Short title: Rudimentary basal spikelets in wheat

High expression of *VRT2* increases the number of rudimentary basal spikelets in wheat

Anna E. Backhaus[1,2], Ashleigh Lister[2], Melissa Tomkins[1], Nikolai M. Adamski[1], James Simmonds[1], Iain Macaulay[2], Richard J. Morris[1], Wilfried Haerty[2], Cristobal Uauy[1]

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

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

One sentence summary: Large transcriptional gradients exist *within* a wheat spike and are associated with rudimentary basal spikelet development, resulting in the characteristic lanceolate shape of wheat spikes.

Author Contribution: (Contributions are based on the CRediT Contributor Roles Taxonomy (https://casrai.org/credit/)
AEB: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft
AL: Investigation, Methodology
MT: Software, Formal analysis, Writing – review & editing
NMA: Resources, Writing – review & editing
JS: Data curation, Formal analysis, Resources, Writing – review & editing
IM: Methodology, Supervision
RJM: Conceptualization, Methodology, Supervision, Writing – review & editing
WH: Conceptualization, Funding acquisition, Supervision, Writing – review & editing
CU: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing

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 Programmes and by the John Innes Foundation. Additional funding was provided by the European Research Council (866328).

Author for contact: Cristobal Uauy (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, 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 correlates 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* MADS-box transcription factors, including *VRT-A2*, are expressed highest in the basal section of the wheat spike and display the opposite expression gradient to flowering E-class *SEP1* 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 new insights into developmental genetics pathways of grass inflorescences.

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

87 mature spike the most basal one or two spikelets are often only formed in a rudimentary
88 manner, with small glumes present but all floral structures remaining immature.

89 Efforts to understand the genetics of wheat spikelet initiation and development have focused
90 on members of the MADS-box transcription factor (TF) family, which play central roles in the
91 flowering gene models (Zhao et al., 2006). Li et al. (2019) showed that MADS-box genes of
92 the *SQUAMOSA*-clade, *VERNALISATION1 (VRN1)*, *FRUITFULL2* (*FUL2*)
93 and *FRUITFULL3* (*FUL3*), have overlapping functions in controlling the timing of the
94 transitions from the vegetative to IM as well as the formation of the terminal spikelet.
95 In *vrn1ful2*-null mutants, the IM remained indeterminate causing the mutants to form more
96 spikelets per spike. However, all lateral spikelets were replaced by leafy shoots in the *vrn1ful2*
97 double and *vrn1ful2ful3* triple mutants (Li et al., 2019). These mutants had increased
98 expression of genes belonging to the *SHORT VEGETATIVE PHASE* (*SVP*) family of MADS-
99 box genes, including *VEGETATIVE TO REPRODUCTIVE TRANSITION 2* (*VRT2*).
100 Subsequent studies determined that overexpression of *VRT2* led to reversion of basal spikelets
101 to spikes and the downregulation of other MADS-box genes required for floral development,
102 including members of the *SEPALLATA1* (*SEP1*) clade (Li et al., 2021). Together, these studies
103 exemplify the importance of the temporal sequence of flowering gene expression for the correct
104 development of the wheat spike.

105 Attempts to unravel the genetic network controlling wheat spike development have focused on
106 these temporal changes in expression patterns across consecutive developmental stages. For
107 example, Li et al. (2018) and Feng et al. (2017) performed transcriptome profiling using pooled
108 samples of multiple complete spikes from six (vegetative to floret differentiation) and four
109 (double ridge to young floret) developmental stages, respectively. In a few cases, studies have
110 examined the expression patterns of individual genes (via quantitative reverse transcription
111 (qRT)-PCR) and found gene expression gradients along the spike. For example, Debernardi et
112 al. (2017) demonstrated that *APETALA2* (*AP2*) is expressed higher in the apical section of
113 wheat spikes than in central or basal sections. This *AP2* expression gradient was associated
114 with morphological changes along the same spike. This study alongside work in barley
115 (Youssef et al., 2017), suggests that gene expression gradients *within* individual developmental
116 stages could be important to further unravel the genetic control of spike development.
117 However, despite its potential biological significance, spatial transcriptome profiles along the
118 spike have yet to be investigated in wheat.

