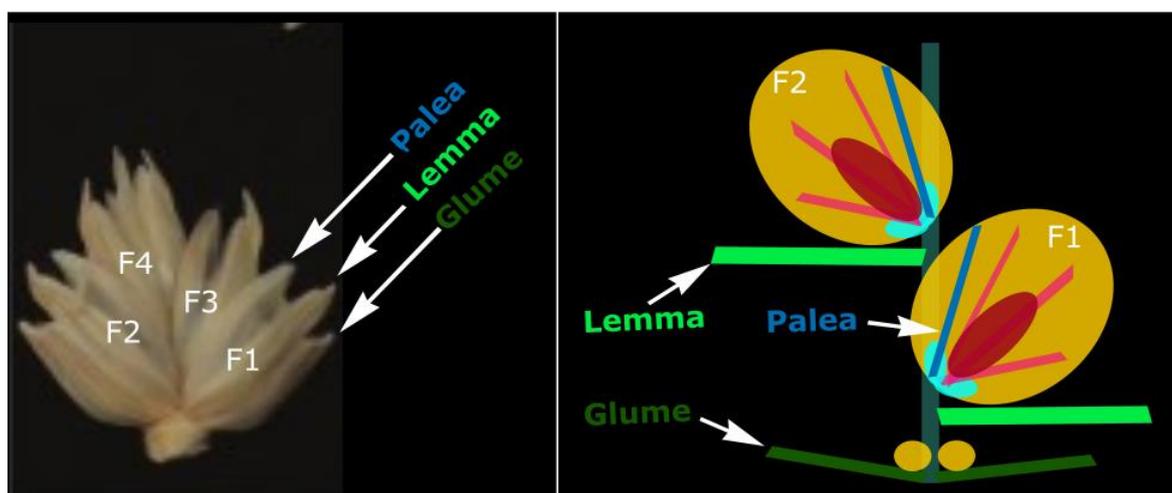


Figure 1.5: The wheat spikelet structure. Left, image of wheat spikelet adapted from Adamski et al. (2021), arrows indicate the right glume and the lemma and palea of the first floret (F1). Right, graphical representation of the wheat spikelet. At the bottom, the two glumes (dark green) are subtending a suppressed axillary meristem (yellow). Above, the first floret is subtended by the lemma (light green), which is outside the floral meristem structure (yellow background). Inside the floral meristem are palea (dark blue), lodicules (light blue), stamen (red), and the carpel (dark red).

In the full *SQUAMOSA*-clade loss-of-function mutant in wheat (*vrn1ful2ful3*), the spike axillary meristems that would develop into spikelets in the wildtype generated fully vegetative tillers and

the leaf ridges were de-repressed and formed leaves (Li et al., 2019). The phenotypic effect is thus similar to the one reported in *Arabidopsis* loss of function mutant of the orthologous genes (Ferrández et al., 2000). In a subsequent publication, Li et al. (2021) showed that in the *vrn1ful2* double mutant, the expression of the MADS-box transcription factors belonging to the *SHORT VEGETATIVE PHASE (SVP)* clade were significantly increased. Fascinatingly, the loss-of-function *svp1vrt2* double mutant developed axillary spikelets or spikes subtended by leaves in the nodes of the elongating stem below the peduncle (Li et al., 2021). The constitutive over-expression of *VRT2* on the other hand led to a reversion of basal spikelets to spike-like structures that form terminal spikelet structures, branched spikelets and severely elongated glume and lemma (Liu et al., 2021). This phenotypic effect can be observed in a variety of mutants and hints at the synflorescence structure of the wheat spike, in which spikelets are merely modified spikes and readily revert back to spike identity, especially in the base of the spike (see *FRIZZY PANICLE* mutant 'Miracle Wheat', (Dobrovolskaya et al., 2015; Poursarebani et al., 2015)). The introgression of a novel SVP allele from *Triticum aestivum* ssp. *polonicum* was reported to increase glume and lemma length significantly, albeit palea length only marginally (Adamski et al., 2021). This divergence in the SVP effect on glume/lemma versus palea further suggests that glume and lemma are controlled by a different genetic network than the floral structures. In summary, the phenotypic effects of the SVP-clade genes affect axillary meristem determinacy across the plant and also growth of the subtending leaves, but not that of floral tissues (i.e., glume and lemma are affected but not palea).

The function of SVPs in wheat appears to be in line with their bud-suppressing activity in trees, in which their expression is downregulated during spring (Falavigna et al., 2019). In wheat, their expression during vegetative development is also responsible for the suppression of axillary buds, however they must be downregulated upon inflorescence formation to allow for the axillary meristems of the spike to adopt spikelet identity. The control and function of SVPs will play a central role in Chapter 2 of this thesis. The effect of these genes on AMs, leaf ridge de-repression and glume and lemma length also support the idea that these organs are not developed from floral meristems and are thus controlled to some extent independently of floral meristem formation.

### 1.3 What we know about rudimentary basal spikelets

#### 1.3.1 Lanceolate shape of the wheat inflorescence

Across cereal inflorescences not all grains are of equal size. In wheat, the most basal and apical spikelets of the inflorescence are smaller and produce less grain than central spikelets. This bi-directional gradient in spikelet fertility is also the cause of the typical lanceolate shape of the wheat spike (Bonnett, 1966). The gradient in spikelet productivity from the centre to the tip can be explained by the later formation of the apical spikelets and is thought to bear evolutionary

advantages. Flowers are sensitive to abiotic stress during development, especially during anthesis (Thakur et al., 2010). Having a gradual development means that short-termed, adverse, abiotic conditions won't affect all flowers and thus increase the chance of having some viable off-spring. Furthermore, the reduced amount of vascular bundles, and thus nutrients, reaching the tip of the spike has also been hypothesised to cause a reduction in apical spikelet size (Whingwiri et al., 1981).

However, the reduced fertility of the basal spikelets cannot be explained by the same factors. The basal spikelets are initiated first and have a high number of vascular bundles reaching them (Whingwiri et al., 1981). Despite this, the most basal spikelets are often very small and produce no grain at all (termed rudimentary spikelets) (Bonnett, 1966). In 1936, Bonnett observed that basal spikelets lack behind their central counter parts in development just after being initiated. The formation of glume, lemma and florets are all delayed in the basal spikelets compared to central. However, the causes of the differences in development between the basal and central spikelets remain unknown.

### 1.3.2 Rudimentary basal spikelets in the literature

Previously, only a few studies have addressed rudimentary basal spikelet formation. In 2015, Guo and colleagues performed a genome wide association study (GWAS) in wheat and recorded the fertility of apical, central and basal spikelets separately, alongside many other traits. They found that multiple loci identified for grain number and floret survival in the central spikelets are shared with basal spikelets and focused subsequently more on differences between central and apical spikelets (Guo et al., 2017). Shi et al. (2018) performed a GWAS study using 212 wheat accessions, grown in five different environments and recorded the number of grains in the three apical spikelets and also in the three most basal spikelets. They found that grain number in basal spikelets was positively correlated with overall grains/spike. Furthermore, they found that the number of grains in basal spikelets can vary greatly between genotypes and environments, indicating a great potential for improving the fertility of basal spikelets (Shi et al., 2018). Similar results were obtained by a previous study that analysed apical and basal spikelet fertility in a diversity panel and found that both traits affect grains per spikelet (Guo et al., 2015). They also identified multiple genetic loci that were specifically related to basal, and not apical, spikelet fertility suggesting that different factors affect the two (Shi et al., 2018). Furthermore, it has been established that basal spikelets are readily aborted under reduced resource availability. For example, Stockman et al. (1983) found that shading pre-anthesis reduced number of grains per basal spikelets over-proportionately. The relationship between resource availability, floret fertility and spikelet survival will be discussed in more detail in Chapter 4.

Regardless of basal spikelet developmental differences not being mentioned specifically, their development is also uniquely affected in a variety of mutants. In barley, for example, the deletion of *BM1* can lead to the development of tillers at the base of spikes at positions where florets would

normally be located but not in the central or apical positions of the spike (Trevaskis et al., 2007). Shaw et al. (2013) noticed that in the late flowering genotypes, caused by loss of function of *PHOTOPERIOD1* (*PPD1*; A, B and D copy), the internode length of basal spikelets was extended and that sometimes the most basal spikelet formed bract leaves below the spikelets. An additional feature of the latest flowering genotypes was an increase in internode length at the base of the spike and, in some plants, the appearance of a leaf-like organ in place of a spikelet at the most basal spike internode (Shaw et al., 2013).

Debernardi et al. (2017) noticed that basal spikelets were differentially affected in mutants of the *Q* gene or in miRNA172 mutants. Both the constitutive overexpression of miR172 and the *Q*-5A loss-of-function mutants showed one to three rudimentary spikelets in the most basal nodes of the spike and empty lemmas in the fully formed basal spikelets (Debernardi et al., 2017). They furthermore were able to show the differences in phenotypic effects of the mutants between the base and apical section of the spike are due to the expression gradient of these genes across the spike. They found miR172 levels were highest in the base of the spike and decreased towards the tip while the opposite trend was observed for *Q* (Debernardi et al., 2017). This study thus introduced the important concept that genes might be gradually expressed across the spike and that molecular signals show variation between the different spikelets. This concept is further explored in Chapter 2 of this thesis and forms an important basis for our understanding of the formation of rudimentary basal spikelets.

## 1.4 Current state of techniques for developmental studies of wheat

### 1.4.1 RNA-seq for macro and micro dissected tissues

An important tool for the investigation of gene functions and genetic networks has been the commercial and relatively low-cost availability of RNA sequencing. In wheat, the use of transcriptomic technologies has lagged behind other crops for multiple years due its large, highly repetitive genome. While rice and maize genome assemblies have been available since 2005 and 2009, respectively, the first whole genome sequence assembly with reliable gene models for wheat was only released in 2018 (IWGSC et al., 2018). The availability of the genome enabled a suite of transcriptomic studies of the wheat spike. Previous to the release, RNA chips and RT-PCR could be used to gain some insight into the expression pattern of a sub-set of genes, one of the earliest studies in wheat used these technologies to investigate the grain transcriptome in 2004 (Altenbach and Kothari, 2004). In the 2010s, publications also used the available sequencing technologies in wheat to obtain RNA sequence data and to use this for reference-free analysis of differentially expressed genes. For example, Zhu et al. (2016) leveraged this technique to study the difference in expression in the young stems and spikes (at glume to stamen primordium stage) of dwarf and tall varieties (Zhu et al., 2016). In 2017, two studies were released that used the early draft genomes

for their wheat spike RNA-seq data analysis. Wang et al. (2017) sequenced the transcriptome of 90 wheat varieties at double ridge stage grown under field conditions (Wang et al., 2017). While Feng et al. (2017) used the newly available tools to sequence whole spikes collected at double ridge, floret meristem, anther primordium and terminal spikelet stage under controlled conditions, identifying important genes of floral meristem development and stamen formation in wheat (Feng et al., 2017). A comprehensive summary of available RNA-seq studies for different development conditions, stress responses, tissues and wheat stages has been developed by Borrill et al. in 2016, who published the wheat expression browser, making multiple different RNA-seq studies easily accessible, requiring no bioinformatic skills (Borrill et al., 2016). Since then, RNA-seq has mostly been used to identify differentially expressed genes between different genotypes, rather than for the analysis of temporal or spatial changes in expression throughout wheat spike development. For example, RNA-seq was used to compare gene expression in the inflorescences of wildtype and *GRAIN NUMBER INCREASE1 (GN1)* mutants (Sakuma et al., 2019) or to compare the spikes expression profiles of normal and multi-ovary genotypes by Guo et al. (2019).

In all these studies, whole spikes were used for the RNA extraction without any further dissections. This bears two main disadvantages: (i) multiple tissues, such as floret, glume, and rachis tissue are mixed together, and (ii) these same tissues are also mixed at multiple developmental timepoints as each spike bears many spikelets at different developmental stages at any given time. In mammalian studies, it is common practice to separate different organs before RNA-extraction to produce more meaningful RNA-seq data. However, this can be technically challenging for smaller tissues. In 1996, the invention of laser capture microdissections for human tissue overcame this bottleneck partially by making it possible to dissect smaller tissue sections, much more precisely than possible by hand, using lasers under the microscope (Emmert-Buck et al., 1996). Combined with downstream RNA-seq analysis, laser capture microdissections became a useful technology in mammalian and later also in plant and crop research. In barley, the technology was recently used to develop a very dense spatial and temporal expression atlas of the barley spike (Thiel et al., 2021).

An important accompanying development to the evolution of spatial tissue dissections was the development of ever improved low-input RNA sequencing. As we want to understand the differences between not only whole plant genotypes but different organs, tissues or even just cells, it becomes increasingly important to be able to sequence RNA from as little input material as possible (Macaulay et al., 2015). In 2018, the field of spatial transcriptomics in plants was yet again revolutionised by the publication of the first spatial transcriptomics protocol (Giacomello et al., 2017). In this new suite of technologies, that have since then evolved rapidly, the tissue does not need to be sectioned anymore prior to sequencing but rather the whole tissue is placed onto barcoded slides that allow for the sequencing of each RNA molecule while retaining the information about where it was expressed in the tissue. In wheat, the use of spatially resolved transcriptomic data is nearly non-existing. The development of new spatial technologies with improved spatial

resolution and their future applications will be discussed further in Chapter 2 as we developed a low input sequencing protocol for wheat spikes to investigate the differences in spikelet development across the spike.

#### 1.4.2 Imaging of the wheat spike

Another crucial technology for the work presented in the thesis was the development of primordia imaging techniques. Microscopy has always been a central tool for biological research and the scientific advancements in plant biology are often closely intertwined with the evolution of microscopic techniques, starting with the imaging of cork cells by Hooke in 1665 (Crouvisier-Urien et al., 2019). Since then, light microscopy has evolved dramatically and today high-quality light transmission stereo microscopes are readily available to most researchers. Within the field of wheat and grass research, stereo microscopy of the developing reproductive organs (such as spikelet, floret and inflorescence) has been used to study their development and identity. Much of what we know today about the developmental progression of the wheat, barley and maize inflorescence stems from the works of Bonnett, Kirby and Waddington during the mid-20<sup>th</sup> century. Their publications first described how the inflorescence primordia are progressively initiated and how spike development is linked to plant growth. Furthermore, they published a developmental scale (Waddington stages), which is extensively used till today (Bonnett, 1966; Kirby and Appleyard, 1981; Waddington et al., 1983).

Besides light microscopy, the development of scanning electron microscopy (SEM) further aided the study of grass inflorescence, allowing for the acquisition of more detailed micrographs. The principles of SEM and transmission electron microscopes (TEM) were developed during the 1920s and 30s and this was predominantly driven by the camera industry (McMullan, 1995). It wasn't until the 1960s that SEM technology became available to research laboratories (McMullan, 1995). The use of SEM to resolve the development of grass inflorescences was championed by Elizabeth Anne Kellogg, who has been a leading figure in the field of plant comparative biology over the last 30 years.

However, imaging the surface of plant organs can only reveal some of the information regarding the organ's structure and development. Thus, it is also common practice to section tissues and apply histological stains, which enable the analysis of sub-surface tissue structures, cell types and development using light microscopy. But it was the invention of confocal laser scanning microscopy (CLSM), a specialized type of fluorescence microscopy, that made optical sectioning and thus the imaging of live tissue possible (Running et al., 1995). Fluorescence microscopy has also been a key technique for cell biology as it can be used to determine the location of fluorescence dye tags (see (Moreno et al., 2006)). However, we will here focus on the use of confocal microscopy for organ development studies. Over 27 years ago Running et al. (1995) first described how CLSM can be

applied to image cell walls of the developing shoot apex of plants (*Arabidopsis*). The method they proposed was simpler than classical histological staining and sectioning of the plant tissue, and allowed for a high throughput acquisition of cellular resolution images of the whole apex (Running et al., 1995). In this first description of the method, the three-dimensional orientation of cells within their original position in the organ were visualised by Propidium Iodide staining cells of fixed apices. Since then, methods to image live plant tissue have been developed, which makes it possible to compare growth and development *in situ* and at cellular resolution. The following surge in studies using CLSM was also aided by the rapid innovations of computational imaging software (Running et al., 1995). Image processing software, such as ImageJ, have enabled the rapid, and systematic, analysis of complex and large image files obtained from CLSM (Collins, 2007). Lastly, the discovery that plant apices, and other tissue, can be grown *ex situ* was exploited in the latest development of CLSM for plant development studies. By growing apices *ex situ*, one can image the same apex multiple times and thus track the patterns of cell growth and division underlying the tissue's developmental path. MorphoGraphX, a sophisticated computational program that makes this type of analysis possible, was developed over the last ten years (Barbier de Reuille et al., 2015). In *Arabidopsis*, major discoveries have been achieved by the live imaging of meristems and subsequent cell lineage tracking using MorphoGraphX (Barbier de Reuille et al., 2015). However, it was thus far not possible to use this technology for wheat floral development studies (or any other grass) as the tissue does not survive *ex situ*, thus limiting the application for CLSM to fixed tissues. CLSM is generally not widely used in wheat research (yet). This is potentially also due to the lack of fundamental research studies, which were held back by the absence of a reference genome until only recently. However, we are now seeing an increase in fundamental studies conducted in wheat and this trend requires the adaptation of methods, such as imaging techniques, to wheat. In Chapter 3, we present a protocol that allows for *ex situ* cultivation of wheat spike apices, which opens the door to CLSM and MorphoGraphX application in wheat.

## 1.5 Thesis Aim

The overall aim of this thesis is to understand the genetic and developmental causes of rudimentary basal spikelets in wheat. For this, we investigated the differences in transcriptional signalling during initiation (Chapter 2), cellular growth deviations thereafter (Chapter 3), and final abortion (Chapter 4) of basal spikelets compared to their central counterparts. Throughout the chapters, we investigated the following questions:

- What are the transcriptional differences between central and basal spikelets?
- Does *SVP* expression negatively affect basal spikelet establishment?
- Is increased basal spikelet abortion caused by their poor development or resource allocation?

## 2 High expression of MADS-box gene *VRT2* increases the number of rudimentary basal spikelets in wheat

All results described in this chapter have been published in the following manuscript (Appendix 1):

Backhaus, A.E., Lister, A., Tomkins, M., Adamski, N.M., Simmonds, J., Macaulay, I., Morris, R.J., Haerty, W. and Uauy, C., 2022. High expression of the MADS-box gene *VRT2* increases the number of rudimentary basal spikelets in wheat. *Plant Physiology*.: <https://doi.org/10.1093/plphys/kiac156>

### 2.1 Chapter Summary

Spikelets are the fundamental building blocks of *Poaceae* inflorescences, and their development and branching patterns determine the various inflorescence architectures and grain yield of grasses. In wheat (*Triticum aestivum* L.), the central spikelets produce the most and largest grains, while spikelet size gradually decreases acro- and basipetally, giving rise to the characteristic lanceolate shape of wheat spikes. The acropetal gradient corresponds with the developmental age of spikelets, however the basal spikelets are developed first, and the cause of their small size and rudimentary development is unclear. Here, we adapted G&T-seq, a low-input transcriptomics approach, to characterise gene expression profiles within spatial sections of individual spikes before and after the establishment of the lanceolate shape. We observed larger differences in gene expression profiles between the apical, central and basal sections of a single spike than between any section belonging to consecutive developmental timepoints. We found that *SVP* (*SHORT VEGETATIVE PHASE*) MADS-box transcription factors, including *VRT-A2* (*VEGETATIVE TO REPRODUCTIVE TRANSITION 2*), are expressed highest in the basal section of the wheat spike and display the opposite expression gradient to flowering E-class *SEP1* (*SEPALLATA 1*) genes. Based on multi-year field trials and transgenic lines, we show that higher expression of *VRT-A2* in the basal sections of the spike is associated with increased numbers of rudimentary basal spikelets. Our results, supported by computational modelling, suggest that the delayed transition of basal spikelets from vegetative to floral developmental programmes results in the lanceolate shape of wheat spikes. This study highlights the value of spatially resolved transcriptomics to gain insights into developmental genetics pathways of grass inflorescences.

### 2.2 Introduction

The arrangement of flowers in individual plants of the same species is highly conserved and follows a systematic and rhythmic pattern. This systematic appearance of flowers is not surprising, as floral architectures are determined by the regular initiation of flower primordia on the flanks of the apical meristem and their rate of initiation and developmental fate are under strong genetic control (Prusinkiewicz et al., 2007). The unifying feature of floral architecture in grasses (*Poaceae*) is the formation of all flowers (termed florets) within spikelets (Kellogg et al., 2013). Spikelets are the

fundamental building blocks of grass inflorescences and their development and branching patterns determine the various inflorescence architectures of grasses (e.g., spikes, panicles). Wheat (*Triticum aestivum*) forms a spike shaped inflorescence, in which sessile spikelets are directly attached to the inflorescence axis (or rachis) in a distichous phyllotaxis (Koppolu and Schnurbusch, 2019). Upon floral transition, the vegetative meristem ceases to initiate leaf primordia and transitions into the inflorescence meristem (IM). During the Double Ridge stage (DR) of wheat spike development, the IM initiates a lower leaf ridge and an upper spikelet ridge (or primordia) during each iteration. Within the inflorescence the upper ridges differentiate into spikelet meristems, while the lower ridges are suppressed upon flowering (Bommert and Whipple, 2018). DR initiation will continue at the IM until the terminal spikelet stage, when IM forms a final spikelet (Koppolu and Schnurbusch, 2019). Spikelet initiation and development has been extensively studied in wheat and other monocot crops, such as rice (*Oryza sativa*), maize (*Zea mays*), and barley (*Hordeum vulgare*), as the number of spikelets per spike is a major determining factor for grain number and thus yield per spike.

Not all spikelets across the wheat spike, however, produce the same amount of grain. The central spikelets produce the most and largest grains, while spikelet size gradually decreases acro- and basipetally. Within a single spike, the most apical and basal spikelets might produce no or only one grain while the central spikelets of the same spike set 3-5 grains. Bonnett (1966) documented that this distinct lanceolate shape of the wheat spike is first established during the Glume Primordia (GP) stage (just after the DR stage). This asynchronous development among the spikelets is maintained throughout the development of the spike. The gradual decrease in spikelet size from the central to apical section of the spike can be explained by the continuous development of new spikelet ridges from the apical inflorescence meristem: the most apical spikelets are the youngest and had the least time to develop. However, basal spikelets are initiated first and it is unclear why they remain smaller than their central counterparts. In the mature spike the most basal one or two spikelets are often only formed in a rudimentary manner, with small glumes present but all floral structures remaining immature.

Efforts to understand the genetics of wheat spikelet initiation and development have focused on members of the MADS-box transcription factor (TF) family, which play central roles in the flowering gene models (Zhao et al., 2006). Li et al. (2019) showed that MADS-box genes of the *SQUAMOSA*-clade, *VERNALISATION 1 (VRN1)*, *FRUITFULL 2 (FUL2)* and *FRUITFULL 3 (FUL3)*, have overlapping functions in controlling the timing of the transitions from the vegetative to IM as well as the formation of the terminal spikelet. In *vrn1ful2*-null mutants, the IM remained indeterminate causing the mutants to form more spikelets per spike. However, all lateral spikelets were replaced by leafy shoots in the *vrn1ful2* double and *vrn1ful2ful3* triple mutants (Li et al., 2019). These mutants had increased expression of genes belonging to the *SHORT VEGETATIVE PHASE (SVP)* family of MADS-box genes, including *VEGETATIVE TO REPRODUCTIVE TRANSITION 2 (VRT2)*. Subsequent

studies determined that overexpression of *VRT2* led to reversion of basal spikelets to spikes and the downregulation of other MADS-box genes required for floral development, including members of the *SEPALLATA 1* (*SEP1*) clade (Li et al., 2021). Together, these studies exemplify the importance of the temporal sequence of flowering gene expression for the correct development of the wheat spike.

Attempts to unravel the genetic network controlling wheat spike development have focused on these temporal changes in expression patterns across consecutive developmental stages. For example, Li et al. (2018) and Feng et al. (2017) performed transcriptome profiling using pooled samples of multiple complete spikes from six (vegetative to floret differentiation) and four (double ridge to young floret) developmental stages, respectively. In a few cases, studies have examined the expression patterns of individual genes (via reverse transcription quantitative PCR (RT-qPCR)) and found gene expression gradients along the spike. For example, Debernardi et al. (2017) demonstrated that *APETALA 2* (*AP2*) is expressed higher in the apical section of wheat spikes than in central or basal sections. This *AP2* expression gradient was associated with morphological changes along the same spike. This study alongside work in barley (Youssef et al., 2017), suggests that gene expression gradients *within* individual developmental stages could be important to further unravel the genetic control of spike development. However, despite its potential biological importance, spatial transcriptome profiles along the spike have yet to be investigated in wheat.

In this study, we aimed to characterise gene expression profiles along the spike during the establishment of the lanceolate shape of the wheat spike from DR to GP. We adapted G&T-seq (Genome and Transcriptome sequencing), a low-input sequencing approach to sequence the transcriptome of the sections. Recently, Giolai et al. (2019) adapted the protocol to identify expression differences across single leaves of *Arabidopsis* (GaST -seq), demonstrating that the G&T-seq method can be readily used for sequencing of hand harvested, small input plant material without the need of previous tissue dissociation or treatment. G&T-seq is thus comparable to methods using laser-micro dissection followed by sequencing to achieve spatially resolved transcriptome wide sequencing data. In comparison, the available transcriptome sequencing methods at higher resolutions (such as single cell RNA-seq or fluorescence-activated cell sorting (FACS)) are not spatially resolved as the complete tissue is dissolved into single cells for barcoding or selection prior to sequencing (Rich-Griffin et al., 2020).

We sequenced the apical, central, and basal sections of individual spikes before (DR) and after (GP) the establishment of the lanceolate shape. Gene expression profiles differed most strongly between spatial sections of the same spike, as opposed to temporal sections (any two sections from different timepoints). Members of the *SVP* gene family were expressed most highly in the basal sections with expression decreasing upwards from the base (acropetally), while members of the *SEP1* gene family showed the opposite expression pattern, i.e. most highly expressed in apical

sections, with expression decreasing towards the base (basipetally). The increased number of rudimentary basal spikelets due to *VRT-A2* misexpression supports the hypothesis that high expression levels of *SVPs* in the basal section delays spikelet establishment, leading to their rudimentary shape in the mature spike. This study highlights that spikelets within the same spike experience significantly different flowering signals due to their consecutive development and spatial position within the spike. Acknowledging these differences can help us gain a better understanding of the genetic flowering pathway of grass inflorescences.

## 2.3 Results

### 2.3.1 Low-input sequencing enables spatial analysis of the wheat spike transcriptome

To investigate transcriptional differences between the apical, central, and basal section of developing wheat spikes, we adapted the low-input G&T sequencing (G&T-seq) method for RNA-seq of small plant tissue sections. G&T has been developed for single-cell RNA and DNA sequencing of mammalian systems (Macaulay et al., 2015) and was previously adapted for *Arabidopsis thaliana* (GaST-seq; Giolai et al., 2019). We collected four individual developing wheat (cv Paragon) spikes at both the double ridge (DR) and glume primordia (GP) stage and hand-dissected them into apical, central, and basal sections (Figure 2.1A).

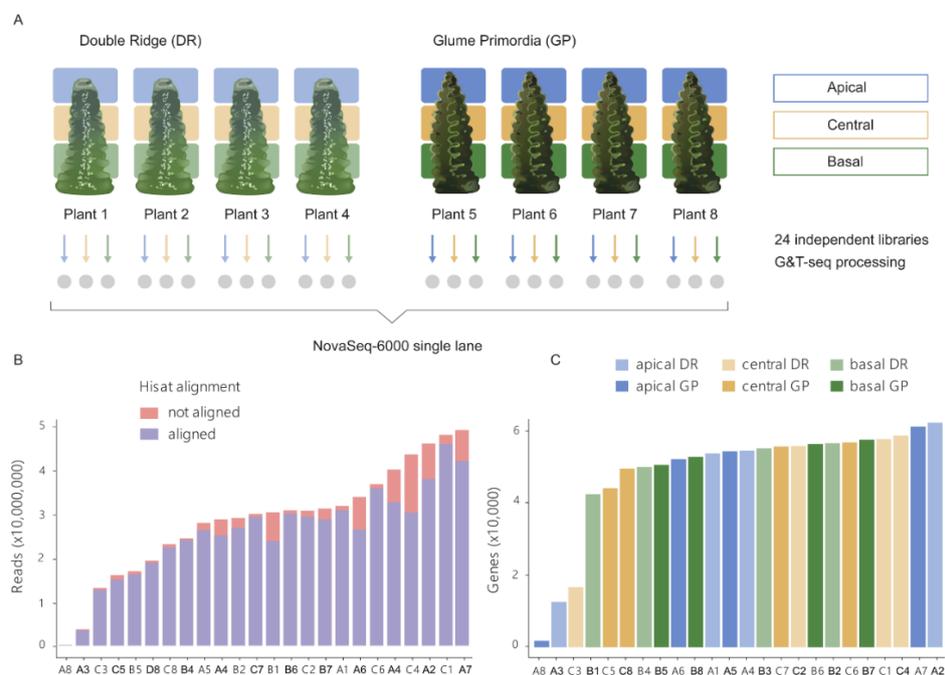


Figure 2.1 Low-input G&T sequencing of developing wheat spikes. (A) Summary diagram of tissue collection and sequencing. Grey circles indicate the 24 individual libraries prepared for sequencing from each individual tissue section dissected from individual spikes. (B) Reads per library after trimming and quality controls (see Methods). Stacked bars indicate the number of reads aligned (blue) and not aligned (red) by HISAT to the RefSeqv1.0 genome. (C) Number of expressed genes (>10 read counts) per library based on tissue section and Waddington developmental stage (DR: Double Ridge; GP: Glume Primordia). In (B and C), the X-axis indicates the ID of each sample which is composed of the tissue section (A: apical, C: central, B: basal) and plant number (1-8) as indicated in (A). Detailed quality control data for each library is provided in Table 2.2.

On average, samples had 28,799,626 reads (coefficient of variation (CV) 43%), of which 90% (CV 8.5%) aligned to the genome post adaptor trimming (Figure 2.1B, Table 2.1). Furthermore, the number of aligned reads and the number of expressed genes per library was largely homogenous among the spatial sections and Waddington stages (Table 2.1).

Table 2.1: Average number of reads aligned to the RefSeqv1.0 genome and expressed genes (>10 read counts) in the three tissue sections and two Waddington developmental stages (DR: Double Ridge; GP: Glume Primordia) (n = 4 biological replicates per tissue section \* developmental stage).

Section	Reads aligned		Genes expressed	
	DR	GP	DR	GP
<b>Apical</b>	26,342,934	23,913,657	44,488	41,152
<b>Central</b>	29,823,023	25,854,540	45,913	50,074
<b>Basal</b>	25,174,482	23,627,494	49,513	52,740

On average, 47,313 genes per library were expressed (>10 read counts) and we found no difference ( $P > 0.56$ , ANOVA) in the number of expressed genes across spatial (apical, central, basal) or between temporal (DR, GP) conditions (Figure 2.1C). We excluded three libraries with low average number of expressed genes (difference greater than five times the standard deviation; Figure 2.1C, Table 2.2) and two libraries because they were strong outliers in the principal component analysis (PCA; Figure 2.2A). In total, 19 RNA-seq libraries (DR: 3 apical, 4 central, 3 basal; GP: 2 apical, 3 central, 4 basal) passed our selection criteria and were used in the subsequent analyses. We identified 91,646 genes being expressed across these 19 libraries.

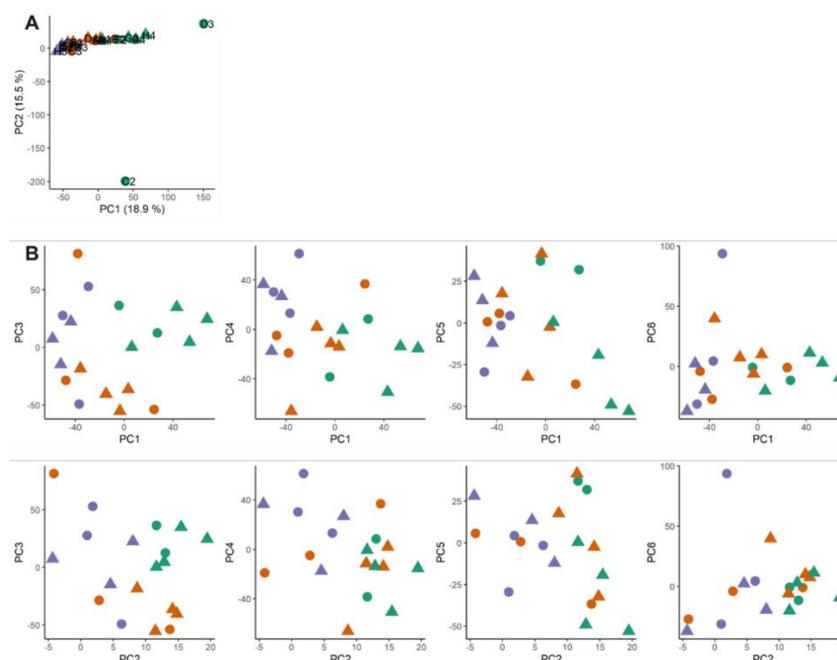


Figure 2.2: Principal Component analysis (PCA) of RNA-seq libraries. (A) Principal Component (PC) analysis of 21 libraries that passed initial quality control (three libraries excluded due to low average number of expressed genes), indicating two outlier libraries (C2&D3) which were subsequently removed from the analysis. (B) PCs of the 19 libraries that passed all controls. Purple = Apical, Orange = Central, Green = Basal; Circle = Double Ridge, Triangle = Glume Primordia stage.

Table 2.2: Summary of raw read number, filtered reads, alignment %, and expressed genes for each of the 24 sequencing libraries created using the G&T-seq method. DR= Double Ridge, GP = Glume Primordia. If a library passed QC, it was based on the number of expressed genes (column I) and PCA results (see Figure 2.2). ID refers to the name used for each library in Figure 2.1.

Library	Tissue	Sample	Plant	ID	Raw read pairs total	Passed (tringalore) reads	Hisat aligned (%)	Expressed genes (Htcounts >10)	Passed QC
A2	DR	apical_DR	1	A_1	33,664,671	32,062,547	96.48	52,244	TRUE
A3	DR	basal_DR	3	B_3	30,295,269	28,976,305	87.31	53,547	TRUE
A4	GP	central_GP	6	C_6	39,099,793	37,073,012	97.23	55,162	TRUE
B2	DR	central_DR	1	C_1	50,420,405	48,321,912	95.51	56,075	TRUE
B3	DR	apical_DR	4	A_4	42,117,378	40,262,975	81.15	53,021	TRUE
B4	GP	basal_GP	6	B_6	33,216,608	31,029,623	97.32	54,516	TRUE
C2	DR	basal_DR	1	B_1	32,554,056	30,596,728	78.7	41,019	FALSE
C3	DR	central_DR	4	C_4	45,516,116	43,762,080	70.07	57,037	TRUE
C4	GP	apical_GP	7	A_7	51,169,262	49,102,394	86.01	59,389	TRUE
D2	DR	apical_DR	2	A_2	48,138,049	46,330,044	82.26	60,489	TRUE
D3	DR	basal_DR	4	B_4	26,472,432	24,714,312	97.89	48,429	FALSE
D4	GP	central_GP	7	C_7	31,773,685	30,269,408	97.47	54,078	TRUE
E2	DR	central_DR	2	C_2	32,293,673	31,058,949	95.18	54,278	TRUE
E3	GP	apical_GP	5	A_5	29,384,704	28,160,317	94.09	52,669	TRUE
E4	GP	basal_GP	7	B_7	33,033,120	31,584,188	91.39	55,825	TRUE
F2	DR	basal_DR	2	B_2	30,668,419	29,202,553	92.89	55,058	TRUE
F3	GP	central_GP	5	C_5	16,920,777	16,339,753	94	42,838	TRUE
F4	GP	apical_GP	8	A_8	358,800	339,630	82.08	1,837	FALSE
G2	DR	apical_DR	3	A_3	5,305,986	4,057,862	90.03	12,197	FALSE
G3	GP	basal_GP	5	B_5	18,033,782	17,161,540	95.22	49,315	TRUE
G4	GP	central_GP	8	C_8	24,422,756	23,395,813	96.21	48,217	TRUE
H2	DR	central_DR	3	C_3	14,598,985	13,525,176	95.48	16,260	FALSE
H3	GP	apical_GP	6	A_6	35,657,275	34,268,070	77.76	50,712	TRUE
H4	GP	basal_GP	8	B_8	20,880,756	19,595,838	97.5	51,304	TRUE

### 2.3.2 Transcriptome-wide differences are largest between the apical and basal sections of the spike

To investigate global differences among the 19 RNA-seq libraries, we performed a principal component analysis (PCA; Figure 2.3A). The first two PCs explained 19% and 16% of the overall variance present in the libraries. We observed that the two PCs separated libraries by the spatial position (apical, central, basal) rather than developmental stage (DR, GP). There was a clear separation between libraries originating from apical and basal spike sections, while libraries from central sections were dispersed between these two clusters (Figure 2.3A). We investigated PC1 to PC6 and found that none of these combinations clustered libraries by developmental stage (Figure 2.2B). Given that we sequenced developing spike sections of single plants, as opposed to the more commonly employed pooling of multiple biological samples, we found as expected some degree of heterogeneity between samples from the same location and stage (Figure 2.3A).

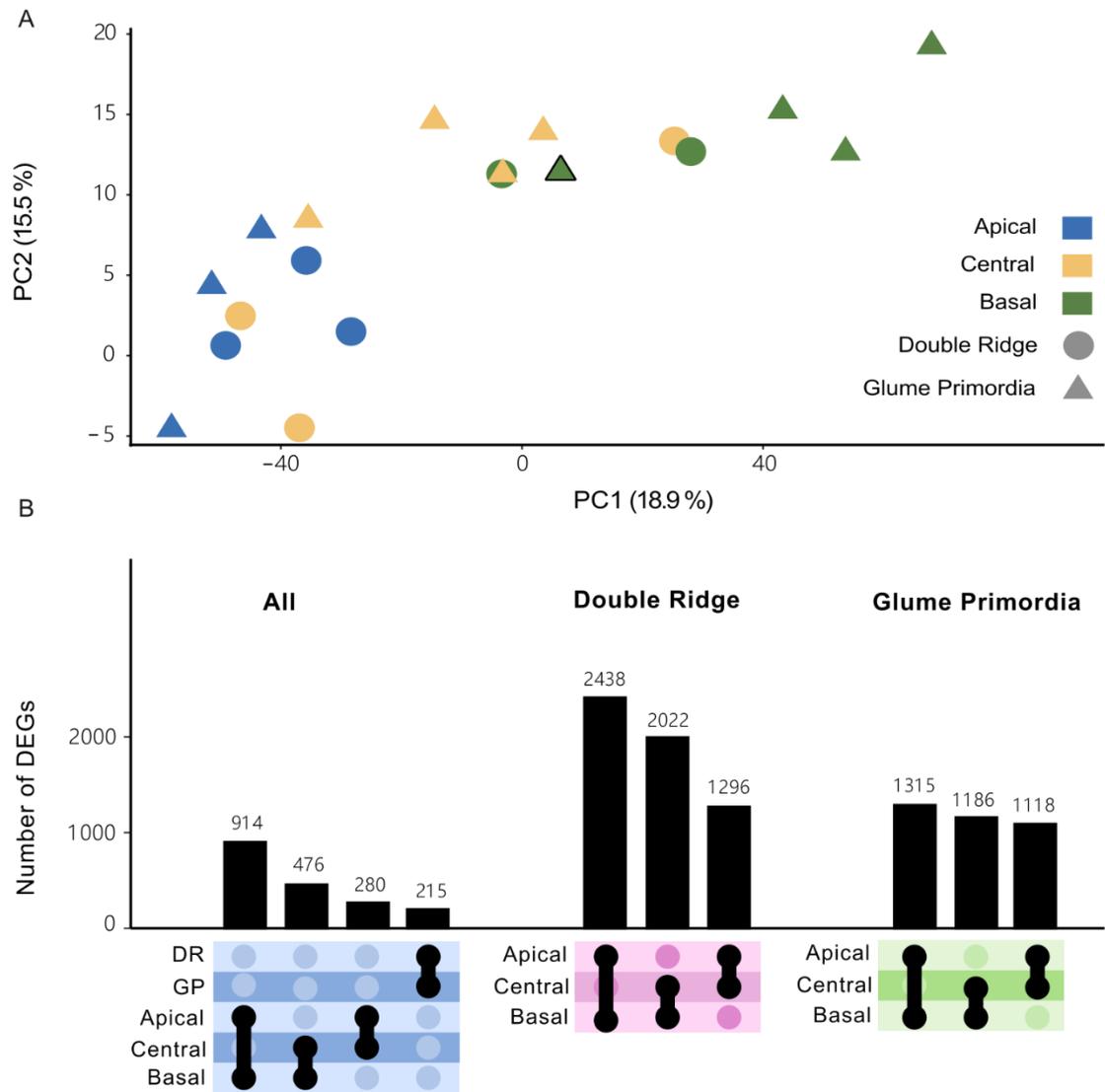


Figure 2.3: Transcriptome wide differences are larger between tissue sections than developmental timepoints (A) Principal component analysis (PCA) on the 19 transcriptome libraries from apical (blue), central (yellow) and basal (green) sections of Double Ridge (DR, circles) and Glume Primordia (GP, triangles) spikes. Black bordered triangle is plant 8 (GP) in which the basal section clustered closer with central-GP sections than the other basal-GP sections. (B) UpSet plot showing the number of differentially expressed genes (DEGs) between spatial sections and Waddington stages.

To investigate this variation further, we quantified changes in gene expression across biological replicates by calculating CVs for each gene (see Methods). The median CV for a gene across the biological replicates was 39% (Figure 2.4) with a Q1-Q3 interquartile range between 24% and 62%. We also calculated the CV per gene for published datasets from Li et al. (2018) and Feng et al. (2017). Both studies sequenced developing wheat spikes at similar developmental stages, pooling many spikes per sample. Li et al. (2018) pooled between 100 to 200 spikes of winter wheat (KN9204) per sample, while Feng et al. (2017) reported pooling of 10 to 50 spikes (cv. Chinese Spring) per sample. In both studies the median CV of a gene was lower (14% and 21%, respectively) than in our study. The larger CVs in our data could be explained by the biological variation that exists between individual plants, which may have been reduced by the pooling of many spikes in both Li et al. (2018) and Feng et al. (2017).

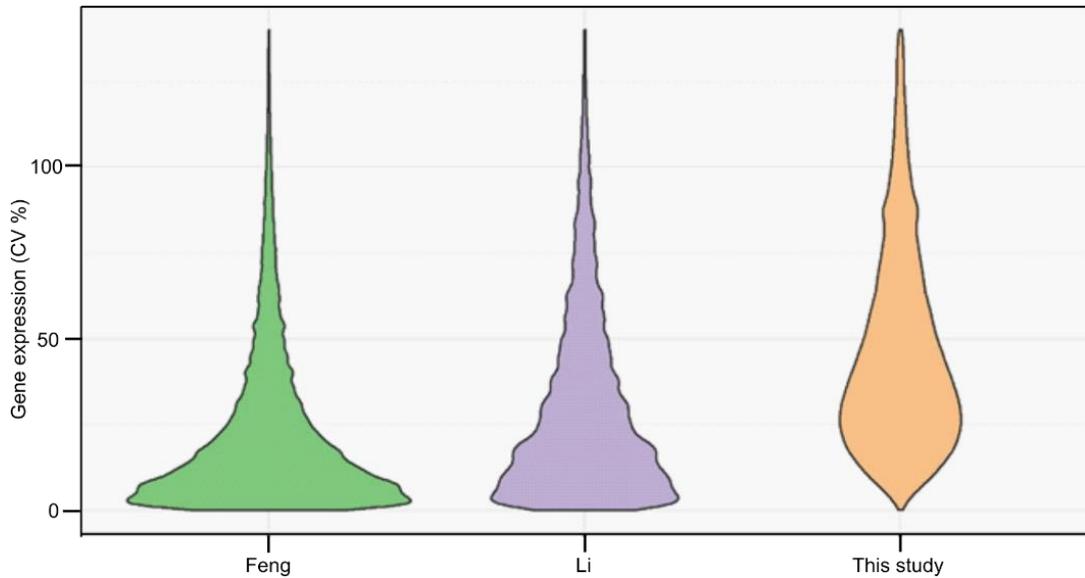


Figure 2.4: Comparison of Coefficient of variation (CV) of gene expression in wheat RNA-seq data sets. CVs of gene expression. CVs among the biological replicates were calculated for all expressed genes in this study (number of biological replicates = 2-4). Feng et al. (2017) and Li et al. (2018) have n=2 biological replicates of pooled samples.

We first analysed differentially expressed genes (DEGs) between the DR and GP stage and between apical, central and basal sections across the two Waddington developmental stages. The number of DEGs between DR and GP (215 genes) was smaller than the number of DEGs identified between the spatial positions, which ranged from 280 DEGs between central and apical sections to 914 DEGs between the apical and basal sections (Figure 2.3B). Next, we compared the apical, basal and central sections within each Waddington developmental stage. We identified more DEGs by comparing the spatial sections within either Waddington stage individually than in the combined analysis. The number of DEGs between apical and basal sections at each stage (DR: 2,438; GP: 1,315) were similar to the number of DEGs between central and basal sections (DR: 2,022; GP: 1,186). The number of DEGs between these sections at DR, however, was nearly double the number of DEGs at GP. In contrast, the number of DEGs between apical and central sections was similar at both stages (DR: 1,296; GP: 1,118), suggesting that the basal section of the spike is most different in the earlier developmental stage. Only 11% of the DEGs were shared between DR and GP in the apical to basal comparison, 7% between the central to basal DEGs, and 5% between apical to central DEGs. In total, we identified 5,353 unique genes as differentially expressed between any of the three sections at either Waddington stage (please see Supplemental Table S2 in original publication (Backhaus et al., 2022) as this table was too large to be re-printed in this thesis). Overall, the number of DEGs was largest between the apical and basal sections, reflecting the strong spatial clustering observed in the PCA graph, but most genes that were differentially expressed across the spike did not maintain this gradient over the two developmental stages. In summary, despite the high biological variation in gene expression in our data compared to previous pooled whole-spike studies, we could detect transcriptome wide differences between the spatial sections of developing wheat spikes.

### 2.3.3 The SVP MADS-box transcription factors have opposing expression profiles to flowering E-class genes

To further investigate the differences in expression across the spike and to identify genes with similar expression patterns, we performed hierarchical and k-means clustering (Figure 2.5). We restricted the clustering to the 5,353 genes identified as differentially expressed across the spike at either one or both Waddington stages (Figure 2.6A). We identified seven non-redundant clusters, each containing between 8% to 21% of the 5,353 DEGs (Figure 2.5A, Figure 2.6A). Both hierarchical and k-means clustering produced highly similar results (Figure 2.5B).

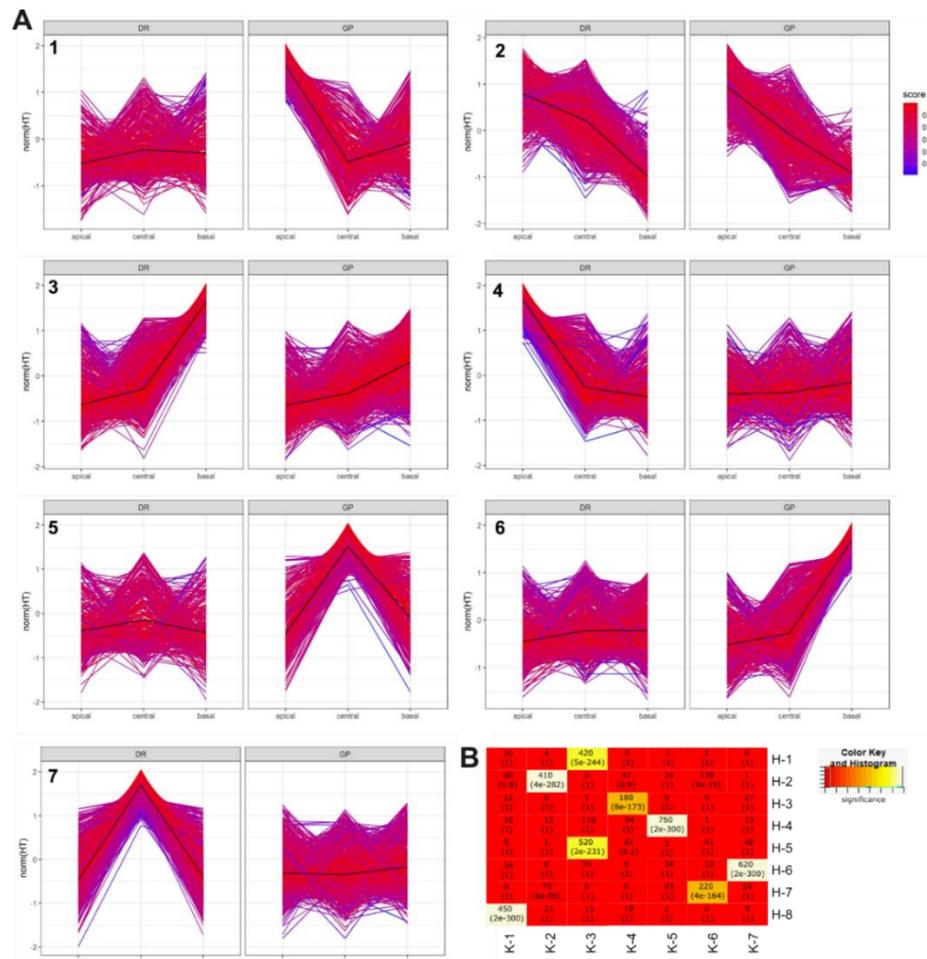


Figure 2.5: Expression patterns of all DEG clusters (n=7) as identified by k-means clustering. (A) Expression pattern of differentially expressed genes (DEGs) in the seven clusters identified by k-means clustering. Score colouring indicates the correlation of gene expression with the centroid pattern of the cluster (black line). Expression values were normalized using DeSeq2, scaled and centred (R(base) function "scale"); see Methods). (B) Correlation of k-means (K) and hierarchical (H) clusters. Upper value is the number of shared genes, lower value is the P-value of the correlation. Calculation was performed using WGCNA "overlapTable" function (Langfelder & Horvath, 2008). Colours = -log(P-value) (red=less significant, white = most significant).

We identified 1,894 genes (35% of DEGs) to be more highly expressed in the apical section, either across both timepoints (503 genes, cluster 1), or only at DR (751 genes, cluster 7) or GP (640 genes, cluster 3). In the central section, 1,362 genes (25%) had higher relative expression at either DR (917 DEGs, Cluster 5) or GP (445 DEGs, Cluster 6). In the basal section, we observed 2,097 genes (39%) being more highly expressed. Cluster 4 contained the most DEGs (1,170) and was characterized by an upregulation of expression in the basal section at both Waddington stages, although this

upregulation was higher at the DR stage. Another 927 genes were upregulated in the basal section, but only at the GP stage (cluster 2).

To further characterize the clusters, we independently tested for enrichment of TF families and gene ontology (GO)-terms relating to biological processes (all GO-terms and TF families in Table 2.3 and 2.4, respectively). Genes that were more highly expressed in the apical section were enriched for the GO terms “reproductive structure development” (GO:0048608) and “floral organ development” (GO:0048437; cluster 1) as well as for HD-Zip\_IV and SRS TF families ( $P < 0.001$ ; cluster 1). In cluster 3 (highly expressed in the apical section at GP; Figure 2.6B) we found no significant enrichment of GO-terms relating to biological processes, but a significant enrichment of MADS\_II TFs ( $P = 0.013$ ). For clusters defined by an increased expression in the central sections (clusters 5) we detected an enrichment for GO-terms related to polyphosphate processes (GO:0006797/0006779;  $P < 0.03$ ), and a significant enrichment of the C2C2\_CO-like TFs at DR ( $P < 0.04$ ). Genes with higher expression in the basal section of the spike (cluster 4; Figure 2.6B) were enriched for a number of GO terms related to photosynthesis, (e.g. GO:0015979;  $P < 0.01$ ) and “negative regulation of flower development” (GO:0009910;  $P < 0.01$ ), as well as for MADS\_II TFs ( $P = 0.08$ ). Cluster 2, which was characterized by higher expression in the basal section only at GP, was enriched for “Jasmonic acid response” (GO:0009753;  $P < 0.001$ ) and Tify ( $P < 0.001$ ) and C2C2\_CO-like TFs ( $P < 0.04$ ).

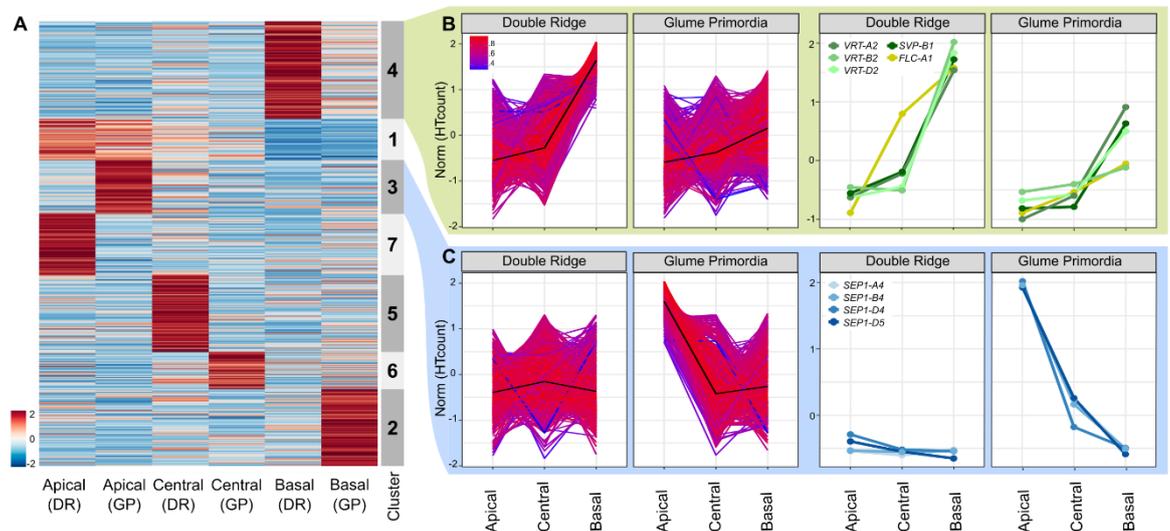


Figure 2.6: Clustering of differentially expressed genes across the wheat spike (A) Normalized expression matrix and K-means clustering of the 5,353 genes differentially expressed across the spike at either one or both Waddington stages (DR = Double Ridge; GP = Glume Primordia). Colours (blue to red) show relative log<sub>2</sub> expression of genes after normalisation. (B) Expression pattern of the 1,170 genes allocated to Cluster 4 (left), and of MADS-box transcription factors in the same cluster (right). Colours indicate how well the gene expression pattern fits the average expression pattern (black line). Red = best fit, Blue = least good fit. (C) Expression pattern of the 640 genes and MADS-box transcription factors of Cluster 3 as arranged in B. Norm = normalized and scaled gene expression. RefSeq1.1 gene IDs and raw expression values of genes shown in the right-hand panels are presented in Supplemental Table S2 (of the published manuscript).

We were interested in further characterizing the expression patterns of the MADS-box TFs as they were significantly enriched in two of the seven clusters and are important in floral transition and development (Becker and Theißen, 2003; Feng et al., 2017). We detected 14 differentially expressed MADS-box TFs in our study, five of which were more highly expressed in the apical section, four of these only at GP stage (cluster 3) and one being consistently expressed across both Waddington stages (cluster 1). In contrast, five MADS-box TFs were more highly expressed in the basal section at both Waddington stages (cluster 4) and another two were more highly expressed in the basal section only at GP (cluster 2). An additional two MADS-box genes were part of the remaining clusters (Table 2.4).