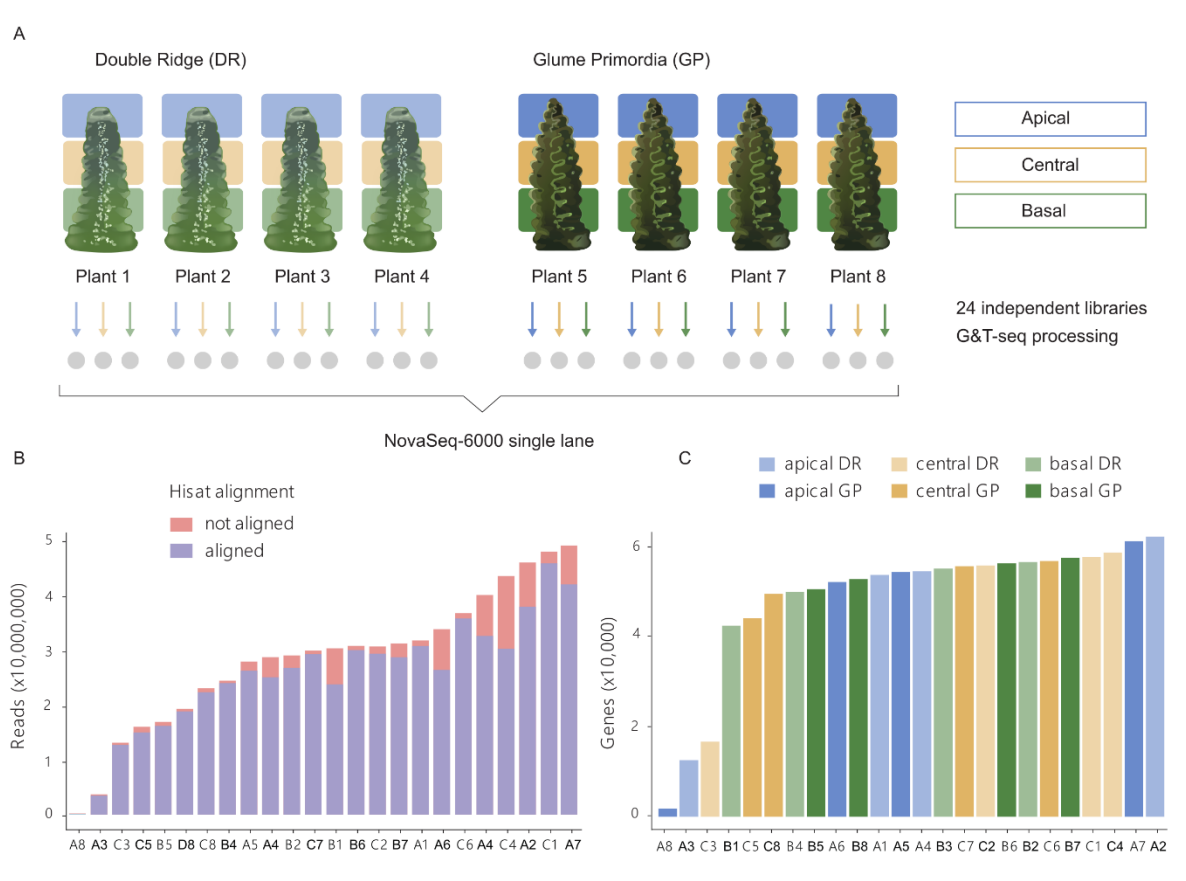
119 In this study, we aimed to characterise gene expression profiles along the spike during the
120 establishment of the lanceolate shape of the wheat spike from DR to GP. We adapted G&T-
121 seq (Genome and Transcriptome sequencing), a low-input sequencing approach to sequence
122 the transcriptome of the sections. Recently, Giolai et al. (2019) adapted the protocol to identify
123 expression differences across single leaves of *Arabidopsis* (GaST -seq), demonstrating that the
124 G&T-seq method can be readily used for sequencing of hand harvested, small input plant
125 material without the need of previous tissue dissociation or treatment. G&T-seq is thus
126 comparable to methods using laser-micro dissection followed by sequencing to achieve
127 spatially resolved transcriptome wide sequencing data. In comparison, the available
128 transcriptome sequencing methods at higher resolutions (such as single cell RNA-seq or
129 fluorescence-activated cell sorting (FACS)) are not spatially resolved as the complete tissue is
130 dissolved into single cells for barcoding or selection prior to sequencing (Rich-Griffin et al.,
131 2020).

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

146 ## Results

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

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



155

156 **Figure 1:** (A) Summary diagram of low-input G&T tissue collection and sequencing approach.
157 Grey circles indicate the 24 individual libraries prepared for sequencing from each individual
158 tissue section dissected from individual spikes. (B) Reads per library after trimming and quality
159 controls (see Methods). Stacked bars indicate the number of reads aligned (blue) and not
160 aligned (red) by HISAT to the RefSeqv1.0 genome. (C) Number of expressed genes (>10 read
161 counts) per library based on tissue section and Waddington developmental stage (DR: Double
162 Ridge; GP: Glume Primordia). In (B and C), the X-axis indicates the ID of each sample which
163 is composed of the tissue section (A: apical, C: central, B: basal) and plant number (1-8) as
164 indicated in (A). Detailed quality control data for each library is provided in Supplemental
165 Table S1.

166

167 On average, samples had 28,799,626 reads (coefficient of variation (CV) 43%), of which 90%
168 (CV 8.5%) aligned to the genome post adaptor trimming (Figure 1B, Table 1). Furthermore,
169 the number of aligned reads and the number of expressed genes per library was largely
170 homogenous among the spatial sections and Waddington stages (Table 1). On average, 47,313
171 genes per library were expressed (>10 read counts) and we found no difference ($P > 0.56$) in
172 the number of expressed genes across spatial (apical, central, basal) or between temporal (DR,
173 GP) conditions (Figure 1C). We excluded three libraries with low average number of expressed
174 genes (difference greater than five times the standard deviation; Figure 1C, Supplemental Table
175 S1) and two libraries because they were strong outliers in the principal component analysis
176 (PCA; Supplemental Figure S1A). In total, 19 RNA-seq libraries (DR: 3 apical, 4 central, 3
177 basal; GP: 2 apical, 3 central, 4 basal) passed our selection criteria and were used in the
178 subsequent analyses. We identified 91,646 genes being expressed across these 19 libraries.

179

180 **Table 1**: Average number of reads aligned to the RefSeqv1.0 genome and expressed genes
181 (>10 read counts) in the three tissue sections and two Waddington developmental stages (DR:
182 Double Ridge; GP: Glume Primordia) (n = 4 biological replicates per tissue section *
183 developmental stage).