In the apical/GP cluster 3 we noticed that all MADS-box genes belonged to the *Triticum aestivum* *SEPALLATA 1* (*SEP1*) group (Figure 2.6B). All three homoeologs of *SEP1-4* (*TraesCS7A02G122000*, *TraesCS7B02G020800*, *TraesCS7D02G120500*) and the D-genome copy of *SEP1-5* (*TraesCS7D02G120600*) were part of this cluster. The *SEP* genes were expressed at relatively low levels at DR (Figure 2.6), but were significantly upregulated at GP, with their transcript levels being highest in the apical section. The increased expression of *SEP1-4* at GP was in agreement with their previously reported expression patterns in tetraploid wheat by Li et al. (2021).

In the contrasting cluster 4 (upregulation in basal sections), we noticed the presence of multiple MADS-box genes belonging to the *SVP* family (Figure 2.6B, right-hand panel), which consists of three genes in wheat (*SVP1*, *VRT2* and *SVP3*). Members of this family are important for the transition from vegetative to floral meristem identity in cereals (Trevaskis et al., 2007). All three homoeologs of *VRT2* (*TraesCS7A02G175200*, *TraesCS7B02G080300*, *TraesCS7D02G176700*) and the B-genome copy of *SVP1* (*TraesCS6B02G343900*) were present in cluster 4. The cluster also contained *FLOWERING LOCUS C* (*FLC-A1*; *TraesCS7A02G260900*), although it was expressed higher in DR-central sections compared to the *SVPs* and had a linear expression gradient at GP. All *SVPs* had very similar expression patterns, being strongly expressed in basal sections only. Expression of *SVPs* was higher in all DR sections compared to the equivalent section in GP. Constitutive over-expression of *SVP*-family members in wheat and barley has been shown to delay or even reverse floral development (Trevaskis et al., 2007; Li et al., 2021). This led to hypothesis that the rudimentary development of basal spikelets was associated with an increase in *VRT2* expression levels.

Table 2.3: Enrichment of Gene Ontology (GO)-terms in the seven identified clusters of DEGs. Enrichment was performed using PLAZA (<https://bioinformatics.psb.ugent.be/plaza>; Van Bel et al., 2017; accessed on 01/04/2021). GO-terms relating to Biological process (BP) and Cellular Compartment (CC) were retained. GO-terms relating to Molecular function (MF) were not analysed. Log2Fold enrichment based on universe of expressed genes; adjusted P value (Benjamini-Hochberg); Subset Ratio is the percentage of genes associated with the GO-term.

Cluster	Identifier	Type	Log2-Enrichment Fold	P-Value	Subset Ratio	Description
1	GO:0022414	BP	0.97	1.61E-02	17.09%	reproductive process
1	GO:0000003	BP	0.97	1.71E-02	17.09%	reproduction
1	GO:0003006	BP	1.13	2.43E-03	16.24%	developmental process involved in reproduction
1	GO:0048608	BP	1.24	1.36E-03	14.96%	reproductive structure development
1	GO:0061458	BP	1.24	1.36E-03	14.96%	reproductive system development
1	GO:0048367	BP	1.27	4.28E-02	10.26%	shoot system development
1	GO:0071695	BP	2.4	4.73E-04	5.98%	anatomical structure maturation
1	GO:0021700	BP	2.31	9.70E-04	5.98%	developmental maturation
1	GO:0048437	BP	2.15	3.74E-02	4.70%	floral organ development
1	GO:0003002	BP	2.26	4.41E-02	4.27%	regionalization
1	GO:0045814	BP	2.91	4.16E-02	2.99%	negative regulation of gene expression, epigenetic
1	GO:0009836	BP	4.25	7.71E-04	2.56%	fruit ripening, climacteric
1	GO:0009835	BP	4.16	1.09E-03	2.56%	fruit ripening
1	GO:0051567	BP	4.08	1.50E-03	2.56%	histone H3-K9 methylation
1	GO:0061647	BP	3.87	3.56E-03	2.56%	histone H3-K9 modification
1	GO:0048825	BP	3.41	2.21E-02	2.56%	cotyledon development
1	GO:0006346	BP	4.38	4.21E-03	2.14%	methylation-dependent chromatin silencing
1	GO:0010158	BP	5.21	3.88E-03	1.71%	abaxial cell fate specification
1	GO:0048497	BP	5.08	5.49E-03	1.71%	maintenance of floral organ identity
1	GO:0090700	BP	5.08	5.49E-03	1.71%	maintenance of plant organ identity
1	GO:0045596	BP	4.73	1.51E-02	1.71%	negative regulation of cell differentiation
1	GO:0010093	BP	4.52	2.71E-02	1.71%	specification of floral organ identity
1	GO:0090701	BP	4.48	3.02E-02	1.71%	specification of plant organ identity
1	GO:0009299	BP	5.73	2.17E-02	1.28%	mRNA transcription
1	GO:0048480	BP	5.73	2.17E-02	1.28%	stigma development
1	GO:0048479	BP	5.41	4.42E-02	1.28%	style development
2	GO:0010446	BP	5.04	1.71E-06	2.05%	response to alkaline pH
2	GO:0071467	BP	5.04	1.71E-06	2.05%	cellular response to pH
2	GO:0071469	BP	5.04	1.71E-06	2.05%	cellular response to alkaline pH
2	GO:1900067	BP	5.04	1.71E-06	2.05%	regulation of cellular response to alkaline pH
2	GO:0009268	BP	4.67	1.14E-05	2.05%	response to pH
2	GO:0009611	BP	2.16	3.26E-05	5.87%	response to wounding
2	GO:0009753	BP	2.16	7.28E-05	5.57%	response to jasmonic acid
4	GO:0055114	BP	0.94	1.43E-05	12.08%	oxidation-reduction process
4	GO:0006979	BP	1.32	7.90E-05	6.34%	response to oxidative stress
4	GO:0015979	BP	1.49	7.09E-03	3.63%	photosynthesis
4	GO:0042743	BP	2.06	4.59E-03	2.27%	hydrogen peroxide metabolic process
4	GO:0042737	BP	1.8	3.74E-02	2.27%	drug catabolic process
4	GO:0017001	BP	2.08	1.81E-02	1.96%	antibiotic catabolic process
4	GO:0042744	BP	2.27	1.07E-02	1.81%	hydrogen peroxide catabolic process
4	GO:0048506	BP	2.73	1.17E-03	1.66%	regulation of timing of meristematic phase transition
4	GO:0048510	BP	2.73	1.17E-03	1.66%	regulation of timing of transition from vegetative to reproductive phase
4	GO:0040034	BP	2.42	9.34E-03	1.66%	regulation of development, heterochronic
4	GO:0009910	BP	2.41	9.92E-03	1.66%	negative regulation of flower development
4	GO:0090344	BP	6.04	8.09E-10	1.06%	negative regulation of cell aging
4	GO:0090342	BP	4.18	1.19E-04	1.06%	regulation of cell aging
4	GO:0007569	BP	4.13	1.54E-04	1.06%	cell aging
4	GO:0009228	BP	4.23	7.44E-03	0.76%	thiamine biosynthetic process
4	GO:0042724	BP	4.23	7.44E-03	0.76%	thiamine-containing compound biosynthetic process
4	GO:0018131	BP	5.81	1.31E-02	0.45%	oxazole or thiazole biosynthetic process
4	GO:0046484	BP	5.81	1.31E-02	0.45%	oxazole or thiazole metabolic process
5	GO:0006797	BP	7.52	2.94E-02	1.48%	polyphosphate metabolic process
5	GO:0006799	BP	7.52	2.94E-02	1.48%	polyphosphate biosynthetic process

Table 2.4: Enrichment of Transcription factor families and MADS-box transcription factor genes in the seven identified clusters of DEGs. Gene Names and rice orthologs based on Schilling et al. (2020).

Cluster	TF families	Hypergeometric Distribution (P-value)	MADS-box genes in cluster		
			RefSeq1.1 ID	Gene Name	Rice ortholog
1	HD-Zip_IV	<0.0001	TraesCS3B02G318300	<i>TaMADS32-B1</i>	<i>OsMADS32</i>
	SRS	<0.0001			
2	Tify	<0.0001	TraesCS4A02G078700	<i>TaSEP1-A1</i>	<i>OsMADS1</i>
	C2C2_CO-like	0.037	TraesCS2A02G311100	<i>TaAGL12-A2</i>	<i>OsMADS33</i>
			TraesCS7A02G122000	<i>TaSEP1-A4</i>	<i>OsMADS5</i>
3	MADS_II	0.013	TraesCS7D02G120500	<i>TaSEP1-D4</i>	<i>OsMADS5</i>
			TraesCS7D02G120600	<i>TaSEP1-D5</i>	<i>OsMADS5</i>
			TraesCS7B02G020800	<i>TaSEP1-B4</i>	<i>OsMADS5</i>
			TraesCS7D02G176700	<i>TaVRT2-D2</i>	<i>OsMADS55</i>
			TraesCS7A02G260900	<i>TaFLC-A1</i>	<i>OsMADS37</i>
4	MADS_II	0.084	TraesCS7A02G175200	<i>TaVRT2-A2</i>	<i>OsMADS55</i>
			TraesCS7B02G080300	<i>TaVRT2-B2</i>	<i>OsMADS55</i>
			TraesCS6B02G343900	<i>TaSVP-B1</i>	<i>OsMADS22</i>
5	C2C2_CO-like	0.036	TraesCS7D02G380300	<i>TaAP3-D1</i>	<i>OsMADS16</i>
6	None	None	TraesCS2D02G418800	<i>TaBS-D2</i>	<i>OsMADS31</i>
7	None	None	None	None	None

#### 2.3.4 SVP expression is higher in basal and peduncle sections and increased across all sections in *T. polonicum* VRT-A2b isogenic lines

To validate the expression pattern of *VRT2* in the individual spike analysis, we performed RT-qPCR on independently collected, pooled spike sections from cv Paragon, carrying the wildtype *VRT-A2a* allele (Figure 2.7A (blue curves); Table 2.5). We included a later timepoint, Terminal Spikelet (TS), which is about 10 days after GP to study how *VRT2* expression changes in later stages. At TS stage, the central spikelets have developed multiple florets primordia. We also included a small part of the peduncle (stem) section just below the spike as an additional spatial section. We focused the expression analysis on the A-genome homoeolog, *VRT-A2*, as its role in spike, glume and grain development of wheat was recently characterised (Adamski et al., 2021; Liu et al., 2021).

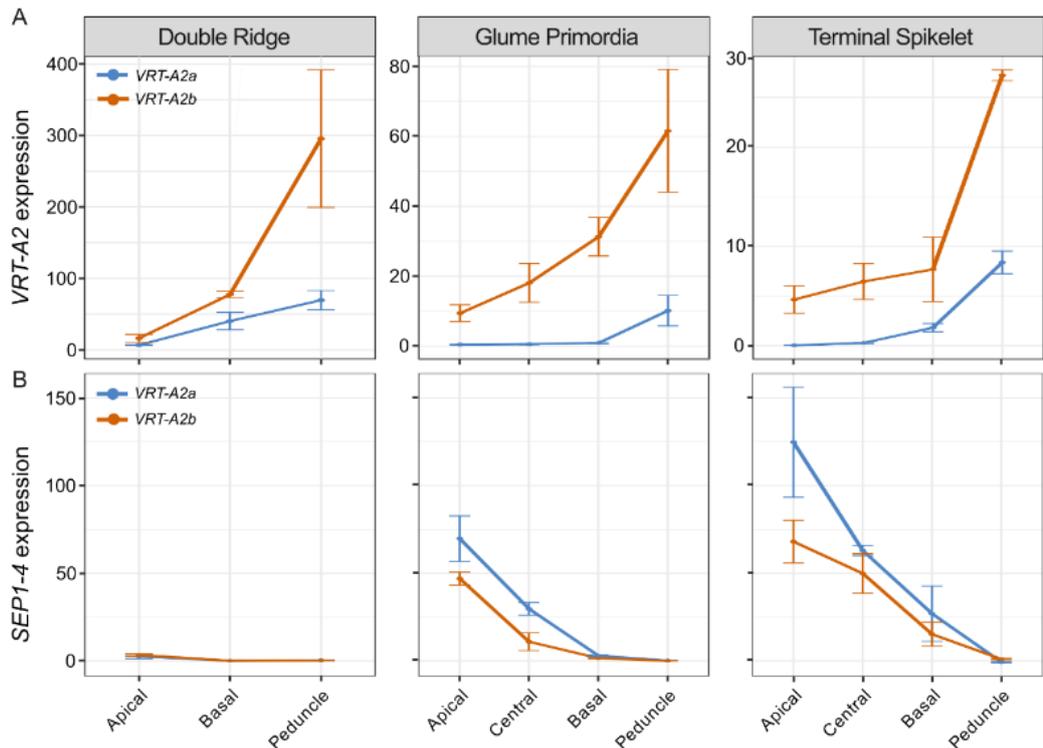


Figure 2.7: Effect of *VRT-A2b* allele on *VRT-A2* and *SEP1-4* expression in the spike. Relative expression ( $2^{\text{ddCT}}$ ) of *VRT-A2* (A) and *SEP1-4* (B) in the different sections of the spike across three timepoints in near isogenic lines (NILs) carrying either the wild-type *VRT-A2a* (blue) or the *VRT-A2b* (orange) allele from *Triticum turgidum* ssp. *polonicum*. The data are shown as mean  $\pm$  SE of gene expression (as determined by qRT-PCR) compared with control gene Actin. N = 3 biological replicates (See Table 2.5 for expression data and Table 2.6 for statistical analysis of gene expression differences).

We identified a significant interaction effect between Waddington stage and spatial section ( $P < 0.0001$ , Table 2.6); we thus analysed the three Waddington stages separately (Figure 2.7). At DR, we were limited to dissecting the spike into apical, basal and peduncle sections, as the small size of the spike meristem did not allow precise dissection of the central section when using multiple (pooled) spikes. At DR, we found *VRT-A2* marginally expressed, with significantly lower expression levels in the apical section compared to the basal ( $P = 0.003$ ) and peduncle sections ( $P = 0.001$ ). Although expression in the peduncle was higher than in the basal section at DR, this was not significant ( $P = 0.116$ ). At GP, *VRT-A2* expression was borderline detectable in the apical, central and basal sections, but expression was significantly higher in the peduncle with respect to the three spike tissues ( $P < 0.001$  for all three comparisons). Lastly, *VRT-A2* expression at TS stage was just detectable and significantly different between all sections ( $P = 6.6\text{E-}06$ ). Overall, expression decreased significantly from DR to GP/TS Waddington stages in the apical ( $P = 0.00015$ ), basal ( $P = 0.0074$ ), and peduncle ( $P = 0.012$ ) sections consistent with the previously reported strong downregulation of *VRT-A2* in the early wheat spike development (Li et al., 2021; Adamski et al., 2021; Liu et al., 2021). This is also consistent with the observed downregulation of *VRT-A2* orthologs upon floral transition in barley (Trevaskis et al., 2007) and rice (Harrop et al., 2016). As observed in the low-input RNA-seq data, the RT-qPCR data confirmed the strong basipetal gradient in *VRT-A2*

expression across the spike at DR and revealed that its expression was even higher within the peduncle.

We hypothesised that the higher expression in the basal section of the wheat spike compared to the central and apical sections is associated with the rudimentary development of the basal spikelets. To test the effect of higher *VRT-A2* expression on basal spikelet development, we analysed the effect of the *Triticum turgidum* ssp. *polonicum* *VRT-A2b* allele on the expression gradient of *VRT-A2* and spike morphology. Adamski et al. (2021) showed that *VRT-A2* in *T. polonicum*, a tetraploid subspecies of wheat, carries a sequence re-arrangement in its first intron. This results in the higher expression of the *T. polonicum* *VRT-A2b* allele, with respect to the wildtype *VRT-A2a* allele, during early spike development. We performed RT-qPCR on a cv Paragon NIL carrying the *VRT-A2b* allele and compared *VRT-A2* expression against the Paragon wildtype NIL described above (Figure 2.7; Table 2.5). Consistent with the results of Adamski et al. (2021), we detected significantly higher expression of *VRT-A2b* compared to the wildtype allele across most of the tissue sections (see Table 2.6 for individual comparisons), and a progressive decrease in *VRT-A2b* expression over time ( $P = 0.031$ ). In contrast to the wildtype NILs, ANOVA did not identify a significant interaction effect between spatial section and Waddington stage in *VRT-A2b* NILs ( $P = 0.18$ ). We thus examined the overall expression patterns and found that across all three developmental stages *VRT-A2b* expression differences were significant ( $P < 0.0001$ ). These results suggest that the basipetal expression gradient in the spike is maintained in the NILs with the *T. polonicum* *VRT-A2b* allele.

We also tested the effect of *VRT-A2* expression levels on *SEP1* expression in the tissue sections of *VRT-A2* NILs. We confirmed that *SEP1-4* expression is only marginally detectable at DR and differences in expression between the apical, basal, and peduncle section are hardly detectable with only marginally higher expression in the apical sections ( $P = 0.015$ ) at this stage. At GP, *SEP1-4* expression is significantly higher towards the tip of the spike ( $P < 0.0001$ ) consistent with the low-input RNA-seq data. Furthermore, *SEP1-4* expression was significantly lower in *VRT-A2b* NILs compared to the wildtype allele across all spike sections ( $P = 0.008$ ), confirming that higher *VRT-A2* expression can negatively affect *SEP1-4* expression. Similar trends were observed at TS, where expression was significantly lower in basal sections ( $P < 0.0001$ ) and the gradient across the spike was maintained in *VRT-A2b* NILs, but expression was overall lower ( $P < 0.003$ ). *SEP1-4* is not expected to be expressed in vegetative tissue such as the peduncle, therefore the lack of expression in this tissue across the three stages indicates that no floral tissue was accidentally sampled as peduncle (Figure 2.7B).

Table 2.5: Relative expression of *VRT-A2* and *SEP1-4* measured in Paragon NILs with either the wild-type (*VRT-A2a*) or *T. polonicum* allele (*VRT-A2b*) using qRT-PCR. Developing spikes from both genotypes were dissected at the Double Ridge (DR), Glume Primordia (GP), and Terminal Spikelet (TS) stage. At each stage, developing spikes were dissected into an apical, central, and basal section, except at DR, where only apical and basal sections were collected. In addition, a section of the peduncle was also collected at each stage. Values represent estimated mean from 2ddCT of 3 independent biological replicates per tissue\*timepoint.

Waddington stage	Spike section	Relative <i>VRT-A2</i> expression			Relative <i>SEP1-4</i> expression		
		<i>VRT-A2a</i>	<i>VRT-A2b</i>	<i>VRT-A2b</i> effect (%)	<i>VRT-A2a</i>	<i>VRT-A2b</i>	<i>VRT-A2b</i> effect (%)
DR	Apical	6.50	15.35	136%	2.63	3.38	29%
DR	Basal	40.08	76.56	91%	0.08	0.11	27%
DR	Peduncle	69.30	294.41	325%	0.35	0.28	-20%
GP	Apical	-0.59	9.38	1695%	69.77	46.98	-33%
GP	Central	-0.48	18.07	3859%	29.68	10.95	-63%
GP	Basal	-0.12	31.28	25697%	2.98	1.56	-48%
GP	Peduncle	9.11	61.59	576%	0.22	0.07	-68%
TS	Apical	0.05	4.71	10056%	124.79	68.08	-45%
TS	Central	0.41	6.53	1500%	63.01	49.89	-21%
TS	Basal	1.51	7.74	414%	26.94	15.23	-43%
TS	Peduncle	8.39	27.28	225%	-0.66	1.03	-256%



### 2.3.5 Misexpression of *VRT-A2b* in *T. polonicum* increases rudimentary basal spikelet numbers

To evaluate if the higher expression of *VRT-A2* in basal spikelets affects their development, we examined the *VRT-A2* NILs (BC<sub>4</sub> and BC<sub>6</sub>) sown as winter crops in four environments. In each field trial, we evaluated the number of rudimentary basal spikelets (RBS), that is spikelets which are reduced in size and do not contain mature grains (Figure 2.8A). The number of RBS was significantly increased in NILs carrying the *VRT-A2b* allele in all four environments ( $P < 0.0001$ , except Morley 2017  $P < 0.01$ ; Figure 2.8B; Table 2.7). The *VRT-A2a* NILs and the recurrent parent Paragon had on average 1.85 RBS, whereas *VRT-A2b* NILs produced on average 2.91 RBS. A similar difference in RBS between the NILs was observed in glasshouse conditions (*VRT-A2b* effect of +1.6 RBS; Table 2.9; Figure 2.8C).

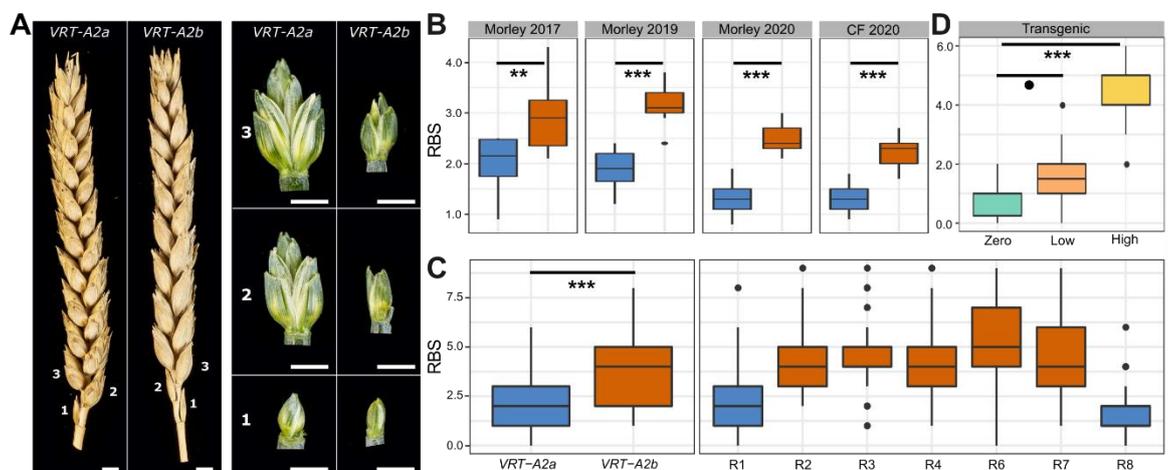


Figure 2.8: Phenotypic difference between *VRT-A2a* (blue) and *VRT-A2b* (orange) on rudimentary basal spikelet numbers (RBS). (A) Mature spikes from the field (left) and dissected basal spikelets at anthesis (right) from the glasshouse. Numbers indicate position along the spike starting at the base. Scale bar = 0.5 cm. (B) Number of RBS per spike from 10-ear samples collected in the field at maturity at Morley (2017, 2019 and 2020) and Church Farm (CF, 2020). 2017/2019 = 10-15 independent replicates per genotype; 2020 = 20 independent replicates per genotype and location. (C) Number of RBS recorded in the glasshouse for the NILs (left panel) and for seven critical recombinant lines (R1-R4, R6-R8; see Table 2.9 for graphical genotype of these lines from Adamski et al. (2021),  $n = 18-20$  plants). (D) RBS per spike recorded for the transgenic lines carrying zero ( $n = 10$  plants), low (1-5,  $n = 20$  plants) or high (9-35,  $n = 10$  plants) copy-number insertions of *VRT-A2b* in cv. Fielder. In B-D, the box represents the middle 50% of data with the borders of the box representing the 25th and 75th percentile. The horizontal line in the middle of the box represents the median. Whiskers represent the minimum and maximum values, unless a point exceeds 1.5 times the interquartile range in which case the whisker represents this value and values beyond this are plotted as single points (outliers). Statistical classifications in (B) and (C) are based on two-way ANOVA tests and in (D) on Dunnett test against the zero copy number lines. P values:  $o \leq 0.1$ ;  $** \leq 0.01$ ;  $*** \leq 0.001$ .

Furthermore, we also recorded the number of RBS in seven homozygous BC<sub>6</sub> recombinant lines used to fine-map *VRT-A2b* by Adamski et al. (2021). The RBS phenotype was mapped in complete linkage with the 50.3 kbp interval containing *VRT-A2* (Figure 2.8C; Table 2.9). This genetic and phenotypic data suggests that the increase in RBS is a pleiotropic effect of the *T. polonicum* *VRT-A2b* allele and supports the hypothesis that misexpression of *VRT-A2* negatively affects spikelet development in the base of the spike. In Paragon, the first (sometimes second) rudimentary basal spikelet fully develops the floral organs of florets one and two (e.g. lemma, palea, stamen, and ovary), however these are severely reduced in size and delayed in development compared to the

florets of central spikelets just before flowering (~Waddington stage 8-10; Figure 2.9). At this stage, the further growth and development of these basal florets is stopped and in the mature spike only the glumes of RBS are visible. In NILs carrying the *VRT-A2b* allele the development of the most basal spikelet is very similar to the wildtype. However, the second, third and sometimes fourth spikelet also display similar signs of reduced development, leading to the larger number of rudimentary basal spikelets (Figure 2.8B, C; Figure 2.9).

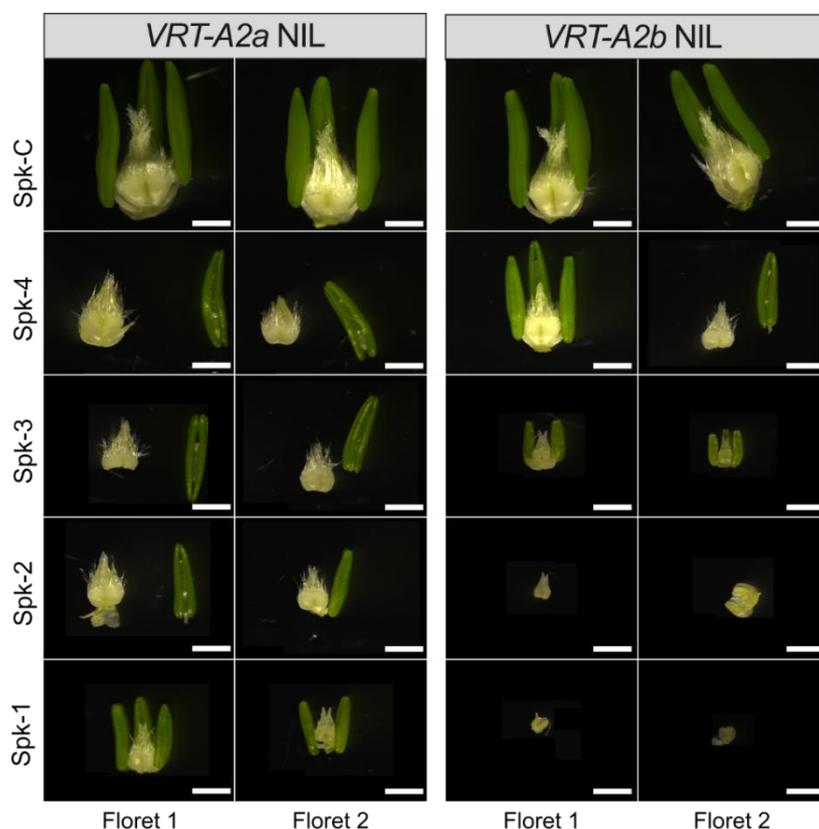


Figure 2.9: Dissected Floret 1 and 2 of basal and central spikelets of BC<sub>6</sub> NILs before anthesis. We dissected florets 1 and 2 of the basal four spikelets (Spk-1 to Spk-4) as well as the floret of the biggest central spikelet (Spk-C) of *VRT-A2* NILs. NILs were grown in glasshouse conditions (2021). Scale = 1mm. Note that not all pictures include all three anthers.

To validate the phenotypic effect of *VRT-A2*, we analysed transgenic wheat lines transformed with the complete genomic *T. polonicum VRT-A2b* sequence (including the native promoter and the intron 1 re-arrangement). Transgenic T<sub>1</sub> lines were classified based on the transgene copy number which was previously shown by Adamski et al. (2021) to be highly correlated with *VRT-A2* expression levels in multiple tissues. We phenotyped lines with zero (n = 2 independent events; 5 plants each), low (1-5 transgene copies; n = 4 independent events; 5 plants each) and high (9-35 transgene copies, n = 2 independent events; 5 plants each) transgene copy number. We identified a significant and stepwise increase in the number of RBS with transgenic copy number, from 0.8 RBS (zero copy) to 1.6 RBS (low copy;  $P = 0.078$  vs zero copy) and to 4.3 RBS (high copy;  $P < 0.0001$  vs zero copy) (Figure 2.8D; Table 2.9). The low copy number lines had an average increase of 0.8 RBS with respect to the zero copy number lines, equivalent to the average difference between the *VRT-A2a* and *VRT-A2b* NILs in the field (*VRT2-A2b* effect of +1.1 RBS). The high copy number lines

produced on average 4.3 RBS, which is higher than the *VRT-A2b* NILs and similar to the number of RBS observed in *T. polonicum* ( $3.75 \pm 0.62$  RBS;  $n = 16$  spikes). The dosage-dependent effects observed in the transgenic lines provide further evidence that elevated expression of *VRT-A2* leads to increased number of rudimentary basal spikelets in polyploid wheat.

Table 2.7: Field evaluations for rudimentary basal spikelets (RBS) in *VRT-A2* NILs. Mean number of rudimentary basal spikelets (RBS) measured in field experiments across four years and two locations using the BC4/BC6 NILs in Paragon background carrying the wildtype (*VRT-A2a*) allele or the introgressed allele from *T. polonicum* (*VRT-A2b*). Each value represents the mean RBS of all plots of all blocks; each plot is the mean RBS of 10 spikes. Standard error, N (number of independent sibling lines for which 5 replicates were grown) and the difference in mean RBS number are also shown.

Year	Location	Allele	mean RBS	Std. error	N	Delta RBS	P-value
2017	Morley	<i>VRT-A2a</i>	1.99	0.196	2	0.93	0.0016
2017	Morley	<i>VRT-A2b</i>	2.92	0.16	3		
2019	Morley	<i>VRT-A2a</i>	1.87	0.093	3	1.32	2.6E-10
2019	Morley	<i>VRT-A2b</i>	3.19	0.093	3		
2020	Morley	<i>VRT-A2a</i>	1.28	0.058	4	1.18	3.1E-15
2020	Morley	<i>VRT-A2b</i>	2.46	0.058	4		
2020	Church Farm	<i>VRT-A2a</i>	1.30	0.051	4	0.91	1.4E-12
2020	Church Farm	<i>VRT-A2b</i>	2.21	0.051	4		

Table 2.8: Simplified table of rudimentary basal spikelet phenotypic data from *VRT-A2* transgenic lines (summarised by copy number class). All lines grown under glasshouse conditions. Classification of transgene copy number are based on data from Adamski et al. (2021). Low copy lines carry between 1 and 5 transgene copies, whereas high copy lines carry between 9 and 35 transgene copies. RBS are the average number of rudimentary basal spikelets (RBS) of the main spike.

Copy Number	Mean RBS
zero	0.8
low	1.6
high	4.3

RefSeqv1.1 Gene model RefSeqv1.0 (bp)	TraesCS7A02G170100 126,178,865	TraesCS7A02G173500 127,838,581	TraesCS7A02G174100 128,148,666	TraesCS7A02G174500 128,404,964	TraesCS7A02G175100 128,783,942	Intergenic 128,792,366	TraesCS7A02G175200 128,830,245	intergenic 128,832,107	128,834,372	TraesCS7A02G175400 128,911,919	TraesCS7A02G175600 128,917,566	Rudimentary basal spikelets	Pairwise
Marker name	S11	S12	S13	S14	S15	S16	S17	S18	S19	S23	S24		
R1	Pol	Pol	Paragon	Paragon	Paragon	Paragon	Paragon	Paragon	Paragon	Paragon	Paragon	2.3 ± 0.16	b
R2	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Paragon	4.2 ± 0.16	cd
R3	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Pol	Paragon	4.6 ± 0.16	de
R4	Paragon	Paragon	Paragon	Paragon	Pol	Pol	Pol	Pol	Pol	Pol	Pol	4.5 ± 0.11	d
R5	Paragon	Paragon	Paragon	Paragon	Pol	Pol	Pol	Pol	Pol	Pol	Pol	NA	-
R6	Paragon	Paragon	Paragon	Paragon	Pol	Pol	Pol	Pol	Pol	Pol	Pol	5.2 ± 0.17	e
R7	Paragon	Paragon	Paragon	Paragon	Paragon	Pol	Pol	Pol	Pol	Pol	Pol	4.5 ± 0.16	de
R8	Paragon	Paragon	Paragon	Paragon	Paragon	Paragon	Paragon	Paragon	Pol	Pol	Pol	1.3 ± 0.15	a
										Paragon NILs (BC4)		2.1 ± 0.15	b
										Polonicum NIL (BC4)		3.7 ± 0.15	c
										VRT-A2b effect		1.60	

Table 2.9: Graphical genotype from Adamski et al. (2021) and RBS phenotype of BC6 recombinant inbred lines (RILs) for VRT-A2. Marker names, physical position in RefSeqv1.0 assembly and associated RefSeqv1.1 gene models are indicated and further described in Adamski et al. (2021). Lines were grown at the John Innes Centre (Glasshouses) and phenotyped for rudimentary basal spikelets. Phenotypic values are means of all spikes of 18-20 biological replicates ± standard error of the means. Pairwise (column P) Tukey comparison ( $p > 0.05$ ). The average of the BC4 parental NILs are shown alongside the VRT-A2b effect. Note that R7 and R8 are the critical recombinants that map the RBS trait to the same location as all phenotypes by Adamski et al. (2021); i.e. completely linked to VRT-A2 (TraesCS7A02G175200 in green).

## 2.4 Discussion

### 2.4.1 High-resolution spatial transcriptomics in crops

We hypothesised that the establishment of the lanceolate shape in wheat spikes could be manifested in gene expression differences between the apical, central and basal sections of a developing spike, as has been shown using RT-qPCR for individual genes in wheat (*AP2*; (Debernardi et al., 2017)) and barley (*VRS2*, Youssef et al. (2017)). However, currently available transcriptome data (e.g., Li et al. (2018) and Feng et al. (2017)) lack the spatial resolution *within* each individual developmental stage to answer this question. This focus on ‘between stage’ comparisons (as opposed to within a single stage) is perhaps related to the technical challenges of dissecting and sectioning young meristems. Given the relatively small size of these spike meristems (0.2 mm length at Transition Stage; 3 mm length at Terminal Spikelet stage), RNA-seq methods require bulking of multiple individuals (usually between 30 and 50 different plants) to accumulate enough tissue for a single RNA-seq sample. If one sought to further section each meristem, this would require even further bulking. While laborious, this is achievable; however, under this scenario, the challenge is to properly stage ~100 plants to an equivalent developmental stage. Furthermore, it can be technically challenging to section these young spikes each time into the exact same apical, central and basal sections. Consequently, the spatial resolution in gene expression within a wheat spike at individual developmental stages has remained largely uncharacterised to date.

To address this challenge, we adapted the G&T method for micro-scale spatial-transcriptomics workflow (Macaulay et al., 2015; Giolai et al., 2019), to conduct RNA-seq of the apical, central and basal sections of individual, hand-dissected wheat spikes. This highly-automated workflow requires low tissue input and allowed us to combine 24 Nextera libraries into a single Illumina NovaSeq lane. For 19 out of the 24 samples the method worked successfully, determined by >20,000 expressed genes per library and the clustering among biological replicates. We found that the number of expressed genes per library was on average similar to the number of genes reported for bulked whole spike RNA-seq samples (Feng et al., 2017; Li et al., 2018). This is consistent with the fact that the hand-dissected sections are composed of a large mixture of different tissues (e.g., rachis, spikelet, and floret primordia) and cell types, which in the equivalent maize ears have distinct expression profiles (Xu et al., 2021). Compared to previous bulk RNA-seq studies in developing wheat spikes, the variation observed here (measured as CV) was high among biological replicates (Figure 2.4). This variation is likely caused by both biological variation (e.g., inherent variation of individual plants) and technical variation (e.g. inaccuracies in sectioning and in the developmental staging of the plant/spike) as well as the number of replicates in our analysis. A minimum of six replicates has been proposed for bulked RNA-seq (Schurch et al., 2016). Our results suggest that the RNA-seq from these small sections would benefit from a higher number of biological replicates, which should be feasible considering the high-throughput method employed for RNA extraction

and library preparation, the low tissue input requirement, and the possibility to pool multiple biological replicates per sequencing lane. Despite some limitations, we could identify over 5,000 DEGs between the spatial sections for subsequent functional analysis.

In addition to G&T-Seq, several other technologies have been proposed for obtaining high resolution transcriptional profiles of plant tissues, for example, single cell RNA-seq (McFaline-Figueroa et al., 2020; Rich-Griffin et al., 2020), FACS, and the isolation of nuclei tagged in specific cell types (INTACT). These methods, however, are not spatially resolved as the complete tissue is dissolved into single cells for barcoding or selection (Rich-Griffin et al., 2020). Thus, these current methodologies do not allow, for example, to investigate whether the cell type composition of spikelets differs across the inflorescence. This would only be possible if spikelets were 'harvested' individually, for example through laser capture microdissection (LCM) before dissolving the tissue further into individual cells. Thiel et al. (2021) recently combined LCM followed by RNA-seq of the distinct lower/leaf ridge and upper/spikelet ridge of barley spikes. This allowed them to identify precise spatio-temporal expression patterns of many genes related to architecture and yield in barley spikes with unprecedented resolution. Looking ahead, increased resolution of Spatial Transcriptomics (currently 100  $\mu\text{m}$ ; (Giacomello et al., 2017)), which quantifies full transcriptomes while maintaining tissue integrity, offers the true prospect of direct localisation and quantification of gene expression. Our results argue strongly for the need of these transcriptome-wide and spatially resolved approaches to advance our biological understanding of fundamental developmental processes in plants.

#### 2.4.2 The composite nature of spikes

Early morphological studies of wheat spike development described that the stronger elongation of central spikelets during their initial establishment (glume primordia stage) first causes the lanceolate shape of the wheat spike (Bonnett, 1966). The continuous formation of primordia at the tip of the spike means that at any given growth stage, spikelets in different developmental stages will be present across the spike (Bonnett, 1966). In this study, we detected more differentially expressed genes between the three spatial sections of the spike (apical, central and basal) than between the two investigated developmental stages (Double Ridge and Glume Primordia). We identified 215 DEGs between the two developmental stages, consistent with Li et al. (2018) who identified 206 DEGs between consecutive stages across a time course of six inflorescence development stages. Feng et al. (2017) identified 753 DEGs between the Double Ridge and Floret Primordia stage, which are further apart in development than the stages used in this study. They also detected fewer DEGs when comparing early stages than between more developed spikes. By contrast, we identified 1,315 and 2,438 unique genes to be differentially expressed between the apical and basal section at DR and GP, respectively. The higher number of DEGs between spatial

sections could be due to the developmental gradients occurring in the three spatial sections, which are revealed by the spatial sampling. These differences would be blurred when comparing whole inflorescences between stages due to the mixture of tissue types and spikelets at different developmental stages. A possible improvement for future transcriptome studies could be the collection of only central sections of the developing spikes or complete spatial sampling as conducted here.

The composite nature of the inflorescence tissues has been acknowledged by studies in maize (ears and tassels), where new meristems are initiated in a stepwise manner. Leiboff and Hake (2019) quantified the meristematic tissue composition of maize and sorghum (*Sorghum bicolor*) tassels. For example, maize tassels in the second stage are mainly composed of spikelet pair meristems, but also contain some meristems in spikelet and inflorescence state. They concluded that the changes in these tissue compositions over time corresponded well with the independently staged transcriptional changes of the tassels. Eveland et al. (2014) showed that the range of developmental ages across the maize ear, if acknowledged, can be used as an advantage in RNA-seq studies. They sequenced the tip, middle, and basal sections of 10-mm long ears independently, aiming to analyse the expression patterns in specific developmental meristem types enriched in these sections (inflorescence, spikelet, and floral meristems, respectively). The dissection of the ear therefore allowed them to study gene expression specifically for each meristematic tissue type rather than for all meristem types in intact ears. In this study, we observed that apically expressed genes are enriched for GO-terms related to “shoot system development” and “maintenance of floral organ identity”. This is consistent with the hypothesis that the apical part of the inflorescence is younger and undergoing early phases of spikelet development initiation compared to the central inflorescence section.

#### 2.4.3 Delayed transition of basal spikelets from vegetative to floral developmental programmes

We detected transcriptional gradients across the spike, with the basal section deviating most strongly from the rest of the spike. We noticed that both *SVP* and *CENTRORADIALIS (CEN)* genes remained highly expressed in the basal section of the spike, whereas their expression was lower in the central and apical sections. *In-situ* hybridisation of these genes also showed that their expression is strongest in vegetative tissue and basal spikelets in early spike development (Li et al., 2021). In contrast, *SEP1-4* and *SEP1-5* genes were expressed in the opposite gradient and showed the strongest expression in apical and central sections of the spike at Glume Primordia stage. Recent studies allow us to interpret these gradients in the context of the early steps of vegetative to floral growth transition. In wheat (Li et al., 2021; Adamski et al., 2021; Liu et al., 2021), rice (Sentoku et

al., 2005; Lee et al., 2008), and barley (Trevaskis et al., 2007), *SVPs* have been characterised to be associated with vegetative growth and are downregulated upon floral transition.

In wheat, the double *SVP* mutant *vrt2svp1* leads to the formation of axillary inflorescences (Li et al., 2021). Similarly, overexpression of *CEN-D2* (*TERMINAL FLOWER 1; TFL1-2D*) in wheat extends the duration of the Double Ridge stage (Wang et al., 2017), whereas loss-of-function mutations in barley *CEN* suggest they repress floral development under short-day conditions (Bi et al., 2019). Double knockout mutants of the MADS-box *SQUAMOSA* genes *vrn1ful2* highlighted that these two genes act as transcriptional repressors of *SVP* and *CEN* genes in early wheat spike development (Li et al., 2019). Furthermore, through a series of genetic and biochemical studies, Li et al. (2021) showed that the downregulation of *SVP* genes is necessary for the formation of flowering promoting MADS-box protein complexes including VRN1, FUL2 and SEP proteins. Hence the coordinated downregulation of *SVPs*, and possibly *CEN* genes, along with the upregulation of *SEP* genes is required for normal floral transition and spikelet development in wheat. Previous studies in rice have found similar expression patterns, as well as mutant effects, of *SVPs* and *SEPs* suggesting a conserved function in flowering transition across the two species (Ren et al., 2016; Wu et al., 2018).

Based on our results, the floral developmental programme across the wheat spike appears to be most advanced in its apical and central sections, while being delayed in the basal sections. We hypothesise that this is due to elevated *VRT2* expression at the base of the spike, which hinders the progression of the flowering programme via *SEP* class flowering genes. Likewise, the higher expression levels of the wheat *CEN2* and *CEN5* homologs at the base are consistent with a delay in floral transition that could interfere with the development of the spikelet primordia. Therefore, although the basal spikelet primordia are initiated first chronologically, their developmental age in terms of the floral programme is delayed with respect to the more recently formed central and apical spikelet primordia. This could explain in part why the spikelet primordia in the basal region of the spike elongate less and develop slower than central spikelets despite being initiated first (Bonnett, 1966). Likewise, the less advanced floral developmental programme could also explain why the overexpression of *SVPs* in barley (*HvBM10*) leads to complete floral reversion in basal but not apical spikelets (Trevaskis et al., 2007).

We hypothesise that *SVPs* need to be downregulated upon floral transition to allow timely establishment and progression of the early spikelet primordia. Failure to do so would delay their development and result in their final rudimentary shape in the mature spike. In line with this hypothesis, we observed increased RBS in genotypes with prolonged and increased *VRT2* expression in a dosage-dependent manner. In our RT-qPCR data, we also observe increased expression of *SVPs* alongside reduced expression of *SEPs* in *VRT-A2b* lines at Double Ridge and Glume Primordia stage. However, we cannot exclude the possibility that the increase in *VRT2*

expression in the *VRT-A2b* lines could also affect basal spikelet development at a later stage of spike developmental (e.g., from Terminal Spikelet stage to anthesis). We are currently using quantitative live imaging to compare cellular growth dynamics of spikelets at different stages of spike development between the *VRT2* NILs.

The finding that the expected downregulation of *SVPs* and *CENs* does not follow the chronological age of the tissues suggests that other gradients across the spike might influence spikelet development. Debernardi et al. (2017) showed that in tetraploid wheat *AP2-5* and *miR172* have consistent and opposing expression gradients across the spike at three consecutive developmental stages. The persistent expression gradient of *AP2-5* supports the idea that expression patterns across the spike, beyond the ones caused by age differences of spikelets, exist. Furthermore, they proposed a model illustrating that the phenotypic effect of mutants across the spike differs due to the existing gradient of expression of this gene (Debernardi et al., 2017). Other examples of mutants with different phenotypic effects across an inflorescence include *vrn1ful2* (Li et al., 2019) in wheat, *tassel sheath1* (*tsh1*, (Whipple et al., 2010)) and *ramosa2* (*ra2*, (Bortiri et al., 2006)) in maize, *SEPALLATA* double mutant *Osmads5Osmads34* in rice (Zhu et al., 2022), as well as *many noded dwarf1* (*mnd1*, (Walla et al., 2020)), *frizzy panicle* (Poursarebani et al., 2015), and *vrs2* in barley, which was also found to be consistently differentially expressed across the spike (Youssef et al., 2017). *VRS2* has been shown to maintain a basal to apical expression pattern across three, post awn initiation developmental timepoints in barley (Youssef et al., 2017). The study of *vrs2* mutants revealed that *VRS2* is furthermore engaged with the basal-apical patterns of auxin, cytokinin, and gibberellin across the spike. While hormonal gradients across the spike in early development have not been studied in great detail in wheat, they have been shown to play crucial roles in floral induction and development in *Arabidopsis* (Reinhardt et al., 2000). Their patterns across the spike should be investigated in future studies addressing developmental differences across the spike.

#### 2.4.4 A model for the regulation of leaf and spikelet ridge outgrowth in the base of the spike

Recently, Meir et al. (2021) proposed that in shoot apical meristems of tomato (*Solanum lycopersicum*), similar to processes during embryonic development, transient programmes are required to inhibit a preceding setup (i.e. vegetative growth), before a new developmental program (flowering) can be initiated. We propose that the altered gene expression and development of the basal spikelets could be a consequence of their initiation during the transient phase between vegetative and floral network shifts and thus being exposed to mixed signals of development. Upon floral transition, the lower (leaf) ridge is suppressed, while the growth of spikelet ridges from the previously suppressed axillary meristems is activated. Development of lower ridges subtending all branching events is suppressed in grass inflorescences upon flowering transition (Whipple et al.,

2010). Li et al. (2019) noticed that this suppression was disrupted in the double *vrn1ful2* and triple *vrn1ful2ful3*-null mutants, which fail to down-regulate *SVP* genes. In these mutants, the upper spikelet meristems generate vegetative structures resembling tillers that are subtended by bracts or leaves originating from the lower leaf ridge.

We observed that genes that were highly expressed in the basal section of the inflorescence (cluster 4) have previously been shown to be expressed specifically in the lower/bract ridge and before or at vegetative to floral transition. This is also supported by the GO-term enrichment of photosynthesis related terms in cluster 4. Our tissue sections do not allow us to distinguish lower and upper ridge tissues, however, the two ridges have been separately collected and sequenced via LCM in barley (Thiel et al., 2021). In this barley dataset, we found a higher expression of *HvVRT2* (*HORVU7Hr1G036130*) and *FLOWERING LOCUS C* (*HvFLC*; *HORVU7Hr1G054320*) in the lower ridge compared to the upper ridge, whereas *HvSVP1* (*HORVU6Hr1G077300*) was also marginally more highly expressed (Figure 2.10). Furthermore, the barley *MND1* gene (*HORVU7Hr1G113480*) has recently been shown to be expressed in leaf primordia and during the Double Ridge stage in the basal region of the spike in barley (Walla et al., 2021), while it is most highly expressed in the vegetative meristem and lower/leaf ridge in the LCM data (Thiel et al., 2021). We observed that in our data, the wheat *MND1* putative orthologs (*TraesCS7A02G506400*, *TraesCS7B02G413900*, *TraesCS7D02G494500*) were significantly more highly expressed in the basal section than the apical section at both DR and GP stage (Supplemental Table S2, original publication). The suppressed leaf ridge (or bract) has been proposed to act as a signalling centre, regulating the fate of the upper spikelet meristem ridge (Whipple, 2017). Insufficient bract suppression during the formation of the basal spikelets might therefore negatively affect initiation and development of spikelets.

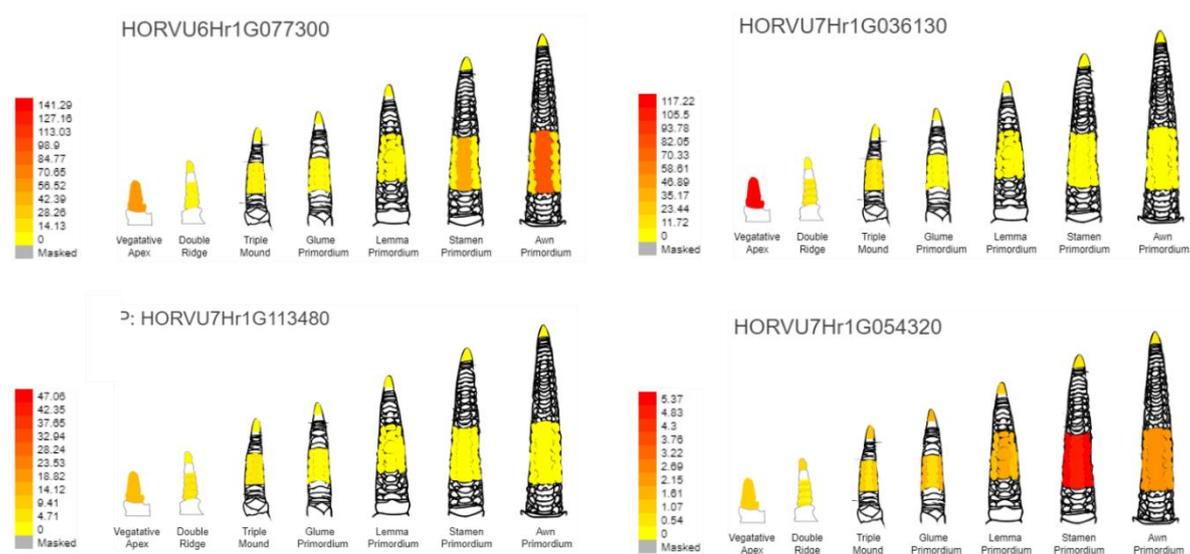


Figure 2.10: Expression of barley genes in Thiel et al. (2021), which are putative orthologous to wheat genes highly expressed in basal spike sections. Expression of barley genes *HvSVP1* (*HORVU6Hr1G077300*), *HvVRT2* (*HORVU7Hr1G036130*), *HvMND1* (*HORVU7Hr1G113480*), and *HvFLC* (*HORVU7Hr1G054320*). The barley genes are expressed more strongly in the lower leaf ridge than the upper. Data from Thiel et al. (2021) and visualized in the barley eFP gene expression browser ([http://bar.utoronto.ca/eplant\\_barley/](http://bar.utoronto.ca/eplant_barley/)). Scale bar = TPM (transcript per million)

At DR, the widest point of the spike is indeed as expected the base and not the central section (Figure 2.1A). The lower ridge is however much less developed in the central section and can be hardly seen in the apical ridges. Interestingly, mutants failing to repress the lower ridge growth, such as *third outer glume1 (trd1*; (Houston et al., 2012)), the barley ortholog of maize *tsh1*, develop large bracts from the lower ridge in basal spikelets, unlike apical spikelets, which do not develop bracts from their lower ridges regardless of the absence of *TRD*. This is reminiscent of the gradient in the strength of the phenotypic effects observed from the top to the base of the inflorescence in multiple *Poaceae* mutants (discussed above). We therefore hypothesise that the basal meristems develop into smaller spikelets and larger bract primordia due to a slow suppression of “vegetative growth signals” (e.g., *SVPs*) and a concomitant slow upregulation of “floral growth signals” (e.g., *SEPs*) upon floral transition. To investigate how a change from vegetative to floral signalling might affect the development of individual meristems, we modelled the genetic interaction of *SVPs* and *SEPs*, as proposed by Li et al. (2021), in the spatial context of a growing spike (Figure 2.11, the model can be accessed via the interactive web-version of the model (Supplemental File S1 in publication)). Under the assumptions that *SVP* suppresses *SEP* expression, *SVP* expression is downregulated upon flowering, and that *SEP* promotes spikelet outgrowth, the model could recapitulate (a) the observed opposing gradients in expression of *SVPs* and *SEPs* along the spike, and (b) the formation of a lanceolate shaped wheat spike with reduced spikelet elongation and stronger bract growth in the most basal spikelets. Thus, whilst this hypothesis will require further investigation and testing, modelling supports its plausibility.

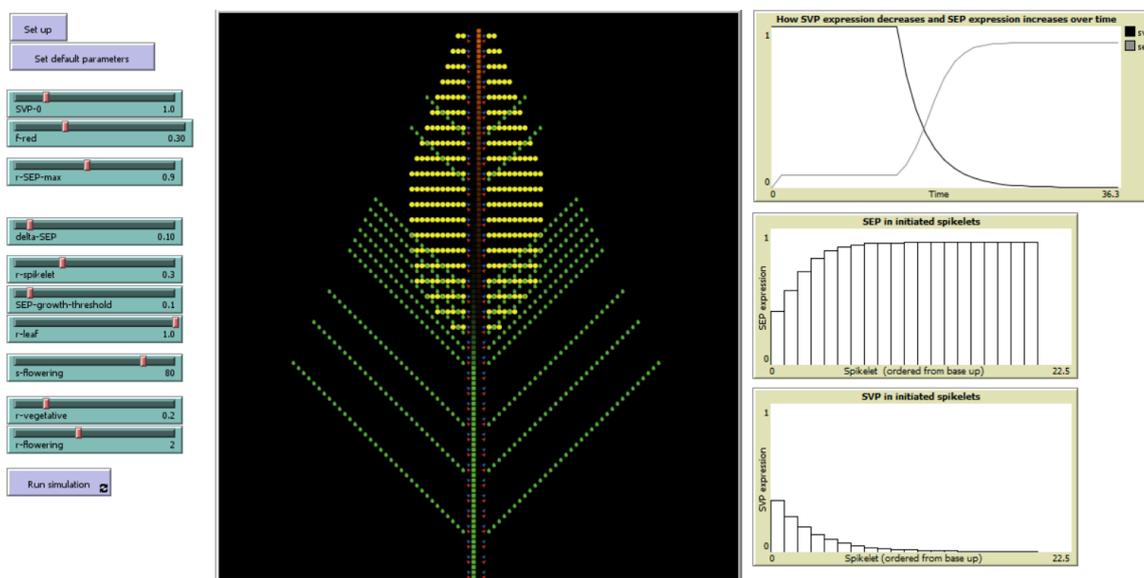


Figure 2.11: Example of Netlogo simulation outcome with default parameters. Green = leaf tissue, Yellow = spike tissue, red = meristem initiation point committed to leaf development. Blue = axillary meristem committed to spikelet/branch. The model uses expression of *SEP* and *SVP* class genes to predict when meristems (red) produce leaf tissue (green) and when they switch to producing spike tissue (yellow). See methods for model parameters and use. Top-right hand graph shows the decrease of *SVP* expression and the increase of *SEP* expression over time. The middle and bottom graphs depict the gradients of *SEP* and *SVP* expression, respectively, from the basal to the apical spikelets. The modelled growth (central graph) shows how basal spikelets are smaller than their central counterparts and have leaf ridge growth due to the higher *SVP* and lower *SEP* expression.

## 2.5 Materials and Methods

### 2.5.1 Plant materials

Hexaploid wheat (*Triticum aestivum*) germplasm used in this study includes wildtype hexaploid wheat cultivar Paragon and P1/VRT2 germplasm described in Adamski et al. (2021) including P1 NILs, recombinants, and T<sub>1</sub> transgenic lines carrying the *T. polonicum* VRT-A2b copy under the native promoter. *T. polonicum* accession T1100002 was obtained from the John Innes Centre Germplasm Resources Unit (<https://www.seedstor.ac.uk/search-infoaccession.php?idPlant=27422>). For field experiments, we used between two to four sibling BC<sub>4</sub>/BC<sub>6</sub> NILs differing for the VRT-A2b allele.

### 2.5.2 Low input RNA sequencing

Paragon seedlings were grown in a single batch in a controlled environment growth chambers in 24-cell seed trays under long-day (16 h light/8 h dark) photoperiods at 300  $\mu\text{mol m}^{-2} \text{s}^{-1}$ , with a day temperature of 20 °C and a night temperature of 15 °C. Inflorescences for Double Ridge (DR) stage were collected 18 days after sowing, while inflorescences for Glume Primordia (GP) stage were collected 22 days after sowing. All plants were grown in “John Innes Cereal Mix” (40% Medium Grade Peat, 40% Sterilized Soil, 20% Horticultural Grit, 1.3  $\text{kg}\cdot\text{m}^{-3}$  PG Mix 14-16-18 + Te Base Fertiliser, 1  $\text{kg}\cdot\text{m}^{-3}$  Osmocote Mini 16-8-11 2 mg + Te 0.02% B, Wetting Agent, 3  $\text{kg}\cdot\text{m}^{-3}$  Maglime, 300  $\text{g}\cdot\text{m}^{-3}$  Exemptor).

Four individual spikes per developmental stage (DR and GP) were dissected into apical, central, and basal sections (1:1:1 ratio) using a stereo microscope (Leica MZ16). Sections were immediately placed into 96-well plates (on ice) containing 10  $\mu\text{L}$  of RLT plus (Qiagen, Hilden, Germany). All instruments and surfaces were cleaned with 80% v/v ethanol, RNase-free water and lastly RNase-out solution after each sample to reduce cross-contamination and RNA degradation. Samples were stored at -80 °C until cDNA preparation, using the G&T-seq method as previously described (Macaulay et al., 2015). cDNA was normalised to 0.2 ng/  $\mu\text{L}$  before Nextera (Illumina, San Diego, CA, USA) library preparation using a Mosquito HV liquid handler (STP, Royston, UK) in a total reaction volume of 4  $\mu\text{L}$  as described in Mora-Castilla et al. (2016). Libraries were pooled by volume and sequenced on a single lane of a NovaSeq 6000 (NVS200S2 flow cell, 100 bp paired-end reads). The raw RNA-seq read libraries used in this study are available from NCBI BioProject PRJNA749586.

### 2.5.3 Bioinformatic analysis

For the RNA-seq analysis, we used the RefSeqv1.0 genome assembly and the RefSeqv1.1 gene annotation ([https://urgi.versailles.inra.fr/download/iwgsc/IWGSC\\_RefSeq\\_Assemblies/v1.0/](https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Assemblies/v1.0/); IWGSC et al. (2018)). Reads were trimmed and adapters were removed using trim-galore v.0.4.2

([https://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)) with settings: “--paired --fastqc --a GGTATCAACGCAGAGT --clip\_R1 20 --clip\_R2 20 --trim-n”. Minimum length of reads retained was set to 50 bp. Reads were aligned to the RefSeqv1.0 genome assembly using HISAT2 v. 2.1.0 (<https://daehwankimlab.github.io/hisat2/>; (Kim et al., 2019)) with the following parameters: “--pen-noncansplice 20 --mp 1,0 --rna-strandness RF”. Alignment files were converted to BAM format, sorted, indexed, filtered, and purged of all none-primary alignments (0x100 flag) using samtools (v. 1.9; Li et al. (2009)). HTSeq v.0.6.1 (<https://htseq.readthedocs.io/en/master/>; (Anders et al., 2015)) was used to count the read numbers mapped to the RefSeqv1.0 gene models.

HT-read count normalization and differential expression analyses were performed using the DESeq2 v.1.28.1 R packages (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>; (Love et al., 2014); RStudio 1.2.5001). Genes with an average expression below 10 HT-count, and which were not expressed (i.e.  $\leq 10$  HT counts) in at least three libraries, were removed from the analysis. Correlation between expressed genes and Waddington stage and/or section was tested by ANOVA. Raw read data from Li et al. (2018) and Feng et al. (2017) were pseudo-aligned using Kallisto Sleuth pipeline (<https://scilifelab.github.io/courses/rnaseq/labs/kallisto>) and the coefficient of variation was calculated for each gene (by condition) using R (RStudio 1.2.5001) dplyr (plyr 1.8.6). Differentially expressed genes (DEGs) between the two Waddington stages (DR and GP) were calculated with the design “~plant + waddington”, while DEGs between the three sections (apical, central, basal) were determined using the design “~section + waddington:plant + waddington”. DEGs among the sections within each Waddington stage were determined with the design “~plant + section”. For each gene, an adjusted *P*-value was computed by DESeq2 (using the Benjamini and Hochberg method (Benjamini and Hochberg, 1995)), and those with an adjusted *P*-value of  $\leq 0.05$  were considered differentially expressed. DESeq2 also computed Log2FoldChanges as well as the associated uncertainty (lfcSE, see (Love et al., 2014) for further detail). The “contrast” function was used to determine pairwise comparison *P*-values. The full set of expression data and comparisons is presented in Supplemental Dataset S1 (that can be accessed via the original publication (Backhaus et al., 2022)). Enrichment of GO-terms was performed using the online tool “PLAZA” (<https://bioinformatics.psb.ugent.be/plaza>; Van Bel et al. (2017)) using the recommended settings, and all enriched GO-terms of Biological function (BF) and Cellular Compartment (CC) were retained. In brief, PLAZA determines the overrepresentation of a certain GO-term in a gene set compared to the genome-wide background frequency (= all expressed genes in this experiment; submitted manually). The significance of over- or underrepresentation is determined using the hypergeometric distribution and the Bonferroni method is applied to correct for multiple testing. Note that enrichment folds are reported in log2 fold scale. Enrichment of TF families (Genes that were annotated as TFs were obtained from [https://opendata.earlham.ac.uk/wheat/under\\_license/toronto/Ramirez-Gonzalez\\_etal\\_2018-06025-Transcriptome-Landscape/data/data\\_tables/](https://opendata.earlham.ac.uk/wheat/under_license/toronto/Ramirez-Gonzalez_etal_2018-06025-Transcriptome-Landscape/data/data_tables/) (Ramirez-Gonzalez et al.,

2018) and MADS-box TFs (based on Schilling et al. (2020)) was performed in R using the `phyper()` function from `stats` package v.4.0.1 to test for Hypergeometric Distribution. All DEGs were scaled and centred using R-base function “`scale`”. All cluster analysis was performed on scaled data using R (`stats`) functions `kmeans` and `hclust`, followed by visualisation through `pheatmap` v.1.0.12 (<https://cran.rstudio.com/web/packages/pheatmap/index.html>). Correlation to centroid cluster shape of each gene expression pattern was calculated using the “`cor`” function from R `stats`.

#### 2.5.4 Reverse transcription quantitative PCR analysis

*P1* NILs were grown in controlled growth chambers in 24-cell seed trays under the same conditions as used in the low input RNA-seq experiment (see above). For each biological replicate, we pooled 30 inflorescences for DR stage, 15 for GP stage, and nine for Terminal Spikelet stage ( $n = 4$  biological replicates per stage). Inflorescences from NILs were dissected using a stereo microscope (Leica MZ16). Inflorescences were dissected into apical, central, basal and peduncle sections (1:1:1:1 ratio). At Double Ridge stage, inflorescences were only dissected into apical, basal and peduncle section as the inflorescences were too small to be accurately dissected into four sections for all 30 plants per biological replicate. Each section was immediately placed into 1.5-mL tubes on dry ice and tubes were snap frozen in liquid nitrogen as soon as all plants for the sample were collected. Samples were stored at  $-80^{\circ}\text{C}$  until needed. Inflorescences were collected within 2-3 hours, 9 hours after the lights came on in the growth chamber. Tissue was homogenized in a TissueLyser II (Cat No.: 85300, QIAGEN) using 3-mm steel beads (Cat No.: 69997, Qiagen); tubes were shaken for 20-s at 28 Hz with dry ice.

All RNA extractions were performed using the RNeasy Plant Mini Kit (Cat No.: 74904, Qiagen) with RLT buffer according to the manufacturer’s protocol followed by RNA ethanol precipitation ([https://projects.iq.harvard.edu/files/hl1lab/files/ethanol-precipitation-of-rna\\_hla.pdf](https://projects.iq.harvard.edu/files/hl1lab/files/ethanol-precipitation-of-rna_hla.pdf)). DNA digestion was performed using the RQ1 RNase-free DNase set (Cat No.: M6101, Promega) according to the manufacturer’s protocol. RNA was reverse transcribed using M-MLV reverse transcriptase (Cat No.: 28025013, Thermofisher) according to the manufacturer’s protocol. For the RT-qPCR reactions, LightCycler 480 SYBR Green I Master Mix (Roche Applied Science, UK) was used according to the manufacturer’s protocol. The reactions were run in a LightCycler 480 instrument (Roche Applied Science, UK) under the following conditions: 10 min at  $95^{\circ}\text{C}$ ; 40 cycles of 10 sec at  $95^{\circ}\text{C}$ , 15 sec at  $62^{\circ}\text{C}$ , 30 sec at  $72^{\circ}\text{C}$ ; dissociation curve from  $60^{\circ}\text{C}$  to  $95^{\circ}\text{C}$  to confirm primer specificity. All reactions were performed with three technical replicates per sample and using *Actin* as the reference gene (Uauy et al., 2006). Relative gene expression was calculated using the  $2^{-\Delta\Delta\text{Ct}}$  method (Livak and Schmittgen, 2001) with a common calibrator so that values are comparable across genes, tissues, and developmental stages. All primers used in RT-PqCR came from Adamski et al. (2021) can be found in Table 2.10.

Table 2.10: List of primers used in qRT-qPCR in this chapter. Primer sequence, amplicon size, annealing temperature, and primer use are listed.

Primer	Primer sequence 5'-3'	Amplicon size	Annealing temperature	Primer use
VRT-A2_F	CCGGCAATTCATGCAACAAATT	216 bp	62°C	qRT-PCR
VRT-A2_R	GAACCGTCATCATTGTCCTGT			
SEP1-4_F	GGCGACAAAGAGCCAACAGT	81bp	66°C	qRT-PCR
SEP1-4_R	TCCAACATCCTGGCAAGACA			
Actin-Fwd	ACCTTCAGTTGCCAGCAAT	91 bp	62°C	qRT-PCR
Actin-Rev	CAGAGTCGAGCACAAATACCAGTTG			

All RT-qPCR data was normalised using a log<sub>2</sub> transformation. A three-way ANOVA including Waddington stage, section, and genotype yielded significant two-way interactions. The differences between sections of the genotypes were therefore further analysed individually for each Waddington stage and genotype. For each of the two genotypes we individually performed Tukey multiple comparison tests to determine differences between the sections within each developmental stage by Tukey multiple comparison test. Differences between the two genotypes were also analysed individually for each Waddington stage. Furthermore, the differences between the genotypes were investigated individually for each section within the Waddington stage if the interaction term was significant (in GP and TS). For all analysis see Table 2.6.

#### 2.5.5 Field experiments and phenotyping

VRT-A2 NILs were evaluated in four field experiments. Three trials were located at The Morley Agricultural Foundation trials site, Morley St Botolph, UK (52°33'15.1"N 1°01'59.2"E) in 2017, 2018 and 2020 and one trial was sown in 2020 at the John Innes Experimental trials site in Norwich, UK (52°37'50.7"N 1°10'39.7"E). In Morley (2017) we analysed two BC<sub>4</sub> lines of VRT-A2a and three BC<sub>4</sub> lines of VRT-A2b. In Morley (2019) we analysed two BC<sub>6</sub> and one BC<sub>4</sub> line per VRT-A2 allele and in Morley and Church Farm 2020 we analysed two BC<sub>6</sub> and two BC<sub>4</sub> lines for each VRT-A2 allele. All experiments were drilled as yield-scale plots (6 m x 1.2 m) and sown by grain number for comparable plant densities aiming for 275 seeds m<sup>-2</sup>. The trials were arranged in a randomised complete block design (RCBD) with five replicates per sibling line per location. A 10-ear grab sample was collected from each plot pre-harvest for the assessment of rudimentary basal spikelet (RBS) numbers and other phenotypes (recorded in Adamski et al., 2021). RBS were defined as spikelets carrying no grain at maturity and counted for each spike individually. To determine the differences between the *P1<sup>POL</sup>* and *P1<sup>WT</sup>* NILs, we performed analysis of variance (ANOVA) on the multiple field trials phenotypic data. For the analysis of individual trials, we used a two-way ANOVA including Genotype + Block performed in R ('car' package version 3.0-10; RStudio 1.2.5001).

### 2.5.6 Glasshouse phenotyping

We evaluated the BC<sub>4</sub> NILs and BC<sub>6</sub>F<sub>3</sub> recombinant lines, as well as *T. polonicum* accession T1100002, under standard glasshouse conditions. 18-20 plants per genotype were grown in 1 L pots containing John Innes Cereal Mix under long day conditions (16 h light, 8 h dark). The genotypes of all plants were confirmed using KASP marker *SP1Pol* (Adamski et al., 2021). We counted the number of rudimentary basal spikelets (RBS) for all tillers of all biological replicates at maturity. To evaluate the differences in RBS between genotypes, we performed a two-way ANOVA analysis and post-hoc multi-pairwise comparisons Sidak test ('car' package version 3.0-10; RStudio 1.2.5001).

### 2.5.7 Phenotyping of transgenic lines

T<sub>1</sub> lines from Adamski et al. (2021) differing for the copy number of the *VRT-A2b* transgenic construct (zero = 0 copies; low = 1-5 copies; high = 9-35 copies) were grown in 1 L pots with John Innes Cereal Mix under 16 h light at 20°C and 8 h dark at 15°C in controlled environment growth chambers. We measured RBS number for the main tiller of all plants at maturity. To determine differences in RBS between the three transgenic classes, we performed analysis of variance (two-way ANOVA; 'car' package version 3.0-10). We performed Dunnett tests to compare the low and high copy lines against the zero copy number controls (RStudio 1.2.5001).

### 2.5.8 Modelling

The computational model of wheat spike shape formation was developed using the multi-agent programming language and modelling environment, Netlogo (Wilensky, 1999). Gene interactions were modelled as previously described (Li et al., 2021). The model can be accessed via the interactive web-version of the model (Supplemental File S1, in publication).

In brief, both spikelets and leaves are initiated with rates that depend on the levels of SEP (*SEPALLATA*). Leaf initiation rates are suppressed by SEP, whereas spikelet initiation requires SEP. The maximum initiation rates are the same for both spikelets and leaves but different before ( $r_{vegetative}$ ) and after ( $r_{flowering}$ ) flowering. Once initiated, the leaves and spikelets grow at a rate defined by the parameters  $r_{leaf}$  and  $r_{spikelet}$ , respectively. Leaf growth does not depend on SVP (*SHORT VEGETATIVE PHASE*) or SEP levels, whereas spikelets only increase in size every iteration if their SEP level is above a given threshold ( $SEP_{growth\_threshold}$ ). Expression of both SVP and SEP only occurs at meristem initiation. After this, the levels of SVP and SEP cannot increase, although SEP is degraded. SVP is not degraded, solely because at this point, nothing is dependent on SVP levels, whilst spikelet growth depends upon SEP levels.

*SVP* expression rates start to decrease, once flowering is triggered, according to:

$$r_{SVP}(t + 1) = r_{SVP}(t)f_{red}$$

where  $r_{SVP}(t)$  is the rate of *SVP* expression at that time step, and  $f_{red}$  is a rate reduction factor.

*SEP* expression depends upon the levels of *SVP* in the meristem in which the initiation points are located, depending on a Hill function (Alon, 2007),

$$r_{SEP}(SVP) = r_{SEP,max} \left( \frac{K_D^n}{K_D^n + SVP^n} \right),$$

where  $r_{SEP,max}$  is the maximum rate of *SEP* expression,  $K_D$  is the binding constant, and  $n$  is the Hill coefficient. The resulting curves for *SVP* and *SEP* expression are shown in Figure 2.11.

*SVP* levels are initiated with the current value of  $r_{SVP}$ . *SEP* levels are initiated using  $r_{SEP}$ , and reduce by degradation rate,  $\delta_{SEP}$ , following

$$\frac{dSEP}{dt} = SEP \cdot \delta_{SEP}$$

## 3 Tracing spikelet primordia growth using MorphoGraphX

### 3.1 Summary

In all plants, above-ground organs are formed on the flanks of the shoot apical meristem (SAM). In flowering plants, the SAM usually initiates exclusively leaf primordia after germination and undergoes transition to floral primordia development after receiving the signal to do so. In grasses, floral meristems can be either formed directly on the main rachis or they are developed on primary/secondary branches. The number, positioning and size of floral meristems are thus controlling the architecture of grass inflorescences. Quantification of how wheat meristems develop, however, does not exist. Here, we adapted the protocol for live cell imaging and subsequent analysis in MorphoGraphX for the wheat spike meristem. MorphoGraphX is a powerful computational tool to study the growth and division of cells in plant organs. However, it has mostly been used to study the development of *Arabidopsis* and its application to other species was limited. We showed that wheat meristems can survive and be imaged *ex situ* when grown on an adapted media, containing gibberellic acid. Furthermore, we present the proof-of-concept heatmap, which demonstrates that cell growth in the developing wheat spike meristem can be analysed using our protocol. We propose to use this method to analyse the growth of the lower leaf ridge in *THIRD OUTER GLUME1* mutants, a gene known to inhibit leaf ridge growth in other grasses. Finally, we will employ this method in the future to determine if over-expression of *SVP* negatively affects spikelet ridge establishment. The quantification of growth in wheat using MorphoGraphX will bring us one step closer to understanding early spike growth and developmental at a cellular level in this major crop.

### 3.2 Introduction

The architecture of plants is the result of the activity and identity of primordia formation by the root and shoot apical meristem (Pautler et al., 2013). The shoot apical meristem (SAM, also termed apex) is responsible for the formation of all above-ground organs. In the centre of the SAM, a zone of undifferentiated pluripotent stem cells is maintained, which are consistently dividing. On the flanks of the SAM, new primordia are initiated and develop into mature organs. All plant-based food products we consume are therefore the products of a shoot or root apical meristem.

The edible grains of the major staple crop wheat (*Triticum aestivum* L.) are produced in specialized branching structures termed spikelets. Each spikelet is directly attached to the main rachis of the inflorescence, termed spike, and contains multiple florets, which when fertilised become grains (Bonnett, 1966). The spikelets are formed early during plant development, just after the apical meristem has transitioned from the vegetative to reproductive development phase. Initially, the apex elongates and enters a stage termed double ridge. During this stage, primordia pairs are initiated in a distichous pattern along the elongating wheat apex. The upper ridge is the axillary

meristem (AM), which obtains spikelet identity. Each spikelet primordium is subtended by a suppressed leaf primordium (the lower ridge). Spikelet initiation continues until the apex tip, containing the SAM, forms the terminal spikelet.

Multiple mutants have been identified that cause variations in spike architecture by altering the number or determinacy of these early primordia. For example, deletion of *THIRD OUTER GLUME 1* (*TRD1*) in grasses leads to the formation of bract leaves below the spikelets of mature spikes in barley (Houston et al., 2012), rice (Wang et al., 2009) and maize (Whipple et al., 2010). *TRD1* is the major suppressor of the lower leaf ridge in grass inflorescences. In *trd1* mutants, the lower leaf ridge continues to grow beyond the double ridge stage, leading to a more leaf-like appearance of these ridges (Houston et al., 2012). In maize and barley, scanning electron micrographs have confirmed qualitatively that the lower ridges in *trd1* mutants appear larger in size, especially in the basal spikelets. However, this phenotype has not been further quantified and it remains unclear when and how *trd1* affects cell growth.

Furthermore, we showed in a previous study that the introgression of the *VEGETATIVE TO REPRODUCTIVE TRANSITION 2* (*VRT-A2*) allele from *Triticum turgidum* ssp. *polonicum* (*VRT-A2b*) leads to increased rudimentary basal spikelet formation (Backhaus et al., 2022). In wheat, the most basal spikelets are commonly very small and produce no or few grains, even though they are initiated first during development. *VRT2* is a known suppressor of floral development and is downregulated upon the transition from vegetative to floral development in wildtype plants. The *VRT-A2b* allele leads to increased expression during the vegetative phase and ectopic expression in the spike, especially in the glumes, due to a deletion and rearrangement in the first intron of the gene (Adamski et al., 2021). We hypothesized that basal spikelets are rudimentary and preferentially aborted during spike development because they are delayed in development immediately after formation. However, no phenotypic effect on basal spikelet primordia initiation or growth by *VRT-A2b* has been established yet. This might be partially due to the relatively small effect of the *VRT-A2b* allele and the limitations of current imaging techniques.

Currently, the developing apical meristem can be either imaged using light microscopy or scanning electron-microscopy (SEM). After dissection of the tissue and removal of all leaves that are covering the apex, both techniques can be used to capture the morphology and size of the developing spike. Light microscopy has much lower resolution but is a relatively cost effective, simple and a high-throughput methodology. The resulting micrographs can be used to determine spike length, age and developmental stage (Waddington et al., 1983). SEM is a useful technique to obtain similar information, however at a much higher resolution, which facilitates earlier and more accurate recognition of primordia and has been, for example, used to reveal the early growth defects in *trd1* mutants (Houston et al., 2012). Both imaging techniques require the dissection of the spike, which subsequently dies. This precludes imaging the same spike multiple times during development and

growth. The growth of a spike between two timepoints can thus only be estimated by imaging multiple, different individuals at subsequent timepoint.

In *Arabidopsis*, major discoveries have been achieved by the live imaging of meristems and subsequent cell lineage tracking using MorphoGraphX (Barbier de Reuille et al., 2015). MorphoGraphX is a software developed to use the confocal image stacks from a timeseries, find corresponding cells between the timepoints, and use this information to analyse cell growth, division, and more. MorphoGraphX delivers its most informative results from live-cell images and this depends on the survival of the tissue of interest *ex situ*. To our knowledge, MorphoGraphX has not been applied to any grass species, or monocot, to track spike or floral growth using live imaging. Most likely, the inability to cultivate floral tissue of these species *ex situ* has been a major hindrance. While *Arabidopsis* meristems, as well as the vegetative meristem of grasses, can be cultivated on simple media plates (containing sucrose) this is not possible for floral tissues which cease to develop and show visible signs of cell death when cultivated on the same media. For *ex situ* cultivation of other plant organs (such as embryos or seedlings) the fine-tuning of hormone concentrations has played an important role to achieve successful cultivation (Hayta et al., 2021). Min et al (2022) also adapted the growth media to contain zeatin (cytokinin hormone) and gibberellic acid (GA3) in their attempts to image *Aquilegia* primordia growth.

The aim of this chapter was to quantify meristematic growth in wheat. We applied MorphoGraphX to understand early developmental growth effects of mutants with altered spike architecture. We found that the addition of gibberellic acid (GA3) enables *ex situ* survival of wheat meristems. We further adapted the MorphoGraphX protocol to wheat and present here a proof-of-concept growth map, tracking the growth of basal wheat spikelets and leaf ridges growing over a 48-hour time window. We tested the protocol for barley and found that under the same conditions, barley floral meristems do not develop *ex situ* and that molecular differences in flowering control must exist between the two closely related grass species. Furthermore, we developed wheat *TRD1* mutants and found that its known role in leaf ridge suppression in barley, maize and rice is also conserved in wheat. We are now able to trace the developmental differences of the early wheat spike primordia in *TRD1* and *VRT2* mutants using MorphoGraphX.

### 3.3 Results

#### 3.3.1 Development of wheat inflorescence meristem *ex situ*

To track cell growth and division of wheat (*Triticum aestivum* L.) inflorescence primordia, we adapted the protocol developed by Kierzkowski et al. (2012) for the analysis of tomato meristem growth using MorphoGraphX (Barbier de Reuille et al., 2015; Strauss et al., 2022). To image plant primordia multiple times during a developmental timeframe, the tissue of interest cannot be

covered by any other organs of the plant, such as leaves, that might obstruct confocal imaging. While previous studies have found that *Arabidopsis* and tomato primordia can be dissected and cultivated on simple agar media (Murashige and Skoog (MS) media plus 1% sucrose for *Arabidopsis* (Kierzkowski et al., 2012; Barbier de Reuille et al., 2015)), we found that wheat floral primordia fail to develop on this medium. In contrast, vegetative wheat meristems transferred to MS or Linsmaier and Skoog (LS) media (plus sucrose) did continue to develop on the media and grow existing and new leaf primordia (data not shown). To optimise the growth conditions of the floral wheat meristems, we changed sugar sources and added plant growth hormones to the media. We hypothesised that gibberellins may be important to induce floral development because of their flowering inducing role in wheat (Razumov et al., 1960; Evans et al., 1995; Pearce et al., 2013) and other species (Lang, 1957). Indeed, we found that the addition of bioactive gibberellic acid (GA3) is vital for the growth and development of wheat floral tissue *ex situ* (Figure 3.1A). We chose a final concentration of 0.02 mg/L GA3 in the media, as we did not observe any difference in meristem growth on media containing 0.01 to 0.08 GA3 mg/L. Future experiments testing the effect of more extreme variations in GA3 concentration on meristem growth would be useful to determine the minimum amount of GA3 required for inflorescence development. The adapted wheat meristem media (wMM) was further optimised by the addition of zeatin, maltose, and vitamins (see protocol for full details of media preparation) and has been used in all subsequent steps.

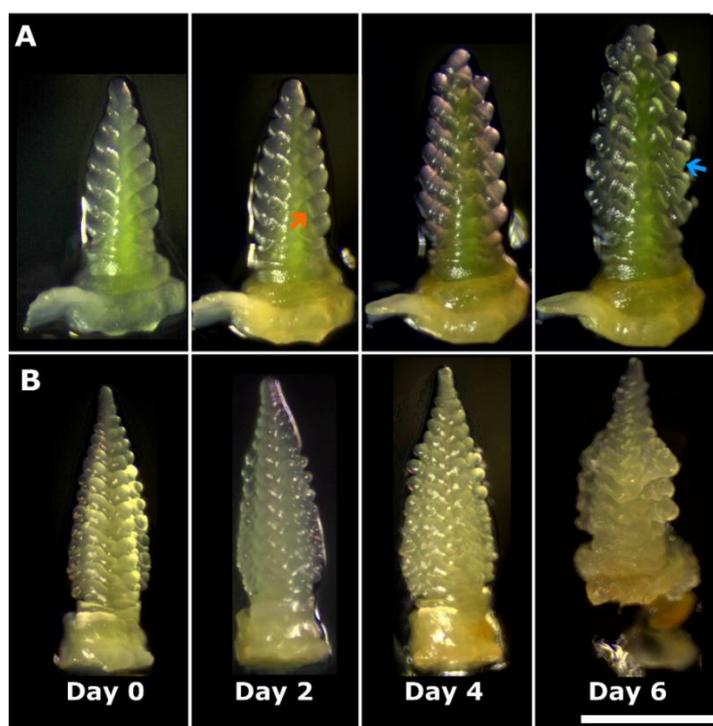


Figure 3.1: Stereomicroscopy images of wheat (A) and barley (B) inflorescence meristem development on wheat meristem media (containing 0.2 mg/L GA3). A) Wheat meristems dissected at spikelet primordium stage (first image from the left) can develop glume (orange arrow), lemma and floret primordia (blue arrow) within 6 days of growing on the media. The rachis of the spike also shows signs of greening over the 6 days. B) Barley meristem grown on the same media over the same timeframe does not form new primordia and shows signs of cell death (shrinking and discoloration) as well as abnormal cell proliferation by day 6. (All images acquired using stereo microscope with magnification x50, bar = 1mm).

Wheat and barley (*Hordeum vulgare L.*) belong to the *Triticeae* tribe and are frequently described as highly similar in inflorescences architecture, as well as primordia developmental pathways (Koppolu and Schnurbusch, 2019). We therefore investigated if the wheat adapted wMM would be adequate for barley meristem development by dissecting barley meristems at the same developmental stage as wheat and transferring them onto the wMM. Barley meristems, however, failed to develop on the wheat adapted wMM (Figure 3.1B). Similar to wheat meristem development in the absence of GA3, barley meristems did not show signs of cell division or growth and hence did not grow or develop existing or new primordia. After six days, the meristems also showed clear signs of cell death and random cell proliferation (Figure 3.1B). This suggests that other signals are necessary to activate the development of the barley meristem *ex situ* than wheat, revealing a potential divergence in important spike development signalling pathways between barley and wheat.

### 3.3.2 TRD1 leaf ridge suppressing function is conserved in wheat

To test our method for live cell imaging of wheat meristems and subsequent analysis in MorphoGraphX, we first analysed a known mutation that affects organ growth at the double ridge stage. *THIRD OUTER GLUME1 (TRD1)* has been previously shown to control lower/leaf ridge growth and mutations in *TRD* orthologs lead to their outgrowth in maize, and rice (Wang et al., 2009; Whipple et al., 2010; Houston et al., 2012). We took advantage of this known phenotypic effect of *trd1* to test the live imaging protocol and to gain new insight into the phenotypic effect of *trd1* on a cellular level in wheat. *TRD1* is furthermore a suitable candidate to test the meristem media as in rice Wang et al. (2009) showed that the function of *NECK LEAF1* (rice *TRD1* ortholog) is independent of the photoperiodic and GA signalling pathways.

To explore the growth effect of *TRD1* on ridge formation in wheat, we employed CRISPR/Cas9 to disrupt the function of all three *TRD1* homoeologs, which are expressed at similar levels during early spike formation with a marginal A-copy bias (Figure 3.2A). We used a tandem array of CRISPR guides targeting the same sequences within the first exon of all three *TRD1* copies (Figure 3.2B; *TraesCS1A02G418200*, *TraesCS1B02G448200*, *TraesCS1D02G425900*). In the T<sub>0</sub> generation, we recovered 18 independent lines with construct insertion copy numbers above 1 (Table 3.1). The maximum number of construct copies was 13, but most of the lines had less than 6 copies (Table 3.1). In the next generation we identified two independent lines displaying the expected phenotypic effect of bract formation in the mature phenotype (Figure 3.2C).

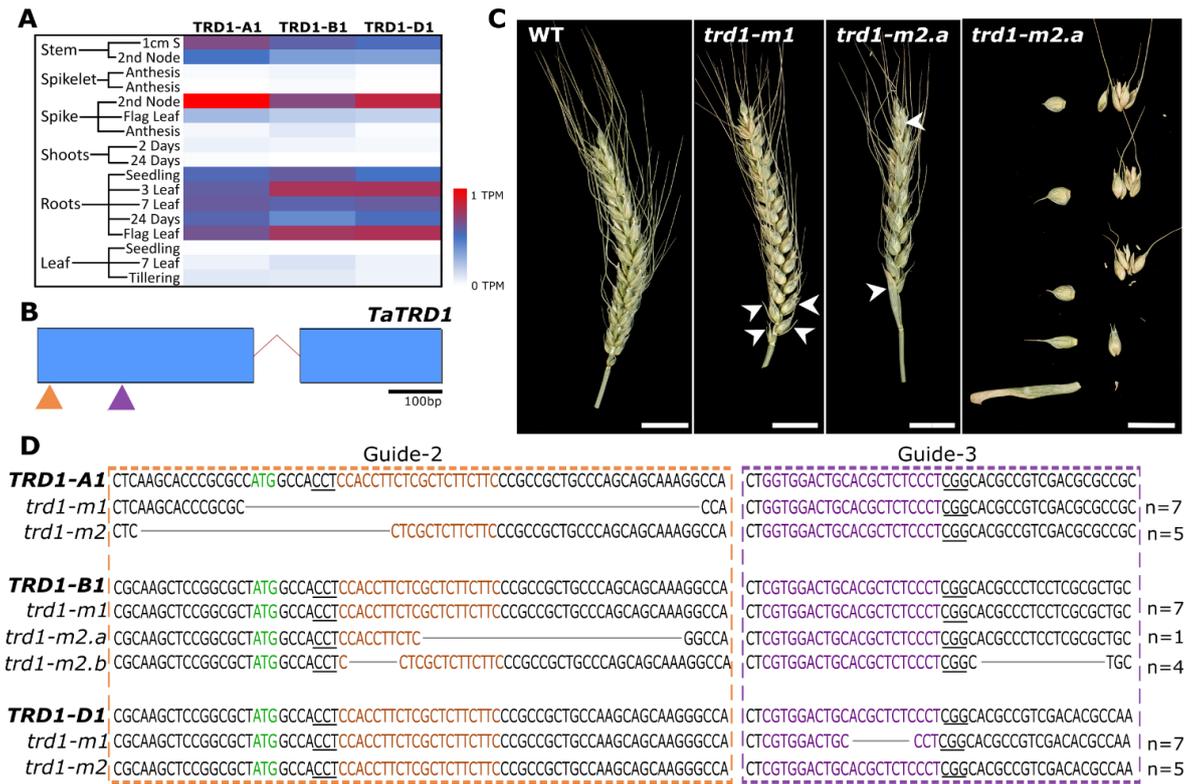


Figure 3.2: THIRD OUTER GLUME 1 (TRD1) mutants in wheat. (A) Expression (transcripts per million, TPM) of TRD1 homoeologs (*TraesCS1A02G418200*, *TraesCS1B02G448200*, *TraesCS1D02G425900*) in wheat (cv. Chinese Spring) from publicly available data (www. wheat-expression.com; (Borrill et al., 2016; Ramirez-Gonzalez et al., 2018)). Expression was summarised by tissue and age using only data from the ‘Chinese Spring developmental timecourse’ dataset. Tissue in which expression was = 0.0 TPM for all homoeologs was excluded from the matrix. (B) Graphic representation of TRD1 gene structure (Blue boxes = exons, red line = intron) based on the gene models in Ensembl. The CRISPR/Cas9 guides used in this study are indicated as orange (Guide-2) and purple (Guide-3) bars. (C) Mature spike phenotypes in the  $T_1$  generation. The wildtype (zero construct copy line identified after transformation) shows no bracts (left panel). All bract leaves in *trd1-1* are indicated by white arrowhead. For *trd1-2a*, only the first and last bract are indicated by white arrowheads, all spikelets between produce visible bracts (right panel). Right panel are the five most basal spikelets dissected from the spike of *trd1-2a*. The most basal bract leaf is the largest and has a more leaf-like appearance than bract leaves of the 2<sup>nd</sup> to 5<sup>th</sup> spikelet. The basal two spikelets were sterile. (D) Sequence of TRD1 wildtype (A, B and D homoeologs) in the region of the guides. Below each track are the alignment results from Sanger sequencing of the two independent mutant lines, *trd1-1* and *trd1-2*.

Table 3.1: Number of CRISPR/Cas9 construct copies detected in  $T_0$  plants from 18 independent events

Construct copies	Events
1	7
2	2
3	1
5	1
6	5
7	1
13	1

Sequencing of the  $T_2$  progeny of the two lines showed that both had homozygous deletions in two of the three *TRD1* homoeologs (Figure 3.2D). In line *trd1-1*, the A genome copy was edited by guide-

2, leading to a 55 bp deletion, while the D copy was edited by guide-3, leading to an 8 bp deletion (Figure 3.2D). In the second independent line *trd1-2*, the A genome copy was also edited by guide-2, leading to a 33 base pair deletion. In the B genome copy, we detected two independent mutations among the progeny of *trd1-2* lines; four of the T<sub>2</sub> seedlings had a deletion in the region of guide-2 while one of the lines was edited by both guides. All *trd1-1* and *trd1-2* mutant plants showed the expected phenotypic effect of bract leaf formation (Figure 3.2C) as well as neck leaf formation (data not shown), as seen in barley, maize and rice. As described in barley, the deletion of bract leaves in *trd1* mutants in wheat is most prominent in the most basal spikelets (Figure 3.1A).

The deletions affected the reading frame of the gene early in the coding region, and deletions were most likely leading to loss of function. The mutations in the A genome copy in both mutant lines lead to the deletion of the start codon as well as the following 50 (*trd1-1*) and 10 (*trd1-2*) base pairs. The mutation of guide-3 in the B-copy of *trd1-2b* leads to frameshift while the mutation of guide-2 in *trd1-2a* leads to the deletion of the 3<sup>rd</sup> till 17<sup>th</sup> amino acid. The D mutation also leads to a frameshift and subsequent amino acid mismatches after the 55<sup>th</sup> amino acid (Figure 3.3). Given the consistent phenotypic data and *in silico* prediction of severe amino acid sequence disruption, we did not further investigate *TRD1* expression or protein structure in the mutants. We are currently developing the homozygous triple mutant of *TRD1* by crossing *trd1-1* and *trd1-2* lines.

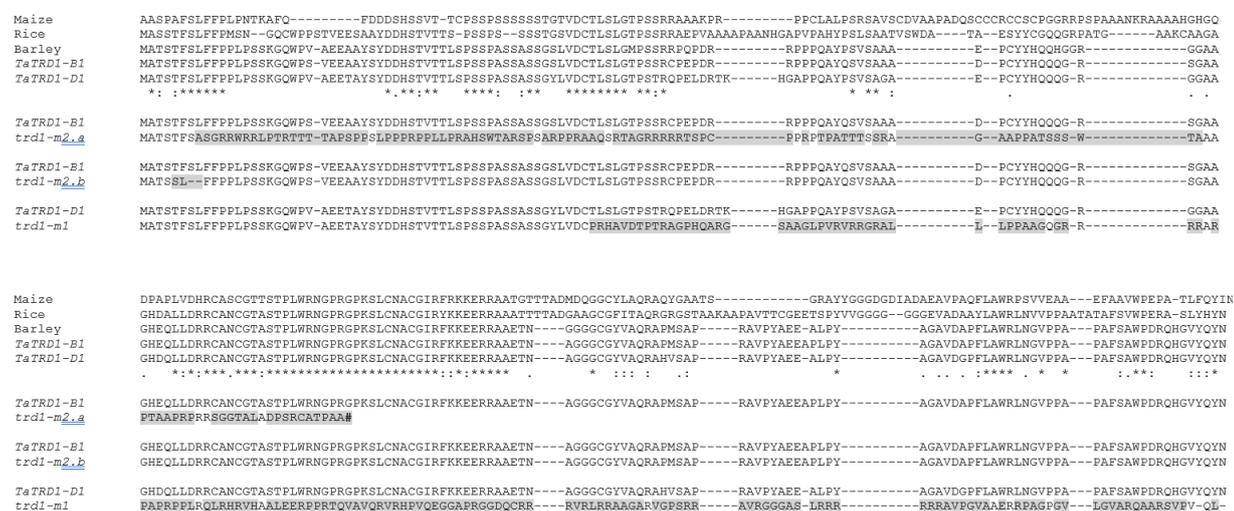


Figure 3.3: Alignment of amino acid sequences of TRD1 wildtype maize, rice, barley orthologs and wheat homoeologs B and D. Conservation of amino acids among wildtype sequences are indicated below (\*=conserved in all sequences, dots indicate partial conservation). Amino acids upstream of MAT (start codons in wheat) in rice and maize were removed to simplify the alignment view. Mutant sequences were only aligned to the respective homoeologs. For the A copy, both mutants led to removal of the start codon (see Figure 3.2) and amino acids were thus not predicted. All amino acids downstream of the first affected amino acid in mutant are highlighted in grey. Amino acids for *trd1-m2.a* predict out-of-frame downstream effects by the deletion and a premature stop code (#).

### 3.3.3 Live imaging of TRD1 mutants confirms enhanced leaf ridge growth

In addition to the mature spike, we also phenotyped the inflorescence development of *trd1* mutants at double ridge stage. We dissected plants at double ridge stage and transferred the spike meristem to the wMM plates. The wildtype and *trd1* mutant are Fielder cultivar and both grew on the media as expected from our previous experiments using cv. Paragon (Figure 3.1A, Figure 3.4A-D). We imaged the inflorescences of wildtype and mutant and could observe enlarged leaf ridges (bracts) in the *trd1* mutant lines in the light microscopy micrographs (Figure 3.4C&D). The phenotypic effect of *trd1* in wheat replicates the effect in barley. Apart from the enlarged leaf bract ridge, we did not observe any other phenotypes in the *trd1* mutant in the micrographs at double ridge stage. We also acquired confocal z-stack images of growing meristems. The resulting z-stacks can be opened in MorphoGraphX to create 2.5D reconstructions of the tissue. The 2.5D projections of the stacks show that the leaf ridges of the *trd1* mutants (Figure 3.4G&H) are larger compared to wildtype Fielder (Figure 3.4F) and Paragon wildtype inflorescences in a comparable growth stage (Figure 3.4E). The only wildtype *TRD1* (cv. Fielder) z-stack we currently had was images from the side, which makes the comparison of ridge length to the mutant complicated, and we therefore included the Paragon z-stack. The next step in this experiment is the analysis of multiple timepoints (at least two) to generate cell growth heat maps. This will allow us to quantify the differences in cell growth, rather than just qualitatively describe the developmental effect of *trd1* mutants. We are currently acquiring these images.

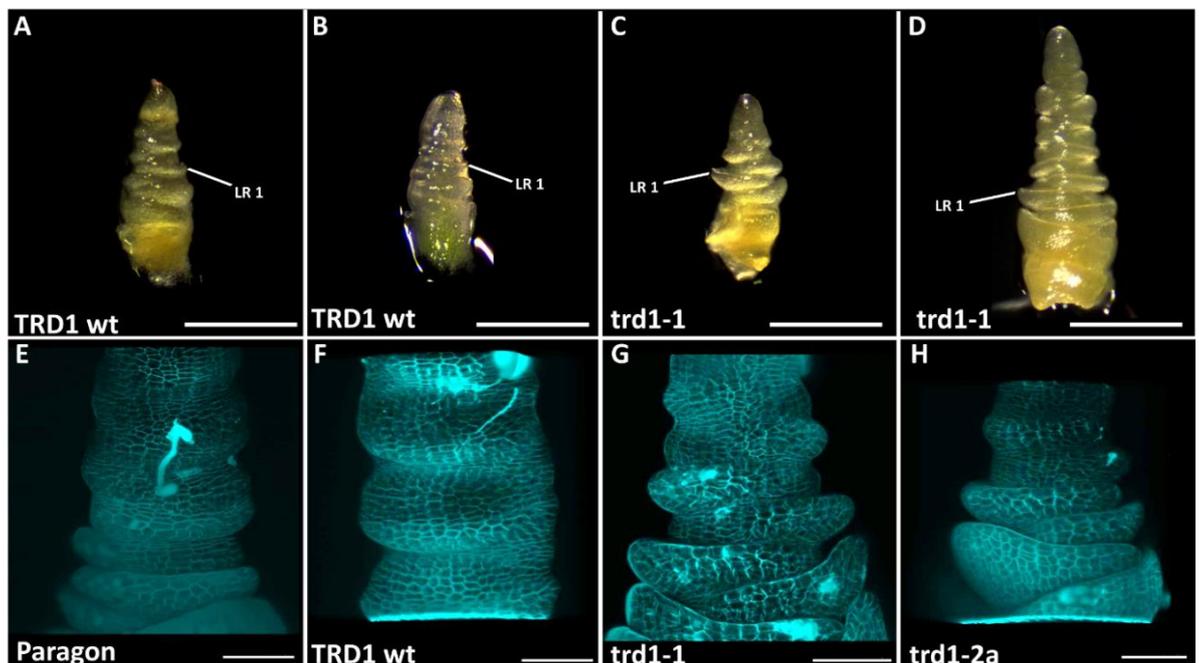


Figure 3.4: TRD1 phenotypic effect at double ridge stage. Light microscopy image of wheat inflorescence imaged while growing *ex situ* on wMM plates for two different *trd1* mutant (C&D) and wildtype (A&B) spike meristems. Note: We are currently working on having time course images of mutant and wildtype on wMM. (E-H) Confocal z-stack images visualised in MorphoGraphX of two *trd1* mutants (G&H, imaged from the front) and wildtype inflorescence (cv. Fielder, sister-lines of *trd1* mutants with 0 construct copies, imaged from the side) imaged on same day as *trd1* mutants but from the side (F) and cv. Paragon wildtype at similar stage (E). Note: We are currently imaging another batch of these to have 3 replications of mutant and wildtype imaged from the front and side. Scale bar = 500  $\mu$ m (A-D), Scale bar = 100  $\mu$ m (E-H)

### 3.3.4 Proof of concept of live cell growth protocol for wheat inflorescence meristem

To test the suitability of our confocal microscope image stacks for MorphoGraphX, we analysed the cell growth of wheat (cv. Paragon, *VRT-A2b* NIL) inflorescences at late double ridge stage. This sample was chosen because the quality of the image stacks was high enough (see protocol). We imaged the sample 48 hours apart and imported both timepoints into MorphoGraphX (Figure 3.5A&B). Neither showed signs of tissue damage or cell death and both had low background noise. We thus proceeded to create the mesh and segment cells at both timepoints (Figure 3.5C&D, see protocol for analysis). We then analysed the growth of cells between the earlier (Figure 3.5C) and the later timepoint (Figure 3.5D) in MorphoGraphX.

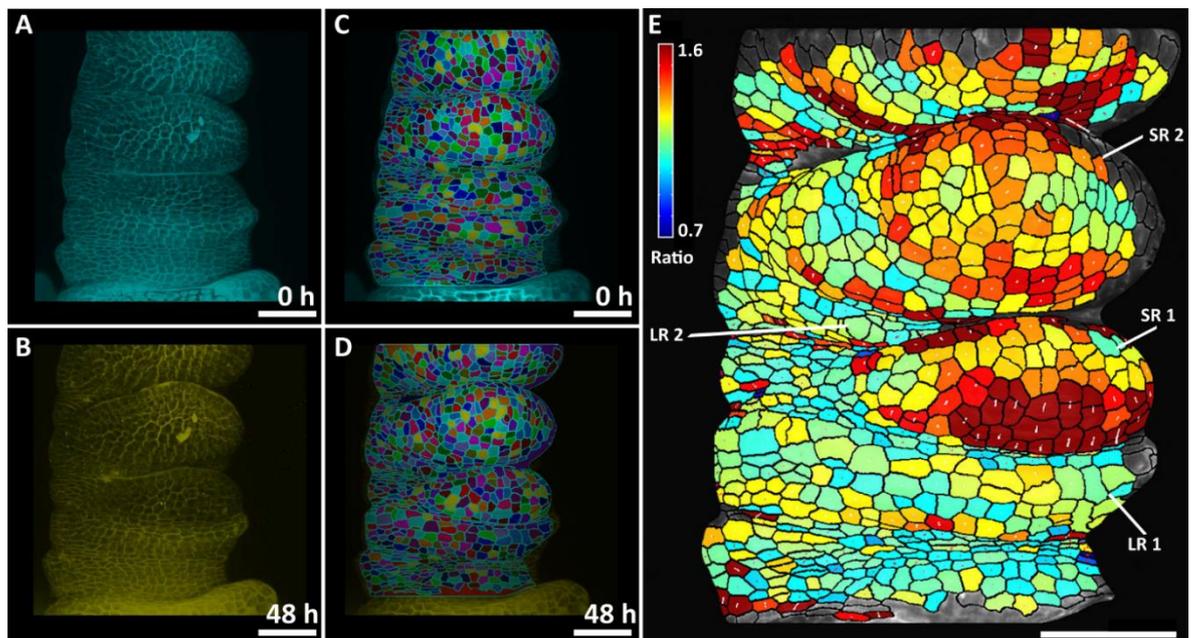


Figure 3.5: Live cell growth analysis of wheat inflorescence. (A) 2.5D projection of confocal images from wheat inflorescence at first imaging timepoint (0 hours). (B) 2.5D projection of same inflorescence, 48 hours later. (C&D) cell segmentation of the two images in MorphoGraphX using manual mesh creation and segmentation. (E) Cell growth (ignoring cell division) between the 0 h and 48 h images of the same inflorescence, visualised on the spike at 48 h. The heat map colours correspond to the amount of growth in the maximum direction. Direction of growth is indicated by white lines (scale bar = 100  $\mu$ m).

The resulting heatmap (Figure 3.5E) shows that the highest cell growth took place in the spikelet ridges (SR), with a growth ratio of 1.6 (growth ratio of 1 being no change in size and ratio < 1 being cell shrinkage). Focusing on the upper spikelet analysed (SR 2), it appears that not all areas of the spikelet grow equally. The centre of the spikelet is growing at a moderate rate. The cells around the centre grow at the highest rate (red to orange), followed by the next outer cells which didn't change in size (light blue). To the left of these cells, there seems to be slightly higher growth in the area that is the side of the spikelet. These areas of growth match the known development of the wheat spikelet at this stage. The spikelet ridge is rapidly expanding, away from the rachis. However, it is also initiating the formation of the glume on the side of the spikelet, which might be the reason for the area of low growth (blue) between the centre of the spikelet and the side as this area will become the groove between the glume and florets. The increased growth on the left of the groove

would thus be indicative of glume initiation. The lower, most basal spikelet of the spike (SR1) is growing at a fast rate and seems to undergo growth at a similar rate or even higher than SR2. However, the growth map also shows that the most basal spikelet is smaller than the spikelet above and is not initiating the glume groove just yet. Furthermore, the leaf ridge (LR) below SR1 is clearly visible and most cells in LR1 have not grown (blue and yellow colour) in the 48h. The leaf ridge of the upper spikelet (LR2) is barely visible at this point and cells are not changing in size.

As we imaged the spike from one side we can only hypothesise that the spikelet outgrowth on the other side is a mirror image of our growth map. We did trial imaging spikes from the front. This results in good analysis of growth in the area of the rachis. However, because the tissue is curved, the quality of the image decreases on the sides where the spikelets are growing, thus giving very poor cell segmentation for the spikelets themselves. Ideally, images of different inflorescences would be taken from the front and side view of the wheat spike. The analysis of spikelet formation and glume initiation in wheat will need to be repeated using more samples and longer time courses to make general assumptions about growth and primordia formation. However, the analysis presented here is a proof-of-concept of our protocol for live cell imaging in wheat and can now be used to investigate inflorescence development, and the effect of genes, such as *TRD1*, upon it.

### 3.3.5 Live cell imaging of *VRT-A2b* basal spikelets

Another application of MorphoGraphX we are currently perusing is the quantification of basal spikelet growth in *VRT-A2b* mutants, which have an increased number of rudimentary basal spikelets in the mature spike. For this, we leveraged the available *VRT-A2b* transgenic material, previously created by Adamski et al. (2021). The transgenic lines carry multiple copies of the Polonicum type *VRT-A2b* allele, leading to the ectopic expression of the gene. In the mature spike, the high copy lines have increased glume length and more rudimentary basal spikelets. We were able to grow *VRT-A2b* transgenic lines *ex situ* on the wMM. We are currently in the process of analysing cell growth and cell division in the most basal spikelets in a 48-hour time window. For this we are preparing growth map analysis of multiple replications for the high-copy and zero-copy lines. The images also demonstrate another problem encountered when growing the meristems on the side to image the spikelets. The lower side is also growing into the agar and thus pushing the meristem over, which makes it difficult to image the same cells in the following time point (Figure 3.6A&B). Furthermore, the second *VRT-A2b* high-copy spike also demonstrates the difficulty of identifying the most basal spikelet ridge. The most basal ridge (1) in the 0-hour sample could also be the suppressed AM of the last leaf ridge, something that will only become obvious in the later imaging timepoints. At double ridge, the transition point between leaf and spikelet ridges is very difficult to be exactly defined. By eye, no differences in growth or spikelet ridge size can be detected between the wildtype and mutant samples, however small differences in developmental age exist

between them and we only have few replications, making any further interpretations difficult at this point (Figure 3.6). Quantifying the differences in growth using MorphoGraphX is a more appropriate strategy to investigate the effect of *VRT-A2b* on early spikelet ridge development.

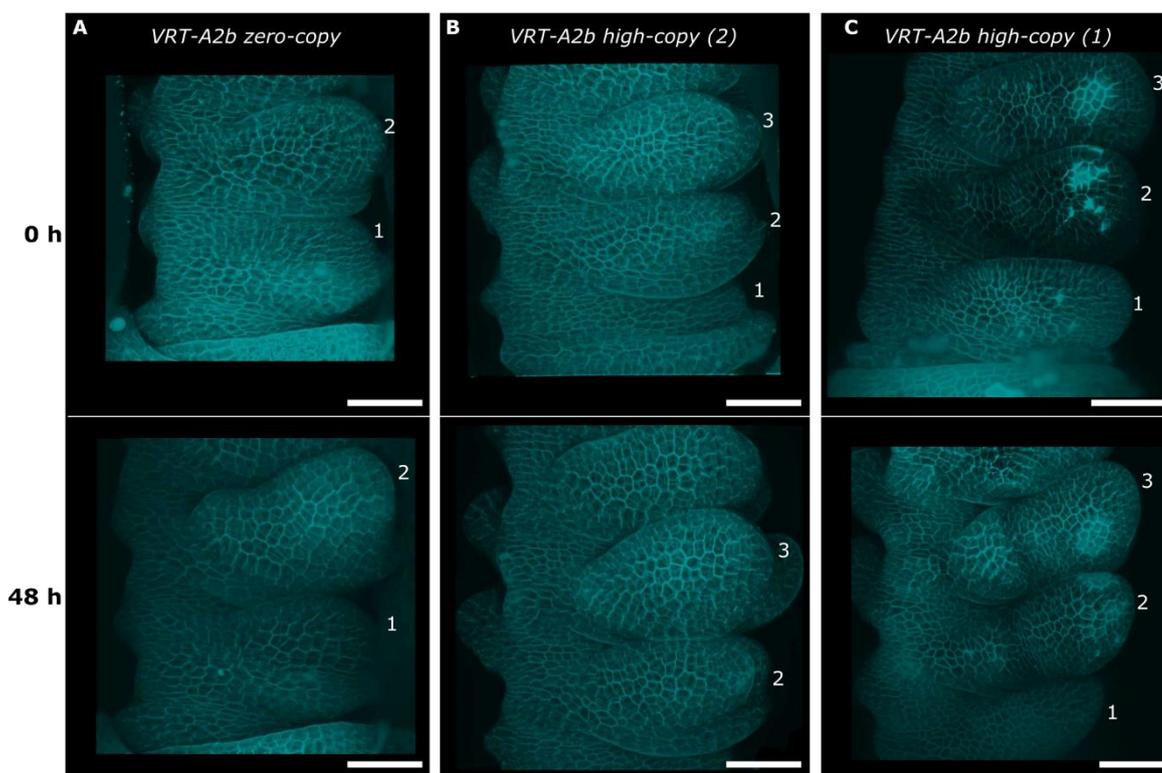


Figure 3.6: Confocal z-stack images of individual spike meristems visualised in MorphoGraphX of *VRT-A2b* zero-copy (A) and *VRT-A2b* high-copy number (B&C) transgenic lines (background cv fielder, see Adamski et al (2021)). We imaged two spikes at 0 h (upper panels) and re-imaged the same spikes again 48 hours later (lower panels). A *VRT-A2b* zero-copy spike was imaged on the same days. All spikes were in early glume primordia stage. Numbers indicate which are the same spikelet ridges in the 0- and 48-hour sample; Scale bar = 100  $\mu$ m.