	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

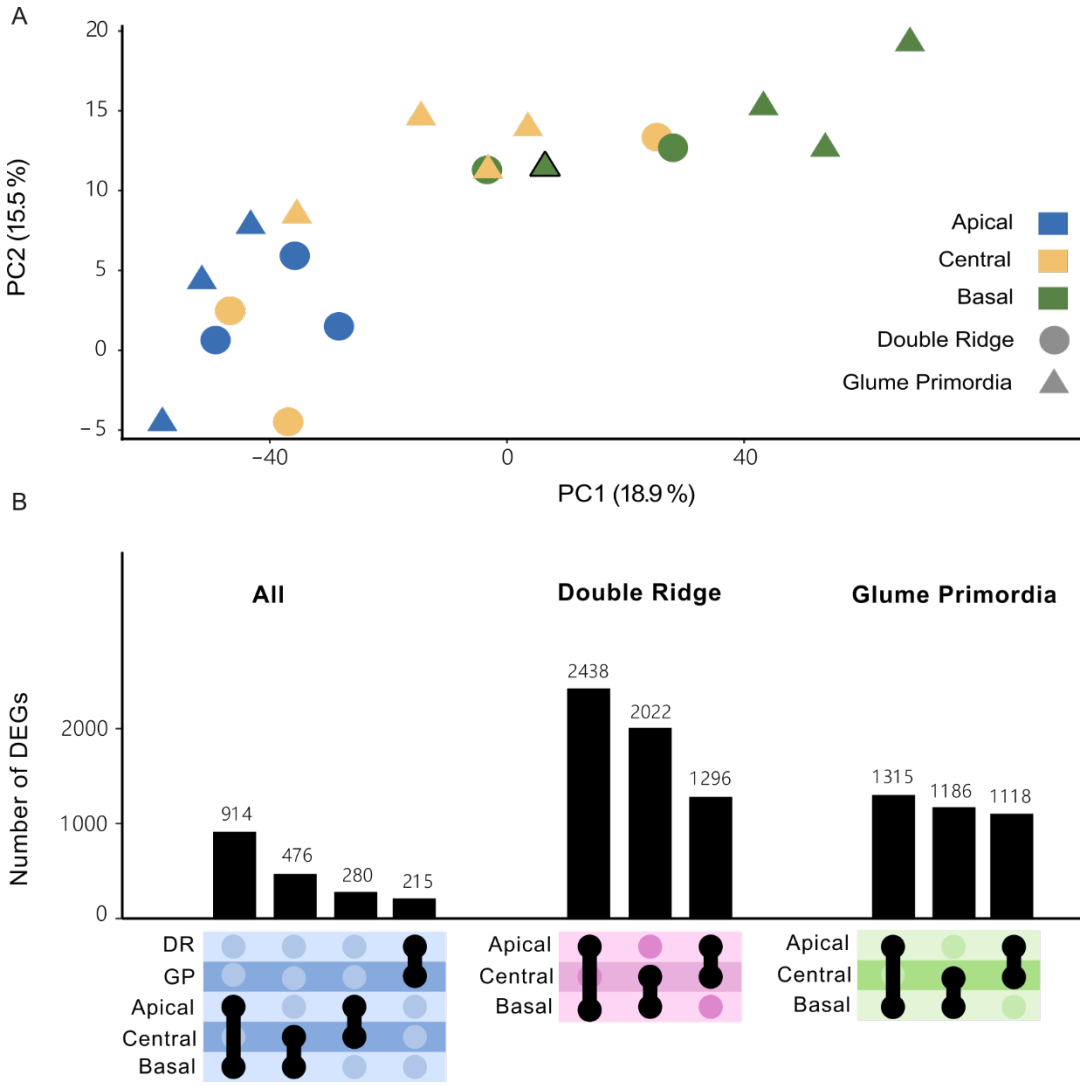
184

185

186 **Transcriptome-wide differences are largest between the apical and basal sections of the**
187 **spike**

188 To investigate global differences among the 19 RNA-seq libraries, we performed a principal
189 component analysis (PCA; Figure 2A). The first two PCs explained 19% and 16% of the overall
190 variance present in the libraries. We observed that the two PCs separated libraries by the spatial
191 position (apical, central, basal) rather than developmental stage (DR, GP). There was a clear
192 separation between libraries originating from apical and basal spike sections, while libraries
193 from central sections were dispersed between these two clusters (Figure 2A). We investigated

194 PC1 to PC6 and found that none of these combinations clustered libraries by developmental
195 stage (Supplemental Figure S1B). Given that we sequenced developing spike sections of single
196 plants, as opposed to the more commonly employed pooling of multiple biological samples,
197 we found as expected some degree of heterogeneity between samples from the same location
198 and stage (Figure 2A).



199

200 **Figure 2:** (A) Principal component analysis (PCA) on the 19 transcriptome libraries from
201 apical (blue), central (yellow) and basal (green) sections of Double Ridge (circles) and Glume
202 Primordia (triangles) spikes. Black bordered triangle is plant 8 (GP) in which the basal section
203 clustered closer with central-GP sections than the other basal-GP sections. (B) UpSet plot
204 showing the number of differentially expressed genes (DEGs) between spatial sections and
205 Waddington stages. Black border indicates Plant 8, which is an outlier of basal GP replicates.

206

207 To investigate this variation further, we quantified changes in gene expression across biological
208 replicates by calculating CVs for each gene (see Methods). The median CV for a gene across

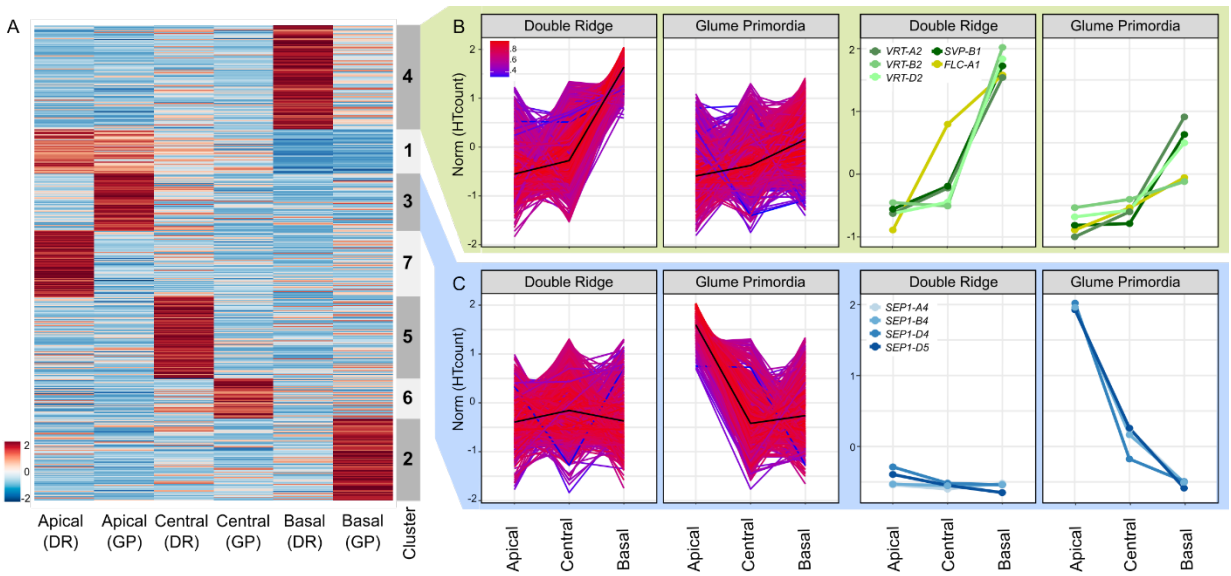
209 the biological replicates was 39% (Supplemental Figure S2) with a Q1-Q3 interquartile range
210 between 24% and 62%. We also calculated the CV per gene for published datasets from Li et
211 al (2018) and Feng et al (2017). Both studies sequenced developing wheat spikes at similar
212 developmental stages, pooling many spikes per sample. Li et al. (2018) pooled between 100 to
213 200 spikes of winter wheat (KN9204) per sample, while Feng et al. (2017) reported pooling of
214 10 to 50 spikes (cv. Chinese Spring) per sample. In both studies the median CV of a gene was
215 lower (14% and 21%, respectively) than in our study. The larger CVs in our data could be
216 explained by the biological variation that exists between individual plants, which may have
217 been reduced by the pooling of many spikes in both Li et al. (2018) and Feng et al. (2017).