## 3.4 Discussion

### 3.4.1 Development of live cell imaging in wheat

The main goal of this study was to use MorphoGraphX in wheat, an important global crop species. MorphoGraphX is a powerful tool to investigate cell growth and development of plant primordia and the effect of genetic factors upon it (Barbier de Reuille et al., 2015). In the model species *Arabidopsis*, MorphoGraphX has helped to explain the differences in leaf cessation between *Arabidopsis thaliana* (simple leaf shape) and *Cardamine hirsuta* (leaflet formation) (Kierzkowski et al., 2019), elucidated the developmental cascade of ovule formation (Vijayan et al., 2021), and enabled researchers to gain a deeper understanding of how organs are formed. However, MorphoGraphX has previously not been used for lineage tracking in grass species and its applications in species besides *Arabidopsis* has been very limited (Kierzkowski et al., 2012; Min et al., 2022). A crucial first step for the applications of MorphoGraphX is the ability to grow the plant organ of interest *ex situ*, so it can be imaged using a confocal microscope. For this, *Arabidopsis* floral

meristems can be dissected from the plant and placed on simple growth media (LS/MS + sucrose), however this media fails to support the growth of floral organs of other species ((Min et al., 2022), this study). Wheat and barley meristems dissected in the vegetative stage can be placed on the simple media and continue to develop existing and new leaf primordia (data not shown). When placing inflorescence meristems on the same media, however, they fail to grow or develop and show signs of cell death after about 1 week (Figure 3.1B). This suggests that other exogenous signals are required for inflorescence development, which are most likely normally supplied by the plants stem or leaves. We found that wheat meristems can grow and develop on media containing gibberellic acid (GA3), confirming that it was absence of developmental signals rather than resource inaccessibility (e.g., sugars) that hindered the growth on the simpler media. We further altered the media to contain zeatin and maltose instead of sucrose, as well as adding vitamins that are commonly added to media used for wheat embryo rescue during transformation assays (Hayta et al., 2019). On the optimised wheat meristem media (wMM), inflorescences successfully developed from double ridge stage and initiated glume, lemma and floret primordia autonomously. If the inflorescences can develop further (i.e., post floret formation) was not tested. However, younger meristems placed on the media either did not develop or continued to initiate leaves, suggesting that further signals are required to achieve the initial vegetative to inflorescence meristem transition. Having a method for wheat meristem growth *ex situ* opens the door for MorphoGraphX and other experiments, such as chemical or mechanical stress or the targeted application of hormones to the wheat inflorescence meristem during growth. In the future, testing if the media can support the *ex situ* cultivation of other organs, such as carpel, stamen or roots will be important.

#### 3.4.2 GA role in wheat vegetative to floral transition

Previously, the live cell imaging process had also been adapted for *Aquilegia*, a non-model species belonging to the buttercup family (*Ranunculaceae*) (Min et al., 2022). To achieve *ex situ* growth, Min et al (2022) also adapted the media to contain zeatin (cytokinin hormone) and gibberellic acid (GA3). Our study also found that the addition of GA3 was vital for *ex situ* cultivation of wheat inflorescences, suggesting that exogenous GA3 application is needed for floral development in both species. In *Arabidopsis*, extensive research has led to the formulation of a developmental model in which gibberellic acids are low in the vegetative apex, upregulated during the transition from vegetative to floral primordia (termed doming), and subsequently need to be downregulated in floral primordia (Yamaguchi et al., 2014; Kinoshita et al., 2020). In wheat, relatively little is known about the role of GA in inflorescence and floral primordia formation and the direct comparison between *Arabidopsis* floral development and grasses is complicated, due to the substantial differences in inflorescence architecture.

Pearce et al. (2013) showed that exogenous GA application accelerates flowering development in wheat. In their study, plants grown under non-inducing (short day) conditions, and carrying the

spring growth habitat *VRN1* (*Vrn1g*) gene, were able to transition from vegetative to floral development and the application of exogenous GA3 accelerated their development (Pearce et al., 2013). They furthermore proposed a model in which the key floral inducing gene *FLOWERING LOCUS T (FT)*, known to be produced in the leaves and transported to the apex (Jaeger and Wigge, 2007), up-regulates GA biosynthetic *GA20ox* genes and down-regulates the catabolic *GA2ox1* gene in the apex (directly or indirectly). It has long been known that GA is able to induce bolting and flowering (Lang, 1957), a function that appears to be conserved in the grasses because exogenous GA application is sufficient to accelerate flowering development in spring varieties grown under long-day conditions (Evans et al., 1995) and in winter varieties that have undergone vernalization (Razumov et al., 1960). Our results are consistent with these results that GA is required for early spike development. To grow the wheat meristems on the media, all leaf primordia are removed, and the meristem is cut off from the stem. Thus, any signal from the leaves/stem are not available to the meristem. Our results highlight GA3 activation as a prerequisite, and vital signal, for spike development. Pearce et al. (2013) suggested that GA is synthesised in the apex. Our experiments would thus suggest that the need for *FT* activation of GA can be circumvented by the direct application of GA3. As proposed in their model, our results furthermore suggest that the induction of *VRN1* is independent of GA and we hypothesise that it needs to be activated before we transfer the spikes to the wMM. Meristems that were transferred to the wMM plates before reaching the double ridge stage failed to develop, suggesting that initial transition of the meristem must have happened *in planta*. In their study, Pearce et al. (2013) applied GA for 18 weeks, leading to reduced spikelet numbers and elongated glume formation in the mature spike, suggesting that excessive or prolonged GA up-regulation interferes with later spike development. This suggests that, analogous to *Arabidopsis*, GA needs to be downregulated at some point during spike development, however, this remains to be investigated.

As the meristems developed glume, lemma and floret primordia in the presence of GA3 on the media plates, the role of GA in *Arabidopsis* floral primordia suppression is either not conserved in wheat, or the differences in wheat inflorescence architecture requires prolonged GA presence. In wheat, the elongation of the vegetative meristem and formation of the double ridges is classically interpreted as the beginning of floral development. However, many more primordia develop between the double ridge stage and true floral primordia formation. The upper ridge of the ridge pairs form first a spikelet primordium, from which glume and lemma primordia are subsequently initiated, while spikelet initiation continuous at the tip of the spike (Bonnett, 1966). Although the identity of these organs is still disputed, glume and lemma can be classified not as floral structures but as modified leaves (Whipple et al., 2004; Kellogg, 2022). Recent classifications suggest that spikelets are modified spikes and that glume and lemma are most likely the reduced/modified leaves in whose axil the axillary meristems are formed (Kellogg, 2022). Thus, no floral meristems would be initiated in the spike until after lemma primordium stage, which might explain why GA3

is not interfering with development from double ridge to lemma primordium. In *Arabidopsis*, the two-phased action of GA has also been connected to its effect on primordia determinacy. Chang et al. (2020) suggested that GA favours indeterminate primordia growth and is downregulated in determinate (floral) tissues. Trying to overlay this model on the wheat spike development, we would hypothesise that GA needs to be upregulated in initiating spikelets until the formation of palea, which are the first structures of the determinate floral meristem. Furthermore, we would also hypothesise that GA needs to be downregulated eventually in the spike tip of wheat, to allow terminal spikelet formation. In barley, the organ determinacy is *vice versa*: The spike is indeterminate while each spikelet is determinate. Interestingly, the wheat adapted wMM did not support the development of barley spikes, which showed no signs of growth and displayed cell death after six days of *ex situ* cultivation. Floral signalling pathways are generally assumed to be highly conserved among grasses, and barley is frequently used as model species for wheat studies. Previous studies by Boden et al. (2014) found that GA is upregulated during meristem transition in barley, and that blocking GA under inductive conditions slows down spike development. Our results highlight that the control of barley spike development differs to that of wheat and *Arabidopsis* as it does not develop on the wMM nor simple LS agar. Potentially, GA application might interfere with the development of barley but not wheat spikelets as they are of determinate nature. However, further investigations will be needed to find the missing molecular signal(s) that needs to be supplied exogenously for barley development *ex situ*. Alternatively, the current media might contain components that inhibit barley development.

#### 3.4.3 MorphoGraphX analysis of live cell growth in wheat

After having established a protocol for live cell imaging in wheat (see Appendix for full protocol), we imaged the developing wheat spikes and analysed the images in MorphoGraphX (Figure 3.5). We were able to generate cell maps from two timepoints and could thus generate cell growth maps of the wheat spike. The wheat spike used for the first, proof-of-concept, growth map was in early glume primordia stage and the spikelet ridges were well established. The growth maps revealed that the spikelet ridge of the 1<sup>st</sup> and 2<sup>nd</sup> most basal spikelet are rapidly growing while the lower/leaf ridges are not growing between the two timepoints.

To test the efficiency of MorphoGraphX to quantify cellular growth defects in different wheat mutants, we developed *trd1* mutants in hexaploid wheat (cv. Fielder). *TRD1* is known to suppress growth of the lower/leaf ridge in barley, maize, and rice (Wang et al., 2009; Whipple et al., 2010; Houston et al., 2012). We found this phenotypic effect of *trd1* highly conserved in wheat (Figure 3.2). At the double ridge stage, the effect of *trd1* can be first observed in barley, where the lower leaf ridge is much bigger than in wildtype (Houston et al., 2012). We observed a similar effect of *trd1* in wheat where the lower leaf ridge appears bigger in the mutant. However, the effect of *trd1* on the growth of lower leaf ridges has not been quantified and it is not clear from micrographs

when growth ceases due to *TRD1*. The analysis of cell growth in *trd1* mutants and wildtype using MorphoGraphX will help us to elucidate the cellular effect and timing of *TRD1* further and make an excellent case study for the use of MorphoGraphX.

We have already acquired live cell images for *TRD1* and *VRT-A2b* mutants (Figure 3.4 and 3.6, respectively). Unfortunately, we ran out of time during this PhD to fully analyse these mutants and wildtype controls in MorphoGraphX. In a week, we can acquire ca. 6-10 individual image stacks as imaging of each sample takes ca. 30 minutes and needs to be repeated every 48-hours. Analysis of each stack is not fully automated and takes 1-2 days per image stack. Not every sample has high enough quality to be analysed, making the MorphoGraphX workflow very time consuming. One common issue is infection of the plate with fungal growth. The addition of plant preservation mixture (PPM) has greatly reduced this issue. Another problem we encountered is the conical shape of the wheat apex. Tissues that are curved become difficult to image and greatly reduces quality, making cell segmentation impossible. When imaging the meristems from the front, the ridges can be clearly seen, however the centre of the spikelet cannot be imaged at high quality (for example *trd1* mutants in Figure 3.4G&H). We overcame this issue by placing the meristems on the side, thus being able to create good cell segmentations for the spikelets (Figure 3.4F). To fully analyse mutants, we would thus recommend going forward to image meristems from the front and side.

In the future, we would also like to employ MorphoGraphX to track cell growth in mutants whose phenotypes cannot be observed using light microscopy. For example, we previously showed that the increased expression of *VRT-A2* in *T. polonicum* (*VRT-A2b* allele) leads to increased rudimentary basal spikelet formation in the mature spike. We could furthermore show that basal spikelets are already aborted at anthesis (Backhaus et al., 2022; Chapter 2) and that around Waddington stage 5.5 (carpel primordia stage) the introgression of *VRT-A2b* leads to a slight delay in development of the spike (Backhaus et al., 2022, BioRxiv; Chapter 4). However, we were unable to trace this phenotype to even earlier timepoints. In the wildtype, *VRT-A2* expression is downregulated upon transition from vegetative to double ridge and its constitutive overexpression leads to leaf and tiller formation along the spike (Li et al., 2021). MorphoGraphX analysis could prove if *VRT-A2b* already affects spikelet establishment in even earlier timepoints during their initial establishment. Furthermore, we are aiming to develop live cell imaging protocols for other tissues, such as the carpel, which would allow for the analysis of cell growth effects of mutations affecting grain size. For example, *GW2* has been shown to increase carpel size as early as five days before anthesis (Simmonds et al., 2016). However, it is not known how early *GW2* acts to affect carpel size. Furthermore, the development of cell membrane reporter lines, such as those available in *Arabidopsis* (Willis et al., 2016), would be a key resource to improve the future usage of MorphoGraphX in wheat. Staining cell walls using Propidium Iodide increases the background noise and negatively affects cell survival, limiting imaging to 3-4 time points for each sample.

We here present a protocol for live cell imaging of wheat inflorescence meristems. A major bottleneck was the inability to grow wheat floral meristems *ex situ* as media adapted for *Arabidopsis* did not support their growth. We found that wheat floral meristems require GA3 to transition through key spikelet developmental stages. The availability of high precision imaging techniques and tools such as MorphoGraphX in wheat, bring us one step closer to understanding early spike growth and developmental in this major crop.

### 3.5 Materials and Methods

#### 3.5.1 Construct assembly

*TRD1* wheat orthologous (*TraesCS1A02G418200*, *TraesCS1B02G448200*, *TraesCS1D02G425900*) were identified by searching for the barley *TRD1* gene (*HORVU1Hr1G091450*) orthologous in Ensembl (<https://plants.ensembl.org/index.html>). Two guides were designed to target the first exon of all three wheat homoeologs; Guide-2 target: CCTCCACCTTCTCGCTCTTCTTC and Guide-3 target: GTGGACTGCACGCTCTCCCTCGG. Guides were chosen based on having no perfect off-target matches and only low off-target matches for the first 10 base pairs. Also low secondary structure formation was assessed using RNAfold webserver (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>)

A modified version of the GoldenGate (MoClo) compatible level 2 vector pGoldenGreenGate-M (pGGG-M) as described in Hayta et al. (2019) was used in this study. The pGGG-AH-L2P2 acceptor plasmid is comprised of the hygromycin resistance gene (*hpt*) containing the *Cat1* intron driven by the rice actin1 (*Act1*) promoter for in planta selection and a LacZ-MCS flanked by two *BsaI* sites at MoClo position 2 with standardized overhangs to accept basic (level 0) components. In brief, the two guide sequences (Guide-2 and Guide-3, Table 3.2) were cloned into pGGG-AH-L2P2 using standard Golden Gate MoClo assembly (Werner et al., 2012), resulting in the Cas9 TRD vector (Figure 3.7). The construct was electroporated into the hypervirulent *Agrobacterium tumefaciens* strain AGL1 (Lazo et al., 1991) containing the helper plasmid pAL155 (additional *VirG* gene). Standard inoculums of *Agrobacterium* (Tingay et al., 1997) were prepared as described in Hayta et al. (2019).

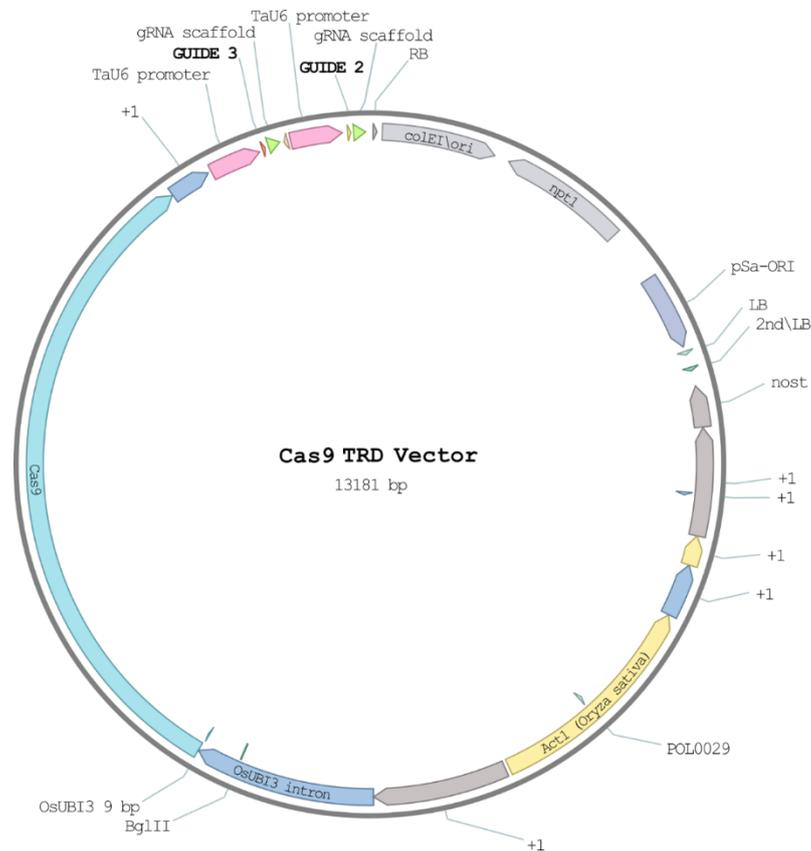


Figure 3.7: Construct plasmid map of the vector containing guide 2 and guide 3 targeted to exon 1 of the TRD1 A, B and D copy in hexaploid wheat. Plasmid map will be available in Addgene upon publication.

### 3.5.2 Wheat transformation

Hexaploid wheat cv 'Fielder' was transformed using the previously described method by Hayta et al. (2019). In brief, under aseptic conditions wheat immature embryos were isolated, pre-treated by centrifugation, inoculated with *A. tumefaciens* AGL1 containing the Cas9 *TRD1* vector and co-cultivated for 3 days. Wheat callus proliferation, shoot regeneration, and rooting were carried out under a stringent hygromycin selection regime before the regenerated plantlets were transferred from *in vitro* to soil and acclimatized. Transgenesis was confirmed by *hpt* gene PCR; transgene copy number analysis was performed using Taqman qPCR and probe (Hayta et al., 2019). The values obtained were used to calculate copy number according to published methods (Livak and Schmittgen, 2001). T<sub>0</sub> plants were grown in controlled climate chambers while all subsequent generations were grown in glasshouse conditions (16h light / 8h dark).

To determine whether the guides edited *TRD1*, we carried out PCR on genomic DNA (gDNA) from T<sub>2</sub> generation plants and Sanger sequenced the resulting amplicon (service provided by <https://www.genewiz.com/>). Primers specific to only one of the three homoeologs copies of the *TRD1* sequence targeted by guide-2 and guide-3 were designed to gain individual amplicons for each homoeolog (Table 3.2). The sequencing primer was common to all three amplicon (Table 3.2). A touchdown PCR was carried out under the following conditions: 30 s at 98 °C, 10 cycles of 10 s at 98 °C, 1 min starting at 10 °C above annealing temperature and decreasing by 1 °C per cycle, and 2

min at 72 °C, a further 25 cycles of 30 s at 98 °C, 1 min at annealing temperature, and 2 min at 72 °C, followed by a final extension at 72 °C. PCR was performed using Q5 High-Fidelity DNA polymerase (New England, M0491S), adding recommended amount of 5X GC-enhancer.

Table 3.2: Primer sequences specific to this study and their amplicon size, annealing temperature, and usage

Primer	Primer sequence 5'-3'	Amplicon size	Annealing temp.	Primer usage
<b>TRD1-A_Fwd</b>	ACCTTCACTCTCCATGCTCG	563 bp	59.7 °C	PCR for Sequencing
<b>TRD1-A_Rev</b>	TGCGCGAGGGTGTATGCT			PCR for Sequencing
<b>TRD1-B_Fwd</b>	CTCCAGTAGCTCACCACCAG	598 bp	59.3 °C	PCR for Sequencing
<b>TRD1-B_Rev</b>	GAACGCGGAGACAAAGTAGAC			PCR for Sequencing
<b>TRD1_common</b>	CCTGCTGCTGGTGGTAGTA	-	-	Sequencing
<b>TRD1-D_Fwd</b>	TGCCTACCTCCACTCTCCA	637 bp	59.3 °C	PCR for Sequencing
<b>TRD-D1_Rev</b>	GCGACAGACAATGCACATACA			PCR for Sequencing
<b>Guide-2 tgca</b>	CTTGGAAGAAGAGCGAGAAGG <u>TGG</u>			Cas9 Guide
<b>Guide-2 aaac</b>	AAACCCACCTTCTCGCTCTT <u>CTTC</u>			Cas9 Guide
<b>Guide-3 cttg</b>	<u>CTTGGTGGACTGCACGCTCTCCCT</u>			Cas9 Guide
<b>Guide-3 aaac</b>	<u>AAACAGGGAGAGCGTGCAGTCCAC</u>			Cas9 Guide

### 3.5.3 Live cell imaging

The protocol for plant growth, live cell imaging of wheat and subsequent analysis in

MorphoGraphX is available as online protocol:

<https://www.protocols.io/private/69A55C01351C11ED98330A58A9FEAC02>

(Copy of protocol: Appendix A2)

## 4 Delayed development of basal spikelets in wheat explains their increased floret abortion and rudimentary nature

### 4.1 Summary

Winter wheat (*Triticum aestivum* L.) breeding efforts have increased grain yield predominantly by raising grain numbers per spikelet, rather than grain weight or spikelet number. However, across a single spike large differences exist in the number of grains per spikelet. The central spikelets produce the highest number of grains in any given genotype while apical and basal spikelets are less productive. Basal spikelets are delayed in development just after initiation and are smaller and less advanced than central spikelets already by the glume primordium stage. However, basal spikelets continue to develop and produce florets until much later in the wheat growth cycle. The precise timings or the cause of their growth cessation, and subsequent abortion, is largely unknown. In this study we investigated the underlying causes of rudimentary basal spikelet abortion. We investigated basal spikelet development in four UK winter wheat varieties as well as a set of near-isogenic lines for *VRT-A2* using shading applications in the field. We propose that basal spikelet abortion is most likely the consequence of complete floret abortion as both occur at the same time and have the same response to shading treatments. Furthermore, we found that the developmental age of florets pre-abortion is an important factor for their likelihood to survive and develop viable seed. Previously, it had been proposed that reduced assimilate availability in the base of the spike leads to increased abortion. Re-analysis of published data alongside data presented here, however, does not support this model. We found that rather than assimilate availability, it is the reduced developmental age of basal florets before abortion that correlates with increased abortion. Using the floret Waddington developmental stage pre-abortion, we were able to predict final grain set per spikelet across the spike, alongside the characteristic gradient in number of grains from basal to central spikelets. We found that advancing past Waddington stage 5.5 seems to be important for floret survival and that most florets in basal spikelets had not reached this stage at the onset of floret abortion. The abortion of all florets could therefore be the reason for their rudimentary appearance in the mature spike, suggesting that basal spikelet abortion is simply the consequence of all florets inside the spikelet being aborted and thus all other spikelet structures (e.g., lemma, rachilla, glume) also ceasing to develop. Future efforts to improve spikelet homogeneity across the spike should thus focus on improved basal spikelet establishment and increasing their development rate pre-abortion.

### 4.2 Introduction

Three of the globally most important staple crops, maize (*Zea mays* L.), rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.), belong to the family of grasses (*Poaceae*). One characteristic feature of the *Poaceae* is their inflorescence, which develops florets inside specialised structures termed spikelets (Kellogg, 2001; Kellogg et al., 2013). Upon floral transition the apical meristem of grasses

elongates and, depending on the species, the spikelets are formed directly on the rachis (e.g., wheat) or on primary (e.g., maize) or secondary (e.g., rice) branches (Kellogg, 2022). Each spikelet can form multiple florets, however in some species the number of florets per spikelet is highly genetically controlled (e.g., maize) while in others the number is also dependent on environmental factors (e.g., wheat) (Bonnett, 1966). Across all grasses the potential yield of a plant depends on the number of inflorescences, number of spikelets and florets per inflorescence, as well as grain weight.

Over the past century, breeding efforts have enhanced yields in winter wheat, predominantly through an increase in grain number, rather than grain weight (Würschum et al., 2018; Voss-Fels et al., 2019; Sakuma and Schnurbusch, 2020). Furthermore, it has been established that the increase in grain numbers was achieved through the improvement of grains per spikelet rather than through an increase in spikelets per inflorescence (termed spike in wheat) (Philipp et al., 2018; Sakuma and Schnurbusch, 2020). However, increasing the number of grains per spike can have negative effects on grain weight, as is the case of 'Miracle Wheat' (Poursarebani et al., 2015). Previous research found that the amount of resource available to the plant can affect the relative growth and development of the initiated spikelets and grains. Thus, trade-offs between the different yield components are created by the 'source-sink' balance (Reynolds et al., 2009).

During wheat spike development, a finite number of spikelets is initiated until terminal spikelet formation, thus the number of spikelets per spike is determined relatively early on in the crop growth cycle (Bonnett, 1966; Kirby and Appleyard, 1981). Each spikelet initiates an indetermined number of floret primordia, of which most are aborted. Wheat spikelets initiate typically many florets (10-12), but only a fraction (typically 3-5) survive abortion and go on to form grains (Sadras and Slafer, 2012). Thus, a higher number of spikelets per spike might not lead to more grains per spike as it can be annulled by increased floret abortion. Furthermore, the weight of individual grains has also been shown to correlated negatively with the number of grains (Sakuma and Schnurbusch, 2020).

Over the past decade, several genes that affect spikelet number, floret abortion and grain weight have been identified in wheat. For example, several genes that increase the number of spikelets initiated have been cloned, such as *FRIZZY PANICLE (FZP)* (Dobrovolskaya et al., 2015; Poursarebani et al., 2015), *WHEAT ORTHOLOG OF APO1 (WAPO-A1)* (Kuzay et al., 2019; Muqaddasi et al., 2019; Voss-Fels et al., 2019), and *TEOSINTE BRANCHED1 (TB1)* (Dixon et al., 2018). However, the introgression of the beneficial alleles into elite material has seldom led to significant increases in yields. For example, the increase in expression of *WAPO1* leads to increased spikelet numbers, which does not translate into yield gains due to increased spikelet abortion (Wittern et al., 2022) and decreased floret survival (Kuzay et al., 2022). Targeting the number of florets rather than spikelets, the reduced-function allele of *GRAIN NUMBER INCREASE 1 (GNI1)* has been shown to

confer yield increases by reducing floret abortion compared to the wildtype *GN1* allele, which functions as a rachilla growth inhibitor (Sakuma et al., 2019). The high frequency (96%) of the increased grain number *GN1* allele among durum wheats suggests that there has been a strong selection pressure for increased grains per spikelet during domestication and breeding (Sakuma et al., 2019). In terms of grain weight, ectopic expression of a semi-dominant allele of the *VEGETATIVE TO REPRODUCTIVE TRANSITION 2 (VRT-A2)* gene from *T. turgidum* ssp. *polonicum* leads to increased grain length, yet does not increase yield with respect to the wildtype *VRT-A2a* allele across multi-year field trials (Adamski et al., 2021).

Within a spike, large differences in the number of grains per spikelet (i.e., spikelet fertility) exist, with the central spikelets producing the highest number of grains in any given genotype. An analysis of grain distributions across the spikes of 210 elite and 180 heritage wheat accessions showed that the number of grains per spikelet has increased in elite material, but that this increase has mostly occurred in the central spikelets, to a lesser extent in apical, and not at all in the most basal spikelets (Philipp et al. 2018). The authors proposed that reducing the variation in spikelet fertility across the spike could be a promising avenue to increase yields and improve grain size homogeneity. However, we have little understanding of the factors that determine these gradients within a spike.

Apical spikelets are initiated last and thus have less time to develop their floret primordia (Bonnett, 1966). The most basal spikelets are initiated first, but they often develop only in a rudimentary form, meaning that they are much smaller than other spikelets, produce no grain and have underdeveloped glumes and lemma. This variation in spikelet development leads to the characteristic lanceolate shape of the wheat spike (Backhaus et al., 2022). Basal spikelets are delayed in development just after initiation and are smaller and less advanced than central spikelets already by the glume primordium stage (Bonnett, 1966). However, the basal spikelets continue to develop and produce florets until later in spike development when basal spikelet abortion occurs. The precise timings or the cause of their growth cessation, and subsequent abortion, is largely unknown in wheat.

A wealth of experimental data has confirmed that across all spikelets resource availability, also termed source strength, is closely linked to floret survival (González et al., 2011; Ferrante et al., 2013). As a way to explain the higher floret abortion of basal spikelets, González et al. (2011) proposed that basal spikelets have poorer resource allocation than central spikelets, although it remains to be established whether basal spikelets do indeed have lower priority in assimilate partitioning than central spikelets. Other factors, such as development stage of florets (Ferrante et al., 2020), vasculature development (Hanif and Langer, 1972) and distance to the rachis (Kadkol and Halloran, 1988) have also been shown to affect floret survival in spikelets, but none of these factors have been studied as the cause of rudimentary basal spikelets.

Multiple studies have used shading applications to reduce photosynthetic activity in the field and shown that pre-anthesis shading reduces yield and also spike and plant dry weight, which is a good indicator of reduced source strength (Fischer, 1985; Savin and Slafer, 1991; Slafer et al., 1994). The effects of shading on altering resource availability are relatively quick (within two days) as determined by measurements of water-soluble carbohydrates (WSC; Stockman et al., 1983). Stockman et al. (1983) also found that shading treatments affected basal spikelet fertility more than apical and central spikelets (using single tiller plants under controlled environment conditions). Furthermore, Slafer et al. (1994) reported the effect of shading on each individual spikelet and showed that the number of basal spikelets with zero fertile florets was increased by 3-4 spikelets under shading conditions. This suggests that not only floret, but also basal spikelet survival, is negatively affected by shading conditions.

In this study, we aimed to characterise the causes of rudimentary basal spikelet development in wheat. We used shading treatments pre-anthesis to reduce resource availability in precise and short time frames that spanned basal spikelet abortion during the crop cycle. We furthermore collected samples after the shading application to assess the effect of shading on sugar concentrations in different spikelet positions across the spike. We also traced the development and number of florets across different spikelet positions to relate pre-anthesis floret development to the probability of floret survival. This study highlights that rudimentary basal spikelets are most likely a consequence of complete floret abortion in basal spikelets. We did not find any evidence for lower assimilate accumulation in the base, but rather that the delayed development of the florets in basal spikelet can explain to a large extent their rudimentary nature.

## 4.3 Results

### 4.3.1 Basal spikelet development ceases two weeks pre-anthesis and is sensitive to resource availability at that time

To investigate when basal spikelet abortion takes place, we applied shading treatments that reduce assimilate availability in field grown wheat plots at defined growth stages (Kemp and Whingwiri, 1980; Stockman et al., 1983). Each shading treatment consisted of ca. 45% light reduction over 12/13 days in field grown plots (Table 4.1). In 2021, we applied two shading treatments; the first treatment (Shading A) started around the stem extension phase, whereas the second treatment (Shading B) was applied one day after removal of Shading A and ended ca. 10 days before anthesis (Figure 4.1A). We applied shading to four UK winter wheat cultivars as well as a set of cv. Paragon near isogenic lines (NILs) carrying either the wildtype *VRT-A2a* or the *T. polonicum* *VRT-A2b* allele (Adamski et al., 2021). *VRT-A2b* has been previously shown to increase the number of rudimentary basal spikelets (RBS) by 1-2 spikelets compared to the Paragon *VRT-A2a* sibling NIL (Backhaus et al., 2022). We found that the early timeframe in 2021 (Shading A) had no effect on the number of RBS formed, however Shading B increased the number of RBS significantly across all genotypes by on

average 1.46 RBS (Figure 4.1C&D). As both shading treatments were applied for the same duration, we hypothesised that basal spikelet abortion is determined between 10 and 22 days before anthesis.

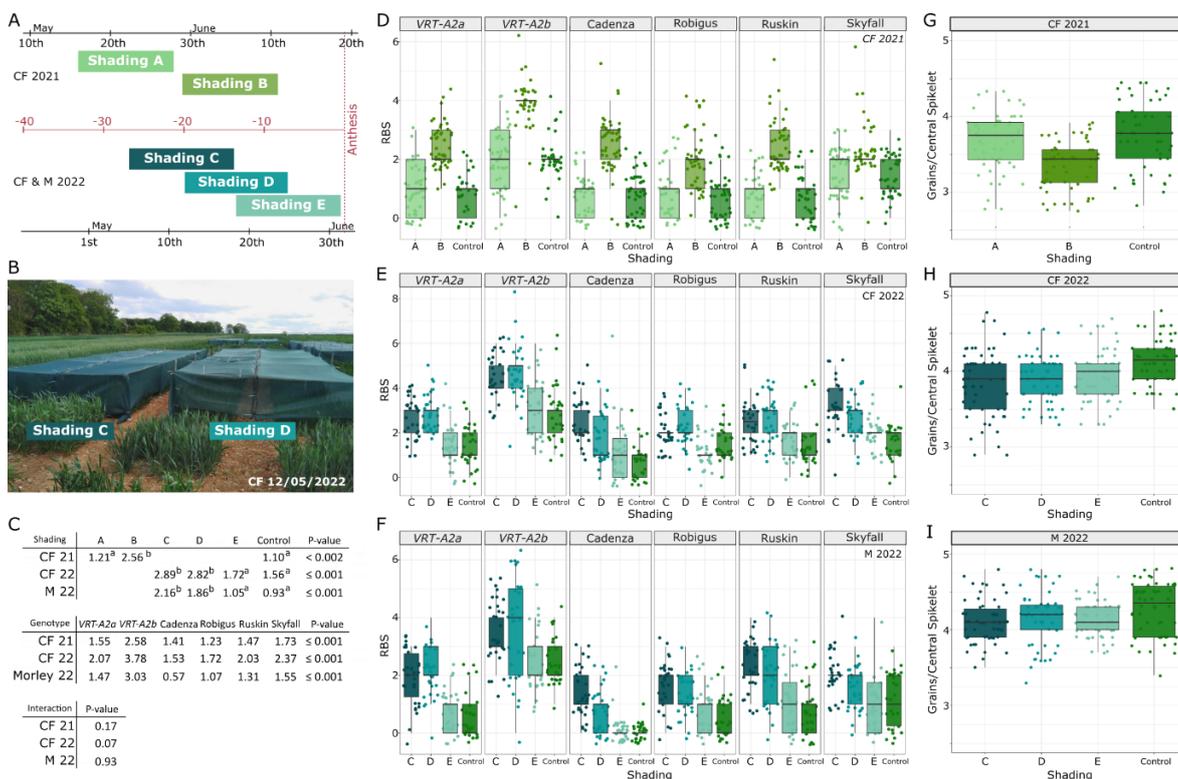


Figure 4.1: Pre-anthesis shading treatment phenotypic effects. (A) Timing of shading applications in Church Farm (CF) in 2021 and CF and Morley (M) in 2022 relative to the average date of anthesis (Table 4.2). Shading A and B were applied for 13 days each (2021), shading in 2022 was applied for 12 days. (B) Image of shading nets for Shading C and D in the field. (C) Estimated means of rudimentary basal spikelet (RBS) for the different shading treatments (top) and genotypes (middle). ANOVA test of significant difference was performed for each trial independently and letters (a-b) indicate LSD results. Interaction between genotype and treatment was non-significant (NS) in CF 2021 and Morley 2022, and borderline NS in CF 2022 (bottom). (D-F) Number of RBS per spike recorded for ten spikes from each block (N=3) in control vs shading applications in CF 2021 (D), CF 2022 (E) and Morley 2022 (F). (G-I) Number of grains per spikelet in the central most spikelets (spikelet position 10-12) from the same data trials. Box represents the middle 50% of data with the borders of the box representing the 25th and 75th percentile. The horizontal line in the middle of the box represents the median. Whiskers represent the minimum and maximum values, unless a point exceeds 1.5 times the interquartile range in which case the whisker represents this value. Points in D-F represent all subsamples (individual RBS measurements from ten to twenty spikes of each of the three blocks), whereas statistical analyses were performed with mean values. In G-I points represent the average number of grains/spikelet of the central three spikelets taken from ten individual spikes from each block. Raw data can be found in Supplemental Dataset S1 (<https://doi.org/10.5281/zenodo.7481986>).

To expand on these results and further investigate when RBS are determined, we replicated the experiment in 2022 across two locations (Church Farm (CF) and Morley (M)). In both locations we applied three shading treatments, which overlapped by one week, allowing us to further narrow down the timing of rudimentary basal spikelet formation (Figure 4.1A). The earliest shading treatment (Shading C) was applied during the stem extension phase and the last shading treatment (Shading E) finished at anthesis. Across all genotypes, Shading C and Shading D significantly increased RBS numbers by 1.33 and 1.26 RBS in CF, respectively, and by 1.23 and 0.93 RBS in Morley, respectively (Figure 4.1E-F). Shading E had no significant effect in CF (Figure 4.1E) and Morley

(Figure 4.1F). This suggests that the last week of Shading C, which was also the first week of Shading D, overlapped the timeframe in which RBS formation is most sensitive to resource limitations, in this case between 10 and ~16 days pre anthesis. Shading E on the other hand was applied after basal spikelet abortion has happened and the number of RBS has been determined. This critical timeframe in 2022 is consistent with the 2021 results and supports the idea that rudimentary basal spikelet formation is linked to a specific growth stage ca. two weeks pre-anthesis.

Table 4.1: Solar radiation (umol/sec/m2) measured in CF 2021 control and shading plots under various light conditions.

Block	25/05/2021; 6pm (low sun)			27/05/2021; 11am (cloudy)			05/06/2021; 11am (sunny)		
	shaded	control	Delta (%)	shaded	control	Delta (%)	shaded	control	Delta (%)
1	129	305	42.3	194	450	43.1	540	1170	46.2
2	130	295	44.1	226	491	46.0	520	1200	43.3
3	130	299	43.5	249	518	48.1	530	1150	46.1
Average	129.67	299.67	<b>43.3</b>	223.00	486.33	<b>45.7</b>	530.00	1173.33	<b>45.2</b>

In addition, we recorded mature plant weight in both Church Farm trials and the number of spikelets/spike in all three trials. In 2021, both shading applications significantly reduced plant weight compared to the control (Table 4.3). In 2022, all shading applications reduced plant weight, although not significantly (Table 4.3). The reduced final plant weight confirms the expected effect of shading on resource availability. Furthermore, the effect of shading on plant weight was equivalent for all shading treatments within the trial and thus any differences in shading effects on RBS between the treatments would be due to timing, rather than intensity. The number of spikelets per spike was not affected by our shading treatments (Table 4.3), which is consistent with the fact that the treatments were applied post the terminal spikelet phase and spikelet initiation was already completed (Kirby and Appleyard, 1981). This also eliminates the possibility that RBS numbers increased due to more spikelets being initiated.

Table 4.2: Flowering dates of all genotypes analysed in this study in Church Farm (CF) trials in 2021 and 2022.

Genotype	CF 2021		CF 2022	
	Control	Control	Shading C	Shading D
VRT-A2a NIL	20/06/2021	03/06/2022	02/06/2022	03/06/2022
VRT-A2b NIL	21/06/2021	04/06/2022	03/06/2022	03/06/2022
Cadenza	19/06/2021	31/05/2022	31/05/2022	31/05/2022
Robigus	20/06/2021	01/06/2022	02/06/2022	02/06/2022
Ruskin	22/06/2021	02/06/2022	04/06/2022	03/06/2022
Skyfall	17/06/2021	30/05/2022	29/05/2022	31/05/2022
Average	19/06/2021	01/06/2022	01/06/2022	01/06/2022

Across the six genotypes, the effect of shading was consistent, and only a borderline non-significant interaction ( $P = 0.07$ ) between genotype and shading was detected in CF 2022 were Shading E also

slightly increased RBS in Ruskin (Figure 4.1E, Table 4.3). Consistent with our previous results, the introgression of *VRT-A2b* increased the number of RBS by 1-2 in all three trials and showed a linear response under shading conditions (Figure 4.1C). Furthermore, some of the significant differences in number of RBS between the winter wheat varieties were consistent across the three trials. For example, across all trials Skyfall had a significantly higher number of RBS (1.6-2.4) than Cadenza (0.6-1.5), which overall had the lowest number of RBS together with Robigus (Figure 4.1C). Adamski et al. (2021) found that the 367 accessions they tested (including hexaploid landraces, durum wheat and 98 UK hexaploid cultivars (including Robigus, Ruskin and Skyfall)), all carried the 563-bp sequence in intron-1 and none had the 160-bp sequence rearrangement found in the *VRT-A2* allele from *T. polonicum*. We therefore hypothesise that the cultivars used in this study, besides the *VRT-A2b* NIL, all carry the wildtype *VRT-A2a* allele, which would suggest that genetic components beyond *VRT-A2* affect RBS in UK winter wheat varieties and that these could be further investigated.

Table 4.3: Statistical analysis of phenotypic data, performed individually for each field experiment (^ = Spikelet position 10-12 were analysed as central; ^^ =In 2021, plants were weighed without spike, in 2022 with spike; CF = Church Farm; RBS = Rudimentary Basal Spikelets; ANOVA performed using split-plot function(R package agricolae); o < 0.1, \* < 0.05, \*\* < 0.01, \*\*\* < 0.001)

Phenotype	Trial	Shading A	Shading B	Shading C	Shading D	Shading E	Control	P-value	sig	P-val (Genotype)	sig	P-val (interaction)	sig
RBS	CF 21	1.21 (a)	2.56 (b)				1.10 (a)	< 0.002	**	≤ 0.001	***	0.17	NS
	CF 22			2.89 (b)	2.82 (b)	1.72 (a)	1.56 (a)	≤ 0.001	***	≤ 0.001	***	0.07	o
	Morley 22			2.16 (b)	1.86 (b)	1.05 (a)	0.93 (a)	≤ 0.001	***	≤ 0.001	***	0.93	NS
Grains /central spikelet^	CF 21	3.67 (a)	3.36 (b)				3.75 (b)	0.007	**	≤ 0.001	***	0.60	NS
	CF 22			3.83	3.88	3.95	4.14	. = 0.052	o	≤ 0.001	***	0.23	NS
	Morley 22			4.11	4.15	4.14	4.27	. = 0.28	NS	≤ 0.001	***	0.74	NS
Plant Weight^^	CF 21	5.9 (b)	5.28 (b)				7.33 (a)	0.006	**	0.03	*	0.40	NS
	CF 22			18.89	18.17	16.60	19.10	0.17	NS	0.01	**	0.95	NS
Spikelets/ Spike (Main)	CF 21	20.36	20.40				20.26	0.94	NS	≤ 0.001	***	0.72	NS
	CF 22			23.73	23.99	24.08	23.89	0.8	NS	≤ 0.001	***	0.75	NS
	Morley 22			24.06	23.86	24.32	23.82	0.56	NS	≤ 0.001	***	0.46	NS

To investigate if the critical timeframe for rudimentary basal spikelet formation coincides with the well-defined timeframe of floret abortion, we recorded the number of grains per spikelet in the central and most fertile spikelets of the spike. In 2021, the number of grains per central spikelet was significantly reduced by Shading B ( $P = 0.007$ ), but not by Shading A. In 2022, the number of grains per spikelet were reduced significantly by shading in CF ( $P = 0.05$ ), but not in Morley ( $P = 0.28$ ). Consistent with 2021, we recorded similar trends in reduced floret fertility by Shading C and D at CF, albeit these effects were non-significant. Across all three trials, the central spikelet fertility was not significantly different between *VRT-A2* NILs. The results from 2021, and to a lesser degree from 2022, suggest that basal spikelet abortion might be happening at the same time as floret abortion in the central spikelets and we hypothesised that both are possibly controlled by the same mechanisms.

#### 4.3.2 Complete floret abortion likely cause of rudimentary basal spikelet formation

To investigate the hypothesis that basal spikelet and floret abortion are determined during the same development phase, we harvested and dissected spikes during shading treatments in both 2022 trials (Morley and CF) and recorded floret number and Waddington development stage in the basal six and central two spikelets. In the control conditions at both locations, the number of florets per spikelet increased in all genotypes from 27 to 20 days pre anthesis, except in Skyfall, which had a similar number of living floret primordia between these two stages. At day 20 pre-anthesis, we recorded the maximum number of florets across the time course, varying between 9-10 (interquartile range), with a maximum of 11 florets/spikelet across all genotypes and both locations (only control plots analysed). Previous studies reported similar numbers of maximum florets in winter wheat (Guo et al., 2016), which suggests that in this experiment the second sampling timepoint (20 days pre-anthesis) overlapped the maximum floret stage of wheat spike development. In the following week, the number of living floret primordia per spikelet decreased drastically by 4-6 (interquartile range). florets per spikelet. From 13 to 0 days pre-anthesis, the number of living floret primordia per spikelet decreased only slightly further. Our results align with the findings of previous studies that described the pattern of floret initiation and abortion over the wheat growth cycle (e.g., (González et al., 2011; Guo et al., 2016)).

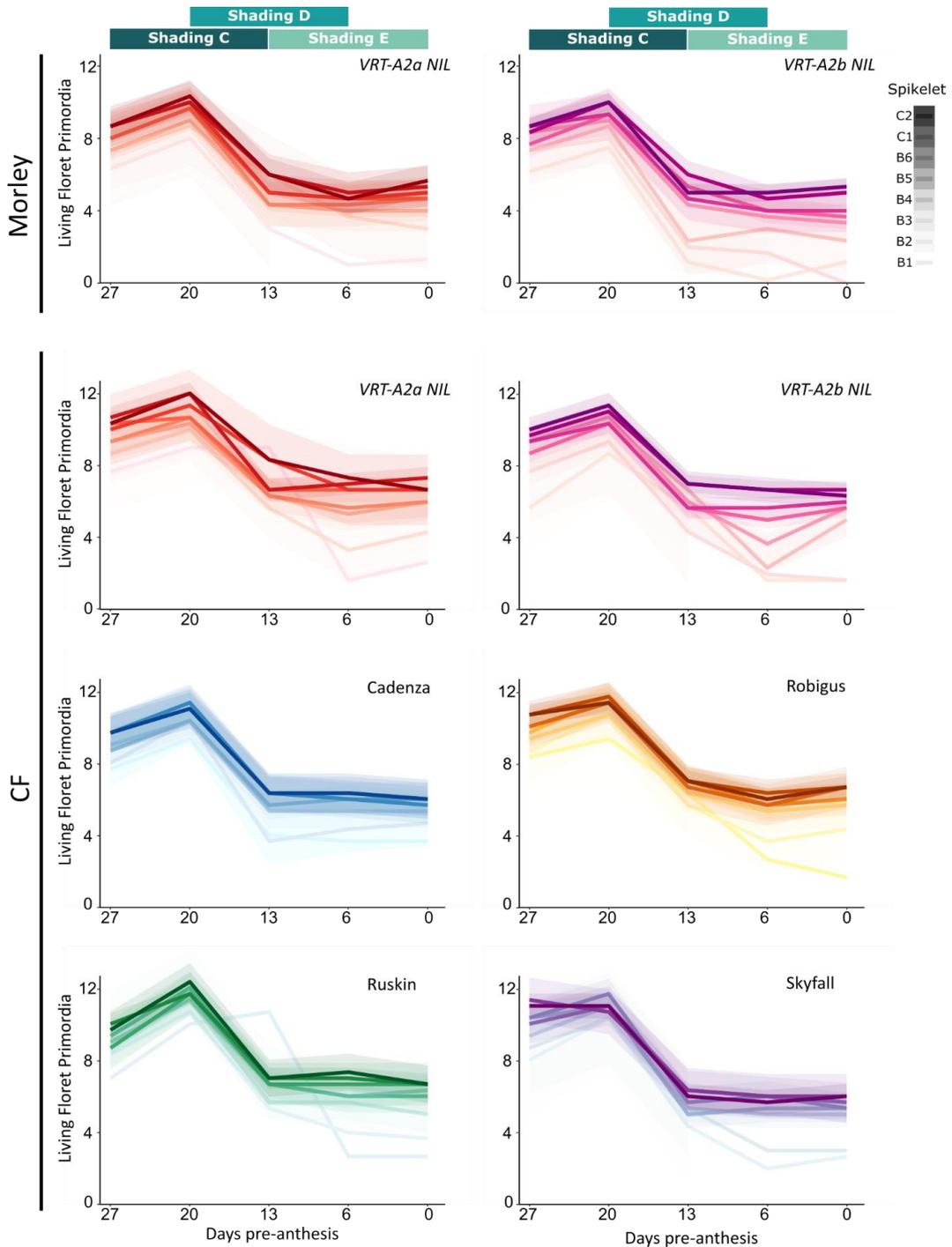


Figure 4.2: Spikelet fertility pre-anthesis in control conditions. We counted the number of living floret primordia per spikelet (y-axis) once per week from 27 to 0 days before anthesis (x-axis), with the first sampling coinciding with the beginning of Shading C. Each week we collected spikes from the control condition (1 spike per block,  $n=3$ ) and dissected the six basal (B1-B6) as well as the two central (C1-C2) spikelets. In CF 2022, we analysed all genotypes, but in Morley 2022 only the VRT-A2 NILs were sampled. Green boxes on top of the graphs represent the timing of the three overlapping shading applications. Colour intensity of the line indicates the spikelet position along the spike (darkest = most central spikelet). Shaded areas represent 95% confidence intervals for each spikelet. Raw data can be found in Supplemental Dataset S2 (<https://doi.org/10.5281/zenodo.7481986>).

Shading C and Shading D both overlapped the critical timeframe of floret abortion (20-13 days pre-anthesis) while Shading E was applied post floret abortion phase. This strengthens our hypothesis that shading only affected RBS numbers when applied during the floret abortion phase (Figure 4.1 D-F), when shading before (Shading A, 2021) or after (Shading E, 2022) the floret abortion phase, the impact of shading on basal spikelet abortion was not significant.

Table 4.4: Number of florets pre-abortion (20 DPA) and the number of florets aborted from 20 DPA to 6 DPA abortion in control, and in control versus Shading D plots. (DPA = Days pre-anthesis; letters = Post-hoc Sidak test significance ( $P < 0.05$ ); ANOVA performed using R base-function; Values estimated mean florets across genotypes; \*  $< 0.05$ , \*\*  $< 0.01$ , \*\*\*  $< 0.001$ )

Spikelet	Florets pre-abortion (20 DPA)	Aborted florets (Control)	Aborted florets (Shading D)	% Abortion (Control)	% Abortion (Shading D)
C2	9.83 (d)	5.12 (ab)	5.54 (a)	52	56
C1	9.58 (cd)	4.75 (a)	5.42 (a)	50	57
B6	9.58 (cd)	5.29 (ab)	5.71 (a)	55	60
B5	9.58 (cd)	5.38 (ab)	6.12 (ab)	56	64
B4	9.29 (cd)	5.67 (ab)	6.25 (ab)	61	67
B3	8.92 (bc)	5.79 (b)	6.67 (b)	65	75
B2	8.5 (b)	6.88 (c)	7.75 (c)	81	91
B1	7.79 (a)	7.12 (c)	7.75 (c)	91	99
P-val	0.001 ***	0.001 ***	0.008 **	-	-

Comparing the number of living floret primordia between basal and central spikelets, we observed that onset of abortion is synchronised across all six basal and the two central spikelets around 20 days pre-anthesis. The number of living floret primordia at/before abortion is relatively similar across all spikelets (overall mean = 9.13; Table 4.4), although the number of living floret primordia is highest in the 2 central and upper 3 basal spikelets (mean= 9.57) and decreases gradually, albeit significantly, from the third basal spikelet (8.92) to the second (8.5) and most basal spikelet (7.79; Table 4.4). Furthermore, floret abortion seems to be more intense in the most basal 2 spikelets, which lose proportionally more florets during the abortion phase than all other spikelets. While central (C1-C2) and upper basal spikelets (B3-B6) lose on average 5.33 florets (or 56% of the initiated florets), the most basal two spikelets abort significantly more florets, 6.88 out of the 8.5 initiated (81%) in the second most, and 7.12 of the 7.79 initiated florets (91%) in the most basal spikelet (Table 4.4). Thus, it is the lower initiation of living floret primordia pre-abortion and the increased loss of florets during abortion that leads to the loss of all florets in basal spikelets in several of the genotypes. The abortion of all florets could be the reason for their rudimentary appearance in the mature spike, suggesting that basal spikelet abortion is simply the consequence of all florets inside the spikelet being aborted and thus all other spikelet structures (lemma, rachilla, glume) also ceasing to develop any further. This would lead to their small and underdeveloped, rudimentary appearance at maturity.

Table 4.5: Sugar concentrations ( $\mu\text{g}/\text{mg}$  tissue weight) after shading treatments across the spike. Values are upper and lower confidence interval.

Sugar	Treatment	Central Spikelet	Basal Spikelet	Central Rachis	Basal Rachis	Peduncle	Internode
Fructose	Control	8.0-10.8	7.4-10.3	7.2-9.9	7.4-10.0	3.0-5.6	0.7-3.4
	Shading B	5.8-8.8	5.3-8.3	3.3-6.0	3.3-5.9	1.5-4.4	0.0-2.3
Glucose	Control	11.4-14.8	10.2-14.8	7.9-10.9	10.3-13.3	11.9-14.9	3.2-6.1
	Shading B	9.6-13.9	9.7-12.8	6.3-9.3	8.7-11.7	9.0-12.2	0.8-3.6

It remains unclear, however, why floret abortion affects basal spikelets more severely. Previous studies hypothesised that basal spikelets would have less priority in assimilate partitioning than the central and apical spikelets (Stockman et al., 1983; González et al., 2011). To test this hypothesis, we measured the concentration of sugars across spikes collected from the control and shading plots at the end of Shading B. We dissected the spikes into basal and central spikelets and rachis and separately collected the peduncle and last internode to be able to investigate differences in sugar assimilation across the spike. We found that shading reduced the sugar concentrations across all sections significantly (Table 4.5, Table 4.6). This is in agreement with the significant reduction in whole plant weight by the shading treatment (Table 4.3) in 2021. However, we did not detect significant differences in sugar concentrations between basal and central spikelets or basal and central rachis for either fructose or glucose (Table 4.6), although basal spikelet and rachis tended to have higher concentrations than central ones. This trend was consistent for fructose and glucose, while sucrose levels were below detection. In 2022, we collected sugar samples again after shading D, and are currently processing the data.

Table 4.6: Statistical analysis of fructose and glucose concentrations ( $\mu\text{g}/\text{mg}$  tissue weight) in dissected spike tissues after Shading B application in shading and control plots. Raw data for table can be found in supplemental dataset S3 (<https://doi.org/10.5281/zenodo.7481986>).

Sugar	Trial	Tissue						P-value	sig
		Central spk.	Basal spk.	Central Rachis	Basal Rachis	Peduncle	Internode		
Fructose	CF2021	8.40 (a)	7.86 (ab)	6.55 (b)	6.60 (b)	3.69 (c)	1.50 (d)	0.004	**
Glucose	CF2021	12.57 (a)	11.64 (ab)	8.59 (c)	11.00 (b)	12.06 (ab)	3.39 (d)	$\leq 0.001$	***
		Treatment				Interaction			
		Control	Shading	P-value	sig			P-value	sig
Fructose	CF2021	6.88 (a)	4.42 (b)	0.0038	**			0.028	*
Glucose	CF2021	10.46 (a)	8.53 (b)	0.0390	*			0.360	NS

4.3.3 Florets in basal spikelets are less developed than same florets in central spikelets  
 Rather than only considering total living floret primordia per spikelet at 20 days pre-anthesis, we wanted to investigate how developmental age of each floret at this timepoint affects their survival chance. To address this, we compared the development stages of each floret from the most basal floret (F1) to the most distal floret from the rachis (F8) across the basal six and central two spikelets (Figure 4.3A; Table 4.7). By taking the mean developmental age of the florets from all genotypes in the CF trial pre-abortion (20 days pre-anthesis), we found that florets in basal spikelets are less developed than their central spikelet counterparts (Figure 4.3A; Table 4.7). For example, floret F1 in the basal spikelets has on average reached Waddington stage 5.4 at this timepoint while the same floret in the central spikelets has on average reached Waddington stage 6.5 (Figure 4.3A; Table 4.7). This additional information is not available when recording only the number of living floret primordia per spikelet as has been the case in previous studies (Stockman et al., 1983; Sibony and Pinthus, 1988; Craufurd and Cartwright, 1989). Comparing the development of the equivalent floret positions (F1 to F1) across the spike reveals that all florets in the basal spikelets (B1) lag behind their central counterparts (C2) by approximately one Waddington stage ( $1.02 \pm 0.04$  SE).