218 We first analysed differentially expressed genes (DEGs) between the DR and GP stage and
219 between apical, central and basal sections across the two Waddington developmental stages.
220 The number of DEGs between DR and GP (215 genes) was smaller than the number of DEGs
221 identified between the spatial positions, which ranged from 280 DEGs between central and
222 apical sections to 914 DEGs between the apical and basal sections (Figure 2B). Next, we
223 compared the apical, basal and central sections within each Waddington developmental stage.
224 We identified more DEGs by comparing the spatial sections within either Waddington stage
225 individually than in the combined analysis. The number of DEGs between apical and basal
226 sections at each stage (DR: 2,438; GP: 1,315) were similar to the number of DEGs between
227 central and basal sections (DR: 2,022; GP: 1,186). The number of DEGs between these sections
228 at DR, however, was nearly double the number of DEGs at GP. In contrast, the number of
229 DEGs between apical and central sections was similar at both stages (DR: 1,296; GP: 1,118),
230 suggesting that the basal section of the spike is most different in the earlier developmental
231 stage. Only 11% of the DEGs were shared between DR and GP in the apical to basal
232 comparison, 7% between the central to basal DEGs, and 5% between apical to central DEGs.
233 In total, we identified 5,353 unique genes as differentially expressed between any of the three
234 sections at either Waddington stage (Supplemental Table S2). Overall, the number of DEGs
235 was largest between the apical and basal sections, reflecting the strong spatial clustering
236 observed in the PCA graph, but most genes that were differentially expressed across the spike
237 did not maintain this gradient over the two developmental stages. In summary, despite the high
238 biological variation in gene expression in our data compared to previous pooled whole-spike
239 studies, we could detect transcriptome wide differences between the spatial sections of
240 developing wheat spikes.

241

242 **The *SVP* MADS-box transcription factors have opposing expression profiles to flowering**
243 **E-class genes**

244 To further investigate the differences in expression across the spike and to identify genes with
245 similar expression patterns, we performed hierarchical and k-means clustering. We restricted
246 the clustering to the 5,353 genes identified as differentially expressed across the spike at either
247 one or both Waddington stages (Figure 3A). We identified seven non-redundant clusters, each
248 containing between 8% to 21% of the 5,353 DEGs (Figure 3A, Supplemental Figure 3A). Both
249 hierarchical and k-means clustering produced highly similar results (Supplemental Figure
250 S3B). We identified 1,894 genes (35% of DEGs) to be more highly expressed in the apical
251 section, either across both timepoints (503 genes, cluster 1), or only at DR (751 genes, cluster
252 7) or GP (640 genes, cluster 3). In the central section, 1,362 genes (25%) had higher relative
253 expression at either DR (917 DEGs, Cluster 5) or GP (445 DEGs, Cluster 6). In the basal
254 section, we observed 2,097 genes (39%) being more highly expressed. Cluster 4 contained the
255 most DEGs (1,170) and was characterized by an upregulation of expression in the basal section
256 at both Waddington stages, although this upregulation was higher at the DR stage. Another 927
257 genes were upregulated in the basal section, but only at the GP stage (cluster 2).

258 

259 **Figure 3**: (A) Normalized expression matrix and K-means clustering of the 5,353 genes
260 differentially expressed across the spike at either one or both Waddington stages. Colours (blue
261 to red) show relative log2 expression of genes after normalisation. (B) Expression pattern of
262 the 1,170 genes allocated to Cluster 4 (left), and of MADS-box transcription factors in the same
263 cluster (right). Colours indicate how well the gene expression pattern fits the average
264 expression pattern (black line). Red = best fit, Blue = least good fit. (C) Expression pattern of
265 the 640 genes and MADS-box transcription factors of Cluster 3 as arranged in B. Norm =
266 normalized and scaled gene expression. RefSeq1.1 gene IDs and raw expression values of
267 genes shown in the right-hand panels are presented in Supplemental Table S2.

268 To further characterize the clusters, we independently tested for enrichment of TF families and
269 gene ontology (GO)-terms (all GO-terms and TF families in Supplemental Table S3 and S4,
270 respectively). Genes that were more highly expressed in the apical section were enriched for
271 the GO terms "reproductive structure development" (GO:0048608) and "floral organ
272 development" (GO:0048437; cluster 1) as well as for HD-Zip_IV and SRS TF families ($P <$
273 0.0001; cluster 1). In cluster 3 (highly expressed in the apical section at GP; Figure 3B) we
274 found no significant enrichment of GO-terms ($P < 0.03$), but a significant enrichment of
275 MADS_II TFs ($P = 0.013$). For clusters defined by an increased expression in the central
276 sections (clusters 5 and 6) we detected an enrichment for GO-terms related to polyphosphate
277 processes (GO:0006797/0006779), but a significant enrichment of the C2C2_CO-like TFs at
278 DR ($P = 0.036$; cluster 5). Genes with higher expression in the basal section of the spike (cluster
279 4; Figure 3C) were enriched for a number of GO terms related to photosynthesis, (e.g.
280 GO:0015979) and "negative regulation of flower development" (GO:0009910), as well as for
281 MADS_II TFs ($P = 0.084$). Cluster 2, which was characterized by higher expression in the
282 basal section only at GP, was enriched for "Jasmonic acid response" (GO:0009753) and Tify
283 and C2C2_CO-like TFs ($P < 0.04$).

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

293 In the apical/GP cluster 3 we noticed that all MADS-box genes belonged to the *Triticum*
294 *aestivum SEPALLATA1* (*TaSEP1*) group (Figure 3C). All three homoeologs of *TaSEP1-4*
295 (*TraesCS7A02G122000, TraesCS7B02G020800, TraesCS7D02G120500)* and the D-genome
296 copy of *TaSEP1-5* (*TraesCS7D02G120600*) were part of this cluster. The *SEP* genes were
297 expressed at relatively low levels at DR (Supplemental Table S2), but were significantly
298 upregulated at GP, with their transcript levels being highest in the apical section. The increased
299 expression of *TaSEP1-4* at GP was in agreement with their previously reported expression
300 patterns in tetraploid wheat by Li et al. (2021).

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

315

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

318 To validate the expression pattern of *VRT2* in the individual spike analysis, we performed qRT-
319 PCR on independently collected, pooled spike sections from cv Paragon, carrying the wildtype
320 *VRT-A2a* allele (Figure 4A, blue curves). We included a later timepoint, Terminal Spikelet
321 (TS), which is about 10 days after GP to study how *VRT2* expression changes in later stages.
322 At TS stage, the central spikelets have developed multiple florets primordia. We also included
323 a small part of the peduncle (stem) section just below the spike as an additional spatial section.
324 We focused the expression analysis on the A-genome homoeolog, *VRT-A2*, as its role in spike,
325 glume and grain development of wheat was recently characterised (Adamski et al., 2021; Liu
326 et al., 2021).

327

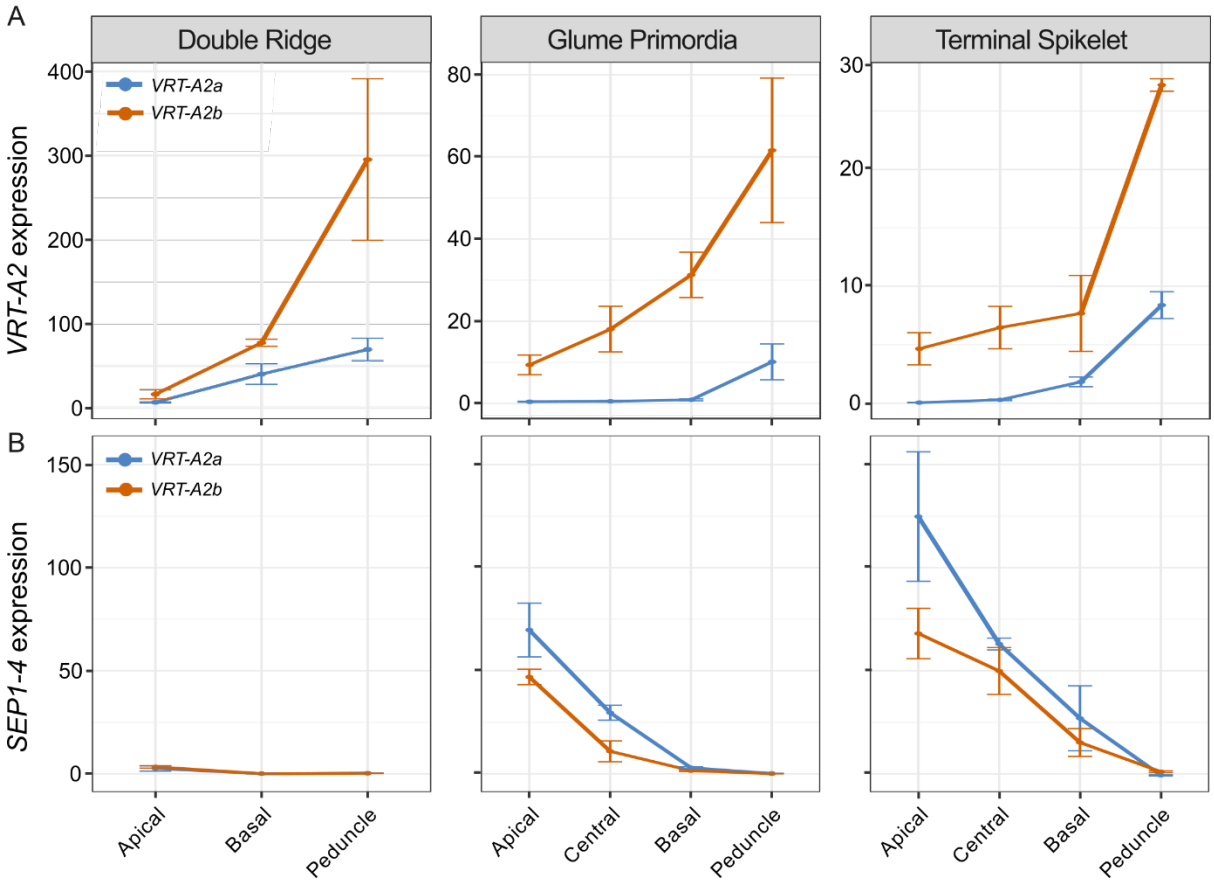


Figure 4: Relative expression (2^{ddCT}) of *VRT-A2* (A) and *SEP1-4* (B) in the different sections of the spike across three timepoints in near isogenic lines (NILs) carrying either the wildtype *VRT-A2a* (blue) or the *VRT-A2b* allele from *Triticum turgidum* ssp. *polonicum* (orange). The data are shown as mean ± 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).

We identified a significant interaction effect between Waddington stage and spatial section; we thus analysed 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

346 to the three spike tissues ($P < 0.001$ for all three comparisons). Lastly, *VRT-A2* expression at
347 TS stage was just detectable and significantly different between all sections ($P = 6.6E-06$).
348 Overall, expression decreased significantly from DR to GP/TS Waddington stages in the apical
349 ($P = 0.00015$), basal ($P = 0.0074$), and peduncle ($P = 0.012$) sections consistent with the
350 previously reported strong downregulation of *VRT-A2* in the early wheat spike development
351 (Li et al., 2021; Adamski et al., 2021; Liu et al., 2021). This is also consistent with the observed
352 downregulation of *VRT-A2* orthologs upon floral transition in barley (Trevaskis et al., 2007)
353 and rice (Harrop et al., 2018). As observed in the low-input RNA-seq data, the qRT-PCR data
354 confirmed the strong basipetal gradient in *VRT-A2* expression across the spike at DR and
355 revealed that its expression was even higher within the peduncle.