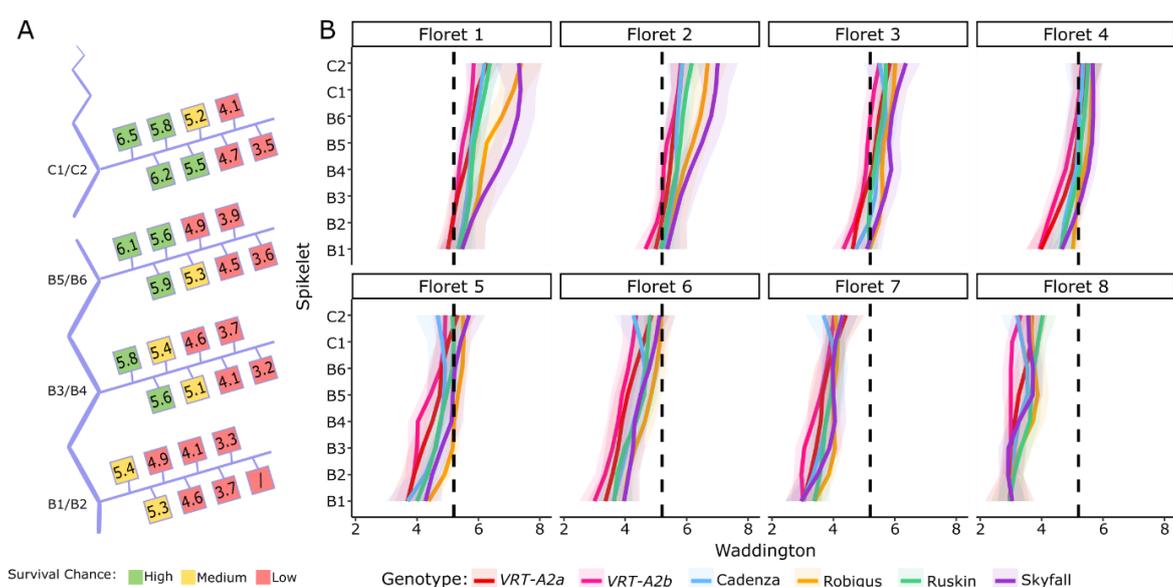


Figure 4.3: Waddington development stage of florets across the spikes 20 days pre-anthesis. (A) Graphical representation of the spike with the basal (B1-B6) and central (C1-C2) spikelets analysed in this study branching off the main rachis (vertical blue). Within each spikelet, the average developmental score of two distichous spikelets across all genotypes (in CF 2022, Control) is represented by the coloured squares from the most proximal floret (F1, left) to the most distal floret (F8, right). Survival chance estimation is indicating by colouring, florets with Waddington stage  $\geq 5.5$  are green, between W5 and W5.5 yellow and below W5 are red. (B) Waddington stage of florets F1 to F8 across the spike, from basal (B1-B6) to central (C1-C2) spikelets. Colours represent genotypes. The black dotted line is positioned at Waddington stage 5.25 (as an arbitrary transition point of survival chance). Shaded area represents 95% confidence interval.

Table 4.7: Average development stage (Waddington) of floret 1-8 in basal (B) six and central (C) two spikelets across all genotypes 20 DPA (Control treatment only). Average was taken of the neighbouring pairs of spikelets (B1&B2, B3&B4, B5&B6, C1&C2)

	<b>B1</b>	<b>B2</b>	<b>Average</b>	<b>B3</b>	<b>B4</b>	<b>Average</b>	<b>B5</b>	<b>B6</b>	<b>Average</b>	<b>C1</b>	<b>C2</b>	<b>Average</b>
Floret 1	5.3	5.5	5.4	5.7	5.8	5.8	6.1	6.2	6.1	6.5	6.6	6.5
Floret 2	5.1	5.4	5.3	5.6	5.6	5.6	5.9	6.0	5.9	6.2	6.2	6.2
Floret 3	4.9	5.1	5.0	5.3	5.4	5.4	5.5	5.6	5.6	5.8	5.8	5.8
Floret 4	4.5	4.6	4.6	5.0	5.2	5.1	5.3	5.4	5.3	5.4	5.5	5.5
Floret 5	3.9	4.3	4.1	4.6	4.7	4.6	4.9	5.1	5.0	5.1	5.2	5.2
Floret 6	3.6	3.7	3.7	4.1	4.2	4.1	4.4	4.6	4.5	4.7	4.8	4.7
Floret 7	3.1	3.4	3.3	3.6	3.7	3.7	3.9	3.9	3.9	4.1	4.1	4.1
Floret 8	NA	NA	NA	3.2	3.3	3.2	3.6	3.6	3.6	3.4	3.6	3.5

It is generally accepted that floret survival after abortion, rather than floret number pre-abortion, is the determining factor of final grains per spikelet (Langer and Hanif, 1973; Fischer and Stockman, 1980; Whingwiri and Stern, 1982; Sibony and Pinthus, 1988; González et al., 2003, 2005; Ferrante et al., 2010). These studies, however, did not consider differences in development pre-abortion. We therefore wanted to test if the differences in developmental age of the basal spikelets might be responsible for the greater floret abortion and subsequently, lower grain numbers. To explore this idea, we hypothesised that the more developed florets are the less likely they are to be aborted and that, therefore, florets need to have reached a minimum Waddington stage to have a high chance of survival. Using data collected at maturity (final grain number per spikelet; Table 4.8), we determined that under field conditions in the 2022 trials, florets beyond Waddington 5.5 at 20 days pre-anthesis had a very high chance of survival whereas florets below Waddington 5.0 had a low chance of survival. To illustrate this concept, we coloured the florets in Figure 3A as either red (low survival chance, Waddington < 5.0), yellow (medium survival chance,  $5.0 \leq$  Waddington < 5.5) or green (high survival chance, Waddington  $\geq$  5.5). Using this criterion, we would predict that across all spikelet positions, florets F6, F7 and F8 have a very low survival chance as they do not pass this threshold, while floret F5 would have a medium chance of survival only in central spikelets and in some genotypes (Figure 4.3A-B). The most basal florets (F1 and F2) would be assumed to have a very high chance of survival in all spikelets except the most basal spikelets (B1 and B2). It is important to note that this the threshold of survival at Waddington 5.5 was chosen as it best fits the actual number of grains per spikelet observed in mature spikes (ca. 4 grains per central spikelet, Figure 4.1H-I).

Table 4.8: T-test comparison of predicted versus actual grain values per spikelet ( $\alpha < 0.1$ , \*  $< 0.05$ , \*\*  $< 0.01$ , \*\*\*  $< 0.001$ )

		Mean (actual)	Mean (5.5)	P (5.5)	Sig	Mean (optimised)	P (optimised)	Sig
Genotype	VRT-A2a	3.09	2.35	0.194		3.12	0.940	
	VRT-A2b	2.4	1.49	0.099	o	2.47	0.903	
	Cadenza	2.89	2.07	0.100	o	2.89	0.986	
	Robigus	3.01	3.08	0.914		3.06	0.942	
	Ruskin	3.03	2.19	0.197		3.01	0.971	
	Skyfall	2.68	3.13	0.530		2.8	0.854	
	C2	4.11	3.79	0.407		4.11	0.997	
Spikelet	C1	4.24	3.62	0.139		3.98	0.200	
	B6	3.81	2.90	0.009	**	3.42	0.004	**
	B5	3.65	2.65	0.021	*	3.17	0.012	*
	B4	3.05	2.07	0.026	*	2.68	0.135	
	B3	2.41	1.67	0.099	o	2.43	0.940	
	B2	1.09	0.97	0.717		1.8	0.043	*
	B1	0.23	0.49	0.155		1.36	0.002	**

#### 4.3.4 Differences in development pre-abortion can predict reduced grain numbers in basal spikelets

The apparent relationship between florets above W5.5 and final grain number raises the question whether the developmental age of florets at the onset of abortion affects their likelihood to survive the abortion process itself. To test this hypothesis, we predicted the number of grains per spikelet for each genotype in control conditions, using the floret data pre-abortion (20 days pre-anthesis). Based on its Waddington stage, we assigned a probability of survival (a cumulative normal distribution function, CDF, with a mean = 5.5, sd = 0.195) for each floret to avoid abortion and produce grain. This allowed us to assign a 'survival probability' to each floret regardless of its position within the spike or spikelet. The sd value was based on the sd for grain number per spikelet in the mature spike. The survival probability increases as development advances and this is captured well by a cumulative normal distribution (or similar saturation functions such as the Hill or logistic function). For example, the survival probability of florets in Waddington stage 5.5 is 0.5, while the survival probability for florets in Waddington stage 6 is nearly 1 (0.99) and for florets in stage 4.5 it is close to 0 ( $1.4 \times 10^{-7}$ ). Summing over these probabilities, rather than counting the number of florets above W5.5, allowed us to compute the expected number of florets that survive whilst accounting for a degree of uncertainty in the survival rate of florets based on the Waddington score.

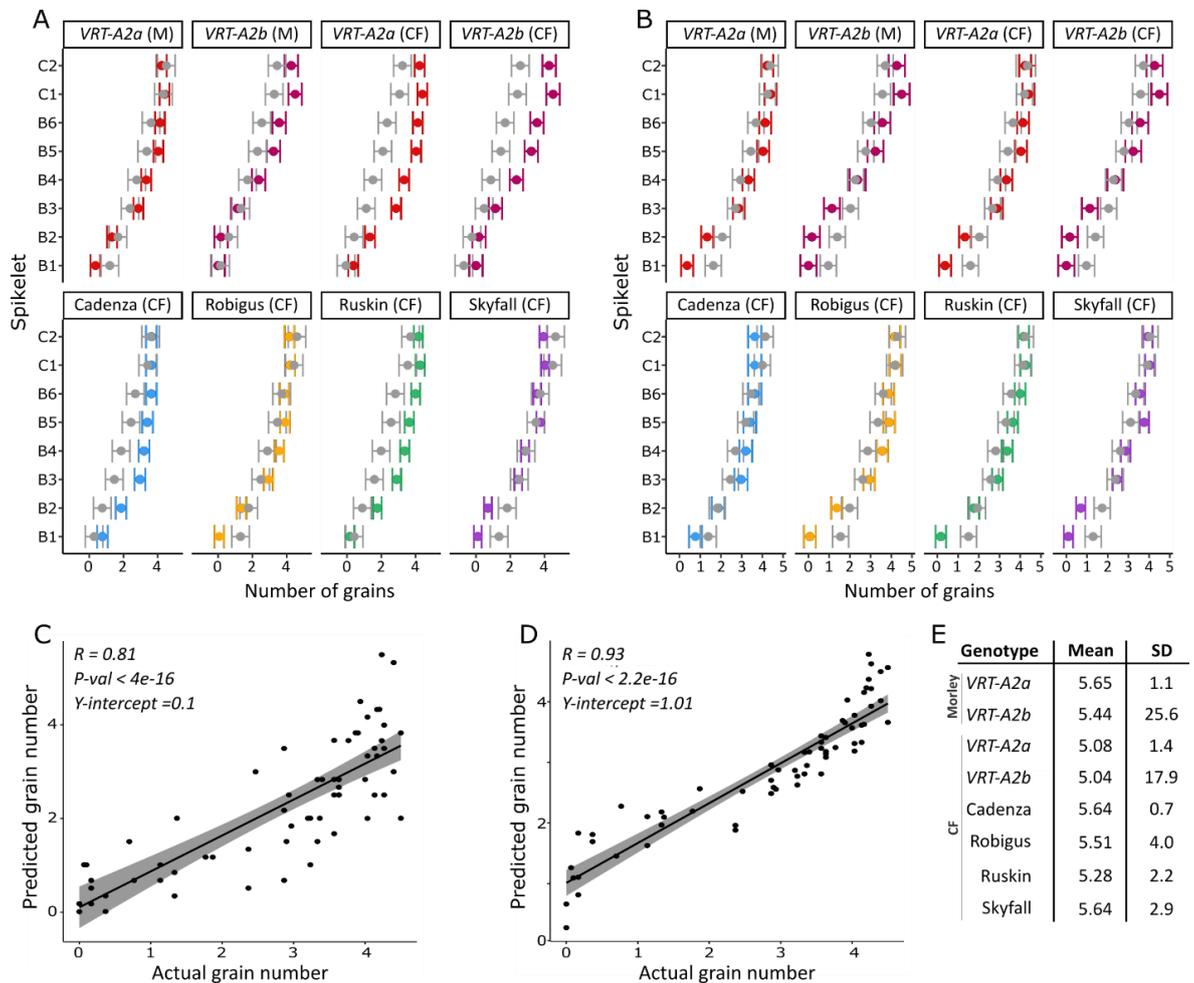


Figure 4.4: Prediction of grains per spikelet from floret development pre-abortion (A) Number of grains per spikelet recorded in mature spike (coloured dots) and predicted number of grains (Gray dots), resulting from the expected number of survived florets for florets 1-8, calculated from a cumulative normal distribution with a half-maximal value at the Waddington stage of 5.5 (mean) and a survival decay characterised by a standard deviation of  $SD = 0.195$  (.). (B) Number of grains predicted per spikelet\*Genotype after optimisation of the mean and standard deviation of the normal distribution for each genotype individually (grey dots) versus actual grain numbers recorded at maturity (coloured dots). Data in A and B is shown confidence interval (bars) plus mean (dot). (C) Linear regression and fit of Actual grain number versus predicted grain number using data from (A). (D) Linear regression and fit of actual versus predicted grain numbers using data predictions shown in (B). (E) Resulting mean and SD values for each Genotype\*Experiment from function optimisation by genotype (Data in panel A&D). B = Basal (spikelet), C = Central (Spikelet). Raw Data for actual grains/spikelet can be found in Supplemental Dataset S1, predicted number of grains/spikelet can be found in Supplemental dataset S4 (<https://doi.org/10.5281/zenodo.7481986>).

Using this method, we calculated the floret survival probability for florets 1 to 8 within a spikelet and then summed up the probabilities, leading to a predicted grain number per spikelet (Figure 4.4A, transparent/red colours). Comparing these values to the actual number of grains per spikelet recorded in mature spikes (Figure 4.4A, solid colours) reveals a close fit. This approach predicted successfully for basal spikelets to have the least grains/spikelet, and for grain numbers to increase steadily towards the central spikelets, as is the case in the mature spike data. The correlation between the actual and predicted grain numbers was high (0.81) and the y-intercept was close to 0 (Figure 4.4B). Using only the number of florets per spikelet pre-abortion, without considering their Waddington stages, fails to predict the gradient across the spike and has consequently a much worse fit to the actual data (Figure 4.5).

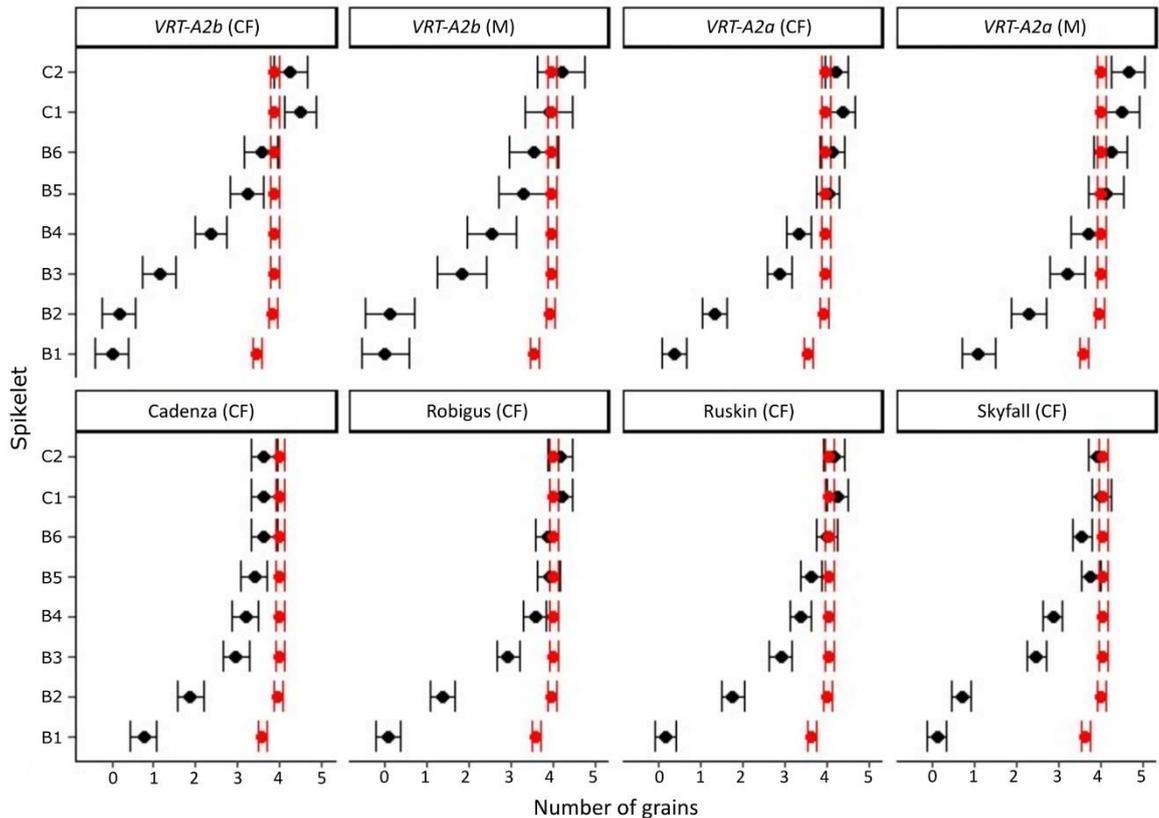


Figure 4.5: Predictions of grains per spikelet using floret count per spikelet pre-abortion (20 DPA) and deducting same number of florets per spikelet (4 florets) across all spike positions. Black = Number of grains per spikelet recorded in mature spike, Red= predicted grains/spikelet. B = Basal (spikelet), C = Central (Spikelet).

Using a t-test, we found that there are no significant differences between the actual and predicted grain values for most of the genotypes. The predictions that fit least well were for Cadenza ( $P = 0.10$ ) and VRT-A2b ( $P = 0.10$ ) (Table 4.8). Similarly, not all spikelets were fitted equally well, and for the 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> spikelet from the base, predictions were significantly lower from the actual grain values (Table 4.8,  $P < 0.05$ ). Next, we explored if the mean and SD of the CDF could be optimised individually for each genotype to achieve a better fit of the actual data. For this we again used the cumulative normal distribution probability function but optimised the mean and SD fully unconstrained using simulated annealing and the downhill simplex method (both gave consistent results). The resulting predictions significantly matched the actual number of grains/spikelet (Figure 4.4C) and improved the overall correlation to  $R = 0.93$  (Figure 4.4D). When comparing the optimised predicted grain values to the actual grain values per genotype, we found no significant nor borderline significant differences. The optimisation led to each genotype having an individual mean (Waddington) stage as turning point for survival probability (Figure 4.4E).

Despite achieving an overall better fit, the correlation between actual and optimised predicted grain values had a higher y-intercept (1.01), which suggests that the approach slightly over-predict low grain values. This was also confirmed by the t-test, which found significant differences between the actual and predicted values for the 1<sup>st</sup> and 2<sup>nd</sup> most basal spikelet (Table 4.8), especially in Robigus and Ruskin. Overprediction of the lower grain values might be due to the optimisation algorithm favouring a better fit of central and upper spikelet values as the errors are larger in

proportion and because we included more spikelets with higher grain values, leading to an unequal distribution in datapoints (Figure 4.4B&D). More data over the full range will allow us to test this hypothesis and validate the proposed model.

## 4.4 Discussion

### 4.4.1 Basal spikelet abortion is likely the consequence of complete floret abortion

In wheat and other grasses, the most basal spikelets are generally less productive than central spikelets and are often only developed in a rudimentary form. Basal spikelets are initiated first, yet the number of grains per spikelet is lowest in the most basal spikelets and increases gradually towards the centre of the spike. Although previous studies have highlighted that basal spikelets are less productive and more readily aborted under stress conditions (Stockman et al., 1983; Savin and Slafer, 1991; Ferrante et al., 2020), the causes of rudimentary basal spikelet formation have not been studied in detail. Previously, we found that basal spikelets are delayed in growth and development immediately after their initiation due to lower expression of spikelet initiation genes than in central spikelets (Backhaus et al., 2022). However, basal spikelets continue to grow and develop throughout the crop cycle and their rudimentary status is finalised later. In this study, we investigated the timing and physiological mechanisms of basal spikelet cessation.

We found that applying resource-limiting shading treatments in the field between 13 to 20 days pre-anthesis significantly increased rudimentary basal spikelet (RBS) numbers by between 0.93 to 1.46 in three independent experiments (Figure 4.1). We therefore hypothesised that basal spikelet abortion is happening during this very defined timeframe and is highly sensitive to resource availability in this period. The timeframe identified for basal spikelet abortion overlaps stem elongation phase and has previously been termed the 'critical period' of wheat development (Fischer, 1985; Siddique et al., 1989; Savin and Slafer, 1991). During the critical period, the stem and spike are undergoing maximal growth and are rapidly accumulating biomass, which has been proposed to be a competitive process between the stem and spike (Fischer and Stockman, 1980; Siddique et al., 1989). Furthermore, floret abortion is also happening during the 10-20 days pre-anthesis. Previous studies showed that shading during the critical period significantly decreases the number of grains per spikelet (Stockman et al., 1983; Savin and Slafer, 1991; Slafer et al., 1994). Thus, through a variety of studies in which floret fertility was manipulated using genetic or environmental factors, the initiation of floret abortion has been tightly linked to the stem extension phase where resource availability is directly, or indirectly, determining the number of florets that will survive abortion. In 2021, we found that shading B significantly reduced central spikelet fertility and in the 2022 experiments, shading C and D had a negative, although not significant, effect on central spikelet fertility (Figure 4.1G-I). Dissections of the spike during the shading treatments in 2022 confirmed that the two shading treatments leading to increased number of RBS (C and D)

overlapped the week of floret abortion, while shading E was applied post the floret abortion phase (Figure 4.2). We therefore concluded that shading increases basal spikelet abortion only if applied during the 'critical period' of spike growth and floret abortion.

As basal spikelet abortion is happening at the same time as floret abortion and is affected similarly to floret abortion by shading, it raises the question if spikelet abortion is simply the consequence of all florets being aborted in these spikelets. The nearly complete abortion of florets in the most basal spikelet across all genotypes supports this idea. Unlike spikelets, the number of florets is indeterminate and each spikelet initiates many floret primordia (in this study on average 9 florets/spikelet), of which most are aborted during the critical phase. Abortion of florets from 6 to 20 days pre-anthesis was strongest in basal spikelets, where on average 91% of the initiated florets are aborted. Contrary to this, in the two central spikelets analysed only 49% and 52% of the florets were aborted during the same time (6-20 days pre-anthesis; Table 4.4). To further understand rudimentary basal spikelet formation, we therefore investigated what causes the disproportionality high abortion of florets in basal spikelets.

4.4.2 Lower resource availability is an unlikely cause for low grain set in basal spikelets  
Previously, Gonzalez (2011) and Stockman (1983) suggested that the increased abortion of florets in the base is due to reduced resource assimilation. This hypothesis is based on the general finding that floret abortion is increased by reducing overall assimilate availability and that basal spikelets have less dry matter weight at anthesis. Furthermore, this hypothesis is supported by the repeated finding that basal spikelets are most reactive to changes in the environment, i.e., losing or gaining over-proportionately more florets when source strength is altered (Stockman et al., 1983; Savin and Slafer, 1991). Stockman et al. (1983) furthermore suggested that the reduced dry matter weight of basal spikelets indicates their lower assimilate priority. However, as dry matter is measured post-abortion, it cannot be determined if the reduced dry matter weight is due to less resource availability or due to the florets being aborted and thus less tissue growth being supported after their abortion. Stockman et al (1983) are, to our knowledge, also the only study that directly measured carbohydrate concentrations in the apical, central and basal sections of the spike (albeit in controlled environments and with single stemmed plants). Confirming results from their previous study on the whole spike (Fischer and Stockman, 1980), they found that soluble carbohydrate concentrations peak in the spike 12 days pre-anthesis in central and apical spikelets (Stockman et al., 1983). Interestingly, carbohydrates peaked 3 days later in the basal (2<sup>nd</sup> and 3<sup>rd</sup>) spikelets. They concluded that florets in basal spikelets are more readily aborted as they are, "the sink of lowest priority in the spike" (Stockman et al., 1983). However, this interpretation ignores the fact that basal spikelets accumulate the highest percentage maximum carbohydrate concentration of 30-40% (of the dry matter weight) at 15 days pre-anthesis, greater than the central and apical spikelets which have maximum carbohydrate concentration of 20-25%. Thus at 12 days pre-anthesis, both basal

and central spikelets have an equivalent concentration of ~20-25%. Our data supports their finding as we didn't detect significant differences in sugar concentrations between the central or basal spikelet or rachis, although we also found that there is a tendency of slightly higher sugar concentrations at the base (Table 4.5). In conclusion, re-analysis of the Stockman et al. (1983) data and our results do not support the hypothesis that basal spikelets have less resource availability. However, both datasets support that shading decreases resource availability to the whole spike, which increases floret abortion across all spikelets (Table 4.5) (Stockman et al., 1983).

#### 4.4.3 Differences in grain set across the spike can be predicted using floret development stages pre-abortion

To understand the causes of increased floret abortion in basal spikelets we divided the question into two parts. One part concerns the factors that initiate floret abortion and determine the overall degree/strength of abortion. As discussed above, the extent of abortion is largely decided by the availability of resources. The second question concerns the order of floret abortion across the spike, which has been shown to always start in the most distal florets of the spikelet and, if resources are limited, moves inwards within a spikelet. Factors proposed to make distal florets more likely to abort include the distance from the rachis (Kadkol and Halloran, 1988), their size, or the developmental age of the floret at the time of abortion (Ferrante et al., 2020). As the most distal florets of the spikelet are also the youngest and smallest, these factors cannot easily be disentangled.

In 2020, Ferrante et al. showed that the improved grain set in a modern cultivar stemmed from its faster rate of floret development pre-abortion, which improved the survival rate of the more distal florets, compared to the traditional cultivar. Furthermore, they found that lower nitrogen levels reduced floret development rates and thus negatively affected floret survival. However, the response of the cultivars was linear, meaning that the improved floret development in the modern cultivar was still beneficial. The study thus connected pre-abortion development of florets to their likelihood of surviving abortion and showed that environmental factors negatively affecting development pre-abortion reduced the survival chance of distal florets (Ferrante et al., 2020). A similar observation had been previously made in barley, where the chance to survive abortion was highly dependent on the development stage attained at the start of floret primordia mortality (Arisnabarreta and Miralles, 2006).

We tested this hypothesis with our data and investigated if reduced development pre-abortion in basal spikelets could be the cause of their increased abortion. We calculated the survival probability of a floret based on its Waddington stage at maximum floret development pre-abortion (20 days pre-anthesis in our 2022 trial, Figure 4.2). Therefore, the survival probability of a floret was independent of its position along the spike or spikelet. We used the sum of the survival probabilities of the first eight florets within a spikelet as prediction of the number of grains per spikelet. We

found that there was a good fit between predicted and actual grain numbers per spikelet if survival probability increased once florets had passed Waddington stage 5.5 (Figure 4.4). The predictions based on Waddington stage 5.5 were able to capture the gradient of grains per spikelet from the centre to the base of the spike and predicted lower grain values in the basal spikelets compared to central for all genotypes (Figure 4.4). This aligns with the result of Ferrante et al. (2020) that development stage of florets at abortion is highly relevant for their likelihood of survival and supports our hypothesis that nearly all florets in basal spikelets are aborted due to their reduced development and not due to reduced resource availabilities.

In our study, the closer a floret was to reaching Waddington stage 5.5, the higher its survival chance became. Waddington stage 5.5 was chosen as the previous naïve analysis of all floret data combined suggested that this stage was a good threshold to match grains per spikelet (Figure 4.3A). Furthermore, we also performed an unconstrained optimisation for each genotype to find the best parameters for the floret survival probability. This allowed us to find the Waddington stage for each genotype that results in the best predictions. Interestingly, these optimised values ranged between 5.0 to 5.7 for the six different genotypes and two locations, highlighting that Waddington stage 5.5 is indeed an important stage to be reached for high floret survival.

Reanalysis of the Ferrante et al. (2020) data also lends support to Waddington stage 5.5 as the onset of floret abortion and important checkpoint. At this stage (equivalent to  $-270\text{ }^{\circ}\text{C d}$  from anthesis in their study), all of the florets that reached 5.5 completed development up to anthesis, and conversely 91% of the florets that did not reach anthesis also didn't reach 5.5 at  $-270\text{ }^{\circ}\text{C d}$  (Figure 4.6). For the different genotypes and conditions investigated in our study and in Ferrante et al (2022) Waddington stage 5.5 gave good predictions, suggesting that it is rather the developmental rate pre-abortion that changes between genotypes and conditions, rather than the Waddington stage checkpoint itself. Further studies would be warranted to elucidate the overall applicability of this hypothesis. Investigations of the molecular mechanisms happening when a floret reaches Waddington stage 5.5 will be important next steps to understand the importance of this stage for floret abortion.

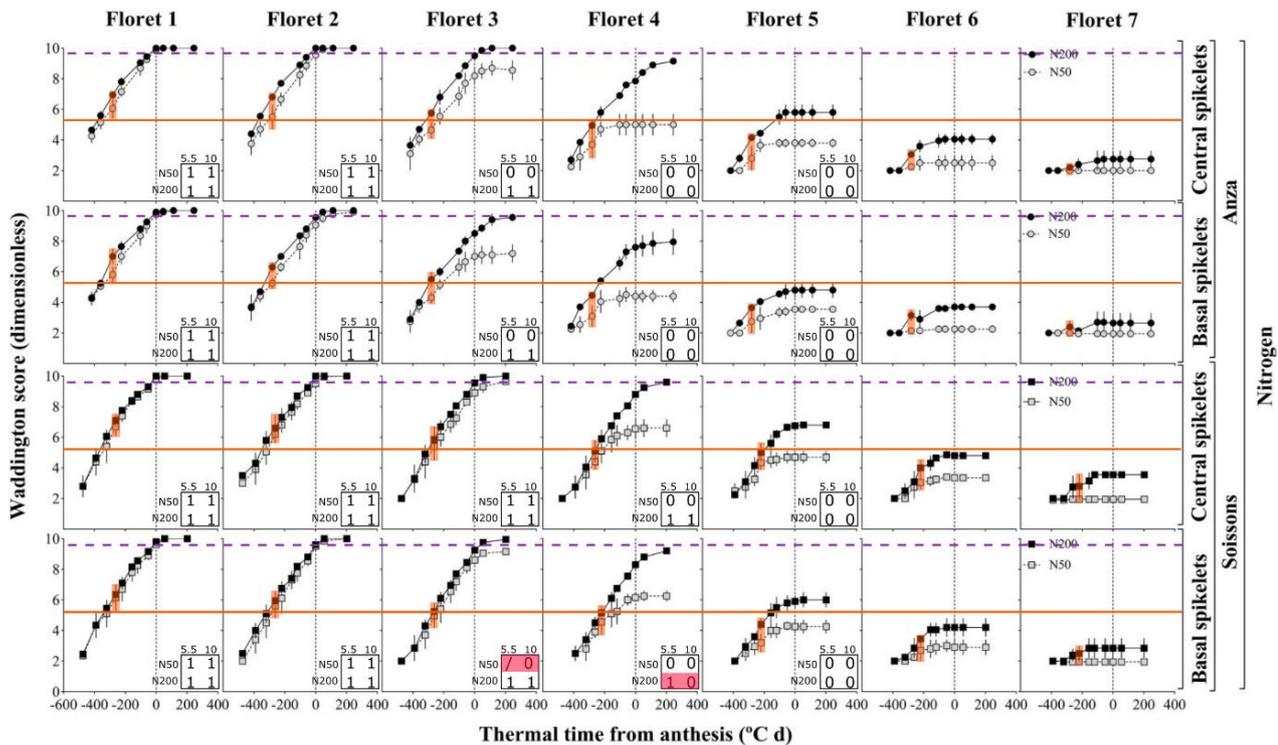


Figure 4.6: Re-analysis of Ferrante et al (2020) Figure 9. Orange bar was added to mark Waddington stage 5.5 and the data points corresponding to maximum floret number stage pre-abortion (circa -270 °C d) were highlighted in orange to determine if florets had reached Waddington stage 5.5 at that time. Purple dotted line indicates cut-off for florets considered to have reached maturity (W10). Out of the 56 floret development traces, 22 reached Waddington stage 10 by the end of the time course and all of these had passed Waddington stage 5.5 at -270 °C d (which corresponds to maximum floret number stage according to Figure 8 (Ferrante et al., 2020)). Of the 34 florets that did not reach Waddington stage 10 at maturity, 91% had not reached W5.5 at -270 °C. Only floret 4 of the basal spikelet in Soisson (high nitrogen) had reached W5.5 but then did not develop until the W10 cut-off (purple line) while the development stage of floret 3 in the same genotype and spikelet at nitrogen level 50 could not be accurately scored. Decisions on floret development are indicated in box on bottom right of each panel. If florets did not reach the stage they were scored 1, if they didn't they were scored 0. Red shading indicates that floret development at W5.5 does not match their development stage at W10. As all florets were infertile beyond floret 5 we stopped analysis beyond floret 5. Original legend from Ferrante et al (2020), "Dynamics of the floret development from floret 1 (F1, floret primordium closest to the rachis) to floret 7 (F7, floret primordium most distal to the rachis) in each of the two spikelet categories considered of the main-shoot through thermal time from anthesis (negative values represent the period before anthesis) in the N experiments for Anza and Soissons. Grey and black symbols correspond to N50 (50 KgN ha<sup>-1</sup>) and N200 (200 KgN ha<sup>-1</sup>). Each data-point is the average of all replicates across two growing seasons and within each replicate the value was the average of 10 (2010-11) and 5 plants (2011-12), bars represent the standard error of the means (not visible in some cases as it was smaller than the body of the symbol)."

The initial and optimised grain predictions were both able to predict the gradient across the spike, however some spikelets along the spike were predicted worse than others. Focusing on the predicted values using the common mean value of 5.5, the grains in central spikelets of some genotypes were under-predicted. Using the optimised predictions, grains per central spikelet were predicted highly accurately, however the predictions for the 1<sup>st</sup> and 2<sup>nd</sup> most basal spikelet were significantly different from the actual grain numbers (Figure 4.4). This suggests that even though the predictions based on Waddington stage pre-abortion can account for the majority of the observed variation, additional factors likely play a role in determining abortion differences in the central versus the basal spikelet.

Despite these potential shortcomings, using this very simple rule we were able to predict the grains per spikelet in basal and central spikelets using a general framework. This suggests that the signalling pathways of floret abortion might be the same in central and basal spikelets. The severe delay in floret development in basal spikelets from spikelet initiation until stem elongation might thus explain rudimentary basal spikelet formation, rather than the previously proposed hypothesis of reduced resource availability in basal spikelets. A similar mechanism has been proposed in barley. Unlike wheat, barley has determinate spikelets and an indeterminate spike and therefore spikelet abortion has been studied in much more detail (Alqudah and Schnurbusch, 2013). Under salt stress conditions, all growth stages pre-abortion are shortened and spikelet growth and development is diminished, leading to increased abortion of apical and basal spikelets (Boussora et al., 2019).

#### 4.4.4 Improvement of basal spikelet fertility through targeting pre-abortion development

It has been proposed that a reduction in the variation in spikelet fertility across the spike could be a promising avenue to increase yields and improve grain size homogeneity in breeding programs of small cereals grains (Arisnabarreta and Miralles, 2006; Philipp et al., 2018). Our results suggest that this would not be possible by reducing abortion, but rather through improving spikelet and floret development pre-abortion. Reducing the amount of abortion either through improved genetics (such as *GN1* introgression; Sakuma et al., (2019)) or increased resource availability leads to increase grains per spikelet. However, this seems to always be in a linear fashion across the spike and would not specifically improve basal spikelets compare to central.

Interestingly, when the survival probability function was left to vary freely during optimisation, the resulting means all fell within a range of 5.0 to 5.7, consistent with our hypothesis that Waddington stage 5.5 is an important developmental stage for floret abortion survival. Overall, the differences in mean values found by optimising the function for each genotype matched the flowering dates of the varieties. The CDF mean Waddington stage value was highest in Skyfall and Cadenza (in CF trial), followed by Robigus, Ruskin, *VRT-A2a* and *VRT-A2b* (Figure 4.4D), which matches the sequence of flowering dates of these genotypes scored in control plots in 2022 CF trials (Table 4.2), with Skyfall flowering first and *VRT-A2b* flowering last.

This suggests that Skyfall florets were more advanced and thus the common mean of 5.5 overpredicted grains. Skyfall also appeared to be slightly ahead of the other genotypes in floret development as it was already at maximum floret potential a week before the other genotypes (Figure 4.2). Genetically mapping the effect of Skyfall on RBS would thus potentially be able to uncover further genes involved in the control of pre-abortion development.

In this study, we included a set of NILs for *VRT-A2*, a MADS-box transcription factor, previously shown to increase RBS numbers under controlled and field conditions (Backhaus et al., 2022). We confirmed that RBS was increased by *VRT-A2b* and furthermore found the allele to have no

interaction with shading (Figure 4.1). Using the same parameters, we were able to predict the grains per spikelet as accurately as for the wildtype *VRT-A2a* NIL. This suggests that the introgression of *VRT-A2b* indeed affects pre-abortion development rather than increasing abortion *per se* in the basal spikelets. This is also supported by the subtle overall delay in development across florets in *VRT-A2b* NILs by 0.21 Waddington stages compared to the wildtype at the onset of abortion (Table 4.9). Optimisation of the cumulative distribution function furthermore came to very similar mean values for both NILs in CF (*VRT-A2a* = 5.08 and *VRT-A2b* = 5.04) as well as in Morley (*VRT-A2a* = 5.65 and *VRT-A2b* = 5.44).

Table 4.9: Average development stage (Waddington) of florets 1-8 in basal (B) six and central (C) two spikelets 20 DPA (Control treatment only) for *VRT-A2* NILs.

	Spikelet B1			Spikelet B2			Spikelet B3			Spikelet B4		
	<i>VRT-A2a</i>	<i>VRT-A2b</i>	Delta									
Floret 1	5.00	5.00	0.00	5.17	5.17	0.00	5.33	5.33	0.00	5.50	5.33	0.17
Floret 2	5.00	4.67	0.33	5.17	5.00	0.17	5.33	5.33	0.00	5.50	5.17	0.33
Floret 3	4.67	4.33	0.33	4.67	4.67	0.00	5.00	5.00	0.00	5.33	5.00	0.33
Floret 4	4.00	4.00	0.00	4.33	4.00	0.33	4.67	4.67	0.00	5.17	4.67	0.50
Floret 5	3.67	3.67	0.00	4.00	4.00	0.00	4.00	4.00	0.00	4.67	4.00	0.67
Floret 6	3.33	3.00	0.33	3.67	3.33	0.33	3.67	3.67	0.00	4.00	3.67	0.33
Floret 7	3.00	3.00	0.00	3.33	3.00	0.33	3.33	3.00	0.33	3.67	3.33	0.33
Floret 8				3.00				3.00		3.33	3.00	0.33
Floret count	7.3	7	0.3	8.3	7.7	0.6	8.7	8	0.7	9	8.3	0.7

	Spikelet B5			Spikelet B6			Spikelet C1			Spikelet C2		
	<i>VRT-A2a</i>	<i>VRT-A2b</i>	Delta									
Floret 1	5.83	5.50	0.33	5.83	5.67	0.17	5.83	5.83	0.00	6.33	5.83	0.50
Floret 2	5.50	5.33	0.17	5.67	5.67	0.00	5.83	5.67	0.17	5.83	5.83	0.00
Floret 3	5.33	5.17	0.17	5.50	5.17	0.33	5.67	5.33	0.33	5.83	5.50	0.33
Floret 4	5.00	5.00	0.00	5.33	5.17	0.17	5.33	5.17	0.17	5.50	5.33	0.17
Floret 5	4.67	4.33	0.33	4.83	5.00	-0.17	5.00	4.67	0.33	5.33	5.00	0.33
Floret 6	4.00	4.00	0.00	4.33	4.00	0.33	4.67	4.33	0.33	4.83	4.33	0.50
Floret 7	3.67	3.67	0.00	3.67	3.67	0.00	4.33	4.00	0.33	4.33	4.00	0.33
Floret 8	3.00	3.00	0.00	3.67	3.00	0.67	3.67	3.00	0.67	3.67	3.33	0.33
Floret count	9	8.7	0.3	9.7	8.7	1	10.3	9.3	1	10.3	9.7	0.6

In Morley, we only collected data for the NILs and the optimisation found a better fit for both NILs using a more advanced Waddington stage mean than in CF (Figure 4.1E). This might suggest that the plants were marginally more advanced in development in Morley, although flowering dates are not available for this experiment to explore this hypothesis. The differences in mean Waddington stages to accomplish better fits for individual genotypes might be mainly correcting for differences that arose because we sampled all genotypes on the same day. Thus, we did not account for developmental differences between the genotypes as there may have been varietal variation, as suggested by the differences in flowering dates (Table 4.2).

Furthermore, our study highlights the importance of the 'critical phase' of wheat development, adding basal spikelet abortion to the traits affected during this phase. Reducing abortion during this phase is a promising avenue for future yield increases and as the spike is particularly sensitive to resource limitation in the 10-20 days pre-anthesis. Management practices might be a promising tool to reduce abortion as precise application of fertiliser at maximum floret stage could reduce abortion.

Ferrante et al. (2020) showed that reduced nitrogen application throughout the growth season (by 75%) slows down the development of florets, which leads to an increase in florets that are not advanced enough to survive abortion. The same effect of reduced nitrogen on floret survival had previously been proposed by Abbate et al. (1995), however they only recorded the effect of nitrogen reduction on grains/m<sup>2</sup> and did not dissect the trait further. It would be interesting to investigate if applying important signalling compounds of resource availability, such as T6P, during the time of floret abortion could reduce abortion rates and thus counter-act increased floret abortion in low nitrogen conditions.

In this study we found that it is the reduced development of floret primordia pre-abortion in basal spikelets that can explain the increased loss of florets during abortion. In several genotypes, all florets are aborted in the basal spikelet, which we propose to be the reason for their rudimentary appearance in the mature spike. Thus, it would be the initial establishment of the basal spikelets and their development rates pre-abortion that need to be targeted to improve homogeneity across the spike.

## 4.5 Materials and Methods

### 4.5.1 Genetic material and plant growth

Wheat germplasm used in this study includes hexaploid UK wheat cultivars Cadenza, Robigus, Ruskin, Skyfall, and near isogenic lines (NILs) differing for the *P1* locus described in Adamski et al. (2021). We used two sibling BC<sub>6</sub> NILs with Paragon as the genetic background, differing for the presence of the wildtype *VRT-A2a* allele or the *T. polonicum* *VRT-A2b* allele. We evaluated cultivars and NILs in three field experiments. One trial was located at The Morley Agricultural Foundation trials site, Morley St Botolph, UK (52°33'15.1"N 1°01'59.2"E) in 2021/22. Two trials were located at the John Innes Centre Experimental trials site in Bawburgh (Church Farm), UK (52°37'50.7"N 1°10'39.7"E) sown in 2020/21 and 2021/22 in two different field locations. We drilled all experiments as 1.2 m<sup>2</sup> plots (1 m x 1.2 m) and we sowed them by grain number for comparable plant densities aiming for 275 seeds\*m<sup>2</sup>. We treated all trials with herbicides and fungicides as needed and we applied between 211-218 kg of nitrogen per hectare and 72-75 kg of sulphur per hectare over the growth season. In the 2021/2022 season at Bawburgh (Church Farm) we only

applied 50kg sulphur per hectare. The experiment was also irrigated once, on the 27<sup>th</sup> of April 2022, at a rate of 12mm per hectare, as the season and field were extremely dry. We arranged plots in a randomized complete block design (RCBD) with a split plot arrangement (main plot = Shading treatment, sub plot= Genotype) and three replications.

#### 4.5.2 Shading treatments

We applied shading by covering the main plots (6 genotypes) with 55% Shade Netting made from long-lasting HDPE tape monofilament threads (LBS Horticulture, product ref: NETS001). Nets were cut to length and supported by metal cones and a bamboo frame, installed ca 20 cm above the crop canopy, while avoiding the netting to touch or constrain stem/spike growth. We pulled nets down and secured on the sides (with reusable zip ties) to reduce light entering from the sides. We measured the relative light penetration of the net using a light meter (Skye Instruments Ltd SKP-200; Table 4.1). In 2021 (CF), we applied two shading treatments for 12 days each. Shading A was applied from 16.05.2021 to 28.05.2021, whereas Shading B was applied from 29.05.2021 to 10.06.2021. We chose dates based on estimation of anthesis happening circa 2 weeks after shading applications. In 2022, we applied three shading treatments and overlapped each other by one week. Shading C was installed on 05.05.2022 in CF and 06.05.2022 in Morley. Shading D was installed on 12.05.2022 and 13.05.2022 in CH and Morley, respectively. Shading E was installed on 19.05.2022 in CF and Morley. We removed all shading treatments 13 days after installation, except shading D in Morley which was removed after 12 days.

#### 4.5.3 Phenotyping

To assess floret development during the 2022 season, we cut one main tiller (below the spike, which can be easily assessed by finding the last internode) from the central area of the plot and placed in an emptied box of 1000 µl tips filled with water. We marked the remaining plant to avoid sampling from the same or neighbouring plants again. For each genotype, we took one spike from each block in the control and shading treatment as described in Figure 4.7. For each spike, we dissected the six most basal and two central (most developed) spikelets, and we determined the number of living floret primordia per spikelet and the Waddington stage of each floret, using the Waddington scale and images from Prieto et al. (2018) as reference.

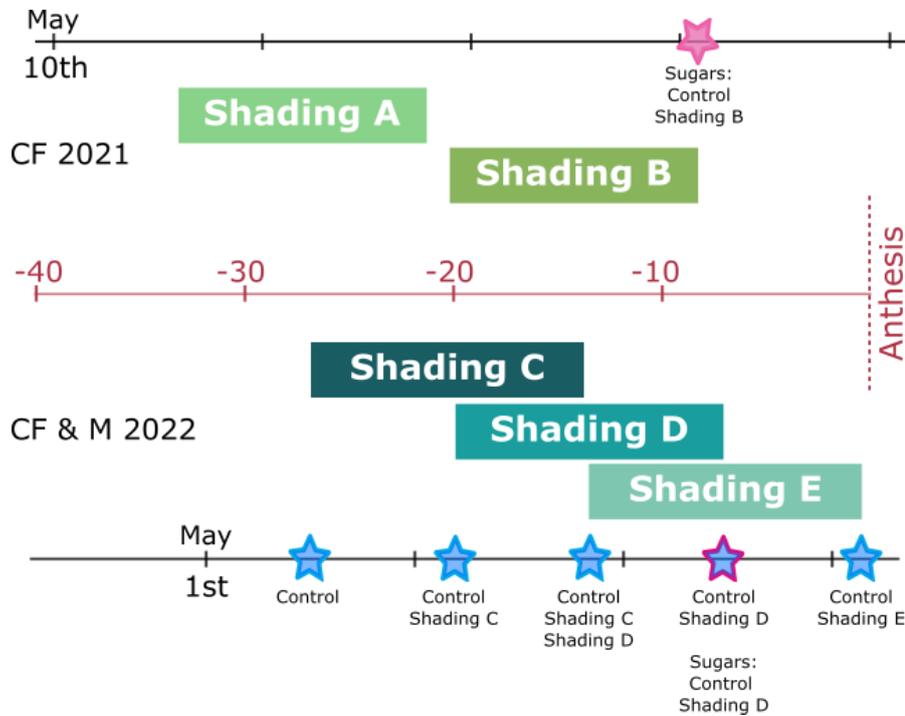


Figure 4.7: Schematic of floret development and sugar sampling in 2021 and 2022. Blue stars indicate floret collection dates, pink star or outline indicate sugar collection dates.

We collected samples at each of the five timepoints depicted in Figure 4.7 and coinciding with the installation/removal of a shading treatment. For the first timepoint (27 days pre anthesis) we collected only spikes from the control plots the same day as Shading C was applied and all subsequent collection timepoints were in one-week intervals. We collected samples each week from the control plots. For Shading C and D, we collected samples one and two weeks after start of the treatment. Samples were only collected after two weeks of Shading E (Figure 4.7).

At the end of the growing season, we hand-harvested mature plants. In Morley we collected grab samples of 10 main spikes from each plot as the ground was too hard to pull plants. For each plot in CF (2021&2022) trials, 10-20 individual plants were pulled from the centre of the 1.2 m<sup>2</sup> plot, which allows for more accurate separation of main and side tillers. Plants were separated, and roots were removed 5 cm above the crown. We assessed plant dry weight before removing all spikes in 2022 but after removing spikes in 2021. We also recorded spikelet number and the number of grains for each spikelet across the main spike. For Morley (2022) we processed spikes in the same manner as in CF trials.

#### 4.5.4 Sugar measurements

Sugar samples were collected by sampling three spikes per plot at the corresponding timepoint. For CF 2021, we sampled at the end of shading B, whereas for CF 2022 and Morley 2022, we collected spikes at the end of shading D from control and shading plots. For each replicate, we dissected spikes into four basal spikelets and rachis, four central spikelet and rachis, apical spikelet and rachis

and internode immediately after dissection and the tissues were snap frozen in liquid nitrogen. Samples were consequently stored at -80°C until further processing.

The method of sugar extraction and measurement is based on Griffiths et al. 2020. Samples were ground using pestle and mortar and ~20 mg of powder (exact weight was recorded) was dissolved in 1.2 mL of ethanol 80% (v/v) in screw-capped tubes. These extracts were mixed thoroughly and incubated for 1 h at 80 °C, mixing halfway through. Extracts were subsequently centrifuged at 12000xg for 1 min and supernatant was collected. We stored samples at -20 °C until assayed.

To perform the assay, we added 5 µL of ethanolic extract to 145 µL of reaction buffer [100 mM HEPES pH 7.4, mM MgCl<sub>2</sub>, 1 mM NAD<sup>+</sup>, 0.5 mM ATP, 1.5 U/µL hexokinase]. All reactions were performed at room temperature in flat-bottomed 96-well microtiter plates and the measurement of each sugar was done by measuring the reduction of NAD<sup>+</sup> to NADH at 340 nm [Varioskan LUX, ThermoFisher]. We first measured baseline absorbance at 340 nm. Subsequently, different enzymes are added sequentially for measurement of either glucose or fructose. The first reaction is initiated by addition of 1.2 U of glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* and incubated for 46 min. Then, absorbance at 340 nm is measured to determine glucose concentration. Subsequently, 0.2 U of phosphoglucosomerase from yeast are added and the reaction is carried out for 46 min, to determine fructose concentration. To correct for variation between runs, a calibration curve for the three sugars was included in each plate and the concentrations were calculated by interpolation. Reactions were performed in triplicates and normalised to the weight of the powder used in extractions.

#### 4.5.5 Data processing and analysis

Using the raw phenotypic data from mature spikes, we calculated the number of rudimentary basal spikelets (RBS), total spikelets, and central spikelet fertility in R. We defined RBS as spikelets carrying no grain and we determined RBS for each spike individually. On average, spikes had ~20-25 spikelets, we therefore calculated the number of grains per central spikelet by averaging the number of grains in the 10-ear samples from the 10<sup>th</sup>, 11<sup>th</sup>, and 12<sup>th</sup> spikelet (from the base). Using the raw floret development scores, we calculated the total number of florets per spikelet by counting the floret Waddington scores per spikelet.

To determine the differences between the genotype and treatments, we performed analysis of variance (ANOVA) on mature plant and floret development phenotypic data. For the analysis of individual trials, we used a split-plot two-way ANOVA including genotype as sub-plot and shading treatment as primary plot performed in R (RStudio 2022.02.0) using the `sp.plot()` function and post-hoc multiple comparisons were performed using the `LSD.test()` ('agricolae' package version 1.3-5; De Mendiburu Delgado (2009)). Floret survival was analysed using R base ANOVA function and post-hoc Sidak test. To analyse differences in sugar concentrations we first performed a preliminary 3-

way ANOVA function in R (RStudio 2022.02.0), which indicated that treatment and tissue had significant effects on glucose and fructose concentrations, however genotype had no significant effect or interactions. We thus excluded genotypic effects from the model, which was also not the focus of this analysis. To analyse the effect of tissue and treatment on glucose and fructose concentrations, we performed split-plot two-way ANOVA including tissue as sub-plot and treatment as primary plot using the `sp.plot()` function and post-hoc multiple comparisons were performed using the `LSD.test()` ('agricolae' package version 1.3-5; De Mendiburu Delgado (2009)). Confidence intervals were calculated using `emmeans` (version 1.8.0). Additional ANOVA and post-hoc tests performed are indicated below each supplemental data.

To predict the survival chances of each floret, we utilised the cumulative distribution function (`pnorm()`, R RStudio 2022.02.0). Initially, we calculated survival probabilities using a fixed mean of 5.5 and standard deviation (SD) of 0.195. We applied this function to florets 1 to 8 for spikes collected 20 days pre-anthesis (i.e., maximum floret stage) in control conditions in CF (all genotypes) and Morley (only NILs). The probabilities of florets 1 to 8 within each spikelet were summed to get the predicted number of grains per spikelet. The idea being that if the survival chance on all florets is very high (nearly 1), the number of grains would equal number of florets per spikelet. We used 5.5 as a mean value for the cumulative distribution function as our hypothesis was that Waddington 5.5 is an important stage, florets beyond this stage thus have a very high survival chance while florets below have very low survival probability. We chose SD = 0.195 as this was the standard deviation of the grains/spikelet dataset from mature spikes.

We used the `optim` function in R, with the choice of optimisers defined by `method="SANN"` (simulated annealing) and `method="Nelder-Mead"` (downhill simplex or Nelder-Mead method), with an objective function (with parameters mean and SD of the cumulative normal distribution) that defines the Euclidean distance between the predicted and measured grain number. The CDF mean and SD were optimised to minimise the difference between predicted and measured grain numbers.

The following supplemental datasets are available via zenodo (<https://doi.org/10.5281/zenodo.7481986>):

**Supplemental Dataset S1:** All raw field data collected in 2021 and 2022 for mature spikes (post-harvest).

**Supplemental Dataset S2:** All developmental scores taken for floret development in the basal six and central two spikelets from spike collected in CF 2022.

**Supplemental Dataset S3:** Raw normalized sugar concentration ( $\mu\text{g}/\text{mg}$  tissue wt) for samples collected in CF 2021 and 2022.

**Supplemental Dataset S4:** Number of grains predicted for each spikelet using CDF, with and without optimisation.

## 5 General Discussion

The overall aim of this thesis was to understand the genetic and developmental causes of rudimentary basal spikelets in wheat. For this, we investigated the differences in transcriptional signalling during initiation (chapter 2), cellular growth deviations thereafter (chapter 3), and final abortion (chapter 4) of basal spikelets compared to their central counterparts. Throughout the chapters, we investigated the following questions:

- What are the transcriptional differences between central and basal spikelets?
- Does *SVP* expression negatively affect basal spikelet establishment?
- Is increased basal spikelet abortion caused by their poor development or resource allocation?

### 5.1 The rise of basal spikelets: Developmental and transcriptional differences in the basal spikelets

#### 5.1.1 Establishment of basal spikelets

Not all spikelets across the wheat spike produce the same amount of grain. The central spikelets produce the most and largest grains, while spikelet size gradually decreases acro- and basipetally. Within a single spike, the most apical and basal spikelets might produce no or only one grain while the central spikelets of the same spike set 3-5 grains. Bonnett (1966) documented that this leads to the distinct lanceolate shape of the wheat spike and is first established during the Glume Primordia stage (just after the initiation of spikelet primordia). This asynchronous development among the spikelets is maintained throughout the development of the spike. While the gradual decrease in spikelet size from the central to apical section of the spike can be explained by the continuous development of new spikelet ridges from the apical inflorescence meristem, basal spikelets are initiated first, and it is unclear why they remain smaller than their central counterparts. In the mature spike the most basal one or two spikelets are often only formed in a rudimentary manner, with small glumes present and all floral structures remaining immature. Improving the homogeneity in grain set across the spike has been proposed as a breeding strategy, however the underlying causes of rudimentary basal spikelet development have not been directly addressed in previous studies (Philipp et al., 2018). Thus, our knowledge of the underlying developmental and transcriptional causes of rudimentary basal spikelets is very limited and has not advanced much since the discoveries made by Bonnett in 1966.

In the second chapter of this thesis, we present the first study directly investigating the transcriptome wide differences that exist across a wheat spike. We found that gene expression

profiles differed most strongly between spatial sections of the same spike, as opposed to temporal sections. We identified 215 DEGs between the two developmental stages, consistent with Li et al. (2018) who identified 206 DEGs between consecutive stages. By contrast, we identified 1,315 and 2,438 unique genes to be differentially expressed between the apical and basal section at DR and GP, respectively. This highlights that spikelets within the same spike experience significantly different flowering and developmental signals due to their consecutive development and spatial position within the spike. Previous studies focus on 'between stage' comparisons, as opposed to within a single stage, thus largely ignoring these spatial differences.

#### 5.1.2 Delayed transition of vegetative to floral signals in base

Through the analysis of gene ontology terms and transcriptional clusters, we noticed that both, *SHORT VEGETATIVE PHASE (SVP)* and *CENTRORADIALIS (CEN)* genes remained highly expressed in the basal section of the spike, whereas their expression was lower in the central and apical sections. We also found that in contrast, E-class flowering genes belonging to the *SEPALLATA (SEP1-4 and SEP1-5)* genes, were expressed in the opposite gradient and showed the strongest expression in apical and central sections of the spike at Glume Primordia stage. This sparked our interest as these genes are important regulators of the vegetative to floral transition. In wheat (Li et al., 2021; Adamski et al., 2021; Liu et al., 2021), rice (Sentoku et al., 2005; Lee et al., 2008), and barley (Trevaskis et al., 2007), *SVPs* have been characterised to be associated with vegetative growth and are downregulated upon floral transition. On the other hand, *SEP* class genes are known to be upregulated at the glume primordia stage and are generally associated with the onset of floral development in tetraploid wheat (Li et al., 2021).

Constitutive over-expression of *SVP*-family members in wheat and barley has been shown to delay or even reverse floral development (Trevaskis et al., 2007; Li et al., 2021). Furthermore, through a series of genetic and biochemical studies, Li et al. (2021) showed that the downregulation of *SVP* genes is necessary for the formation of flowering promoting MADS-box protein complexes including *VRN1*, *FUL2* and *SEP* proteins. Hence the coordinated downregulation of *SVPs*, and possibly *CEN* genes, along with the upregulation of *SEP* genes is required for normal floral transition and spikelet development in wheat. Based on our expression data for *SVPs* and this information, we formed the hypothesis that the delay in development at spikelet onset in the basal spikelets might be caused by the prolonged/increased expression of *SVPs*. This hypothesis was furthermore supported by the finding that increased expression of one of the *SVP* genes (*VRT-A2*), caused by the *VRT-A2b* allele originating from *Triticum turgidum* ssp. *polonicum*, leads to an increased formation of rudimentary basal spikelets (Backhaus et al 2022).

Within the grasses, *SVPs* have mainly been described as suppressors of the floral transition through inhibiting the switch from vegetative to floral development. However, our results suggest that they

might play a more active role in the direct suppression of spikelet ridge development, rather than just timing of floral transition. This difference in *SVP* function would also explain the phenotypes described by Li et al. (2021). They showed that in the double *svp1vrt2* mutant in tetraploid wheat, unlike in the wildtype, spikes or spikelet structures are developed from the axillary meristems in the internode axils below the spike. These axillary meristems are normally suppressed and fail to form any organs. In wildtype wheat these axillary meristems do develop rudimentary and can be observed when the outer leaves are removed but fail to develop beyond a primordial, double ridge like stage. This suggests that *SVPs* might play a key role in the suppression of axillary meristems, after tiller formation, and before spikelet formation. *SVPs* have also been researched extensively in perennial trees. The *SVP* class genes in kiwi, peach, apple, poplar and other trees have been shown to play an important role in bud dormancy and their expression sharply decreases at the beginning of bud release due to favourable conditions (Falavigna et al., 2019). However, in wheat and grasses the suppression of spikelet primordia by *SVPs* has yet to be confirmed directly. To investigate this effect, we developed a protocol for live cell imaging in wheat. We hypothesised that if *SVPs* negatively affect spikelet initiation, we should be able to observe differences in cell size and growth between the wildtype and *VRT-A2b* mutants. However, using available microscopy techniques it is not possible to quantify differences in primordia growth accurately. Using the images acquired from live cell imaging using confocal microscopy, cell growth and division can however be quantified using MorphoGraphX. Having this tool now available in wheat (see Chapter 3) means that we can directly quantify the differences in spikelet ridge growth between the wildtype and *VRT-A2b* mutants. This analysis will be important to establish if indeed the onset of basal spikelets is disrupted in the *VRT-A2b* mutant or if later events during spike development cause the increased number of rudimentary basal spikelets.

### 5.1.3 The wheat spike as a cascade of phytomer

Lastly, our interpretation of results throughout this project was guided by the hypothesis that wheat development should be studied as a sequel of phytomer initiation and suppression events. The axillary meristems in wheat are initially active (e.g., tillers), then suppressed (across internodes), and again activated in the spike (e.g., spikelets). A better understanding of these transitions in axillary meristem activation and suppression will help us to further decipher developmental networks in wheat. Some of the identities of tissues in the spike are still disputed, however we would argue that the spike itself is indeed a cascade of phytomers. Initially, the wheat apical meristem forms fully developed leaves and the accompanying axillary meristems develop into tillers. As the internodes begin to elongate, leaves are still formed but the axillary meristems now appear to be suppressed. Subsequently within the spike, each spikelet ridge is subtended by a bract, which is a suppressed leaf. Each spikelet ridge is an axillary meristem that is not suppressed and has an indeterminate identity, similar to tillers. The indeterminate nature of spikelet primordia

is also evident from the different mutants, that found spikelets to readily revert to tillers in the absence of floral signals (Trevaskis et al., 2007; Li et al., 2021). Within each spikelet, the glume and lemma can also be interpreted as bracts, but the axillary meristems of glume are suppressed/absent and the lemma AMs have true floral identity and are determinate. In this context, the different phenotypic effects of *SVPs* (elongated glume and lemma, release of axillary meristems and increased rudimentary spikelets) could possibly be unified by the concept of *SVPs* favouring the growth of the bract/leaf and suppressing axillary meristems. Future experiments elucidating the transcriptional and developmental identity of the specific spike tissues will be needed to confirm this hypothesis. Also the screening of further mutants for phenotypic effects across all 'leaf' tissues (leaves, bracts, glume, lemma) or all axillary meristems (tillers, internodes, and spikelets) would yield important answers to this question.

## 5.2 New technologies to investigate primordia development in grasses

### 5.2.1 Spatially resolved sequencing technologies for wheat

In the past, spatially resolved sequencing of plant organs has not been a readily available technique in plant genetics. Commonly, whole organs (leaves, inflorescences, florets) are collected and pooled into one sample for RNA extraction and sequencing. This approach is adequate when comparing gene expression patterns between different organs to identify if a gene is, for example, root or shoot specific. However, the sequencing of whole tissues or organs can only reveal little about the function of genes within the organs and is also only of limited use when trying to compare organ development over time. Grass inflorescences are complex structure that contain multiple, different primordia and cell types. Thus far RNA sequencing data has mostly been collected from whole, intact inflorescence meristems (e.g., Feng et al. (2018)).

The composite nature of the maize inflorescence tissues was acknowledged by a study of the maize ears and tassels, where new meristems are initiated in a stepwise manner. Eveland et al. (2014) showed that the range of developmental ages across the maize ear, if acknowledged, can be used as an advantage in RNA-seq studies. They sequenced the tip, middle, and basal sections of 10-mm long ears independently, aiming to analyse the expression patterns in specific developmental meristem types enriched in these sections (inflorescence, spikelet, and floral meristems, respectively). The dissection of the ear therefore allowed them to study gene expression specifically for each meristematic tissue type rather than for all meristem types in intact ears.

We hypothesised that the establishment of the lanceolate shape in wheat spikes could be manifested in gene expression differences between the apical, central and basal sections of a developing spike, as has been shown using RT-qPCR for individual genes in wheat (*AP2*; Debernardi (2017)) and barley (*VRS2*, Youssef et al. (2017)). One of the only studies in wheat that dissected the

developing wheat spike into apical, central and basal sections showed that in tetraploid wheat *APETALA (AP2-5)* and miR172 have consistent and opposing expression gradients across the spike at three consecutive developmental stages (Debernardi et al., 2017). The persistent expression gradient of *AP2-5* supports the idea that expression patterns across the spike, beyond the ones caused by age differences of spikelets, exist that cannot be detected when sequencing whole/intact tissue samples. However, currently available transcriptome data (e.g., Li et al. (2018) and Feng et al. (2017)) lack the spatial resolution *within* each individual developmental stage to answer this question. This focus on 'between stage' comparisons (as opposed to within a single stage) is perhaps related to the technical challenges of dissecting and sectioning young meristems.

In the second chapter of this thesis, we presented a novel, low-input sequencing approach to sequence the transcriptome of the sections. Recently, Giolai et al. (2019) adapted the protocol to identify expression differences across single leaves of *Arabidopsis* (GaST-seq), demonstrating that the G&T-seq method can be readily used for sequencing of hand harvested, small input plant material without the need of previous tissue dissociation or treatment. G&T-seq is thus comparable to methods using laser-micro dissection followed by sequencing to achieve spatially resolved transcriptome wide sequencing data. We used the novel method to sequence the apical, central, and basal sections of individual spikes before and after the establishment of the lanceolate shape. The major benefit of the G&T-seq approach was the ability to obtain sufficient RNA from very low-input tissue samples. G&T-seq is able to do this by employing streptavidin-coupled magnetic beads that are coated in oligoDT primer baits that can be used to purify RNA from the sample with minimal loss. A very similar method had been successfully employed by Meir et al. (2021) in the same year to sequence individual tomato meristems that are transitioning from vegetative to floral development.

Looking ahead, increased resolution of Spatial Transcriptomics, which quantifies full transcriptomes while maintaining tissue integrity, offers a great prospect for the localisation and quantification of gene expression. Using this method, the tissue of interest is sectioned and hybridised to sequencing slides, which allows for the barcoding of all RNA molecules released onto the slide. Thus, the origin of the molecule can be traced back after sequencing. In plants, the currently most advanced method for Spatial Transcriptomics was published in 2017 by Giacomello et al., however their resolution is currently limited to 100  $\mu\text{m}$ . Much higher resolution has been achieved by other spatial transcriptomic methods, such as Slide-Seq (Rodrigues et al., 2019; Stickels et al., 2021) and Stereo-Seq (Liu et al., 2022; Xia et al., 2022). These techniques have primarily been developed for application in mammalian samples and in order to reach their full potential in plant samples, the methods will need to be further adapted to plant tissue. However, their ability to identify gene expression patterns at sub-cellular resolution in relatively large organ sections will make this effort worthwhile. Our results argue strongly for the need of transcriptome-wide and spatially resolved

approaches to advance our biological understanding of fundamental developmental processes in plants.

### 5.2.2 Novel microscopy to analyse cell growth

We hypothesised in the second chapter, that the basal meristems develop into smaller spikelets and larger bract primordia due to a slow suppression of “vegetative growth signals” (e.g., *SVPs*) and a concomitant slow upregulation of “floral growth signals” (e.g., *SEPs*) upon floral transition. We proposed that it is the prolonged expression of *SVPs* in the basal section of the spike that reduced spikelet elongation and promote stronger bract growth in the most basal spikelets. As we also had access to *SVP* mutants with ectopic expression of *VRT2*, we proposed to test this hypothesis by comparing the growth of basal spikelets under normal and increased *VRT2* expression. However, using the classical methods of stereo light microscopy, differences in growth of organs, especially if the organs are very young and small, cannot be visualised. Another major bottleneck to the approach is that the same meristem cannot be imaged multiple times as the dissection leads to its death. We therefore proposed to cultivate spike meristems *ex situ*, and image cell growth using confocal microscopy, a method commonly used for developmental studies in *Arabidopsis*. Analysis of cell growth of living tissue bears the advantage that the growth rate of an organ can be quantified using MorphoGraphX or similar image analysis tools (de Reuille et al., 2015).

In the third chapter, we describe the protocol we developed that allows for *ex situ* cultivation of wheat meristems. The *ex situ* cultivation of the floral wheat meristems enables the optical sectioning of the meristems using confocal microscopy of living tissue. The acquired images can be subsequently used in MorphoGraphX, which we demonstrated in a proof-of-concept heatmap created for one meristem (Chapter 3). One drawback of this method is that the analysis of the images in MorphoGraphX is time consuming. The background noise is relatively high as we are constraint to visualising the cell walls using a fluorescent dye (propidium iodide) as we do not currently have cell wall reporter lines available. Using reporter lines has the advantage that only the tissue of interest has fluorescent signal (propidium iodide also stains dead cells and the agar) and furthermore allows for more imaging timepoints as the dye is toxic to the tissue. In the future, the development of cell wall reporter lines in wheat would thus be highly beneficial. The use of MorphoGraphX in wheat is furthermore not only limited to the spike. The method could be used in the future to image cell growth in other tissues, such as the carpel, stamen, or roots. Each tissue might require a re-optimisation of the media. Preliminary trials with carpels showed that these can be successfully cultivated on the same wMM as spike meristems. Other future applications of this method might also include the perturbation of normal growth on the media. For example, different hormones or growth signals as well as mechanical or chemical stresses can be easily and directly applied to the media. In previous studies, hormones such as gibberellic acid were applied to the whole plant by injecting GA3 into the stem (Boden et al., 2014). However, hormones are known to have many different effects on different tissues and their application to the whole plant instead of

directly to the spike might cause various side-effects. Furthermore, the application of hormones or stress can easily be removed in the *ex situ* growth conditions by either removing the stress or moving the meristems onto a fresh medium plate. We hope that the protocol developed in chapter 3 will find versatile applications and be of use to researchers in wheat developmental studies. Over the last 5-10 years fundamental research has become possible in wheat, we hope that the *ex situ* growth protocol and MorphoGraphX will further aid these developments (<https://www.protocols.io/private/69A55C01351C11ED98330A58A9FEAC02>).

### 5.2.3 Importance of field-based physiological experiments

The smaller, more fine-scaled and in-depth transcriptome and microscopy techniques discussed previously are important new tools to gain better understanding of wheat development. They aid the analysis of gene expression data and cell growth that might cause the reduced development of basal spikelets, among other phenotypes. However, it is equally important to not lose sight of more classical field experiments and their usefulness in wheat research. Firstly, it is immensely important to focus even fundamental research in wheat on topics that have a good chance to be of practical relevance. Fine-mapping, cloning and studying a specific gene in depth is a time and resource consuming task. Within the field of wheat research, we would thus suggest that these efforts are best spend on genes that are related to yield, stress resilience or disease resistance. For example, the mapping and cloning of genes such as *GW2*, *GNI1*, *VRT2* or *Rht13* resulted in practical applications as well as the discovery of novel biological mechanisms (Simmonds et al., 2016; Sakuma et al., 2019; Adamski et al., 2021; Borrill et al., 2022). Once a gene is cloned, perfect functional markers can be developed that aid the accurate and fast introgression of these genes into new cultivars. It also allows breeders to genotypically screen their material for the beneficial allele of the gene and see if they already have the gene in their breeding material. But the discovery of the underlying gene and mutation in all cases also broadened our understanding of the crop or its genome. For example, only recently the causal gene for the height reducing *Rht13* gene was found to be surprisingly a classical disease resistance gene (Borrill et al., 2022). *VRT2* was found to be the responsible gene for the long glume phenotype in *polonicum* accessions, however the causal deletion is not in the exon but in the first intron, leading to subtle ectopic expression of the gene (Adamski et al., 2021). To study these new genes or alleles of interest, it is important to employ and efficiently use field trials. Wheat cannot be only grown in the laboratory or glasshouse and the full complexity and effect of genetic factors, and the environment, must be at least at some point during the experimental journey evaluated in the field. Field research was the backbone of all crop research until only ~60 years ago. Future research must also harness the benefits from field trials.

Furthermore, the use of field experiments can also add insight into crop physiology otherwise difficult to achieve. One strength of field experiments is that relatively high number of genotypes,

conditions and replications can be tested at a relatively low cost. In this thesis, we made use of field-based experiments to further investigate when basal spikelets are determined to be rudimentary. As described above, basal spikelets are delayed in development from initiation onwards. However, the spikelets do continue to develop until much later in the crop cycle when they are eventually aborted. From our earlier experiments we could thus not determine if the setback during initiation causes rudimentary basal spikelets, or if it is only the later abortion that determines their number. To address this concern, we evaluated the floret development of six different genotypes in the field. We grew the genotypes under normal conditions but also applied targeted shading (to reduce resource availability) for ca. 12 days in three different timeframes. We were thus able to determine during which time the basal spikelets are most sensitive to resource limitations and were able to narrow down this time to ca. 10 days pre-anthesis. Furthermore, we were able to show that most likely it is not lower resource availability in the basal spikelets but rather their delayed development that causes the increased abortion of basal compared to central spikelets. Thus, this very different and completely independent study in the field lends support to our previous hypothesis of early spikelet establishment being the cause of rudimentary basal spikelets.

### 5.3 The fall of basal spikelets: Increased floret abortion in basal spikelets

To determine the later fate of the initially less developed basal spikelets we traced their development from initiation (glume primordium stage) until flowering (chapter 4). Initial trials in the glasshouse in 2018 suggested that basal spikelets lag behind in development by ca. 1 Waddington stage compared to central spikelets. We observed that this lag is very consistent throughout development and the basal spikelets do not 'catch up' in development (data not shown). However, it did appear that more florets are aborted in the basal spikelets than in central, raising the question what causes the increased abortion in the base.

To formally investigate the later development and abortion of spikelets we designed a field experiment in which we used shading netting to reduce resource availability in short, and very precise timeframes throughout development. We found that the number of rudimentary basal spikelets only significantly increased when shading was applied ca. 10 days pre-anthesis. This suggests that during this time basal spikelets, or at least the florets within, are actively aborted and that this abortion is sensitive to resource availability. We could furthermore show that this coincides with the time of floret abortion in central spikelets, suggesting that complete floret abortion is the underlying cause of basal spikelet abortion. Previous studies suggested that abortion of florets in the basal spikelets is more severe due to worse allocation of resources (in form of sugars) to basal spikelets. However, neither our data nor the re-interpretation of data by Stockman et al. (1983) support this hypothesis as we did not find any differences in sugar concentrations across the spike.

Alternatively, Ferrante et al. (2020) suggested that the probability of a floret to survive abortion depends on its developmental age. They developed this hypothesis based on the observation that reduced nitrogen fertilizer leads to delayed floret development and thus increase in floret abortion. Similarly, they also showed that floret development pre-abortion is more advanced in the modern variety with increased grains per spikelet. Analysis of the developmental progress of florets in our experiment also suggested that their age (i.e. Waddington developmental stage) at abortion is an important factor for their survival probability. Florets below a given threshold age had a much lower chance of survival, regardless of their position within the spikelet or spike. As the youngest florets are also the most distal within one spikelet it can be difficult to decipher these two factors. However, because we also analysed the basal spikelets, we could show that even the most proximal florets in the first or second position are likely to be aborted due to their delayed development compared to the equivalent florets in central spikelets. Lastly, we also found that before abortion the overall development and number of florets is marginally lower in the *VRT-A2b* NILs (that have increased rudimentary basal spikelets) compared to the wildtype *VRT-A2a*. These field results lend further support to our previous hypothesis, that it is the poorer establishment of basal spikelets rather than increased abortion in the base (compared to central section) that causes increased rudimentary basal spikelet formation in *VRT-A2b* NILs. Albeit this hypothesis awaits final confirmation through MorphoGraphX analysis of basal spikelet growth during initiation. Lastly, this suggests that abortion of florets across the spike follows the same rules and that if one would want to increase basal spikelet productivity in relation to central spikelets, it is their initiation that needs to be targeted.

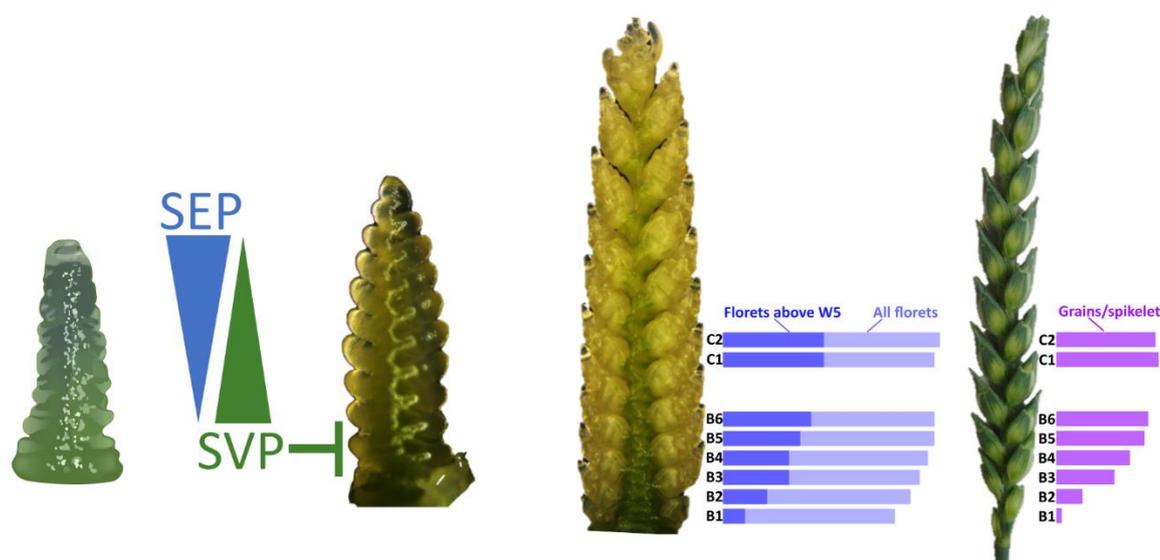


Figure 5.1: Summarising figure of wheat spike meristem development from initiation to mature spike, visualising the major findings of this thesis. Gradients in expression of genes across the spike due to a slow transition from vegetative to floral networks cause basal spikelet to lag in development. At abortion, this delay in development leads to increased abortion of florets in basal spikelets, causing the low number of grains in the mature spike. Meristems and spikes are not shown at scale.

## 5.4 Concluding Statement

Overall, this thesis has been the first investigation specifically focused on understanding the causes of rudimentary basal spikelet formation in wheat. Our results show that basal spikelet development differs from central spikelets since their initiation. Bonnett (1966) already described that basal spikelets are smaller and underdeveloped compared to central spikelets immediately after their initiation. We found that this delay in initiation is most likely due to them being formed while the meristem is transitioning from a vegetative to floral signalling network. This transition is not immediate and thus the first formed (i.e., basal) spikelets receive mixed signals. Using novel transcriptomic approaches, we were able to identify genetic factors directly involved in this transition of signalling networks and showed that *VRT2* expression negatively affects basal spikelets, as predicted by our model developed in NetLogo. To gain a complete understanding of basal spikelet development, we also investigated their abortion and found that is most likely the complete abortion of the underdeveloped florets in basal spikelets that finally causes their rudimentary shape. This hypothesis awaits further confirmation by our analysis of spikelet ridge growth using MorphoGraphX.

Rudimentary basal spikelets are a rarely studied phenotype in wheat even though it is directly linked to spike yield and overall plant productivity. Deciphering their unique developmental pathway has helped us to understand their initiation and abortion. However, studying them has also allowed us to develop novel sequencing and imaging techniques in wheat and, more generally, has allowed us to develop a better understanding of wheat spike development.

## 6 References

- Abbate PE, Andrade FH, Culot JP** (1995) The effects of radiation and nitrogen on number of grains in wheat. *The Journal of Agricultural Science* **124**: 351-360
- Adamski NM, Borrill P, Brinton J, Harrington SA, Marchal C, Bentley AR, Bovill WD, Cattivelli L, Cockram J, Contreras-Moreira B** (2020) A roadmap for gene functional characterisation in crops with large genomes: lessons from polyploid wheat. *Elife* **9**: e55646
- Adamski NM, Simmonds J, Brinton JF, Backhaus AE, Chen Y, Smedley M, Hayta S, Florio T, Crane P, Scott P, Pieri A, Hall O, Barclay JE, Clayton M, Doonan JH, Nibau C, Uauy C** (2021) Ectopic expression of *Triticum polonicum* *VRT-A2* underlies elongated glumes and grains in hexaploid wheat in a dosage-dependent manner. *The Plant Cell* **33(7)**: 2296-2319
- Alon U** (2007) An introduction to systems biology: design principles of biological circuits. Chapman & Hall/CRC, London
- Alqudah AM, Schnurbusch T** (2013) Awn primordium to tipping is the most decisive developmental phase for spikelet survival in barley. *Functional Plant Biology* **41**: 424-436
- Altenbach SB, Kothari KM** (2004) Transcript profiles of genes expressed in endosperm tissue are altered by high temperature during wheat grain development. *Journal of Cereal Science* **40**: 115-126
- Anders S, Pyl PT, Huber W** (2015) HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* **31**
- Arisnabarreta S, Miralles DJ** (2006) Floret development and grain setting in near isogenic two- and six-rowed barley lines (*Hordeum vulgare* L.). *Field Crops Research* **96**: 466-476
- Backhaus AE, Lister A, Tomkins M, Adamski NM, Simmonds J, Macaulay I, Morris RJ, Haerty W, Uauy C** (2022) High expression of the MADS-box gene *VRT2* increases the number of rudimentary basal spikelets in wheat. *Plant Physiology* **189**: 1536-1552
- Barbier de Reuille P, Routier-Kierzkowska A-L, Kierzkowski D, Bassel GW, Schüpbach T, Tauriello G, Bajpai N, Strauss S, Weber A, Kiss A, Burian A, Hofhuis H, Sapala A, Lipowczan M, Heimlicher MB, Robinson S, Bayer EM, Basler K, Koumoutsakos P, Roeder AHK, Aegerter-Wilmsen T, Nakayama N, Tsiantis M, Hay A, Kwiatkowska D, Xenarios I, Kuhlemeier C, Smith RS** (2015) MorphoGraphX: A platform for quantifying morphogenesis in 4D. *eLife* **4**: e05864
- Barkworth ME, Bothmer Rv** (2009) Scientific names in the Triticeae. *In* *Genetics and Genomics of the Triticeae*. Springer, pp 3-30
- Becker A, Theißen G** (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution* **29**: 464-489
- Benjamini Y, Hochberg Y** (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc* **57**
- Bentley A** (2022) Broken bread—avert global wheat crisis caused by invasion of Ukraine. *Nature* **603**: 551-551
- Bi X, van Esse W, Mulki MA, Kirschner G, Zhong J, Simon R, von Korff M** (2019) *CENTRORADIALIS* interacts with *FLOWERING LOCUS T*-like genes to control floret development and grain number. *Plant physiology* **180**: 1013-1030
- Boden SA, Weiss D, Ross JJ, Davies NW, Trevaskis B, Chandler PM, Swain SM** (2014) *EARLY FLOWERING3* regulates flowering in spring barley by mediating gibberellin production and *FLOWERING LOCUS T* expression. *The Plant Cell* **26**: 1557-1569
- Bommert P, Satoh-Nagasawa N, Jackson D, Hirano H-Y** (2005) Genetics and evolution of inflorescence and flower development in grasses. *Plant and Cell Physiology* **46**: 69-78
- Bommert P, Whipple C** (2018) Grass inflorescence architecture and meristem determinacy. *Seminars in Cell & Developmental Biology* **79**: 37-47
- Bonnett OT** (1966) Inflorescences of maize, wheat, rye, barley, and oats: their initiation and development. *Bulletin (University of Illinois (Urbana-Champaign campus) Agricultural Experiment Station)* **721**
- Borrill P, Mago R, Xu T, Ford B, Williams SJ, Derkx A, Bovill WD, Hyles J, Bhatt D, Xia X** (2022) An autoactive NB-LRR gene causes Rht13 dwarfism in wheat. *bioRxiv*

- Borrill P, Ramirez-Gonzalez R, Uauy C** (2016) expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform *Plant Physiology* **170**: 2172-2186
- Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S** (2006) *ramosa2* Encodes a LATERAL ORGAN BOUNDARY Domain Protein That Determines the Fate of Stem Cells in Branch Meristems of Maize. *The Plant Cell* **18**: 574-585
- Boussora F, Allam M, Guasmi F, Ferchichi A, Rutten T, Hansson M, Youssef HM, Börner A** (2019) Spike developmental stages and ABA role in spikelet primordia abortion contribute to the final yield in barley (*Hordeum vulgare L.*). *Botanical Studies* **60**: 13
- Chang W, Guo Y, Zhang H, Liu X, Guo L** (2020) Same actor in different stages: genes in shoot apical meristem maintenance and floral meristem determinacy in *Arabidopsis*. *Frontiers in Ecology and Evolution* **8**: 89
- Ciaffi M, Paolacci AR, Tanzarella OA, Porceddu E** (2011) Molecular aspects of flower development in grasses. *Sexual Plant Reproduction* **24**: 247-282
- Clark LG, Judziewicz EJ** (1996) The grass subfamilies *Anomochlooideae* and *Pharoideae* (*Poaceae*). *Taxon*: 641-645
- Coen ES, Meyerowitz EM** (1991) The war of the whorls: genetic interactions controlling flower development. *Nature* **353**: 31-37
- Collins TJ** (2007) ImageJ for microscopy. *Biotechniques* **43**: S25-S30
- Connor HE** (1979) Breeding systems in the grasses: a survey. *New Zealand Journal of Botany* **17**: 547-574
- Craufurd P, Cartwright P** (1989) Effect of photoperiod and chlormequat on apical development and growth in a spring wheat (*Triticum aestivum*) cultivar. *Annals of Botany* **63**: 515-525
- Crouvisier-Urien K, Chanut J, Lagorce A, Winckler P, Wang Z, Verboven P, Nicolai B, Lherminier J, Ferret E, Gougeon RD, Bellat J-P, Karbowski T** (2019) Four hundred years of cork imaging: New advances in the characterization of the cork structure. *Scientific Reports* **9**: 19682
- De Mendiburu Delgado F** (2009) Una herramienta de análisis estadístico para la investigación agrícola.
- de Reuille PB, Routier-Kierzkowska A-L, Kierzkowski D, Bassel GW, Schüpbach T, Tauriello G, Bajpai N, Strauss S, Weber A, Kiss A** (2015) MorphoGraphX: A platform for quantifying morphogenesis in 4D. *Elife* **4**: e05864
- Debernardi JM, Greenwood JR, Jean Finnegan E, Jernstedt J, Dubcovsky J** (2020) *APETALA 2*-like genes *AP2L2* and *Q* specify lemma identity and axillary floral meristem development in wheat. *The Plant Journal* **101**: 171-187
- Debernardi JM, Lin H, Chuck G, Faris JD, Dubcovsky J** (2017) microRNA172 plays a crucial role in wheat spike morphogenesis and grain threshability. *Development* **144**: 1966-1975
- Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, Mellers G, Ramm K, Cavanagh C, Swain SM, Boden SA** (2018) *TEOSINTE BRANCHED1* Regulates inflorescence architecture and development in bread wheat (*Triticum aestivum*). *The Plant cell* **30**: 563-581
- Dobrovolskaya O, Pont C, Sibout R, Martinek P, Badaeva E, Murat F, Chosson A, Watanabe N, Prat E, Gautier N** (2015) *FRIZZY PANICLE* drives supernumerary spikelets in bread wheat. *Plant Physiology* **167**: 189-199
- Doebley JF, Gaut BS, Smith BD** (2006) The Molecular Genetics of Crop Domestication. *Cell* **127**: 1309-1321
- Doyle JJ, Gaut BS** (2000) *Plant molecular evolution*, Vol 42. Kluwer Academic Publishers
- Dreni L, Osnato M, Kater MM** (2013) The ins and outs of the rice *AGAMOUS* subfamily. *Molecular Plant* **6**: 650-664
- Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA, Liotta LA** (1996) Laser capture microdissection. *Science* **274**: 998-1001
- Evans L, Blundell C, King R** (1995) Developmental responses by tall and dwarf isogenic lines of spring wheat to applied gibberellins. *Functional Plant Biology* **22**: 365-371
- Eveland AL, Goldshmidt A, Pautler M, Morohashi K, Liseron-Monfils C, Lewis MW, Kumari S, Hiraga S, Yang F, Unger-Wallace E** (2014) Regulatory modules controlling maize inflorescence architecture. *Genome research* **24**: 431-443
- Falavigna VdS, Guitton B, Costes E, Andrés F** (2019) I want to (bud) break free: the potential role of *DAM* and *SVP*-like genes in regulating dormancy cycle in temperate fruit trees. *Frontiers in plant science* **9**: 1990

- FAO W, IFAD, UNICEF, WFP** (2018) The state of food security and nutrition in the world 2018: building climate resilience for food security and nutrition, Rome
- Feng N, Song G, Guan J, Chen K, Jia M, Huang D, Wu J, Zhang L, Kong X, Geng S, Liu J, Li A, Mao L** (2017) Transcriptome Profiling of Wheat Inflorescence Development from Spikelet Initiation to Floral Patterning Identified Stage-Specific Regulatory Genes. *Plant Physiology* **174**: 1779-1794
- Ferrándiz C, Gu Q, Martienssen R, Yanofsky MF** (2000) Redundant regulation of meristem identity and plant architecture by *FRUITFULL*, *APETALA1* and *CAULIFLOWER*. *Development* **127**: 725-734
- Ferrante A, Savin R, Slafer GA** (2010) Floret development of durum wheat in response to nitrogen availability. *Journal of Experimental Botany* **61**: 4351-4359
- Ferrante A, Savin R, Slafer GA** (2013) Is floret primordia death triggered by floret development in durum wheat? *Journal of experimental botany* **64**: 2859-2869
- Ferrante A, Savin R, Slafer GA** (2020) Floret development and spike fertility in wheat: differences between cultivars of contrasting yield potential and their sensitivity to photoperiod and soil N. *Field Crops Research* **256**: 107908
- Fischer R** (1985) Number of kernels in wheat crops and the influence of solar radiation and temperature. *The Journal of Agricultural Science* **105**: 447-461
- Fischer R, Stockman Y** (1980) Kernel number per spike in wheat (*Triticum aestivum* L.): responses to preanthesis shading. *Functional Plant Biology* **7**: 169-180
- Fischer TRA** (2022) History of Wheat Breeding: A Personal View. In MP Reynolds, H-J Braun, eds, *Wheat Improvement: Food Security in a Changing Climate*. Springer International Publishing, pp 17-30
- Giacomello S, Salmén F, Terebieniec BK, Vickovic S, Navarro JF, Alexeyenko A, Reimegård J, McKee LS, Mannapperuma C, Bulone V** (2017) Spatially resolved transcriptome profiling in model plant species. *Nature plants* **3**: 17061
- Giolai M, Verweij W, Lister A, Heavens D, Macaulay I, Clark MD** (2019) Spatially resolved transcriptomics reveals plant host responses to pathogens. *Plant Methods* **15**: 114
- Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C** (2010) Food Security: The Challenge of Feeding 9 Billion People. *Science* **327**: 812-818
- González FG, Miralles DJ, Slafer GA** (2011) Wheat floret survival as related to pre-anthesis spike growth. *Journal of Experimental Botany* **62**: 4889-4901
- González FG, Slafer GA, Miralles DJ** (2003) Floret development and spike growth as affected by photoperiod during stem elongation in wheat. *Field Crops Research* **81**: 29-38
- González FG, Slafer GA, Miralles DJ** (2005) Photoperiod during stem elongation in wheat: is its impact on fertile floret and grain number determination similar to that of radiation? *Functional Plant Biology* **32**: 181-188
- Grass Phylogeny Working Group G, Barker NP, Clark LG, Davis JI, Duvall MR, Guala GF, Hsiao C, Kellogg EA, Linder HP, Mason-Gamer RJ** (2001) Phylogeny and subfamilial classification of the grasses (*Poaceae*). *Annals of the Missouri Botanical Garden*: 373-457
- Guo J, Zhang G, Song Y, Ma S, Niu N, Wang J** (2019) Comparative transcriptome profiling of multi-ovary wheat under heterogeneous cytoplasm suppression. *Scientific Reports* **9**: 8301
- Guo J, Zhang Y, Shi W, Zhang B, Zhang J, Xu Y, Cheng X, Cheng K, Zhang X, Hao C** (2015) Association analysis of grain-setting rates in apical and basal spikelets in bread wheat (*Triticum aestivum* L.). *Frontiers in Plant Science* **6**: 1029
- Guo Z, Chen D, Alqudah AM, Röder MS, Ganai MW, Schnurbusch T** (2017) Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. *New Phytologist* **214**: 257-270
- Guo Z, Slafer GA, Schnurbusch T** (2016) Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat. *Journal of Experimental Botany* **67**: 4221-4230
- Hanif M, Langer RHM** (1972) The vascular system of the spikelet in wheat (*Triticum aestivum*). *Annals of Botany* **36**: 721-727

- Harrop TW, Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, Kater MM** (2016) Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *The Plant Journal* **86**: 75-88
- Hayta S, Smedley MA, Clarke M, Forner M, Harwood WA** (2021) An efficient Agrobacterium-mediated transformation protocol for hexaploid and tetraploid wheat. *Current Protocols* **1**: e58
- Hayta S, Smedley MA, Demir SU, Blundell R, Hinchliffe A, Atkinson N, Harwood WA** (2019) An efficient and reproducible Agrobacterium-mediated transformation method for hexaploid wheat (*Triticum aestivum* L.). *Plant Methods* **15**: 1-15
- Houston K, Druka A, Bonar N, Macaulay M, Lundqvist U, Franckowiak J, Morgante M, Stein N, Waugh R** (2012) Analysis of the barley bract suppression gene *TRD1*. *Theoretical and Applied Genetics* **125**: 33-45
- IWGSC, Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, Pozniak CJ, Choulet F, et al.** (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**
- Jaeger KE, Wigge PA** (2007) FT protein acts as a long-range signal in *Arabidopsis*. *Current biology* **17**: 1050-1054
- Judziewicz EJ, Soderstrom TR** (1989) Morphological, anatomical, and taxonomic studies in *Anomochloa* and *Streptochoeta* (*Poaceae: Bambusoideae*). *Smithsonian Contributions to Botany*
- Kadkol G, Halloran G** (1988) Development and floret survival in wheat ears with supernumerary spikelets. *Annals of Botany* **62**: 559-568
- Kellogg E, Camara P, Rudall P, Ladd P, Malcomber S, Whipple C, Doust A** (2013) Early inflorescence development in the grasses (*Poaceae*). *Frontiers in Plant Science* **4**
- Kellogg EA** (2001) Evolutionary history of the grasses. *Plant physiology* **125**: 1198-1205
- Kellogg EA** (2022) Genetic control of branching patterns in grass inflorescences. *The Plant Cell* **34**: 2518-2533
- Kemp D, Whingwiri E** (1980) Effect of tiller removal and shading on spikelet development and yield components of the main shoot of wheat and on the sugar concentration of the ear and flag leaf. *Functional Plant Biology* **7**: 501-510
- Khush GS** (2001) Green revolution: the way forward. *Nature reviews genetics* **2**: 815-822
- Kierzkowski D, Nakayama N, Routier-Kierzkowska A-L, Weber A, Bayer E, Schorderet M, Reinhardt D, Kuhlemeier C, Smith RS** (2012) Elastic Domains Regulate Growth and Organogenesis in the Plant Shoot Apical Meristem. *Science* **335**: 1096-1099
- Kierzkowski D, Runions A, Vuolo F, Strauss S, Lymbouridou R, Routier-Kierzkowska A-L, Wilson-Sánchez D, Jenke H, Galinha C, Mosca G** (2019) A growth-based framework for leaf shape development and diversity. *Cell* **177**: 1405-1418. e1417
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL** (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**: 907-915
- Kinoshita A, Vayssières A, Richter R, Sang Q, Roggen A, van Driel AD, Smith RS, Coupland G** (2020) Regulation of shoot meristem shape by photoperiodic signaling and phytohormones during floral induction of *Arabidopsis*. *Elife* **9**: e60661
- Kirby E, Appleyard M** (1987) Development and structure of the wheat plant. *In* *Wheat breeding*. Springer, pp 287-311
- Kirby EJM, Appleyard M** (1981) *Cereal development guide*. Cereal Unit, National Agricultural Centre., Stoneleigh, Warwickshire
- Koppolu R, Schnurbusch T** (2019) Developmental pathways for shaping spike inflorescence architecture in barley and wheat. *Journal of Integrative Plant Biology* **61**: 278-295
- Kuzay S, Lin H, Li C, Chen S, Woods DP, Zhang J, Lan T, von Korff M, Dubcovsky J** (2022) *WAP0-A1* is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PLoS genetics* **18**: e1009747
- Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-Guedira G, DeWitt N, Peters Haugrud A, Faris JD, Akhunov E, Bai G, Dubcovsky J** (2019) Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics* **132**: 2689-2705

- Lang A** (1957) The effect of gibberellin upon flower formation. *Proceedings of the National Academy of Sciences* **43**: 709-717
- Langer RHM, Hanif M** (1973) A study of floret development in wheat (*Triticum aestivum* L.). *Annals of Botany* **37**: 743-751
- Laurie DA** (2009) Developmental and Reproductive Traits in the *Triticeae*. In GJ Muehlbauer, C Feuillet, eds, *Genetics and Genomics of the Triticeae*. Springer US, New York, NY, pp 591-609
- Lazo GR, Stein PA, Ludwig RA** (1991) A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Bio/technology* **9**: 963-967
- Li C, Lin H, Chen A, Lau M, Jernstedt J, Dubcovsky J** (2019) Wheat *VRN1*, *FUL2* and *FUL3* play critical and redundant roles in spikelet development and spike determinacy. *Development* **146**
- Li K, Debernardi JM, Li C, Lin H, Zhang C, Jernstedt J, Korff Mv, Zhong J, Dubcovsky J** (2021) Interactions between SQUAMOSA and SHORT VEGETATIVE PHASE MADS-box proteins regulate meristem transitions during wheat spike development. *The Plant Cell* **33**: 3621-3644
- Li Y, Fu X, Zhao M, Zhang W, Li B, An D, Li J, Zhang A, Liu R, Liu X** (2018) A Genome-wide View of Transcriptome Dynamics During Early Spike Development in Bread Wheat. *Scientific Reports* **8**: 15338
- Liu J, Chen Z, Wang Z, Zhang Z, Xie X, Wang Z, Chai L, Song L, Cheng X, Feng M, Wang X, Liu Y, Hu Z, Xing J, Su Z, Peng H, Xin M, Yao Y, Guo W, Sun Q, Liu J, Ni Z** (2021) Ectopic expression of *VRT-A2* underlies the origin of *Triticum polonicum* and *Triticum petropavlovskyi* with long outer glumes and grains. *Molecular Plant*
- Liu Y, Li C, Han Y, Li R, Cui F, Zhang H, Su X, Liu X, Xu G, Wan S** (2022) Spatial transcriptome analysis on peanut tissues shed light on cell heterogeneity of the peg. *Plant biotechnology journal* **20**: 1648-1650
- Livak KJ, Schmittgen TD** (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup>. *Methods* **25**
- Lombardo F, Yoshida H** (2015) Interpreting lemma and palea homologies: a point of view from rice floral mutants. *Frontiers in Plant Science* **6**: 61
- Love A** (1984) *Conspectus of the Triticeae*. *Fedd. Report*. **95**: 425-521
- Love MI, Huber W, Anders S** (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550
- Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, Goolam M, Saurat N, Coupland P, Shirley LM, Smith M, Van der Aa N, Banerjee R, Ellis PD, Quail MA, Swerdlow HP, Zernicka-Goetz M, Livesey FJ, Ponting CP, Voet T** (2015) G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* **12**: 519-522
- Malcomber ST, Preston JC, Reinheimer R, Kossuth J, Kellogg EA** (2006) Developmental Gene Evolution and the Origin of Grass Inflorescence Diversity. In *Advances in Botanical Research*, Vol 44. Academic Press, pp 425-481
- McFaline-Figueroa JL, Trapnell C, Cuperus JT** (2020) The promise of single-cell genomics in plants. *Current opinion in plant biology* **54**: 114-121
- McMullan D** (1995) Scanning electron microscopy 1928-1965. *Scanning* **17**: 175-185
- Meir Z, Aviezer I, Chongloi GL, Ben-Kiki O, Bronstein R, Mukamel Z, Keren-Shaul H, Jaitin D, Tal L, Shalev-Schlosser G, Harel TH, Tanay A, Eshed Y** (2021) Dissection of floral transition by single-meristem transcriptomes at high temporal resolution. *Nature Plants* **7**: 800-813
- Min Y, Conway SJ, Kramer EM** (2022) Quantitative live imaging of floral organ initiation and floral meristem termination in *Aquilegia*. *Development* **149**: dev200256
- Moreno N, Bougourd S, Haseloff J, Feijó JA** (2006) Imaging plant cells. In *Handbook of biological confocal Microscopy*. Springer, pp 769-787
- Muqaddasi QH, Brassac J, Koppolu R, Plieske J, Ganai MW, Röder MS** (2019) *TaAPO-A1*, an ortholog of rice *ABERRANT PANICLE ORGANIZATION 1*, is associated with total spikelet number per spike in elite European hexaploid winter wheat (*Triticum aestivum* L.) varieties. *Scientific reports* **9**: 1-12

- Murai K, Miyamae M, Kato H, Takumi S, Ogihara Y** (2003) *WAP1*, a wheat *APETALA1* homolog, plays a central role in the phase transition from vegetative to reproductive growth. *Plant and Cell Physiology* **44**: 1255-1265
- O'Connor DL, Runions A, Sluis A, Bragg J, Vogel JP, Prusinkiewicz P, Hake S** (2014) A division in PIN-mediated auxin patterning during organ initiation in grasses. *PLoS Computational Biology* **10**: e1003447
- Pautler M, Tanaka W, Hirano H-Y, Jackson D** (2013) Grass Meristems I: Shoot Apical Meristem Maintenance, Axillary Meristem Determinacy and the Floral Transition. *Plant and Cell Physiology* **54**: 302-312
- Pearce S, Vanzetti LS, Dubcovsky J** (2013) Exogenous Gibberellins Induce Wheat Spike Development under Short Days Only in the Presence of *VERNALIZATION1*. *Plant Physiology* **163**: 1433-1445
- Philipp N, Weichert H, Bohra U, Weschke W, Schulthess AW, Weber H** (2018) Grain number and grain yield distribution along the spike remain stable despite breeding for high yield in winter wheat. *PloS one* **13**: e0205452-e0205452
- Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawroński P, Bini F, Govind G, Rutten T, Sakuma S, Tagiri A** (2015) The genetic basis of composite spike form in barley and 'Miracle-Wheat'. *Genetics* **201**: 155-165
- Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawroński P, Bini F, Govind G, Rutten T, Sakuma S, Tagiri A, Wolde GM, Youssef HM, Battal A, Ciannamea S, Fusca T, Nussbaumer T, Pozzi C, Börner A, Lundqvist U, Komatsuda T, Salvi S, Tuberosa R, Uauy C, Sreenivasulu N, Rossini L, Schnurbusch T** (2015) The Genetic Basis of Composite Spike Form in Barley and 'Miracle-Wheat'. *Genetics* **201**: 155-165
- Prieto P, Ochagavía H, Savin R, Griffiths S, Slafer GA** (2018) Dynamics of floret initiation/death determining spike fertility in wheat as affected by *Ppd* genes under field conditions. *Journal of Experimental Botany* **69**: 2633-2645
- Ramirez-Gonzalez RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, van Ex F, Pasha A, Khedikar Y, Robinson SJ, Cory AT, Florio T, Concia L, Juery C, Schoonbeek H, Steuernagel B, Xiang D, Ridout CJ, Chalhoub B, Mayer KFX, Benhamed M, Latrasse D, Bendahmane A, Wulff BBH, Appels R, Tiwari V, Datta R, Choulet F, Pozniak CJ, Provart NJ, Sharpe AG, Paux E, Spannagl M, Brautigam A, Uauy C, Int Wheat Genome S** (2018) The transcriptional landscape of polyploid wheat. *Science* **361**: 662-+
- Ray DK, Ramankutty N, Mueller ND, West PC, Foley JA** (2012) Recent patterns of crop yield growth and stagnation. *Nature communications* **3**: 1-7
- Razumov V, Lmsart R, Tan K** (1960) The effect of gibberellic acid on the vernalization of plants. *Bot. Zh* **45**: 1732-1738
- Reinhardt D, Mandel T, Kuhlemeier C** (2000) Auxin Regulates the Initiation and Radial Position of Plant Lateral Organs. *The Plant Cell* **12**: 507-518
- Reinhardt D, Pesce E-R, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C** (2003) Regulation of phyllotaxis by polar auxin transport. *Nature* **426**: 255-260
- Ren D, Rao Y, Leng Y, Li Z, Xu Q, Wu L, Qiu Z, Xue D, Zeng D, Hu J** (2016) Regulatory role of OsMADS34 in the determination of glumes fate, grain yield, and quality in rice. *Frontiers in plant science* **7**: 1853
- Reynolds M, Foulkes MJ, Slafer GA, Berry P, Parry MA, Snape JW, Angus WJ** (2009) Raising yield potential in wheat. *Journal of experimental botany* **60**: 1899-1918
- Reynolds MP, Braun H-J** (2022) *Wheat improvement: food security in a changing climate*. Springer Nature
- Reynolds MP, Pellegrineschi A, Skovmand B** (2005) Sink-limitation to yield and biomass: a summary of some investigations in spring wheat. *Annals of Applied Biology* **146**: 39-49
- Rich-Griffin C, Stechemesser A, Finch J, Lucas E, Ott S, Schäfer P** (2020) Single-cell transcriptomics: a high-resolution avenue for plant functional genomics. *Trends in plant science* **25**: 186-197
- Rodrigues SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, Welch J, Chen LM, Chen F, Macosko EZ** (2019) Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**: 1463-1467

- Rudall PJ, Stuppy W, Cunniff J, Kellogg EA, Briggs BG** (2005) Evolution of reproductive structures in grasses (*Poaceae*) inferred by sister-group comparison with their putative closest living relatives, *Ecdeiocoleaceae*. *American Journal of Botany* **92**: 1432-1443
- Running MP, Clark SE, Meyerowitz EM** (1995) Confocal microscopy of the shoot apex. *In* *Methods in cell biology*, Vol 49. Elsevier, pp 217-229
- Sadras VO, Slafer GA** (2012) Environmental modulation of yield components in cereals: Heritabilities reveal a hierarchy of phenotypic plasticities. *Field Crops Research* **127**: 215-224
- Sajo M, Longhi-Wagner H, Rudall P** (2007) Floral development and embryology in the early-divergent grass *Pharus*. *International Journal of Plant Sciences* **168**: 181-191
- Sajo MG, Longhi-Wagner H, Rudall P** (2008) Reproductive morphology of the early-divergent grass *Streptochaeta* and its bearing on the homologies of the grass spikelet. *Plant Systematics and Evolution* **275**: 245-255
- Sakuma S, Golan G, Guo Z, Ogawa T, Tagiri A, Sugimoto K, Bernhardt N, Brassac J, Mascher M, Hensel G, Ohnishi S, Jinno H, Yamashita Y, Ayalon I, Peleg Z, Schnurbusch T, Komatsuda T** (2019) Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proceedings of the National Academy of Sciences* **116**: 5182-5187
- Sakuma S, Schnurbusch T** (2020) Of floral fortune: tinkering with the grain yield potential of cereal crops. *New Phytologist* **225**: 1873-1882
- Sanchez-Garcia M, Royo C, Aparicio N, Martín-Sánchez J, Alvaro F** (2013) Genetic improvement of bread wheat yield and associated traits in Spain during the 20th century. *The Journal of Agricultural Science* **151**: 105-118
- Savin R, Slafer GA** (1991) Shading effects on the yield of an Argentinian wheat cultivar. *The Journal of Agricultural Science* **116**: 1-7
- Schilling S, Kennedy A, Pan S, Jermiin LS, Melzer R** (2020) Genome-wide analysis of MIKC-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytologist* **225**: 511-529
- Schurch NJ, Schofield P, Gierliński M, Cole C, Sherstnev A, Singh V, Wrobel N, Gharbi K, Simpson GG, Owen-Hughes T** (2016) How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *Rna* **22**: 839-851
- Shang Q, Wang Y, Tang H, Sui N, Zhang X, Wang F** (2021) Genetic, hormonal, and environmental control of tillering in wheat. *The Crop Journal* **9**: 986-991
- Shaw LM, Turner AS, Herry L, Griffiths S, Laurie DA** (2013) Mutant alleles of *Photoperiod-1* in wheat (*Triticum aestivum* L.) that confer a late flowering phenotype in long days. *PLoS One* **8**: e79459
- Shen C, Li G, Dreni L, Zhang D** (2021) Molecular control of carpel development in the grass family. *Frontiers in Plant Science* **12**: 635500
- Shi W, Yue L, Cheng J, Guo J, Li L, Xie K, Dong J, Xu Y, Guo J, Zhou M** (2018) A genome-wide associate study reveals favorable alleles conferring apical and basal spikelet fertility in wheat (*Triticum aestivum* L.). *Molecular Breeding* **38**: 1-12
- Sibony M, Pinthus M** (1988) Floret initiation and development in spring wheat (*Triticum aestivum* L.). *Annals of Botany* **61**: 473-479
- Siddique K, Kirby E, Perry M** (1989) Ear: stem ratio in old and modern wheat varieties; relationship with improvement in number of grains per ear and yield. *Field crops research* **21**: 59-78
- Simmonds J, Scott P, Brinton J, Mestre TC, Bush M, del Blanco A, Dubcovsky J, Uauy C** (2016) A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theoretical and Applied Genetics* **129**: 1099-1112
- Slafer G, Calderini D, Miralles D, Dreccer M** (1994) Preanthesis shading effects on the number of grains of three bread wheat cultivars of different potential number of grains. *Field Crops Research* **36**: 31-39
- Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, Arlotta P, Macosko EZ, Chen F** (2021) Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nature biotechnology* **39**: 313-319