356 We hypothesised that the higher expression in the basal section of the wheat spike compared
357 to the central and apical sections is associated with the rudimentary development of the basal
358 spikelets. To test the effect of higher *VRT-A2* expression on basal spikelet development, we
359 analysed the effect of the *Triticum turgidum* ssp. *polonicum VRT-A2b* allele on the expression
360 gradient of *VRT-A2* and spike morphology. Adamski et al. (2021) showed that *VRT-A2* in *T.*
361 *polonicum*, a tetraploid subspecies of wheat, carries a sequence re-arrangement in its first
362 intron. This results in the higher expression of the *T. polonicum VRT-A2b* allele, with respect
363 to the wildtype *VRT-A2a* allele, during early spike development. We performed qRT-PCR on
364 a cv Paragon NIL carrying the *VRT-A2b* allele and compared *VRT-A2* expression against the
365 Paragon wildtype NIL described above (Figure 4). Consistent with the results of Adamski et
366 al. (2021), we detected significantly higher expression of *VRT-A2b* compared to the wildtype
367 allele across most of the tissue sections (see Supplemental Table S6 for individual
368 comparisons), and a progressive decrease in *VRT-A2b* expression over time ($P = 0.031$). In
369 contrast to the wildtype NILs, ANOVA did not identify a significant interaction effect between
370 spatial section and Waddington stage in *VRT-A2b* NILs ($P = 0.18$). We thus examined the
371 overall expression patterns and found that across all three developmental stages *VRT-A2b*
372 expression differences were significant ($P < 0.0001$). These results suggest that the basipetal
373 expression gradient in the spike is maintained in the NILs with the *T. polonicum VRT-A2b*
374 allele.

375 We also tested the effect of *VRT-A2* expression levels on *SEP1* expression in the tissue sections
376 of *VRT-A2* NILs. We confirmed that *SEP1-4* expression is only marginally detectable at DR
377 and differences in expression between the apical, basal, and peduncle section are hardly
378 detectable with only marginally higher expression in the apical sections ($P = 0.015$) at this

379 stage. At GP, *SEP1-4* expression is significantly higher towards the tip of the spike ($P < 0.0001$)
380 consistent with the low-input RNA-seq data. Furthermore, *SEP1-4* expression was significantly
381 lower in *VRT-A2b* NILs compared to the wildtype allele across all spike sections ($P = 0.008$),
382 confirming that higher *VRT-A2* expression can negatively affect *SEP1-4* expression. Similar
383 trends were observed at TS, where expression was significantly lower in basal sections ($P <$
384 0.0001) and the gradient across the spike was maintained in *VRT-A2b* NILs, but expression
385 was overall lower ($P < 0.003$). *SEP1-4* is not expected to be expressed in vegetative tissue such
386 as the peduncle, therefore the lack of expression in this tissue across the three stages indicates
387 that no floral tissue was accidently sampled as peduncle (Fig. 4B).

388

389 **Misexpression of *VRT-A2b* in *T. polonicum* increases rudimentary basal spikelet numbers**
390 To evaluate if the higher expression of *VRT-A2* in basal spikelets affects their development, we
391 examined the *VRT-A2* NILs (BC_4 and BC_6) sown as winter crops in four environments. In each
392 field trial, we evaluated the number of rudimentary basal spikelets (RBS), that is spikelets
393 which are reduced in size and do not contain mature grains (Figure 5A). The number of RBS
394 was significantly increased in NILs carrying the *VRT-A2b* allele in all four environments ($P <$
395 0.0001, except Morley 2017 $P < 0.01$; Figure 5B; Supplemental Table S7). The *VRT-A2a* NILs
396 and the recurrent parent Paragon had on average 1.85 RBS, whereas *VRT-A2b* NILs produced
397 on average 2.91 RBS. A similar difference in RBS between the NILs was observed in
398 glasshouse conditions (*VRT-A2b* effect of +1.6 RBS; Supplemental Table S8; Figure 5C).

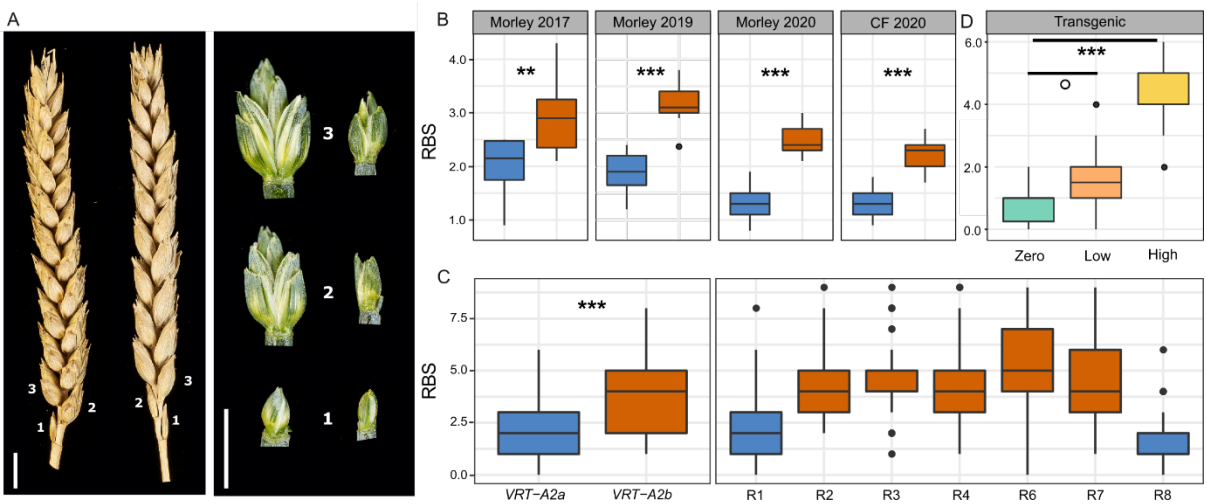


Figure 5: Phenotypic difference between *VRT-A2a* (blue) and *VRT-A2b* (orange) on rudimentary basal spikelet numbers (RBS). (A) Mature spikes from the field (left) and dissected basal spikelets at anthesis (right) from the glasshouse. Numbers indicate position along the spike starting at the base. Scale bar = 1 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). (C) Number of RBS recorded in the glasshouse for the NILs (left panel) 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). (D) RBS per spike recorded for the transgenic lines carrying zero, low (1-5) or high (9-35) 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). *P* values: o ≤0.1; ** ≤ 0.01; *** ≤ 0.001.