- Stockman Y, Fischer R, Brittain E** (1983) Assimilate supply and floret development within the spike of wheat (*Triticum aestivum* L.). *Functional Plant Biology* **10**: 585-594
- Strauss S, Runions A, Lane B, Eschweiler D, Bajpai N, Trozzi N, Routier-Kierzkowska A-L, Yoshida S, Rodrigues da Silveira S, Vijayan A, Tofanelli R, Majda M, Echevin E, Le Gloanec C, Bertrand-Rakusova H, Adibi M, Schneitz K, Bassel GW, Kierzkowski D, Stegmaier J, Tsiantis M, Smith RS** (2022) Using positional information to provide context for biological image analysis with MorphoGraphX 2.0. *eLife* **11**: e72601
- Thakur P, Kumar S, Malik JA, Berger JD, Nayyar H** (2010) Cold stress effects on reproductive development in grain crops: an overview. *Environmental and Experimental Botany* **67**: 429-443
- Theissen G, Saedler H** (2001) Floral quartets. *Nature* **409**: 469-471
- Thiel J, Koppolu R, Trautewig C, Hertig C, Kale S, Erbe S, Mascher M, Himmelbach A, Rutten T, Esteban E** (2021) Transcriptional landscapes of floral meristems in barley. *Science Advances* **7**: eabf0832
- Tingay S, McElroy D, Kalla R, Fieg S, Wang M, Thornton S, Brettell R** (1997) Agrobacterium tumefaciens-mediated barley transformation. *The Plant Journal* **11**: 1369-1376
- Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES** (2003) MADS box genes control vernalization-induced flowering in cereals. *Proceedings of the National Academy of Sciences* **100**: 13099-13104
- Trevaskis B, Tadege M, Hemming MN, Peacock WJ, Dennis ES, Sheldon C** (2007) Short Vegetative Phase-Like MADS-Box Genes Inhibit Floral Meristem Identity in Barley. *Plant Physiology* **143**: 225-235
- Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J** (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* **314**: 1298-1301
- Van Bel M, Diels T, Vancaester E, Kreft L, Botzki A, Van de Peer Y, Coppens F, Vandepoele K** (2017) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Research* **46**: D1190-D1196
- Vegetti AC, Weberling F** (1996) The structure of the paracladial zone in *Poaceae*. *Taxon* **45**: 453-460
- Vijayan A, Tofanelli R, Strauss S, Cerrone L, Wolny A, Strohmeier J, Kreshuk A, Hamprecht FA, Smith RS, Schneitz K** (2021) A digital 3D reference atlas reveals cellular growth patterns shaping the *Arabidopsis* ovule. *eLife* **10**: e63262
- Voss-Fels KP, Stahl A, Wittkop B, Lichthardt C, Nagler S, Rose T, Chen T-W, Zetzsche H, Seddig S, Majid Baig M, Ballvora A, Frisch M, Ross E, Hayes BJ, KHayden MJ, Ordon F, Leon J, Cage H, Friedt W, Stutzel H, Snowdon RJ** (2019) Breeding improves wheat productivity under contrasting agrochemical input levels. *Nature Plants* **5**: 706-714
- Waddington S, Cartwright P, Wall P** (1983) A quantitative scale of spike initial and pistil development in barley and wheat. *Annals of Botany* **51**: 119-130
- Walla A, Wilma van Esse G, Kirschner GK, Guo G, Brünje A, Finkemeier I, Simon R, von Korff M** (2020) An Acyl-CoA N-Acyltransferase Regulates Meristem Phase Change and Plant Architecture in Barley. *Plant Physiology* **183**: 1088-1109
- Wang C, Yang X, Li G** (2021) Molecular Insights into Inflorescence Meristem Specification for Yield Potential in Cereal Crops. *International Journal of Molecular Sciences* **22**: 3508
- Wang L, Yin H, Qian Q, Yang J, Huang C, Hu X, Luo D** (2009) *NECK LEAF 1*, a GATA type transcription factor, modulates organogenesis by regulating the expression of multiple regulatory genes during reproductive development in rice. *Cell research* **19**: 598-611
- Wang Y, Yu H, Tian C, Sajjad M, Gao C, Tong Y, Wang X, Jiao Y** (2017) Transcriptome Association Identifies Regulators of Wheat Spike Architecture. *Plant Physiology* **175**: 746-757
- Weatherwax P** (1923) *The story of the maize plant*. University of Chicago Press
- Weberling F** (1992) *Morphology of flowers and inflorescences*. CUP Archive
- Werner S, Engler C, Weber E, Gruetzner R, Marillonnet S** (2012) Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioengineered* **3**: 38-43
- Whingwiri E, Stern W** (1982) Floret survival in wheat: significance of the time of floret initiation relative to terminal spikelet formation. *The Journal of Agricultural Science* **98**: 257-268
- Whingwiri EE, Kuo J, Stern WR** (1981) The Vascular System in the Rachis of a Wheat Ear. *Annals of Botany* **48**: 189-202

- Whipple CJ** (2017) Grass inflorescence architecture and evolution: the origin of novel signaling centers. *New phytologist* **216**: 367-372
- Whipple CJ, Ciceri P, Padilla CM, Ambrose BA, Bandong SL, Schmidt RJ** (2004) Conservation of B-class floral homeotic gene function between maize and *Arabidopsis*.
- Whipple CJ, Hall DH, DeBlasio S, Taguchi-Shiobara F, Schmidt RJ, Jackson DP** (2010) A Conserved Mechanism of Bract Suppression in the Grass Family. *The Plant Cell* **22**: 565-578
- Wilensky U** (1999) NetLogo, Evanston, IL: Center for connected learning and computer-based modeling, Northwestern University
- Willis L, Refahi Y, Wightman R, Landrein B, Teles J, Huang KC, Meyerowitz EM, Jönsson H** (2016) Cell size and growth regulation in the *Arabidopsis thaliana* apical stem cell niche. *Proceedings of the National Academy of Sciences* **113**: E8238-E8246
- Wittern LM, Barrero JM, Bovill WD, Verbyla KL, Hughes T, Swain SM, Steed G, Webb AA, Gardner K, Greenland A** (2022) Overexpression of the *WAP0-A1* gene increases the number of spikelets per spike in bread wheat. *Scientific reports* **12**: 1-15
- Wolde GM, Mascher M, Schnurbusch T** (2019) Genetic modification of spikelet arrangement in wheat increases grain number without significantly affecting grain weight. *Molecular Genetics and Genomics* **294**: 457-468
- Wu D, Liang W, Zhu W, Chen M, Ferrándiz C, Burton RA, Dreni L, Zhang D** (2018) Loss of LOFSEP transcription factor function converts spikelet to leaf-like structures in rice. *Plant Physiology* **176**: 1646-1664
- Würschum T, Leiser WL, Langer SM, Tucker MR, Longin CFH** (2018) Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long-term genetic trends of grain yield components. *Theoretical and Applied Genetics* **131**: 2071-2084
- Xia K, Sun H-X, Li J, Li J, Zhao Y, Chen L, Qin C, Chen R, Chen Z, Liu G** (2022) The single-cell stereoseq reveals region-specific cell subtypes and transcriptome profiling in *Arabidopsis* leaves. *Developmental Cell* **57**: 1299-1310. e1294
- Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu Z, Wang L, Fox N, Wang X, Drenkow J, Luo A, Char SN, Yang B, Sylvester AW, Gingeras TR, Schmitz RJ, Ware D, Lipka AE, Gillis J, Jackson D** (2021) Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Developmental Cell* **56**: 557-568.e556
- Yamaguchi N, Winter CM, Wu M-F, Kanno Y, Yamaguchi A, Seo M, Wagner D** (2014) Gibberellin acts positively then negatively to control onset of flower formation in *Arabidopsis*. *Science* **344**: 638-641
- Yoshida H** (2012) Is the lodicule a petal: molecular evidence? *Plant science* **184**: 121-128
- Youssef HM, Eggert K, Koppolu R, Alqudah AM, Poursarebani N, Fazeli A, Sakuma S, Tagiri A, Rutten T, Govind G, Lundqvist U, Graner A, Komatsuda T, Sreenivasulu N, Schnurbusch T** (2017) *VRS2* regulates hormone-mediated inflorescence patterning in barley. *Nature Genetics* **49**: 157-161
- Zhu W, Yang L, Wu D, Meng Q, Deng X, Huang G, Zhang J, Chen X, Ferrándiz C, Liang W** (2022) Rice *SEPALLATA* genes *OsMADS5* and *OsMADS34* cooperate to limit inflorescence branching by repressing the *TERMINAL FLOWER1*-like gene *RCN4*. *New Phytologist* **233**: 1682-1700
- Zhu X-X, Li Q-Y, Shen C-C, Duan Z-B, Yu D-Y, Niu J-S, Ni Y-J, Jiang Y-M** (2016) Transcriptome Analysis for Abnormal Spike Development of the Wheat Mutant *dms*. *PLOS ONE* **11**: e0149287

## Appendices

### Appendix 1

#### **High expression of MADS-box gene *VRT2* increases the number of rudimentary basal spikelets in wheat**

**Backhaus, A.E., Lister, A., Tomkins, M., Adamski, N.M., Simmonds, J., Macaulay, I., Morris, R.J., Haerty, W. and Uauy, C., 2022.** High expression of the MADS-box gene *VRT2* increases the number of rudimentary basal spikelets in wheat. *Plant Physiology*.: <https://doi.org/10.1093/plphys/kiac156>

# High expression of the MADS-box gene *VRT2* increases the number of rudimentary basal spikelets in wheat

Anna E. Backhaus <sup>1</sup>, Ashleigh Lister <sup>2</sup>, Melissa Tomkins <sup>1</sup>, Nikolai M. Adamski <sup>1</sup>, James Simmonds <sup>1</sup>, Iain Macaulay <sup>2</sup>, Richard J. Morris <sup>1</sup>, Wilfried Haerty <sup>2</sup> and Cristobal Uauy <sup>1,\*†</sup>

<sup>1</sup> John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK

<sup>2</sup> Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, UK

\*Author for correspondence: [cristobal.uauy@jic.ac.uk](mailto:cristobal.uauy@jic.ac.uk)

<sup>†</sup>Senior author.

Contributions are based on the CRediT Contributor Roles Taxonomy (<https://casrai.org/credit/>). A.E.B. is responsible for conceptualization, methodology, data curation, formal analysis, investigation, visualization, writing—original draft. A.L. involved in investigation and methodology. M.T. involved in software, formal analysis, writing—review, and editing. N.M.A. involved in resources, writing—review, and editing. J.S. is responsible for data curation, formal analysis, resources, writing—review, and editing. I.M. involved in methodology and supervision. R.J.M. responsible for conceptualization, methodology, supervision, writing—review, and editing. W.H. involved in conceptualization, funding acquisition, supervision, writing—review, and editing. C.U. involved in conceptualization, funding acquisition, methodology, project administration, supervision, writing—review, and editing. The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plphys/pages/general-instructions>) is: Cristobal Uauy ([cristobal.uauy@jic.ac.uk](mailto:cristobal.uauy@jic.ac.uk)).

## Abstract

Spikelets are the fundamental building blocks of *Poaceae* inflorescences, and their development and branching patterns determine the various inflorescence architectures and grain yield of grasses. In wheat (*Triticum aestivum*), the central spikelets produce the most and largest grains, while spikelet size gradually decreases acropetally and basipetally, giving rise to the characteristic lanceolate shape of wheat spikes. The acropetal gradient corresponds with the developmental age of spikelets; however, the basal spikelets are developed first, and the cause of their small size and rudimentary development is unclear. Here, we adapted G&T-seq, a low-input transcriptomics approach, to characterize gene expression profiles within spatial sections of individual spikes before and after the establishment of the lanceolate shape. We observed larger differences in gene expression profiles between the apical, central, and basal sections of a single spike than between any section belonging to consecutive developmental time points. We found that *SHORT VEGETATIVE PHASE* MADS-box transcription factors, including *VEGETATIVE TO REPRODUCTIVE TRANSITION 2* (*VRT-A2*), are expressed highest in the basal section of the wheat spike and display the opposite expression gradient to flowering E-class *SEPALLATA1* genes. Based on multi-year field trials and transgenic lines, we show that higher expression of *VRT-A2* in the basal sections of the spike is associated with increased numbers of rudimentary basal spikelets. Our results, supported by computational modeling, suggest that the delayed transition of basal spikelets from vegetative to floral developmental programs results in the lanceolate shape of wheat spikes. This study highlights the value of spatially resolved transcriptomics to gain insights into developmental genetics pathways of grass inflorescences.

Received February 01, 2022. Accepted March 13, 2022. Advance access publication April 4, 2022

© The Author(s) 2022. Published by Oxford University Press on behalf of American Society of Plant Biologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

## Introduction

The arrangement of flowers in individual plants of the same species is highly conserved and follows a systematic and rhythmic pattern. This systematic appearance of flowers is not surprising, as floral architectures are determined by the regular initiation of flower primordia on the flanks of the apical meristem and their rate of initiation and developmental fate are under strong genetic control (Prusinkiewicz et al., 2007). The unifying feature of floral architecture in grasses (*Poaceae*) is the formation of all flowers (termed florets) within spikelets (Kellogg et al., 2013). Spikelets are the fundamental building blocks of grass inflorescences and their development and branching patterns determine the various inflorescence architectures of grasses (e.g. spikes or panicles). Wheat (*Triticum aestivum*) forms a spike-shaped inflorescence, in which sessile spikelets are directly attached to the inflorescence axis (or rachis) in a distichous phyllotaxis (Koppolu and Schnurbusch, 2019). Upon floral transition, the vegetative meristem ceases to initiate leaf primordia and transitions into the inflorescence meristem (IM). During the Double Ridge (DR) stage of wheat spike development, the IM initiates a lower leaf ridge and an upper spikelet ridge (or primordia) during each iteration. Within the inflorescence the upper ridges differentiate into spikelet meristems, while the lower ridges are suppressed upon flowering (Bommert and Whipple, 2018). DR initiation will continue at the IM until the terminal spikelet stage, when IM forms a final spikelet (Koppolu and Schnurbusch, 2019). Spikelet initiation and development have been extensively studied in wheat and other monocot crops, such as rice (*Oryza sativa*), maize (*Zea mays*), and barley (*Hordeum vulgare*), as the number of spikelets per spike is a major determining factor for grain number and thus yield per spike.

Not all spikelets across the wheat spike, however, produce the same amount of grain. The central spikelets produce the most and largest grains, while spikelet size gradually decreases acropetally and basipetally. Within a single spike, the most apical and basal spikelets might produce no or only one grain while the central spikelets of the same spike set 3–5 grains. Bonnett (1966) documented that this distinct lanceolate shape of the wheat spike is first established during the Glume Primordia (GP) stage (just after the DR stage). This asynchronous development among the spikelets is maintained throughout the development of the spike. The gradual decrease in spikelet size from the central to apical section of the spike can be explained by the continuous development of new spikelet ridges from the apical IM: the most apical spikelets are the youngest and had the least time to develop. However, basal spikelets are initiated first and it is unclear why they remain smaller than their central counterparts. In the mature spike the most basal one or two spikelets are often only formed in a rudimentary manner, with small glumes present but all floral structures remaining immature.

Efforts to understand the genetics of wheat spikelet initiation and development have focused on members of the

MADS-box transcription factor (TF) family, which play central roles in the flowering gene models (Zhao et al., 2006). Li et al. (2019) showed that MADS-box genes of the SQUAMOSA-clade, *VERNALISATION 1* (*VRN1*), *FRUITFULL 2* (*FUL2*), and *FUL3*, have overlapping functions in controlling the timing of the transitions from the vegetative to IM as well as the formation of the terminal spikelet. In *vrn1ful2*-null mutants, the IM remained indeterminate causing the mutants to form more spikelets per spike. However, all lateral spikelets were replaced by leafy shoots in the *vrn1ful2* double and *vrn1ful2ful3* triple mutants (Li et al., 2019). These mutants had increased expression of genes belonging to the *SHORT VEGETATIVE PHASE* (*SVP*) family of MADS-box genes, including *VEGETATIVE TO REPRODUCTIVE TRANSITION 2* (*VRT2*). Subsequent studies determined that overexpression of *VRT2* led to reversion of basal spikelets to spikes and the downregulation of other MADS-box genes required for floral development, including members of the *SEPALLATA 1* (*SEP1*) clade (Li et al., 2021). Together, these studies exemplify the importance of the temporal sequence of flowering gene expression for the correct development of the wheat spike.

Attempts to unravel the genetic network controlling wheat spike development have focused on these temporal changes in expression patterns across consecutive developmental stages. For example, Li et al. (2018) and Feng et al. (2017) performed transcriptome profiling using pooled samples of multiple complete spikes from six (vegetative to floret differentiation) and four (DR to young floret) developmental stages, respectively. In a few cases, studies have examined the expression patterns of individual genes (via reverse transcription–quantitative PCR [RT–qPCR]) and found gene expression gradients along the spike. For example, Debernardi et al. (2017) demonstrated that *APETALA 2* (*AP2*) is expressed higher in the apical section of wheat spikes than in central or basal sections. This *AP2* expression gradient was associated with morphological changes along the same spike. This study alongside work in barley (Youssef et al., 2017) suggests that gene expression gradients within individual developmental stages could be important to further unravel the genetic control of spike development. However, despite its potential biological importance, spatial transcriptome profiles along the spike have yet to be investigated in wheat.

In this study, we aimed to characterize gene expression profiles along the spike during the establishment of the lanceolate shape of the wheat spike from DR to GP. We adapted Genome and Transcriptome sequencing (G&T-seq), a low-input sequencing approach to sequence the transcriptome of the sections. Recently, Giolai et al. (2019) adapted the protocol to identify expression differences across single leaves of *Arabidopsis* (*Arabidopsis thaliana*; GaST-seq), demonstrating that the G&T-seq method can be readily used for sequencing of hand harvested, small input plant material without the need of previous tissue dissociation or treatment. G&T-seq is thus comparable to methods using laser-

micro dissection followed by sequencing to achieve spatially resolved transcriptome-wide sequencing data. In comparison, the available transcriptome sequencing methods at higher resolutions (such as single-cell RNA-seq or fluorescence-activated cell sorting [FACS]) are not spatially resolved as the complete tissue is dissolved into single cells for barcoding or selection prior to sequencing (Rich-Griffin et al., 2020).

We sequenced the apical, central, and basal sections of individual spikes before (DR) and after (GP) the establishment of the lanceolate shape. Gene expression profiles differed most strongly between spatial sections of the same spike, as opposed to temporal sections (any two sections from different time points). Members of the *SVP* gene family were expressed most highly in the basal sections with expression decreasing upwards from the base (acropetally), while members of the *SEP1* gene family showed the opposite expression pattern, that is, most highly expressed in apical sections, with expression decreasing toward the base (basipetally). The increased number of rudimentary basal spikelets (RBSs) due to *VRT-A2* misexpression supports the hypothesis that high expression levels of *SVPs* in the basal section delays spikelet establishment, leading to their rudimentary shape in the mature spike. This study highlights that spikelets within the same spike experience significantly different flowering signals due to their consecutive development and spatial position within the spike. Acknowledging these differences can help us gain a better understanding of the genetic flowering pathway of grass inflorescences.

## Results

### Low-input sequencing enables spatial analysis of the wheat spike transcriptome

To investigate transcriptional differences between the apical, central, and basal section of developing wheat spikes, we adapted the low-input GG&T-seq method for RNA-seq of small plant tissue sections. G&T has been developed for single-cell RNA and DNA sequencing of mammalian systems (Macaulay et al., 2015) and was previously adapted for *Arabidopsis thaliana* (GaST-seq; Giolai et al., 2019). We collected four individual developing wheat (cv Paragon) spikes at both the DR and GP stage and hand-dissected them into apical, central, and basal sections (Figure 1A).

On average, samples had 28,799,626 reads (coefficient of variation [CV] 43%), of which 90% (CV 8.5%) aligned to the genome post adaptor trimming (Figure 1B and Table 1). Furthermore, the number of aligned reads and the number of expressed genes per library was largely homogeneous among the spatial sections and Waddington stages (Table 1). On average, 47,313 genes per library were expressed (>10 read counts) and we found no difference ( $P > 0.56$ , analysis of variance [ANOVA]) in the number of expressed genes across spatial (apical, central, and basal) or between temporal (DR, GP) conditions (Figure 1C). We excluded three libraries with low average number of expressed genes (difference >5 times the standard deviation;

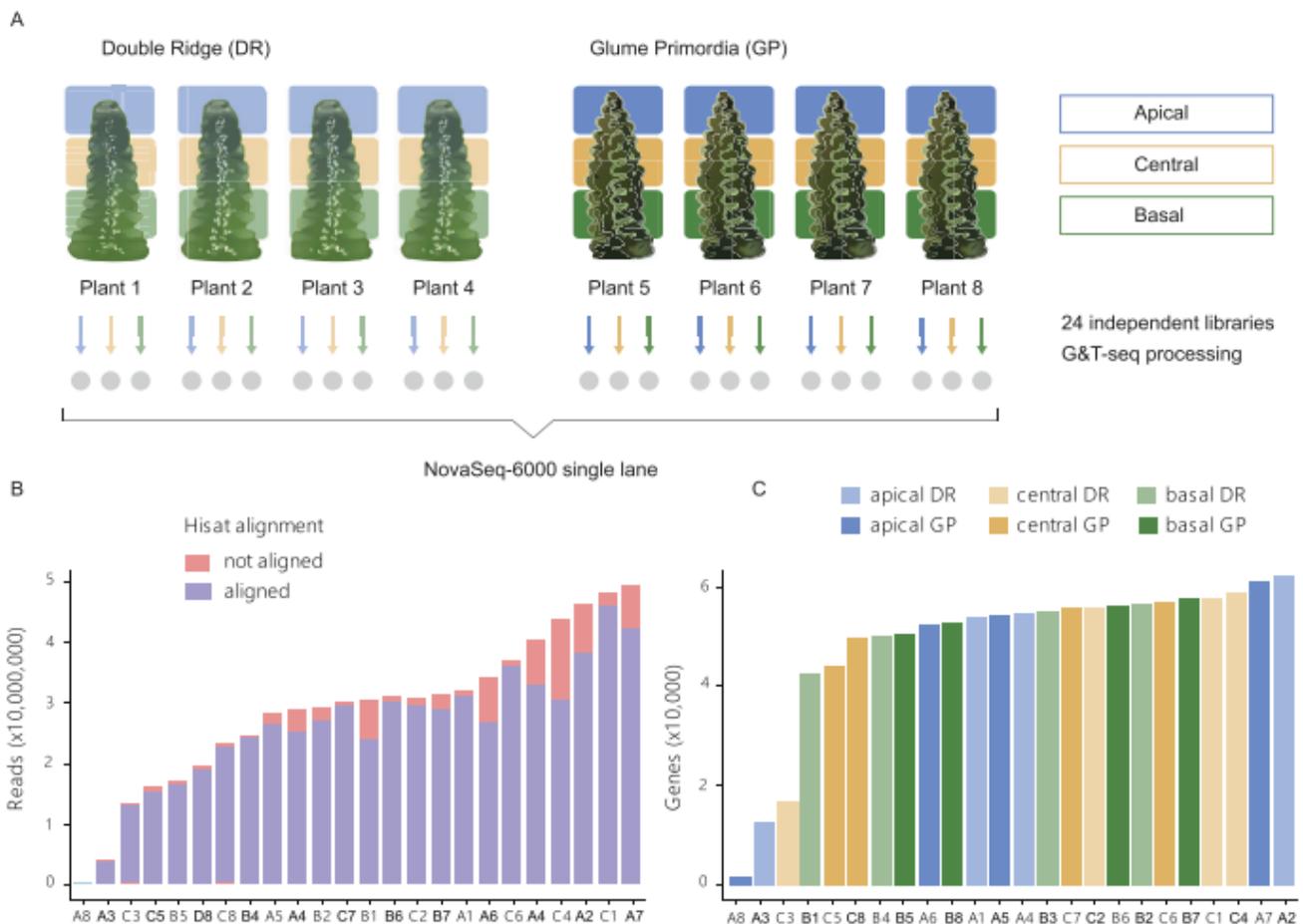
Figure 1C; Supplemental Table S1) and two libraries because they were strong outliers in the principal component analysis (PCA; Supplemental Figure S1A). In total, 19 RNA-seq libraries (DR: 3 apical, 4 central, and 3 basal; GP: 2 apical, 3 central, and 4 basal) passed our selection criteria and were used in the subsequent analyses. We identified 91,646 genes being expressed across these 19 libraries.

### Transcriptome-wide differences are largest between the apical and basal sections of the spike

To investigate global differences among the 19 RNA-seq libraries, we performed a PCA (Figure 2A). The first two PCs explained 19% and 16% of the overall variance present in the libraries. We observed that the two PCs separated libraries by the spatial position (apical, central, and basal) rather than developmental stage (DR and GP). There was a clear separation between libraries originating from apical and basal spike sections, while libraries from central sections were dispersed between these two clusters (Figure 2A). We investigated PC1–PC6 and found that none of these combinations clustered libraries by developmental stage (Supplemental Figure S1B). Given that we sequenced developing spike sections of single plants, as opposed to the more commonly employed pooling of multiple biological samples, we found as expected some degree of heterogeneity between samples from the same location and stage (Figure 2A).

To investigate this variation further, we quantified changes in gene expression across biological replicates by calculating CVs for each gene (see “Materials and Methods”). The median CV for a gene across the biological replicates was 39% (Supplemental Figure S2) with a Q1–Q3 interquartile range between 24% and 62%. We also calculated the CV per gene for published datasets from Li et al. (2018) and Feng et al. (2017). Both studies sequenced developing wheat spikes at similar developmental stages, pooling many spikes per sample. Li et al. (2018) pooled between 100 and 200 spikes of winter wheat (KN9204) per sample, while Feng et al. (2017) reported pooling of 10–50 spikes (cv. Chinese Spring) per sample. In both studies, the median CV of a gene was lower (14% and 21%, respectively) than in our study. The larger CVs in our data could be explained by the biological variation that exists between individual plants, which may have been reduced by the pooling of many spikes in both Li et al. (2018) and Feng et al. (2017).

We first analyzed differentially expressed genes (DEGs) between the DR and GP stage and between apical, central and basal sections across the two Waddington developmental stages. The number of DEGs between DR and GP (215 genes) was smaller than the number of DEGs identified between the spatial positions, which ranged from 280 DEGs between central and apical sections to 914 DEGs between the apical and basal sections (Figure 2B). Next, we compared the apical, basal, and central sections within each Waddington developmental stage. We identified more DEGs by comparing the spatial sections within either Waddington



**Figure 1** Low-input G&T sequencing of developing wheat spikes. **A**, Summary diagram of tissue collection and sequencing. Gray circles indicate the 24 individual libraries prepared for sequencing from each individual tissue section dissected from individual spikes. **B**, Reads per library after trimming and quality controls (see “Materials and methods”). Stacked bars indicate the number of reads aligned (blue) and not aligned (red) by HISAT to the RefSeq version 1.0 genome. **C**, Number of expressed genes (> 10 read counts) per library based on tissue section and Waddington developmental stage. In (**B** and **C**), the X-axis indicates the ID of each sample which is composed of the tissue section (A: apical, C: central, B: basal) and plant number (1–8) as indicated in (**A**). Detailed quality control data for each library is provided in [Supplemental Table S1](#).

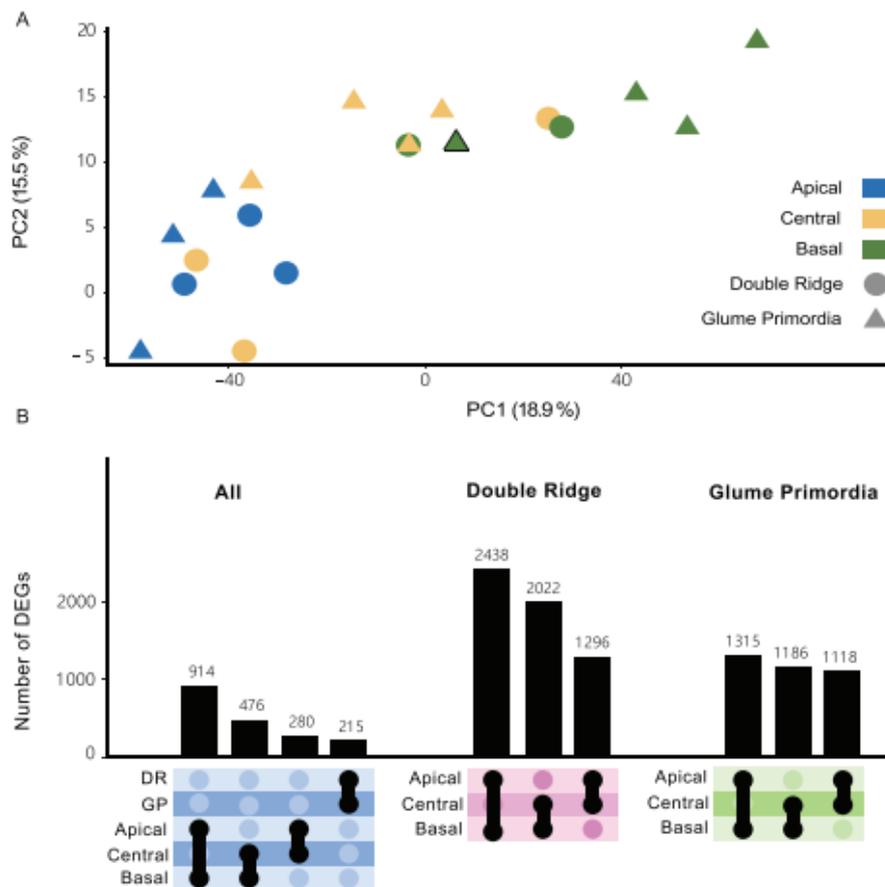
**Table 1** Average number of reads aligned to the RefSeqv1.0 genome and expressed genes (> 10 read counts) in the three tissue sections and two Waddington developmental stages

Section	Reads Aligned		Genes Expressed	
	DR	GP	DR	GP
Apical	26,342,934	23,913,657	44,488	41,152
Central	29,823,023	25,854,540	45,913	50,074
Basal	25,174,482	23,627,494	49,513	52,740

(n = 4 biological replicates per tissue section per developmental stage).

stage individually than in the combined analysis. The number of DEGs between apical and basal sections at each stage (DR: 2,438; GP: 1,315) was similar to the number of DEGs between central and basal sections (DR: 2,022; GP: 1,186). The number of DEGs between these sections at DR, however, was nearly double the number of DEGs at GP. In contrast, the number of DEGs between

apical and central sections was similar at both stages (DR: 1,296; GP: 1,118), suggesting that the basal section of the spike is most different in the earlier developmental stage. Only 11% of the DEGs were shared between DR and GP in the apical to basal comparison, 7% between the central to basal DEGs, and 5% between apical to central DEGs. In total, we identified 5,353 unique genes as differentially expressed between any of the three sections at either Waddington stage ([Supplemental Table S2](#)). Overall, the number of DEGs was largest between the apical and basal sections, reflecting the strong spatial clustering observed in the PCA graph, but most genes that were differentially expressed across the spike did not maintain this gradient over the two developmental stages. In summary, despite the high biological variation in gene expression in our data compared to previous pooled whole-spike studies, we could detect transcriptome-wide differences between the spatial sections of developing wheat spikes.



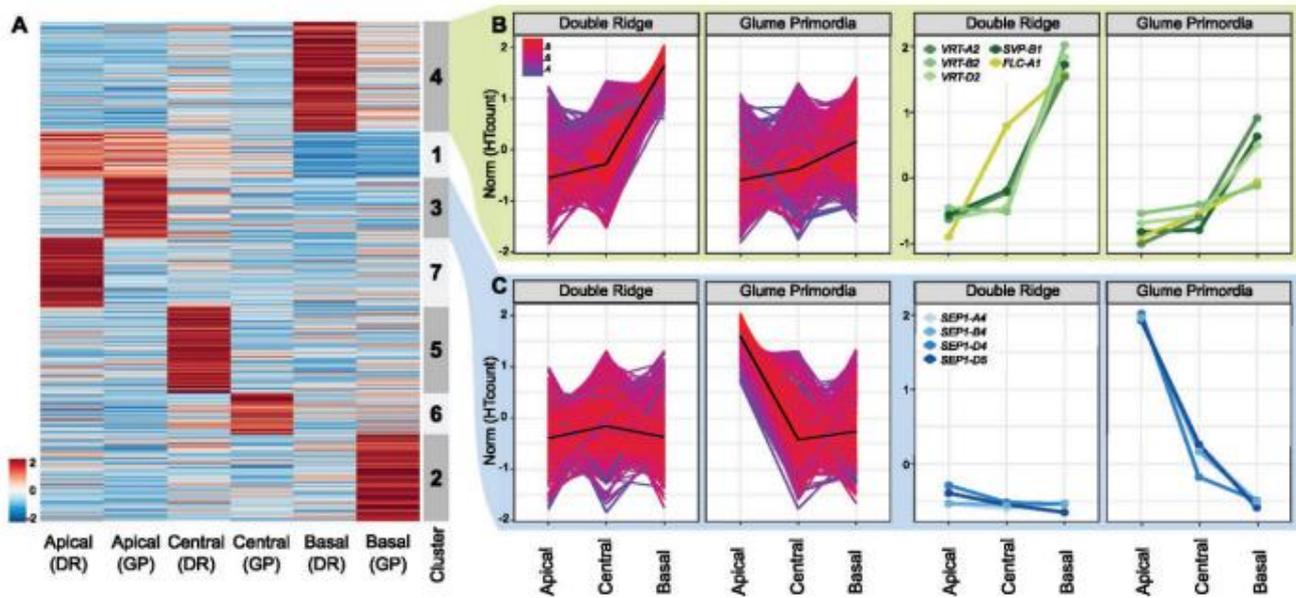
**Figure 2** Transcriptome-wide differences are larger between tissue sections than developmental time points. A, PCA on the 19 transcriptome libraries from apical (blue), central (yellow), and basal (green) sections of DR (circles) and GP (triangles) spikes. Black bordered triangle is plant 8 (GP) in which the basal section clustered closer with central-GP sections than the other basal-GP sections. B, UpSet plot showing the number of DEGs between spatial sections and Waddington stages.

### The SVP MADS-box TFs have opposing expression profiles to flowering E-class genes

To further investigate the differences in expression across the spike and to identify genes with similar expression patterns, we performed hierarchical and k-means clustering. We restricted the clustering to the 5,353 genes identified as differentially expressed across the spike at either one or both Waddington stages (Figure 3A). We identified seven non-redundant clusters, each containing between 8% and 21% of the 5,353 DEGs (Figure 3A; Supplemental Figure 3A). Both hierarchical and k-means clustering produced highly similar results (Supplemental Figure S3B). We identified 1,894 genes (35% of DEGs) to be more highly expressed in the apical section, either across both time points (503 genes, Cluster 1), or only at DR (751 genes, Cluster 7) or GP (640 genes, Cluster 3). In the central section, 1,362 genes (25%) had higher relative expression at either DR (917 DEGs, Cluster 5) or GP (445 DEGs, Cluster 6). In the basal section, we observed 2,097 genes (39%) being more highly expressed. Cluster 4 contained the most DEGs (1,170) and was characterized by an upregulation of expression in the basal section

at both Waddington stages, although this upregulation was higher at the DR stage. Another 927 genes were upregulated in the basal section, but only at the GP stage (Cluster 2).

To further characterize the clusters, we independently tested for enrichment of TF families and gene ontology (GO) terms relating to biological processes (all GO terms and TF families in Supplemental Tables S3 and S4, respectively). Genes that were more highly expressed in the apical section were enriched for the GO terms “reproductive structure development” (GO:0048608) and “floral organ development” (GO:0048437; Cluster 1) as well as for HD-Zip<sub>IV</sub> and SRS TF families ( $P < 0.001$ ; Cluster 1). In Cluster 3 (highly expressed in the apical section at GP; Figure 3B), we found no significant enrichment of GO terms relating to biological processes, but a significant enrichment of MADS<sub>II</sub> TFs ( $P = 0.013$ ). For clusters defined by an increased expression in the central sections (Clusters 5), we detected an enrichment for GO terms related to polyphosphate processes (GO:0006797/0006779;  $P < 0.03$ ), and a significant enrichment of the C2C2\_CO-like TFs at DR ( $P < 0.04$ ). Genes with higher expression in the basal section of the spike (Cluster 4; Figure 3C) were enriched for a number



**Figure 3** Clustering of DEGs across the wheat spike. A, Normalized expression matrix and K-means clustering of the 5,353 genes differentially expressed across the spike at either one or both Waddington stages. Colors (blue to red) show relative log<sub>2</sub> expression of genes after normalization. B, Expression pattern of the 1,170 genes allocated to Cluster 4 (left), and of MADS-box TFs in the same cluster (right). Colors indicate how well the gene expression pattern fits the average expression pattern (black line). Red = best fit, Blue = least good fit. C, Expression pattern of the 640 genes and MADS-box TFs of Cluster 3 as arranged in (B). Norm = normalized and scaled gene expression. RefSeq version 1.1 gene IDs and raw expression values of genes shown in the right-hand panels are presented in [Supplemental Table S2](#).

of GO terms related to photosynthesis (e.g. GO:0015979;  $P < 0.01$ ) and “negative regulation of flower development” (GO:0009910;  $P < 0.01$ ), as well as for MADS-II TFs ( $P = 0.08$ ). Cluster 2, which was characterized by higher expression in the basal section only at GP, was enriched for “jasmonic acid response” (GO:0009753;  $P < 0.001$ ) and Tify ( $P < 0.001$ ) and C2C2\_CO-like TFs ( $P < 0.04$ ).

We were interested in further characterizing the expression patterns of the MADS-box TFs as they were significantly enriched in two of the seven clusters and are important in floral transition and development (Becker and Theißen, 2003; Feng et al., 2017). We detected 14 differentially expressed MADS-box TFs in our study, 5 of which were more highly expressed in the apical section, 4 of these only at GP stage (Cluster 3), and 1 being consistently expressed across both Waddington stages (Cluster 1). In contrast, five MADS-box TFs were more highly expressed in the basal section at both Waddington stages (Cluster 4) and another two were more highly expressed in the basal section only at GP (Cluster 2). An additional two MADS-box genes were part of the remaining clusters ([Supplemental Table S4](#)).

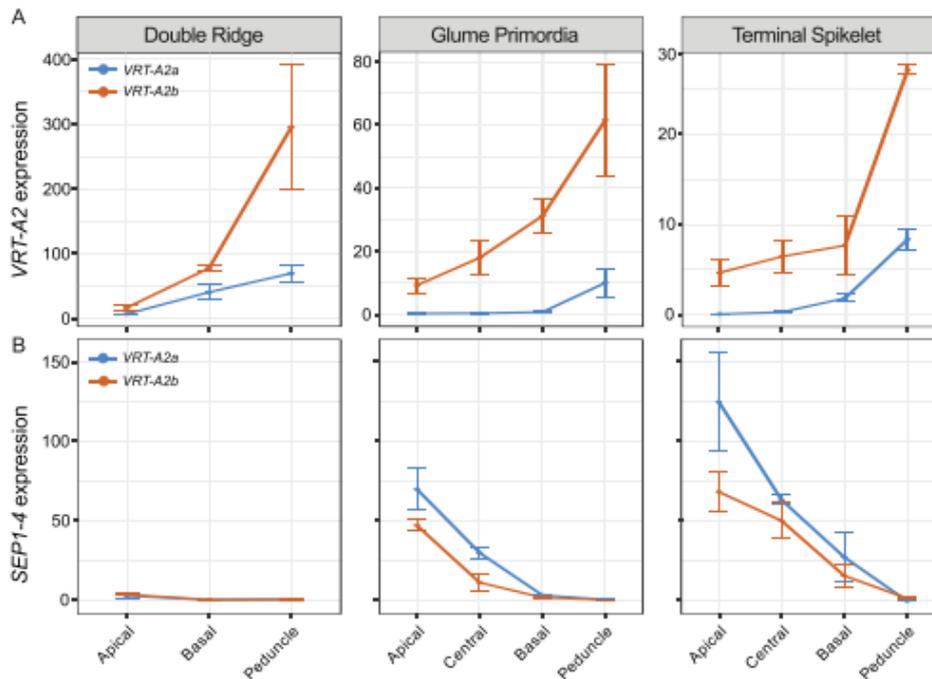
In the apical/GP Cluster 3, we noticed that all MADS-box genes belonged to the *T. aestivum* SEP1 group ([Figure 3C](#)). All three homoeologs of SEP1-4 (*TraesCS7A02G122000*, *TraesCS7B02G020800*, and *TraesCS7D02G120500*) and the D-genome copy of SEP1-5 (*TraesCS7D02G120600*) were part of this cluster. The SEP genes were expressed at relatively low levels at DR ([Supplemental Table S2](#)), but were significantly upregulated at GP, with their transcript levels being highest in the apical section. The increased expression of SEP1-4 at

GP was in agreement with their previously reported expression patterns in tetraploid wheat by Li et al. (2021).

In the contrasting Cluster 4 (upregulation in basal sections), we noticed the presence of multiple MADS-box genes belonging to the SVP family ([Figure 3B, right](#)), which consists of three genes in wheat (SVP1, VRT2, and SVP3). Members of this family are important for the transition from vegetative to floral meristem identity in cereals (Trevaskis et al., 2007). All three homoeologs of VRT2 (*TraesCS7A02G175200*, *TraesCS7B02G080300*, and *TraesCS7D02G176700*) and the B-genome copy of SVP1 (*TraesCS6B02G343900*) were present in Cluster 4. The cluster also contained *FLOWERING LOCUS C* (*FLC-A1*; *TraesCS7A02G260900*), although it was expressed higher in DR-central sections compared to the SVPs and had a linear expression gradient at GP. All SVPs had very similar expression patterns, being strongly expressed in basal sections only. Expression of SVPs was higher in all DR sections compared to the equivalent section in GP. Constitutive over-expression of SVP-family members in wheat and barley has been shown to delay or even reverse floral development (Trevaskis et al., 2007; Li et al., 2021). This led to hypothesis that the rudimentary development of basal spikelets was associated with an increase in VRT2 expression levels.

#### SVP expression is higher in basal and peduncle sections and increased across all sections in *T. polanicum* VRT-A2b isogenic lines

To validate the expression pattern of VRT2 in the individual spike analysis, we performed reverse transcription quantitative PCR (RT-qPCR) on independently collected, pooled



**Figure 4** Effect of *VRT-A2b* allele on *VRT-A2* and *SEP1-4* expression in the spike. Relative expression ( $2^{-\Delta\Delta CT}$ ) of *VRT-A2* (A) and *SEP1-4* (B) in the different sections of the spike across three time points in NILs carrying either the wild-type *VRT-A2a* (blue) or the *VRT-A2b* allele from *Triticum turgidum* ssp. *polonicum* (orange). The data are shown as mean  $\pm$  SE of gene expression compared with control gene *Actin*.  $N = 3$  biological replicates (see Supplemental Table S5 for expression data and Supplemental Table S6 for statistical analysis of gene expression differences).

spike sections from cv Paragon, carrying the wild-type *VRT-A2a* allele (Figure 4A (blue curves); Supplemental Table S5). We included a later time point, Terminal Spikelet (TS), which is about 10 d after GP to study how *VRT2* expression changes in later stages. At TS stage, the central spikelets have developed multiple florets primordia. We also included a small part of the peduncle (stem) section just below the spike as an additional spatial section. We focused the expression analysis on the A-genome homoeolog, *VRT-A2*, as its role in spike, glume, and grain development of wheat was recently characterized (Adamski et al., 2021; Liu et al., 2021).

We identified a significant interaction effect between Waddington stage and spatial section ( $P < 0.0001$ ; Supplemental Table S6); we thus analyzed the three Waddington stages separately (Figure 4). At DR, we were limited to dissecting the spike into apical, basal, and peduncle sections, as the small size of the spike meristem did not allow precise dissection of the central section when using multiple (pooled) spikes. At DR, we found *VRT-A2* marginally expressed, with significantly lower expression levels in the apical section compared to the basal ( $P = 0.003$ ) and peduncle sections ( $P = 0.001$ ). Although expression in the peduncle was higher than in the basal section at DR, this was not significant ( $P = 0.116$ ). At GP, *VRT-A2* expression was borderline detectable in the apical, central and basal sections, but expression was significantly higher in the peduncle with respect to the three spike tissues ( $P < 0.001$  for all three comparisons). Lastly, *VRT-A2* expression at TS stage

was just detectable and significantly different between all sections ( $P = 6.6E-06$ ). Overall, expression decreased significantly from DR to GP/TS Waddington stages in the apical ( $P = 0.00015$ ), basal ( $P = 0.0074$ ), and peduncle ( $P = 0.012$ ) sections consistent with the previously reported strong downregulation of *VRT-A2* in the early wheat spike development (Adamski et al., 2021; Li et al., 2021; Liu et al., 2021). This is also consistent with the observed downregulation of *VRT-A2* orthologs upon floral transition in barley (Trevaskis et al., 2007) and rice (Harrop et al., 2016). As observed in the low-input RNA-seq data, the RT-qPCR data confirmed the strong basipetal gradient in *VRT-A2* expression across the spike at DR and revealed that its expression was even higher within the peduncle.

We hypothesized that the higher expression in the basal section of the wheat spike compared to the central and apical sections is associated with the rudimentary development of the basal spikelets. To test the effect of higher *VRT-A2* expression on basal spikelet development, we analyzed the effect of the *Triticum turgidum* ssp. *polonicum* *VRT-A2b* allele on the expression gradient of *VRT-A2* and spike morphology. Adamski et al. (2021) showed that *VRT-A2* in *T. polonicum*, a tetraploid subspecies of wheat, carries a sequence re-arrangement in its first intron. This results in the higher expression of the *T. polonicum* *VRT-A2b* allele, with respect to the wild-type *VRT-A2a* allele, during early spike development. We performed RT-qPCR on a cv Paragon near isogenic line (NIL) carrying the *VRT-A2b* allele and compared *VRT-A2* expression against the Paragon wild-type NIL

described above (Figure 4; Supplemental Table S5). Consistent with the results of Adamski et al. (2021), we detected significantly higher expression of *VRT-A2b* compared to the wild-type allele across most of the tissue sections (see Supplemental Table S6 for individual comparisons), and a progressive decrease in *VRT-A2b* expression over time ( $P = 0.031$ ). In contrast to the wild-type NILs, analysis of variance (ANOVA) did not identify a significant interaction effect between spatial section and Waddington stage in *VRT-A2b* NILs ( $P = 0.18$ ). We thus examined the overall expression patterns and found that across all three developmental stages *VRT-A2b* expression differences were significant ( $P < 0.0001$ ). These results suggest that the basipetal expression gradient in the spike is maintained in the NILs with the *T. polonicum* *VRT-A2b* allele.

We also tested the effect of *VRT-A2* expression levels on *SEP1* expression in the tissue sections of *VRT-A2* NILs. We confirmed that *SEP1-4* expression is only marginally detectable at DR and differences in expression between the apical, basal, and peduncle section are hardly detectable with only marginally higher expression in the apical sections ( $P = 0.015$ ) at this stage. At GP, *SEP1-4* expression is significantly higher toward the tip of the spike ( $P < 0.0001$ ) consistent with the low-input RNA-seq data. Furthermore, *SEP1-4* expression was significantly lower in *VRT-A2b* NILs compared to the wild-type allele across all spike sections ( $P = 0.008$ ), confirming that higher *VRT-A2* expression can negatively affect *SEP1-4* expression. Similar trends were observed at TS, where expression was significantly lower in basal sections ( $P < 0.0001$ ) and the gradient across the spike was maintained in *VRT-A2b* NILs, but expression was overall lower ( $P < 0.003$ ). *SEP1-4* is not expected to be expressed in vegetative tissue such as the peduncle; therefore, the lack of expression in this tissue across the three stages indicates that no floral tissue was accidentally sampled as peduncle (Figure 4B).

### Misexpression of *VRT-A2b* in *T. polonicum* increases RBS numbers

To evaluate if the higher expression of *VRT-A2* in basal spikelets affects their development, we examined the *VRT-A2* NILs (BC<sub>4</sub> and BC<sub>6</sub>) sown as winter crops in four environments. In each field trial, we evaluated the number of RBS, that is spikelets which are reduced in size and do not contain mature grains (Figure 5A). The number of RBS was significantly increased in NILs carrying the *VRT-A2b* allele in all four environments ( $P < 0.0001$ , except Morley, 2017  $P < 0.01$ ; Figure 5B; Supplemental Table S7). The *VRT-A2a* NILs and the recurrent parent Paragon had on average 1.85 RBS, whereas *VRT-A2b* NILs produced on average 2.91 RBS. A similar difference in RBS between the NILs was observed in glasshouse conditions (*VRT-A2b* effect of +1.6 RBS; Supplemental Table S8; Figure 5C).

Furthermore, we also recorded the number of RBS in seven homozygous BC<sub>6</sub> recombinant lines used to fine-map

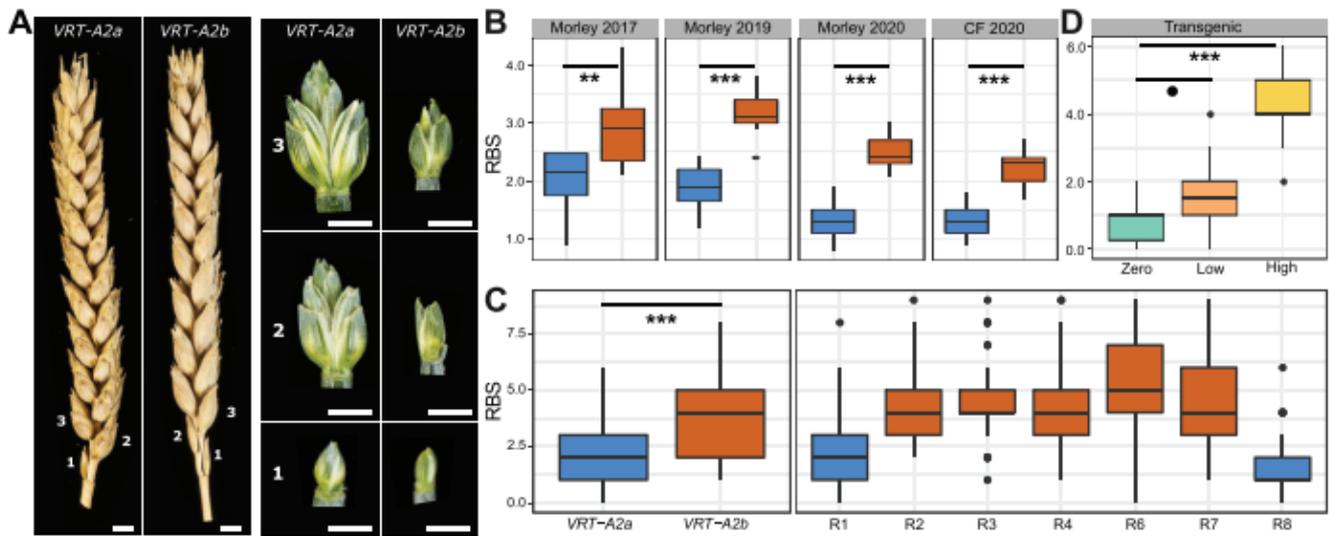
*VRT-A2b* by Adamski et al. (2021). The RBS phenotype was mapped in complete linkage with the 50.3-kbp interval containing *VRT-A2* (Figure 5C; Supplemental Table S8). This genetic and phenotypic data suggests that the increase in RBS is a pleiotropic effect of the *T. polonicum* *VRT-A2b* allele and supports the hypothesis that misexpression of *VRT-A2* negatively affects spikelet development in the base of the spike. In Paragon, the first (sometimes second) RBS fully develops the floral organs of florets one and two (e.g. lemma, palea, stamen, and ovary); however, these are severely reduced in size and delayed in development compared to the florets of central spikelets just before flowering (~Waddington stage 8–10; Supplemental Figure S4). At this stage, the further growth and development of these basal florets is stopped and in the mature spike only the glumes of RBS are visible. In NILs carrying the *VRT-A2b* allele, the development of the most basal spikelet is very similar to the wild-type. However, the second, third, and sometimes fourth spikelet also display similar signs of reduced development, leading to the larger number of RBSs (Figure 5, B and C; Supplemental Figure S4).

To validate the phenotypic effect of *VRT-A2*, we analyzed transgenic wheat lines transformed with the complete genomic *T. polonicum* *VRT-A2b* sequence (including the native promoter and the intron 1 re-arrangement). Transgenic T<sub>1</sub> lines were classified based on the transgene copy number which was previously shown by Adamski et al. (2021) to be highly correlated with *VRT-A2* expression levels in multiple tissues. We phenotyped lines with zero ( $n = 2$  independent events; 5 plants each), low (1–5 transgene copies;  $n = 4$  independent events; 5 plants each) and high (9–35 transgene copies,  $n = 2$  independent events; 5 plants each) transgene copy number. We identified a significant and stepwise increase in the number of RBS with transgenic copy number, from 0.8 RBS (zero copy) to 1.6 RBS (low copy;  $P = 0.078$  versus zero copy) and to 4.3 RBS (high copy;  $P < 0.0001$  versus zero copy) (Figure 5D; Supplemental Table S9). The low copy number lines had an average increase of 0.8 RBS with respect to the zero copy number lines, equivalent to the average difference between the *VRT-A2a* and *VRT-A2b* NILs in the field (*VRT-A2b* effect of +1.1 RBS). The high copy number lines produced on average 4.3 RBS, which is higher than the *VRT-A2b* NILs and similar to the number of RBS observed in *T. polonicum* ( $3.75 \pm 0.62$  RBS;  $n = 16$  spikes). The dosage-dependent effects observed in the transgenic lines provide further evidence that elevated expression of *VRT-A2* leads to increased number of RBSs in polyploid wheat.

## Discussion

### High-resolution spatial transcriptomics in crops

We hypothesized that the establishment of the lanceolate shape in wheat spikes could be manifested in gene expression differences between the apical, central and basal sections of a developing spike, as has been shown using RT-qPCR for individual genes in wheat (AP2; Debernardi, 2017)



**Figure 5** Phenotypic difference between VRT-A2a (blue) and VRT-A2b (orange) on RBS numbers. A, Mature spikes from the field (left) and dissected basal spikelets at anthesis (right) from the glasshouse. Numbers indicate position along the spike starting at the base. Scale bar = 0.5 cm. B, Number of RBS per spike from 10-ear samples collected in the field at maturity at Morley (2017, 2019, and 2020) and Church Farm (CF, 2020). Year 2017/2019 = 10–15 independent replicates per genotype; 2020 = 20 independent replicates per genotype and location. C, Number of RBS recorded in the glasshouse for the NILs (left) and for seven critical recombinant lines (R1–R4, R6–R8; see Supplemental Table S8 for graphical genotype of these lines from Adamski et al. (2021), n = 18–20 plants). D, RBS per spike recorded for the transgenic lines carrying zero (n = 10 plants), low (1–5, n = 20 plants), or high (9–35, n = 10 plants) copy-number insertions of VRT-A2b in cv. Fielder. In (B–D), the box represents the middle 50% of data with the borders of the box representing the 25th and 75th percentile. The horizontal line in the middle of the box represents the median. Whiskers represent the minimum and maximum values, unless a point exceeds 1.5 times the interquartile range in which case the whisker represents this value, and values beyond this are plotted as single points (outliers). Statistical classifications in (B) and (C) are based on two-way ANOVA tests and in (D) on Dunnett's test against the zero-copy number lines. P-values:  $^{\circ} \leq 0.1$ ;  $^{**} \leq 0.01$ ;  $^{***} \leq 0.001$ .

and barley (VRS2; Youssef et al., 2017). However, currently available transcriptome data (e.g. Li et al., 2018 and Feng et al., 2017) lack the spatial resolution within each individual developmental stage to answer this question. This focus on “between stage” comparisons (as opposed to within a single stage) is perhaps related to the technical challenges of dissecting and sectioning young meristems. Given the relatively small size of these spike meristems (0.2-mm length at Transition Stage; 3-mm length at Terminal Spikelet stage), RNA-seq methods require bulking of multiple individuals (usually between 30 and 50 different plants) to accumulate enough tissue for a single RNA-seq sample. If one sought to further section each meristem, this would require even further bulking. While laborious, this is achievable; however, under this scenario, the challenge is to properly stage approximately 100 plants to an equivalent developmental stage. Furthermore, it can be technically challenging to section these young spikes each time into the exact same apical, central, and basal sections. Consequently, the spatial resolution in gene expression within a wheat spike at individual developmental stages has remained largely uncharacterized to date.