Furthermore, we also recorded the number of RBS in seven homozygous BC_6 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) 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; 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

427 are visible. In NILs carrying the *VRT-A2b* allele the development of the most basal spikelet is
428 very similar to the wildtype. However, the second, third and sometimes fourth spikelet also
429 display similar signs of reduced development, leading to the larger number of rudimentary
430 basal spikelets (Figure 5B, C; Supplemental Figure S4).

431 To validate the phenotypic effect of *VRT-A2*, we analysed transgenic wheat lines transformed
432 with the complete genomic *T. polonicum VRT-A2b* sequence (including the native promoter
433 and the intron 1 re-arrangement). Transgenic T_1 lines were classified based on the transgene
434 copy number which was previously shown by Adamski et al. (2021) to be highly correlated
435 with *VRT-A2* expression levels in multiple tissues. We phenotyped lines with zero (n = 2
436 independent events; 5 plants each), low (1-5 transgene copies; n = 4 independent events; 5
437 plants each) and high (9-35 transgene copies, n = 2 independent events; 5 plants each) transgene
438 copy number. We identified a significant and stepwise increase in the number of RBS with
439 transgenic copy number, from 0.8 RBS (zero copy) to 1.6 RBS (low copy; $P = 0.078$ vs zero
440 copy) and to 4.3 RBS (high copy; $P < 0.0001$ vs zero copy) (Figure 5D; Supplemental Table
441 S9). The low copy number lines had an average increase of 0.8 RBS with respect to the zero
442 copy number lines, equivalent to the average difference between the *VRT-A2a* and *VRT-A2b*
443 NILs in the field (*VRT2-A2b* effect of +1.1 RBS). The high copy number lines produced on
444 average 4.3 RBS, which is higher than the *VRT-A2b* NILs and similar to the number of RBS
445 observed in *T. polonicum* (3.75 ± 0.62 RBS; n = 16 spikes). The dosage-dependent effects
446 observed in the transgenic lines provide further evidence that elevated expression of *VRT-A2*
447 leads to increased number of rudimentary basal spikelets in polyploid wheat.

448 # DISCUSSION

449

450 **High-resolution spatial transcriptomics in crops**

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

468 To address this challenge, we adapted the G&T method for micro-scale spatial-transcriptomics
469 workflow (Macaulay et al., 2015; Giolai et al., 2019), to conduct RNA-seq of the apical, central
470 and basal sections of individual, hand-dissected wheat spikes. This highly-automated workflow
471 requires low tissue input and allowed us to combine 24 Nextera libraries into a single Illumina
472 NovaSeq lane. For 19 out of the 24 samples the method worked successfully, determined by
473 >20,000 expressed genes per library and the clustering among biological replicates. We found
474 that the number of expressed genes per library was on average similar to the number of genes
475 reported for bulked whole spike RNA-seq samples (Feng et al., 2017; Li et al., 2018). This is
476 consistent with the fact that the hand-dissected sections are composed of a large mixture of
477 different tissues (e.g., rachis, spikelet, and floret primordia) and cell types, which in the
478 equivalent maize ears have distinct expression profiles (Xu et al., 2021). Compared to previous
479 bulk RNA-seq studies in developing wheat spikes, the variation observed here (measured as
480 CV) was high among biological replicates (Supplemental Figure S2). This variation is likely

481 caused by both biological variation (e.g., inherent variation of individual plants) and technical
482 variation (e.g. inaccuracies in sectioning and in the developmental staging of the plant/spike)
483 as well as the number of replicates in our analysis. A minimum of six replicates has been
484 proposed for bulked RNA-seq (Schurch et al., 2016). Our results suggest that the RNA-seq
485 from these small sections would benefit from a higher number of biological replicates, which
486 should be feasible considering the high-throughput method employed for RNA extraction and
487 library preparation, the low tissue input requirement, and the possibility to pool multiple
488 biological replicates per sequencing lane. Despite some limitations, we could identify over
489 5,000 DEGs between the spatial sections for subsequent functional analysis.

490 In addition to G&T-Seq, several other technologies have been proposed for obtaining high
491 resolution transcriptional profiles of plant tissues, for example, single cell RNA-seq (McFaline-
492 Figueroa et al., 2020; Rich-Griffin et al., 2020), FACS, and the isolation of nuclei tagged in
493 specific cell types (INTACT). These methods, however, are not spatially resolved as the
494 complete tissue is dissolved into single cells for barcoding or selection (Rich-Griffin et al.,
495 2020). Thus, these current methodologies do not allow, for example, to investigate whether the
496 cell type composition of spikelets differs across the inflorescence. This would only be possible
497 if spikelets were 'harvested' individually, for example through laser capture microdissection
498 (LCM) before dissolving the tissue further into individual cells. Thiel et al. (2021) recently
499 combined LCM followed by RNA-seq of the distinct lower/leaf ridge and upper/spikelet ridge
500 of barley spikes. This allowed them to identify precise spatio-temporal expression patterns of
501 many genes related to architecture and yield in barley spikes with unprecedented resolution.
502 Looking ahead, increased resolution of Spatial Transcriptomics (currently 100 µm;
503 (Giacomello et al., 2017)), which quantifies full transcriptomes while maintaining tissue
504 integrity, offers the true prospect of direct localisation and quantification of gene expression.
505 Our results argue strongly for the need of these transcriptome-wide and spatially resolved
506 approaches to advance our biological understanding of fundamental developmental processes
507 in plants.

508

509 **The composite nature of spikes**

510 Early morphological studies of wheat spike development described that the stronger elongation
511 of central spikelets during their initial establishment (glume primordia stage) first causes the
512 lanceolate shape of the wheat spike (Bonnett, 1966). The continuous formation of primordia at

513 the tip of the spike means that at any given growth stage, spikelets in different developmental
514 stages will be present across the spike (Bonnett, 1966). In this study, we detected more
515 differentially expressed genes between the three spatial sections of the spike (apical, central
516 and basal) than between the two investigated developmental stages (Double Ridge and Glume
517 Primordia). We identified 215 DEGs between the two developmental stages, consistent with
518 Li et al. (2018) who identified 206 DEGs between consecutive stages across a time course of
519 six inflorescence development stages. Feng et al. (2017) identified 753 DEGs between the
520 Double Ridge and Floret Primordia stage, which are further apart in development than the
521 stages used in this study. They also detected fewer DEGs when comparing early stages than
522 between more developed spikes. By contrast, we identified 1,315 and 2,438 unique genes to be
523 differentially expressed between the apical and basal section at DR and GP, respectively. The
524 higher number of DEGs between spatial sections could be due to the developmental gradients
525 occurring in the three spatial sections, which are revealed by the spatial sampling. These
526 differences would be blurred when comparing whole inflorescences between stages due to the
527 mixture of tissue types and spikelets at different developmental stages. A possible improvement
528 for future transcriptome studies could be the collection of only central sections of the
529 developing spikes or complete spatial sampling as conducted here.