To address this challenge, we adapted the G&T method for micro-scale spatial-transcriptomics workflow (Macauley et al., 2015; Giolai et al., 2019), to conduct RNA-seq of the apical, central, and basal sections of individual, hand-

dissected wheat spikes. This highly automated workflow requires low tissue input and allowed us to combine 24 Nextera libraries into a single Illumina NovaSeq lane. For 19 out of the 24 samples, the method worked successfully, determined by more than 20,000 expressed genes per library and the clustering among biological replicates. We found that the number of expressed genes per library was on average similar to the number of genes reported for bulked whole spike RNA-seq samples (Feng et al., 2017; Li et al., 2018). This is consistent with the fact that the hand-dissected sections are composed of a large mixture of different tissues (e.g. rachis, spikelet, and floret primordia) and cell types, which in the equivalent maize ears have distinct expression profiles (Xu et al., 2021). Compared to previous bulk RNA-seq studies in developing wheat spikes, the variation observed here (measured as CV) was high among biological replicates (Supplemental Figure S2). This variation is likely caused by both biological variation (e.g. inherent variation of individual plants) and technical variation (e.g. inaccuracies in sectioning and in the developmental staging of the plant/spike) as well as the number of replicates in our analysis. A minimum of six replicates has been proposed for bulked RNA-seq (Schurch et al., 2016). Our results suggest that the RNA-seq from these small sections would benefit from a higher number of biological replicates, which should be feasible considering the high-throughput method

employed for RNA extraction and library preparation, the low tissue input requirement, and the possibility to pool multiple biological replicates per sequencing lane. Despite some limitations, we could identify over 5,000 DEGs between the spatial sections for subsequent functional analysis.

In addition to G&T-Seq, several other technologies have been proposed for obtaining high-resolution transcriptional profiles of plant tissues, for example, single-cell RNA-seq (McFaline-Figueroa et al., 2020; Rich-Griffin et al., 2020), fluorescence-activated cell sorting (FACS), and the isolation of nuclei tagged in specific cell types. These methods, however, are not spatially resolved as the complete tissue is dissolved into single cells for barcoding or selection (Rich-Griffin et al., 2020). Thus, these current methodologies do not allow, for example, to investigate whether the cell-type composition of spikelets differs across the inflorescence. This would only be possible if spikelets were “harvested” individually, for example through laser capture microdissection (LCM) before dissolving the tissue further into individual cells. Thiel et al. (2021) recently combined LCM followed by RNA-seq of the distinct lower/leaf ridge and upper/spikelet ridge of barley spikes. This allowed them to identify precise spatio-temporal expression patterns of many genes related to architecture and yield in barley spikes with unprecedented resolution. Looking ahead, increased resolution of Spatial Transcriptomics (currently 100  $\mu\text{m}$ ; Giacomello et al., 2017), which quantifies full transcriptomes while maintaining tissue integrity, offers the true prospect of direct localization and quantification of gene expression. Our results argue strongly for the need of these transcriptome-wide and spatially resolved approaches to advance our biological understanding of fundamental developmental processes in plants.

### The composite nature of spikes

Early morphological studies of wheat spike development described that the stronger elongation of central spikelets during their initial establishment (GP stage) first causes the lanceolate shape of the wheat spike (Bonnett, 1966). The continuous formation of primordia at the tip of the spike means that at any given growth stage, spikelets in different developmental stages will be present across the spike (Bonnett, 1966). In this study, we detected more DEGs between the three spatial sections of the spike (apical, central, and basal) than between the two investigated developmental stages (DR and GP). We identified 215 DEGs between the two developmental stages, consistent with Li et al. (2018) who identified 206 DEGs between consecutive stages. Feng et al. (2017) identified 753 DEGs between the DR and Floret Primordia stage, which are further apart in development than the stages used in this study. They also detected fewer DEGs when comparing early stages than between more developed spikes. In contrast, we identified 1,315 and 2,438 unique genes to be differentially expressed between the apical and basal section at DR and GP, respectively. The higher number of DEGs between spatial sections could be due to the developmental gradients occurring in the three

spatial sections, which are revealed by the spatial sampling. These differences would be blurred when comparing whole inflorescences between stages due to the mixture of tissue types and spikelets at different developmental stages. A possible improvement for future transcriptome studies could be the collection of only central sections of the developing spikes or complete spatial sampling as conducted here.

The composite nature of the inflorescence tissues has been acknowledged by studies in maize (ears and tassels), where new meristems are initiated in a stepwise manner. Leiboff and Hake (2019) quantified the meristematic tissue composition of maize and sorghum (*Sorghum bicolor*) tassels. For example, maize tassels in the second stage are mainly composed of spikelet pair meristems, but also contain some meristems in spikelet and inflorescence state. They concluded that the changes in these tissue compositions over time corresponded well with the independently staged transcriptional changes of the tassels. Eveland et al. (2014) showed that the range of developmental ages across the maize ear, if acknowledged, can be used as an advantage in RNA-seq studies. They sequenced the tip, middle, and basal sections of 10-mm long ears independently, aiming to analyze the expression patterns in specific developmental meristem types enriched in these sections (inflorescence, spikelet, and floral meristems, respectively). The dissection of the ear, therefore, allowed them to study gene expression specifically for each meristematic tissue type rather than for all meristem types in intact ears. In this study, we observed that apically expressed genes are enriched for GO terms related to “shoot system development” and “maintenance of floral organ identity.” This is consistent with the hypothesis that the apical part of the inflorescence is younger and undergoing early phases of spikelet development initiation compared to the central inflorescence section.

### Delayed transition of basal spikelets from vegetative to floral developmental programs

We detected transcriptional gradients across the spike, with the basal section deviating most strongly from the rest of the spike. We noticed that both *SVP* and *CENTRORADIALIS* (*CEN*) genes remained highly expressed in the basal section of the spike, whereas their expression was lower in the central and apical sections. In situ hybridization of these genes also showed that their expression is strongest in vegetative tissue and basal spikelets in early spike development (Li et al., 2021). In contrast, *SEP1-4* and *SEP1-5* genes were expressed in the opposite gradient and showed the strongest expression in apical and central sections of the spike at GP stage. Recent studies allow us to interpret these gradients in the context of the early steps of vegetative to floral growth transition. In wheat (Adamski et al., 2021; Li et al., 2021; Liu et al., 2021), rice (Sentoku et al., 2005; Lee et al., 2008), and barley (Trevaskis et al., 2007), *SVPs* have been characterized to be associated with vegetative growth and are downregulated upon floral transition.

In wheat, the double *SVP* mutant *vrt2svp1* leads to the formation of axillary inflorescences (Li et al., 2021). Similarly, overexpression of *CEN-D2* (*TERMINAL FLOWER 1*) in wheat extends the duration of the DR stage (Wang et al., 2017), whereas loss-of-function mutations in barley *CEN* suggest they repress floral development under short-day conditions (Bi et al., 2019). Double knockout mutants of the MADS-box *SQUAMOSA* genes *vrn1ful2* highlighted that these two genes act as transcriptional repressors of *SVP* and *CEN* genes in early wheat spike development (Li et al., 2019). Furthermore, through a series of genetic and biochemical studies, Li et al. (2021) showed that the downregulation of *SVP* genes is necessary for the formation of flowering promoting MADS-box protein complexes including *VRN1*, *FUL2*, and *SEP* proteins. Hence the coordinated downregulation of *SVPs*, and possibly *CEN* genes, along with the upregulation of *SEP* genes is required for normal floral transition and spikelet development in wheat. Previous studies in rice have found similar expression patterns, as well as mutant effects, of *SVPs* and *SEPs* suggesting a conserved function in flowering transition across the two species (Ren et al., 2016; Wu et al., 2018).

Based on our results, the floral developmental program across the wheat spike appears to be most advanced in its apical and central sections, while being delayed in the basal sections. We hypothesize that this is due to elevated *VRT2* expression at the base of the spike, which hinders the progression of the flowering program via *SEP* class flowering genes. Likewise, the higher expression levels of the wheat *CEN2* and *CEN5* homologs at the base are consistent with a delay in floral transition that could interfere with the development of the spikelet primordia. Therefore, although the basal spikelet primordia are initiated first chronologically, their developmental age in terms of the floral program is delayed with respect to the more recently formed central and apical spikelet primordia. This could explain in part why the spikelet primordia in the basal region of the spike elongate less and develop slower than central spikelets despite being initiated first (Bonnett, 1966). Likewise, the less advanced floral developmental program could also explain why the overexpression of *SVPs* in barley (*HvBM10*) leads to complete floral reversion in basal but not apical spikelets (Trevaskis et al., 2007).

We hypothesize that *SVPs* need to be downregulated upon floral transition to allow timely establishment and progression of the early spikelet primordia. Failure to do so would delay their development and result in their final rudimentary shape in the mature spike. In line with this hypothesis, we observed increased RBS in genotypes with prolonged and increased *VRT2* expression in a dosage-dependent manner. In our RT-qPCR data, we also observe increased expression of *SVPs* alongside reduced expression of *SEPs* in *VRT-A2b* lines at DR and GP stage. However, we cannot exclude the possibility that the increase in *VRT2* expression in the *VRT-A2b* lines could also affect basal spikelet development at a later stage of spike developmental (e.g.

from Terminal Spikelet stage to anthesis). We are currently using quantitative live imaging to compare cellular growth dynamics of spikelets at different stages of spike development between the *VRT2* NILs.

The finding that the expected downregulation of *SVPs* and *CENs* does not follow the chronological age of the tissues suggests that other gradients across the spike might influence spikelet development. Debernardi et al. (2017) showed that in tetraploid wheat *AP2-5* and *miR172* have consistent and opposing expression gradients across the spike at three consecutive developmental stages. The persistent expression gradient of *AP2-5* supports the idea that expression patterns across the spike, beyond the ones caused by age differences of spikelets, exist. Furthermore, they proposed a model illustrating that the phenotypic effect of mutants across the spike differs due to the existing gradient of expression of this gene (Debernardi et al., 2017). Other examples of mutants with different phenotypic effects across an inflorescence include *vrn1ful2* (Li et al., 2019) in wheat, *tassel sheath1* (*tsh1*; Whipple et al., 2010), and *ramosa2* (Bortiri et al., 2006) in maize, *SEP* double mutant *Osmads5Osmads34* in rice (Zhu et al., 2022), as well as *many noded dwarf1* (*mnd1*; Walla et al., 2020), *frizzy panicle* (Poursarebani et al., 2015), and *vrs2* in barley, which was also found to be consistently differentially expressed across the spike (Youssef et al., 2017). *VRS2* has been shown to maintain a basal to apical expression pattern across three, post awn initiation developmental time points in barley (Youssef et al., 2017). The study of *vrs2* mutants revealed that *VRS2* is furthermore engaged with the basal-apical patterns of auxin, cytokinin, and gibberellin across the spike. While hormonal gradients across the spike in early development have not been studied in great detail in wheat, they have been shown to play crucial roles in floral induction and development in *Arabidopsis* (Reinhardt et al., 2000). Their patterns across the spike should be investigated in future studies addressing developmental differences across the spike.

### A model for the regulation of leaf and spikelet ridge outgrowth in the base of the spike

Recently, Meir et al. (2021) proposed that in shoot apical meristems of tomato (*Solanum lycopersicum*), similar to processes during embryonic development, transient programs are required to inhibit a preceding setup (i.e. vegetative growth), before a new developmental program (flowering) can be initiated. We propose that the altered gene expression and development of the basal spikelets could be a consequence of their initiation during the transient phase between vegetative and floral network shifts and thus being exposed to mixed signals of development. Upon floral transition, the lower (leaf) ridge is suppressed, while the growth of spikelet ridges from the previously suppressed axillary meristems is activated. Development of lower ridges subtending all branching events is suppressed in grass inflorescences upon flowering transition (Whipple et al., 2010). Li

et al. (2019) noticed that this suppression was disrupted in the double *vrn1ful2* and triple *vrn1ful2ful3*-null mutants, which fail to downregulate *SVP* genes. In these mutants, the upper spikelet meristems generate vegetative structures resembling tillers that are subtended by bracts or leaves originating from the lower leaf ridge.

We observed that genes that were highly expressed in the basal section of the inflorescence (Cluster 4) have previously been shown to be expressed specifically in the lower/bract ridge and before or at vegetative to floral transition. This is also supported by the GO term enrichment of photosynthesis related terms in Cluster 4. Our tissue sections do not allow us to distinguish lower and upper ridge tissues; however, the two ridges have been separately collected and sequenced via LCM in barley (Thiel et al., 2021). In this barley dataset, we found a higher expression of *HvVRT2* (HORVU7Hr1G036130) and *FLOWERING LOCUS C* (*HvFLC*; HORVU7Hr1G054320) in the lower ridge compared to the upper ridge, whereas *HvSVP1* (HORVU6Hr1G077300) was also marginally more highly expressed (Supplemental Figure S5). Furthermore, the barley *MND1* gene (HORVU7Hr1G113480) has recently been shown to be expressed in leaf primordia and during the DR stage in the basal region of the spike in barley (Walla et al., 2020), while it is most highly expressed in the vegetative meristem and lower/leaf ridge in the LCM data (Thiel et al., 2021). We observed that in our data, the wheat *MND1* putative orthologs (*TraesCS7A02G506400*, *TraesCS7B02G413900*, and *TraesCS7D02G494500*) were significantly more highly expressed in the basal section than the apical section at both DR and GP stage (Supplemental Table S2). The suppressed leaf ridge (or bract) has been proposed to act as a signaling center, regulating the fate of the upper spikelet meristem ridge (Whipple, 2017). Insufficient bract suppression during the formation of the basal spikelets might therefore negatively affect initiation and development of spikelets.

At DR, the widest point of the spike is indeed as expected the base and not the central section (Figure 1A). The lower ridge is, however, much less developed in the central section and can be hardly seen in the apical ridges. Interestingly, mutants failing to repress the lower ridge growth, such as *third outer glume1* (*trd1*; Houston et al., 2012), the barley ortholog of maize *tsh1*, develop large bracts from the lower ridge in basal spikelets, unlike apical spikelets, which do not develop bracts from their lower ridges regardless of the absence of *TRD*. This is reminiscent of the gradient in the strength of the phenotypic effects observed from the top to the base of the inflorescence in multiple *Poaceae* mutants (discussed above). We, therefore, hypothesize that the basal meristems develop into smaller spikelets and larger bract primordia due to a slow suppression of “vegetative growth signals” (e.g. *SVPs*) and a concomitant slow upregulation of “floral growth signals” (e.g. *SEPs*) upon floral transition. To investigate how a change from vegetative to floral signaling might affect the development of individual meristems, we modeled the genetic interaction of *SVPs* and *SEPs*, as proposed by Li et al. (2021), in the spatial context of a growing

spike (Supplemental File S1). Under the assumptions that *SVP* suppresses *SEP* expression, *SVP* expression is downregulated upon flowering, and that *SEP* promotes spikelet outgrowth, the model could recapitulate (1) the observed opposing gradients in expression of *SVPs* and *SEPs* along the spike and (2) the formation of a lanceolate shaped wheat spike with reduced spikelet elongation and stronger bract growth in the most basal spikelets. Thus, whilst this hypothesis will require further investigation and testing, modeling supports its plausibility.

## Materials and methods

### Plant materials

Hexaploid wheat (*T. aestivum*) germplasm used in this study includes wild-type hexaploid wheat cultivar Paragon and *P1/VRT2* germplasm described in Adamski et al. (2021) including *P1* NILs, recombinants, and *T<sub>1</sub>* transgenic lines carrying the *T. polonicum* *VRT-A2b* copy under the native promoter. *Triticum polonicum* accession T1100002 was obtained from the John Innes Center Germplasm Resources Unit (<https://www.seedstor.ac.uk/search-infoaccession.php?idPlant=27422>). For field experiments, we used between two to four sibling BC<sub>4</sub>/BC<sub>6</sub> NILs differing for the *VRT-A2b* allele.

### Low-input RNA sequencing

Paragon seedlings were grown in a single batch in a controlled environment growth chambers in 24-cell seed trays under long-day (16-h light/8-h dark) photoperiods at 300  $\mu\text{mol m}^{-2} \text{s}^{-1}$ , with a day temperature of 20°C and a night temperature of 15°C. Inflorescences for DR stage were collected 18 d after sowing, while inflorescences for GP stage were collected 22 d after sowing. All plants were grown in “John Innes Cereal Mix” (40% Medium Grade Peat, 40% Sterilized Soil, 20% Horticultural Grit, 1.3 kg·m<sup>-3</sup> PG Mix 14-16-18 + Te Base Fertilizer, 1 kg·m<sup>-3</sup> Osmocote Mini 16-8-11 2 mg + Te 0.02% B, Wetting Agent, 3 kg·m<sup>-3</sup> Maglime, 300 g·m<sup>-3</sup> Exemptor).

Four individual spikes per developmental stage (DR and GP) were dissected into apical, central, and basal sections (1:1:1 ratio) using a stereo microscope (Leica MZ16). Sections were immediately placed into 96-well plates (on ice) containing 10  $\mu\text{L}$  of RLT plus (Qiagen, Hilden, Germany). All instruments and surfaces were cleaned with 80% v/v ethanol, RNase-free water and lastly RNase-out solution after each sample to reduce cross-contamination and RNA degradation. Samples were stored at -80°C until cDNA preparation, using the G&T-seq method as previously described (Macaulay et al., 2015). cDNA was normalized to 0.2 ng/ $\mu\text{L}$  before Nextera (Illumina, San Diego, CA, USA) library preparation using a Mosquito HV liquid handler (STP, Royston, UK) in a total reaction volume of 4  $\mu\text{L}$  as described in Mora-Castilla et al. (2016). Libraries were pooled by volume and sequenced on a single lane of a NovaSeq 6000 (NVS200S2 flow cell, 100-bp paired-end reads).

### Bioinformatic analysis

For the RNA-seq analysis, we used the RefSeq version 1.0 genome assembly and the RefSeq version 1.1 gene annotation ([https://urgi.versailles.inra.fr/download/iwgs/IWGSC\\_RefSeq\\_Assemblies/v1.0/](https://urgi.versailles.inra.fr/download/iwgs/IWGSC_RefSeq_Assemblies/v1.0/); IWGSC et al., 2018). Reads were trimmed and adapters were removed using trim-galore version 0.4.2 ([https://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)) with settings: “-paired -fastqc -a GGTATC AACGCAGAGT -clip\_R1 20 -clip\_R2 20 -trim-n.” Minimum length of reads retained was set to 50 bp. Reads were aligned to the RefSeq version 1.0 genome assembly using HISAT2 version 2.1.0 (<https://daehwankimlab.github.io/hisat2/>; Kim et al., 2019) with the following parameters: “-pen-noncansplice 20 -mp 1.0 -rna-strandness RF.” Alignment files were converted to BAM format, sorted, indexed, filtered, and purged of all non-primary alignments (0 × 100 flag) using samtools (version 1.9; Li et al., 2009). HTSeq version 0.6.1 (<https://htseq.readthedocs.io/en/master/>; Anders et al., 2015) was used to count the read numbers mapped to the RefSeq version 1.0 gene models.

HT-read count normalization and differential expression analyses were performed using the DESeq2 version 1.28.1 R packages (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>; Love et al., 2014; RStudio 1.2.5001). Genes with an average expression below 10 HT-count, and which were not expressed (i.e.  $\leq 10$  HT counts) in at least three libraries, were removed from the analysis. Correlation between expressed genes and Waddington stage and/or section was tested by ANOVA. Raw read data from Li et al. (2018) and Feng et al. (2017) were pseudo-aligned using Kallisto Sleuth pipeline (<https://scilifelab.github.io/courses/rnaseq/labs/kallisto>) and the CV was calculated for each gene (by condition) using R (RStudio 1.2.5001) ddply (plyr 1.8.6). DEGs between the two Waddington stages (DR and GP) were calculated with the design “~plant + section,” while DEGs between the three sections (apical, central, and basal) were determined using the design “~section + waddington:plant + waddington.” DEGs among the sections within each Waddington stage were determined with the design “~plant + section.” For each gene, an adjusted *P*-value was computed by DESeq2 (using the using Benjamini and Hochberg method; Benjamini and Hochberg, 1995), and those with an adjusted *P*-value of  $\leq 0.05$  were considered differentially expressed. DESeq2 also computed Log2FoldChanges as well as the associated uncertainty (Log2FoldChanges (lfcSE), see Love et al. (2014) for further detail). The “contrast” function was used to determine pairwise comparison *P*-values. The full set of expression data and comparisons is presented in Supplemental Dataset S1. Enrichment of GO terms was performed using the online tool “PLAZA” (<https://bioinformatics.psb.ugent.be/plaza>; Van Bel et al., 2017) using the recommended settings, and all enriched GO terms of Biological function and Cellular Compartment were retained. In brief, PLAZA determines the overrepresentation of a certain GO term in a gene set compared to the genome-wide background frequency (= all expressed genes in this experiment; submitted manually).

The significance of over- or underrepresentation is determined using the hypergeometric distribution and the Bonferroni method is applied to correct for multiple testing. Note that enrichment folds are reported in log<sub>2</sub> fold scale. Enrichment of TF families (Genes that were annotated as TFs were obtained from [https://opendata.earlham.ac.uk/wheat/under\\_license/toronto/Ramirez-Gonzalez\\_et\\_al\\_2018-06025-Transcriptome-Landscape/data/data\\_tables/](https://opendata.earlham.ac.uk/wheat/under_license/toronto/Ramirez-Gonzalez_et_al_2018-06025-Transcriptome-Landscape/data/data_tables/); Ramirez-Gonzalez et al., 2018) and MADS-box TFs (based on Schilling et al., 2020) was performed in R using the phyper() function from stats package version 4.0.1 to test for Hypergeometric Distribution. All DEGs were scaled and centered using R-base function “scale.” All cluster analysis was performed on scaled data using R (stats) functions kmeans and hclust, followed by visualization through pheatmap version 1.0.12 (<https://cran.rstudio.com/web/packages/pheatmap/index.html>). Correlation to centroid cluster shape of each gene expression pattern was calculated using the “cor” function from R stats.

### RT-qPCR analysis

P1 NILs were grown in controlled growth chambers in 24-cell seed trays under the same conditions as used in the low input RNA-seq experiment (see above). For each biological replicate, we pooled 30 inflorescences for DR stage, 15 for GP stage, and 9 for Terminal Spikelet stage ( $n = 4$  biological replicates per stage). Inflorescences from NILs were dissected using a stereo microscope (Leica MZ16). Inflorescences were dissected into apical, central, basal, and peduncle sections (1:1:1:1 ratio). At DR stage, inflorescences were only dissected into apical, basal, and peduncle section as the inflorescences were too small to be accurately dissected into four sections for all 30 plants per biological replicate. Each section was immediately placed into 1.5-mL tubes on dry ice and tubes were snap frozen in liquid nitrogen as soon as all plants for the sample were collected. Samples were stored at  $-80^{\circ}\text{C}$  until needed. Inflorescences were collected within 2–3 h, 9 h after the lights came on in the growth chamber. Tissue was homogenized in a TissueLyser II (Cat No.: 85300; QIAGEN, Hilden, Germany) using 3-mm steel beads (Cat No.: 69997; Qiagen); tubes were shaken for 20-s at 28 Hz with dry ice.

All RNA extractions were performed using the RNeasy Plant Mini Kit (Cat No.: 74904; Qiagen) with RLT buffer according to the manufacturer’s protocol followed by RNA ethanol precipitation ([https://projects.iq.harvard.edu/files/hla\\_lab/files/ethanol-precipitation-of-rna\\_hla.pdf](https://projects.iq.harvard.edu/files/hla_lab/files/ethanol-precipitation-of-rna_hla.pdf)). DNA digestion was performed using the RQ1 RNase-free DNase set (Cat No.: M6101; Promega, Madison, WI, USA) according to the manufacturer’s protocol. RNA was reverse transcribed using M-MLV reverse transcriptase (Cat No.: 28025013; ThermoFisher, Waltham, MA, USA) according to the manufacturer’s protocol. For the RT-qPCR reactions, LightCycler 480 SYBR Green I Master Mix (Roche Applied Science, UK) was used according to the manufacturer’s protocol. The reactions were run in a LightCycler 480 instrument (Roche Applied Science, UK) under the following conditions: 10 min

at 95°C; 40 cycles of 10 s at 95°C, 15 s at 62°C, 30 s at 72°C; dissociation curve from 60°C to 95°C to confirm primer specificity. All reactions were performed with three technical replicates per sample and using *Actin* as the reference gene (Uauy et al., 2006). Relative gene expression was calculated using the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen, 2001) with a common calibrator so that values are comparable across genes, tissues, and developmental stages. All primers used in RT-qPCR came from Adamski et al. (2021) can be found in Supplemental Table S10.

All RT-qPCR data were normalized using a log2 transformation. A three-way ANOVA including Waddington stage, section, and genotype yielded significant two-way interactions. The differences between sections of the genotypes were therefore further analyzed individually for each Waddington stage and genotype. For each of the two genotypes we individually performed Tukey's multiple comparison tests to determine differences between the sections within each developmental stage by Tukey's multiple comparison test. Differences between the two genotypes were also analyzed individually for each Waddington stage. Furthermore, the differences between the genotypes were investigated individually for each section within the Waddington stage if the interaction term was significant (in GP and TS). For all analysis see Supplemental Table S6.

### Field experiments and phenotyping

VRT-A2 NILs were evaluated in four field experiments. Three trials were located at The Morley Agricultural Foundation trials site, Morley St Botolph, UK (52°33'15.1"N 1°01'59.2"E) in 2017, 2018 and 2020 and one trial was sown in 2020 at the John Innes Experimental trials site in Norwich, UK (52°37'50.7"N 1°10'39.7"E). In Morley (2017), we analyzed two BC<sub>4</sub> lines of VRT-A2a and three BC<sub>4</sub> lines of VRT-A2b. In Morley (2019), we analyzed two BC<sub>6</sub> and one BC<sub>4</sub> line per VRT-A2 allele and in Morley and Church Farm (CF) 2020, we analyzed two BC<sub>6</sub> and two BC<sub>4</sub> lines for each VRT-A2 allele. All experiments were drilled as yield-scale plots (6 × 1.2 m) and sown by grain number for comparable plant densities aiming for 275 seeds m<sup>-2</sup>. The trials were arranged in a randomized complete block design with five replicates per sibling line per location. A 10-ear grab sample was collected from each plot preharvest for the assessment of RBS numbers and other phenotypes (recorded in Adamski et al., 2021). RBS was defined as spikelets carrying no grain at maturity and counted for each spike individually. To determine the differences between the  $P1^{POL}$  and  $P1^{WT}$  NILs, we performed ANOVA on the multiple field trials phenotypic data. For the analysis of individual trials, we used a two-way ANOVA including Genotype + Block performed in R ("car" package version 3.0-10; RStudio 1.2.5001).

### Glasshouse phenotyping

We evaluated the BC<sub>4</sub> NILs and BC<sub>6</sub>F<sub>3</sub> recombinant lines, as well as *T. polonicum* accession T1100002, under standard glasshouse conditions. About 18–20 plants per genotype were grown in 1 L pots containing John Innes Cereal Mix

under long-day conditions (16-h light, 8-h dark). The genotypes of all plants were confirmed using KASP marker *SP1Pol* (Adamski et al., 2021). We counted the number of RBSs for all tillers of all biological replicates at maturity. To evaluate the differences in RBS between genotypes, we performed a two-way ANOVA analysis and post-hoc multi-pairwise comparisons Sidak test ("car" package version 3.0-10; RStudio 1.2.5001).

### Phenotyping of transgenic lines

T<sub>1</sub> lines from Adamski et al. (2021) differing for the copy number of the VRT-A2b transgenic construct (zero = 0 copies; low = 1–5 copies; high = 9–35 copies) were grown in 1 L pots with John Innes Cereal Mix under 16-h light at 20°C and 8-h dark at 15°C in controlled environment growth chambers. We measured RBS number for the main tiller of all plants at maturity. To determine differences in RBS between the three transgenic classes, we performed ANOVA (two-way ANOVA; "car" package version 3.0-10). We performed Dunnett tests to compare the low and high copy lines against the zero copy number controls (RStudio 1.2.5001).

### Modeling

The computational model of wheat spike shape formation was developed using the multi-agent programming language and modeling environment, Netlogo (Wilensky, 1999). Gene interactions were modeled as previously described (Li et al., 2021). The model can be accessed via the interactive web-version of the model (Supplemental File S1).

In brief, both spikelets and leaves are initiated with rates that depend on the levels of SEP. Leaf initiation rates are suppressed by SEP, whereas spikelet initiation requires SEP. The maximum initiation rates are the same for both spikelets and leaves but different before ( $r_{vegetative}$ ) and after ( $r_{flowering}$ ) flowering. Once initiated, the leaves and spikelets grow at a rate defined by the parameters  $r_{leaf}$  and  $r_{spikelet}$ , respectively. Leaf growth does not depend on SVP or SEP levels, whereas spikelets only increase in size every iteration if their SEP level is above a given threshold ( $SEP_{growth\_threshold}$ ). Expression of both SVP and SEP only occurs at meristem initiation. After this, the levels of SVP and SEP cannot increase, although SEP is degraded. SVP is not degraded, solely because at this point, nothing is dependent on SVP levels, whilst spikelet growth depends upon SEP levels.

SVP expression rates start to decrease, once flowering is triggered, according to:

$$r_{SVP}(t+1) = r_{SVP}(t)f_{red}$$

where  $r_{SVP}(t)$  is the rate of SVP expression at that time step, and  $f_{red}$  is a rate reduction factor.

SEP expression depends upon the levels of SVP in the meristem in which the initiation points are located, depending on a Hill function (Alon, 2007),

$$r_{\text{SEP}}(\text{SVP}) = r_{\text{SEP,max}} \left( \frac{K_D^n}{K_D^n + \text{SVP}^n} \right),$$

where  $r_{\text{SEP,max}}$  is the maximum rate of SEP expression,  $K_D$  is the binding constant, and  $n$  is the Hill coefficient. The resulting curves for SVP and SEP expression are shown in Supplemental Figure S6.

SVP levels are initiated with the current value of  $r_{\text{SVP}}$ . SEP levels are initiated using  $r_{\text{SEP}}$ , and reduce by degradation rate,  $\delta_{\text{SEP}}$ , following

$$\frac{d\text{SEP}}{dt} = \text{SEP} \cdot \delta_{\text{SEP}}.$$

### Data availability

The raw RNA-seq read libraries used in this study are available from NCBI BioProject PRJNA749586.

### Accession numbers

Accession numbers for the sequence data are listed throughout the study and in Supplemental Table S2, and are based on gene nomenclature from the IWGSC RefSeq version 1.0 annotation.

### Supplemental data

The following materials are available in the online version of this article.

**Supplemental Figure S1.** PCA of RNA-seq libraries.

**Supplemental Figure S2.** Comparison of CV of gene expression in wheat RNA-seq data sets.

**Supplemental Figure S3.** Expression patterns of all DEG clusters ( $n = 7$ ) as identified by k-means clustering.

**Supplemental Figure S4.** Dissected Floret 1 and 2 of basal and central spikelets of BC<sub>6</sub> NILs before anthesis.

**Supplemental Figure S5.** Expression of barley genes in Thiel et al. (2021), which are putative orthologs to wheat genes highly expressed in basal spike sections.

**Supplemental Figure S6.** Example of Netlogo simulation outcome with default parameters.

**Supplemental Table S1.** Quality control measurements of all 24 RNA-seq libraries.

**Supplemental Table S2.** Summary of normalized gene expression, statistical analyses, and k-means clustering for all DEGs ( $P_{\text{adj}} \leq 0.05$ ).

**Supplemental Table S3.** Enrichment of GO terms in the seven identified clusters of DEGs.

**Supplemental Table S4.** Enrichment of TF families and MADS-box TF genes in the seven identified clusters of DEGs.

**Supplemental Table S5.** Relative expression of VRT-A2 and SEP1-4 measured in Paragon NILs with either the wild-type (VRT-A2a) or *T. polonicum* allele (VRT-A2b).

**Supplemental Table S6.** Statistical analysis of RT-qPCR data from VRT-A2 NIL spike sections (Supplemental Table S5).

**Supplemental Table S7.** Field evaluations for RBSs in VRT-A2 NILs.

**Supplemental Table S8.** Graphical genotype from Adamski et al. (2021) and RBS phenotype of BC<sub>6</sub> recombinant inbred lines for VRT-A2.

**Supplemental Table S9.** RBS phenotypic data from VRT-A2 transgenic lines.

**Supplemental Table S10.** List of primers used in RT-qPCR in this study.

**Supplemental Dataset S1.** Summary of normalized gene expression, statistical analyses, and k-means clustering for all expressed genes.

**Supplemental File S1.** Interactive web-version of the wheat spike model.

### Acknowledgments

We thank Phil Robinson from Scientific Photography at John Innes Centre (JIC) for taking high-quality images of spikes and spikelets and Tobin Florio (<http://flobzbox-science.com/>) for scientific illustrations for Figure 1. We thank the JIC Field Experimentation and Horticultural Services teams for technical support in field and glasshouse experiments and the JIC Bioimaging facility and staff for their services and technical support.

### Funding

This work was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) through the grant BB/S016945/1, the Designing Future Wheat (BB/P016855/1) and Genes in the Environment (BB/P013511/1) Institute Strategic Programs and by the John Innes Foundation. Additional funding was provided by the European Research Council (866328).

**Conflict of interest statement.** The authors declare no competing interest.

### References

- Adamski NM, Simmonds J, Brinton JF, Backhaus AE, Chen Y, Smedley M, Hayta S, Florio T, Crane P, Scott P, et al. (2021) 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
- Alon U (2007) An Introduction to Systems Biology: Design Principles of Biological Circuits. Chapman & Hall/CRC, London
- Anders S, Pyl PT, Huber W (2015) HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169
- Becker A, Theißen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol* **29**: 464–489
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Method* **57**: 289–300
- Bi X, van Esse W, Mulki MA, Kirschner G, Zhong J, Simon R, von Korff M (2019) CENTRORADIALIS interacts with FLOWERING LOCUS T-like genes to control floret development and grain number. *Plant Physiol* **180**: 1013–1030
- Bommert P, Whipple C (2018) Grass inflorescence architecture and meristem determinacy. *Semin Cell Dev Biol* **79**: 37–47

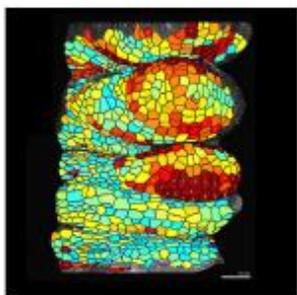
- Bonnett OT (1966) Inflorescences of Maize, Wheat, Rye, Barley, and Oats: Their Initiation and Development. Bulletin (University of Illinois (Urbana-Champaign campus) Agricultural Experiment Station), Champaign, IL, p 721
- Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S (2006) *ramosa2* Encodes a LATERAL ORGAN BOUNDARY domain protein that determines the fate of stem cells in branch meristems of maize. *Plant Cell* **18**: 574–585
- Debernardi JM, Lin H, Chuck G, Faris JD, Dubcovsky J (2017) microRNA172 plays a crucial role in wheat spike morphogenesis and grain threshability. *Development* **144**: 1966–1975
- Eveland AL, Goldshmidt A, Pautler M, Morohashi K, Liseron-Monfils C, Lewis MW, Kumari S, Hiraga S, Yang F, Unger-Wallace E (2014) Regulatory modules controlling maize inflorescence architecture. *Genome Res* **24**: 431–443
- Feng N, Song G, Guan J, Chen K, Jia M, Huang D, Wu J, Zhang L, Kong X, Geng S, et al. (2017) Transcriptome profiling of wheat inflorescence development from spikelet initiation to floral patterning identified stage-specific regulatory genes. *Plant Physiol* **174**: 1779–1794
- Giacomello S, Salmén F, Terebieniec BK, Vickovic S, Navarro JF, Alexeyenko A, Reimegård J, McKee LS, Mannapperuma C, Bulone V (2017) Spatially resolved transcriptome profiling in model plant species. *Nat Plants* **3**: 17061
- Giolai M, Verweij W, Lister A, Heavens D, Macaulay I, Clark MD (2019) Spatially resolved transcriptomics reveals plant host responses to pathogens. *Plant Methods* **15**: 114
- Harrop TWR., Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, Kater MM (2016) Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *Plant J* **86**: 75–88
- Houston K, Druka A, Bonar N, Macaulay M, Lundqvist U, Franckowiak J, Morgante M, Stein N, Waugh R (2012) Analysis of the barley bract suppression gene *Trd1*. *Theor Appl Genet* **125**: 33–45
- IWGSC, Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, Pozniak CJ, Choulet F, et al. (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**: eear7191
- Kellogg E, Camara P, Rudall P, Ladd P, Malcomber S, Whipple C, Doust A (2013) Early inflorescence development in the grasses (Poaceae). *Front Plant Sci* **4**: 250
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**: 907–915
- Koppolu R, Schnurbusch T (2019) Developmental pathways for shaping spike inflorescence architecture in barley and wheat. *J Integr Plant Biol* **61**: 278–295
- Lee S, Choi SC, An G (2008) Rice SVP-group MADS-box proteins, OsMADS22 and OsMADS55, are negative regulators of brassinosteroid responses. *Plant J* **54**: 93–105
- Leiboff S, Hake S (2019) Reconstructing the transcriptional ontogeny of maize and sorghum supports an inverse hourglass model of inflorescence development. *Curr Biol* **29**: 3410–3419. e3413
- Li C, Lin H, Chen A, Lau M, Jernstedt J, Dubcovsky J (2019) Wheat *VRN1*, *FUL2* and *FUL3* play critical and redundant roles in spikelet development and spike determinacy. *Development* **146**: dev175398
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPP (2009) The sequence alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079
- Li K, Debernardi JM, Li C, Lin H, Zhang C, Jernstedt J, von Korff M, Zhong J, Dubcovsky J (2021) Interactions between SQUAMOSA and SHORT VEGETATIVE PHASE MADS-box proteins regulate meristem transitions during wheat spike development. *Plant Cell* **33**: 3621–3644
- Li Y, Fu X, Zhao M, Zhang W, Li B, An D, Li J, Zhang A, Liu R, Liu X (2018) A genome-wide view of transcriptome dynamics during early spike development in bread wheat. *Sci Rep* **8**: 15338
- Liu J, Chen Z, Wang Z, Zhang Z, Xie X, Wang Z, Chai L, Song L, Cheng X, Feng M, et al. (2021) Ectopic expression of *VRT-A2* underlies the origin of *Triticum polonicum* and *Triticum petropavlovskyi* with long outer glumes and grains. *Mol Plant* **14**: 1472–1488
- Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta CT$  Method. *Methods* **25**: 402–408
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550
- Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, Goolam M, Saurat N, Coupland P, Shirley LM, et al. (2015) G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* **12**: 519–522
- McFaline-Figueroa JL, Trapnell C, Cuperus JT (2020) The promise of single-cell genomics in plants. *Curr Opin Plant Biol* **54**: 114–121
- Meir Z, Aviezer I, Chongloi GL, Ben-Kiki O, Bronstein R, Mukamel Z, Keren-Shaul H, Jaitin D, Tal L, Shalev-Schlosser G, et al. (2021) Dissection of floral transition by single-meristem transcriptomes at high temporal resolution. *Nat Plants* **7**: 800–813
- Mora-Castilla S, To C, Vaezeslami S, Morey R, Srinivasan S, Dumdie JN, Cook-Andersen H, Jenkins J, Laurent LC (2016) Miniaturization technologies for efficient single-cell library preparation for next-generation sequencing. *J Lab Automat* **21**: 557–567
- Poursarebani N, Seidensticker T, Koppolu R, Trautewig C, Gawronski P, Bini F, Govind G, Rutten T, Sakuma S, Tagiri A, et al. (2015) The genetic basis of composite spike form in barley and ‘miracle-wheat’. *Genetics* **201**: 155–165
- Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E (2007) Evolution and development of inflorescence architectures. *Science* **316**: 1452–1456
- Ramirez-Gonzalez RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, van EF, Pasha A, et al. (2018) The transcriptional landscape of polyploid wheat. *Science* **361**: eaar6089
- Reinhardt D, Mandel T, Kuhlemeier C (2000) Auxin regulates the initiation and radial position of plant lateral organs. *Plant Cell* **12**: 507–518
- Ren D, Rao Y, Leng Y, Li Z, Xu Q, Wu L, Qiu Z, Xue D, Zeng D, Hu J (2016) Regulatory role of OsMADS34 in the determination of glumes fate, grain yield, and quality in rice. *Front Plant Sci* **7**: 1853
- Rich-Griffin C, Stechemesser A, Finch J, Lucas E, Ott S, Schäfer P (2020) Single-cell transcriptomics: a high-resolution avenue for plant functional genomics. *Trend Plant Sci* **25**: 186–197
- Schilling S, Kennedy A, Pan S, Jermini LS, Melzer R (2020) Genome-wide analysis of MIKC-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neo-functionalization. *New Phytologist* **225**: 511–529
- Schurh NJ, Schofield P, Gierliński M, Cole C, Sherstnev A, Singh V, Wrobel N, Gharbi K, Simpson GG, Owen-Hughes T (2016) How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA* **22**: 839–851
- Sentoku N, Kato H, Kitano H, Imai R (2005) 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 Genom* **273**: 1–9
- Thiel J, Koppolu R, Trautewig C, Hertig C, Kale S, Erbe S, Mascher M, Himmelbach A, Rutten T, Esteban E (2021) Transcriptional landscapes of floral meristems in barley. *Sci Adv* **7**: eabf0832
- Trevaskis B, Tadege M, Hemming MN, Peacock WJ, Dennis ES, Sheldon C (2007) *Short vegetative phase-like* MADS-box genes inhibit floral meristem identity in barley. *Plant Physiol* **143**: 225–235

- Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J** (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* **314**: 1298–1301
- Van Bel M, Diels T, Vancaester E, Kreft L, Botzki A, Van de Peer Y, Coppens F, Vandepoel K** (2017) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res* **46**: D1190–D1196
- Waddington S, Cartwright P, Wall P** (1983) A quantitative scale of spike initial and pistil development in barley and wheat. *Ann Bot* **51**: 119–130
- Walla A, Wilma van Esse G, Kirschner GK, Guo G, Brünje A, Finkemeier I, Simon R, von Korff M** (2020) An acyl-CoA *N*-Acyltransferase regulates meristem phase change and plant architecture in barley1. *Plant Physiol* **183**: 1088–1109
- Wang Y, Yu H, Tian C, Sajjad M, Gao C, Tong Y, Wang X, Jiao Y** (2017) Transcriptome association identifies regulators of wheat spike architecture. *Plant Physiol* **175**: 746–757
- Whipple CJ** (2017) Grass inflorescence architecture and evolution: the origin of novel signaling centers. *New Phytologist* **216**: 367–372
- Whipple CJ, Hall DH, DeBlasio S, Taguchi-Shiobara F, Schmidt RJ, Jackson DP** (2010) A conserved mechanism of bract suppression in the grass family. *Plant Cell* **22**: 565–578
- Wilensky U** (1999) NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. <http://ccl.northwestern.edu/netlogo/>
- Wu D, Liang W, Zhu W, Chen M, Ferrándiz C, Burton RA, Dreni L, Zhang D** (2018) Loss of LOFSEP transcription factor function converts spikelet to leaf-like structures in rice. *Plant Physiol* **176**: 1646–1664
- Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu Z, Wang L, Fox N, et al.** (2021) Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Dev Cell* **56**: 557–568.e556
- Youssef HM, Eggert K, Koppolu R, Alqudah AM, Poursarebani N, Fazeli A, Sakuma S, Tagiri A, Rutten T, Govind G, et al.** (2017) VRS2 regulates hormone-mediated inflorescence patterning in barley. *Nat Genet* **49**: 157–161
- Zhao T, Ni Z, Dai Y, Yao Y, Nie X, Sun Q** (2006) Characterization and expression of 42 MADS-box genes in wheat (*Triticum aestivum* L.). *Mol Genet Genom* **276**: 334
- Zhu W, Yang L, Wu D, Meng Q, Deng X, Huang G, Zhang J, Chen X, Ferrándiz C, Liang W, et al.** (2022) Rice SEPALLATA genes OsMADS5 and OsMADS34 cooperate to limit inflorescence branching by repressing the TERMINAL FLOWER1-like gene RCN4. *New Phytol* **233**: 1682–1700

## Appendix 2

Protocol for ex situ wheat meristem cultivation, imaging and basic MorphoGraphX analysis

<https://www.protocols.io/private/69A55C01351C11ED98330A58A9FEAC02>



## Wheat live meristem imaging

Anna Elisabeth Backhaus<sup>1</sup>, [yuanrong.pei](#)<sup>1</sup>,  
Katie Long<sup>1</sup>, Sadiye Hayta<sup>1</sup>, [cristobal.uauy](#)<sup>1</sup>

<sup>1</sup>John Innes Centre



Anna Elisabeth Backhaus  
[John Innes Centre](#)

COMMENTS 0

WORKS FOR ME 2

### PROTOCOL INFO

Anna Elisabeth Backhaus, yuanrong.pei, Katie Long, Sadiye Hayta, cristobal.uauy . Wheat live meristem imaging. [protocols.io](#)  
<https://protocols.io/view/wheat-live-meristem-imaging-cgpztpv6>

### CREATED

Sep 15, 2022

### LAST MODIFIED

Nov 29, 2022

### PROTOCOL INTEGER ID

70105

### MATERIALS TEXT

- "John Innes Cereal Mix" (40% Medium Grade Peat, 40% Sterilized Soil, 20% Horticultural Grit, 1.3 kg·m<sup>-3</sup> PG Mix 14-16-18 + Te Base Fertiliser, 1 kg·m<sup>-3</sup> Osmocote Mini 16-8-11 2 mg + Te 0.02% B, Wetting Agent, 3 kg·m<sup>-3</sup> Maglime, 300 g·m<sup>-3</sup> Exemptor).
- CuSO<sub>4</sub> (5mM)
- Zeatin (0.5 mg/mL)
- GA<sub>3</sub> (1 mg/mL)
- 100x vitamin stock solution (100 mg/L thiamine HCl (Sigma T3902), 35 g/L myo-inositol (Sigma 17508), 69 g/L proline (Sigma P5607))
- 4.3 g/L MS medium (Duchefa M0221)
- 30 g/L maltose (Melford M22000)
- 1.0 g/L casein hydrolysate (Sigma C7290)
- 3.5 g/L phytigel (Sigma P8169)
- 70% Ethanol
- 5% Bleach (Sigma)
- ddH<sub>2</sub>O
- [PPM](#) (0.1% volume)
- Propidium Iodine (PI, 0.1%)

## Plant Growth 3w

- 1 Germinate wheat seeds by placing them on damp filter paper (tap water) in a Petri dish for 48 hours at 4 °C followed by 48 hours at room temperature .

- 2 Sow seeds about 1-2 cm deep in appropriate soil, in this study we used "John Innes Cereal Mix" in 24-cell seed trays. Seedlings can be grown in any required conditions, we grew all seedlings in (heated) glasshouses under long-day conditions (16h light / 8h dark). 20m
- 3 Grow enough seeds so that you can dissect plants every 3-4 days to check the developmental stage. In this study wheat plants (cv. Paragon) were grown for 21 days to reach double ridge stage but this might differ between genotypes and conditions.

## Meristem Media preparation 1h

### 4 Stock preparations:

1.  $\text{CuSO}_4$ , 5mM: Dissolve 125 mg of  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  (Sigma C3036) in a total volume of 100 mL  $\text{H}_2\text{O}$ .
2. Zeatin (0.5 mg/mL): Dissolve 5 mg zeatin (Duchefa Z0917) in a few drops of 1 N NaOH and make up the volume to 10 mL  $\text{H}_2\text{O}$ .
3.  $\text{GA}_3$  (1 mg/mL): Dissolve 10 mg  $\text{GA}_3$  (Sigma G7645), in a few drops of EtOH and make up the volume to 10 mL with  $\text{H}_2\text{O}$ . Filter sterilize, divide into 1 mL aliquots, and store at  $-20^\circ\text{C}$ .
4. 100x vitamin stock solution: 100 mg/L thiamine HCl (Sigma T3902), 35 g/L myo-inositol (Sigma 17508), 69 g/L proline (Sigma P5607). Filter sterilize and store at  $4^\circ\text{C}$ .

### 5 Media preparation materials:

- 4.3 g/L MS medium (Duchefa M0221)
- 30 g/L maltose (Melford M22000)
- 1.0 g/L casein hydrolysate (Sigma C7290)
- 10 mL 100x vitamin stock
- 0.2 mg/L  $\text{GA}_3$
- 0.4 mg/L zeatin
- 1.25 mg/L  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$
- 3.5 g/L phytigel (Sigma P8169)
- Sterilize the media with a vacuum pump with bottle top filters (SARTOLAB BT50).
- Use ultra-pure water for all stocks and the medium.

### 6 For 500 mL meristem medium:

1. Prepare 1.75 g/250 L phytigel (Sigma P8169) in a 1 L screw cap glass media bottle and autoclave at  $121^\circ\text{C}$  for 15 min, 15 psi. You can prepare them in advance and store them at room temperature.
2. Weigh 2.15 g/L MS medium (Duchefa M0221), 15 g/L maltose, 0.5 g/L casein hydrolysate, and add 250 mL of ultra-pure water in a 500 mL beaker.
3. Stir the medium and adjust the medium to pH 5.8 using 1N NaOH.
4. Heat the filter-sterilized medium and autoclaved phytigel to  $60^\circ\text{C}$  in an oven.
5. Add 5 mL 100x vitamin stock, 0.5  $\mu\text{L}$   $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , 100  $\mu\text{L}$   $\text{GA}_3$  400  $\mu\text{L}$  zeatin slowly under sterile conditions in a laminar flow hood and mix well.
6. Pour the warm medium into a bottle containing the warm phytigel and mix well.
7. Pour in 50 mm deep Petri plates quickly. Media plates and tubes can be stored at  $10^\circ\text{C}$  and used for up to 2 weeks.

30m

## Sterile Meristem Collection

7 Prepare for the meristem collection by gathering all materials in the sterile workbench:

- Forceps and knives
- Bunsen burner
- 70% Ethanol in Falcon tube
- 5% Bleach (Sigma) in Falcon
- ddH<sub>2</sub>O
- sterile Petri dishes
- Meristem Media plates
- Sterile Falcon tubes
- Glass beaker
- Stereo microscope

### Safety information

Clean all surfaces and microscope by wiping them with 70% ethanol. Surface sterilize all knives and forceps.

- 8 Remove seedlings from the soil by cutting the roots just below the seeds to avoid damaging the meristem. Wash the seedlings and remaining roots under running water and remove all roots carefully. Make sure no soil is left on the plants. Cut all leaves and shoots above the area of the meristem. You can process multiple plants at the same time (we use about 10-15 plants at a time).
- 9 Transfer all plant shoots to the sterile bench in clean Petri dishes. Wash shoots in 70% ethanol for 1 minute and rinse in sterile water once. Next, wash shoots in 5% bleach for 2 minutes and rinse in sterile H<sub>2</sub>O three times.
- 10 Using only sterile forceps and knives, transfer the first shoot to a sterile Petri dish. Under the microscope, start cutting away the outer layers of leaves until the meristem is exposed. Remove all leaf primordia below the meristem as they can otherwise grow on the plate and obstruct imaging. Next, cut off the meristem carefully, leaving a few mm of tissue below (the cut cells will die and thus give strong PI signals during imaging). Transfer the meristem to the center of the meristem media plate, laying them flat onto the media surface. It is important the the tissue you want to image is facing upwards and has not been damaged or touched by the dissection. Cut a small groove into the media where the meristem will be placed, this is very important to reduce movement during the following staining and imaging process. Close lid and seal with tape, label lid appropriately.



Meristem placed on media plate

- 11 Continue with the next meristem. Note: Multiple meristems can be grown on the same plate. However, this increases the risk of contamination and is unpractical when planning to use confocal imaging. We would only recommend this for light microscopy imaging as the plate lids do not need to be opened in this case.

Place meristem plates in growth chamber (16h light/8h dark) at 21 °C or alternatively on your work bench.

10m

## Meristem Imaging and staining

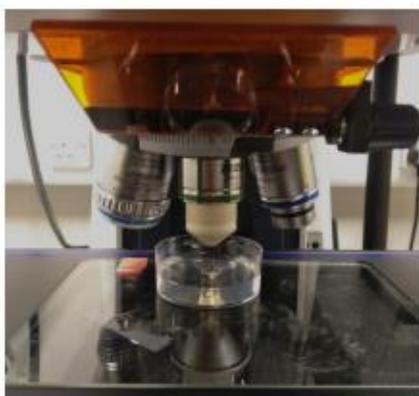
- 12 For stereo microscopy, simply place the plates under a stereomicroscope while keeping the lid closed. This greatly reduces infection. The glass lid will reduce image quality and can be removed if this is of concern. This technique was used to image whole spikes and check primordia formation and growth during the development of the method.
- 13 For confocal microscopy, the cell walls need to be stained using Propidium Iodine (PI) unless cell membrane-specific fluorescence reporter lines are available. In our study, all samples were PI stained.
- 14 Thaw PI aliquots in the dark, using either black Eppendorf or tin foil. Stain samples by opening the plate lid in a sterile environment. Using a pipette, drip 0.1% PI solution on the meristem until it is fully immersed. Stain each meristem for 1:30 minutes. Wash the stain off by flooding the plate with sterile ddH<sub>2</sub>O three times. After, fill the plate again with water so that it doesn't spill over once the confocal lense is dipped inside. NOTE: To reduce contamination of plates we added plant preservative mixture ([PPM](#), 0.1% volume) to the water used for imaging and washing.

### Safety information

40m

## Confocal microscopy

- 15 To image cell growth and development, the cell walls of the plant need to be imaged using confocal microscopy. Here we used the Zeiss LSM 880 confocal microscope equipped with a water immersion objective (W Plan-Apochromat 20×/1.0 DIC UV-IR M27 75 mm). We imaged the meristems every 48 hours, depending on the tissue the imaging frequency needs to be adapted. Ideally, tissues are imaged after every one to two cell divisions.



Meristem on media placed in a confocal microscope (supported by glass insert but two glass slides can also be used in standard stage set-up)

- 16 To image a sample, place it directly into the microscope after staining. Slide two empty glass slides into the stage to support the media plate. After imaging, remove all water from the plate by tipping it out and removing residual water with a pipette. Seal the lids again and return plates to the growth room.
- 17 A 514 nm laser was used for excitation, and emissions were collected between 587-677 nm. Scans were frame averaged twice and z-sections taken at 0.6  $\mu\text{m}$  intervals. The first step has to be 1-2  $\mu\text{m}$  above the last detectable PI signal. The last z-stack should be chosen as low as possible where PI signal can be perceived and cell walls seen. Mostly we worked with a range of 200-250  $\mu\text{m}$ . Images were taken at 1-1.6 x zoom (0.8x usually worked) and frame size was set to "optimum" with image speed at "maximum" to reduce scanning time. The digital gain was set to 1 with Gain (Master) at about 700-800. Laser power was set to about 2 but kept as low as possible.

1h

## Basic MorphoGraphX analysis

- 18 Guidance on how to use MorphoGraphX can be found in the handbook <https://morphographx.files.wordpress.com/2021/08/mgxusermanual.pdf> and also in the youtube tutorials published by theith lab (<https://www.youtube.com/@richardsmithlab6238>). Briefly, we first create the mesh structure by running 'gaussian blur stack' twice and using the voxel tool to remove any background signal around the sample. Next, we applied the 'CNN' prediction (with default settings and Vijayan dataset selection) to improve the cell wall signal of the samples. After CNN, we remove further noise using the voxel tool and then performed 'edge detect' and mesh creation using 'marching cube surface'. The created mesh can be

cleaned by removing the top and bottom as well as any undesired area of the mesh by highlighting and deleting these (this will also reduce the size of the mesh and increase processing speed). Next, we sub-divided the mesh twice and smoothed it 5-10 times. Finally, the cell wall signal is projected onto the mesh with project signal set to 2-6 (this needs to be adapted for each sample using values between 0.5 to 9).

- 19 Once a high-quality mesh was created, we segmented cells using either the auto-segmentation (threshold1.1) or we segmented cells by hand (recommended). To segment cells by hand, mark each individual cell using the color label tool and run 'watershed segmentation' followed by 'fix corners classic'.
- 20 After we created the mesh and cell maps for two timepoints we performed parent-child labelling. For this, the mesh of two timepoints is loaded into MorphoGraphX and moved just above each other. By selecting the pointer tool and then clicking on the same cell in the earlier and later timepoint cells are labelled as parent-child pair. Once all cells are parent labelled growth maps can be created in MorphoGraphX.