530 The composite nature of the inflorescence tissues has been acknowledged by studies in maize
531 (ears and tassels), where new meristems are initiated in a stepwise manner. Leiboff and Hake
532 (2019) quantified the meristematic tissue composition of maize and sorghum tassels. For
533 example, maize tassels in the second stage are mainly composed of spikelet pair meristems,
534 but also contain some meristems in spikelet and inflorescence state. They concluded that the
535 changes in these tissue compositions over time correlated well with the independently staged
536 transcriptional changes of the tassels. Eveland et al. (2014) showed that the range of
537 developmental ages across the maize ear, if acknowledged, can be used as an advantage in
538 RNA-seq studies. They sequenced the tip, middle, and basal sections of 10-mm long ears
539 independently, aiming to analyse the expression patterns in specific developmental meristem
540 types enriched in these sections (inflorescence, spikelet, and floral meristems, respectively).
541 The dissection of the ear therefore allowed them to study gene expression specifically for each
542 meristematic tissue type rather than for all meristem types in intact ears. In this study, we
543 observed that apically expressed genes are enriched for GO-terms related to “shoot system
544 development” and “maintenance of floral organ identity”. This is consistent with the hypothesis

545 that the apical part of the inflorescence is younger and undergoing early phases of spikelet
546 development initiation compared to the central inflorescence section.

547

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

561 In wheat, the double *SVP* mutant *vrt2svp1* leads to the formation of axillary inflorescences (Li
562 et al., 2021). Similarly, overexpression of *CEN-D2* (*TaTFL1-2D*) in wheat extends the duration
563 of the Double Ridge stage (Wang et al., 2017), whereas loss-of-function mutations in barley
564 *CEN* suggest they repress floral development under short-day conditions (Bi et al., 2019).
565 Double knockout mutants of the MADS-box *SQUAMOSA* genes *vrn1ful2* highlighted that
566 these two genes act as transcriptional repressors of *SVP* and *CEN* genes in early wheat spike
567 development (Li et al., 2019). Furthermore, through a series of genetic and biochemical studies,
568 Li et al. (2021) showed that the downregulation of *SVP* genes is necessary for the formation of
569 flowering promoting MADS-box protein complexes including VRN1, FUL2 and SEP proteins.
570 Hence the coordinated downregulation of *SVPs*, and possibly *CEN* genes, along with the
571 upregulation of *SEP* genes is required for normal floral transition and spikelet development in
572 wheat. Previous studies in rice have found similar expression patterns, as well as mutant effects,
573 of *SVPs* and *SEPs* suggesting a conserved function in flowering transition across the two
574 species (Ren et al., 2016; Wu et al., 2018).

575 Based on our results, the floral developmental programme across the wheat spike appears to be
576 most advanced in its apical and central sections, while being delayed in the basal sections. We

577 hypothesise that this is due to elevated *VRT2* expression at the base of the spike, which hinders
578 the progression of the flowering programme via *SEP* class flowering genes. Likewise, the
579 higher expression levels of the wheat *CEN2* and *CEN5* homologs at the base are consistent
580 with a delay in floral transition that could interfere with the development of the spikelet
581 primordia. Therefore, although the basal spikelet primordia are initiated first chronologically,
582 their developmental age in terms of the floral programme is delayed with respect to the more
583 recently formed central and apical spikelet primordia. This could explain in part why the
584 spikelet primordia in the basal region of the spike elongate less and develop slower than central
585 spikelets despite being initiated first (Bonnett, 1966). Likewise, the less advanced floral
586 developmental programme could also explain why the overexpression of *SVPs* in barley
587 (*HvBM10*) leads to complete floral reversion in basal but not apical spikelets (Trevaskis et al.,
588 2007).

589 We hypothesise that *SVPs* need to be downregulated upon floral transition to allow timely
590 establishment and progression of the early spikelet primordia. Failure to do so would delay
591 their development and result in their final rudimentary shape in the mature spike. In line with
592 this hypothesis, we observed increased RBS in genotypes with prolonged and increased *VRT2*
593 expression in a dosage-dependent manner. In our qRT-PCR data, we also observe increased
594 expression of *SVPs* alongside reduced expression of *SEPs* in *VRT-A2b* lines at Double Ridge
595 and Glume Primordia stage. However, we cannot exclude the possibility that the increase in
596 *VRT2* expression in the *VRT-A2b* lines could also affect basal spikelet development at a later
597 stage of spike developmental (e.g., from Terminal Spikelet stage to anthesis). We are currently
598 using quantitative live imaging to compare cellular growth dynamics of spikelets at different
599 stages of spike development between the *VRT2* NILs.

600 The finding that the expected downregulation of *SVP*s and *CEN*s does not follow the
601 chronological age of the tissues suggests that other gradients across the spike might influence
602 spikelet development. Debernardi et al. (2017) showed that in tetraploid wheat *AP2-5* and
603 miR172 have consistent and opposing expression gradients across the spike at three
604 consecutive developmental stages. The persistent expression gradient of *AP2-5* supports the
605 idea that expression patterns across the spike, beyond the ones caused by age differences of
606 spikelets, exist. Furthermore, they proposed a model illustrating that the phenotypic effect of
607 mutants across the spike differs due to the existing gradient of expression of this gene
608 (Debernardi et al., 2017). Other examples of mutants with different phenotypic effects across
609 an inflorescence include *vrn1ful2* (Li et al., 2019) in wheat, *tassel sheath1* (*tsh1*, (Whipple et

610 al., 2010)) and *ramosa2* (*ra2,* (Bortiri et al., 2006)) in maize, *SEPALLATA* double mutant
611 *Osmads5Osmads34* in rice (Zhu et al. 2021), as well as *many noded dwarf1* (*mnd1,* (Walla et
612 al., 2020)), *frizzy panicle* (Poursarebani et al., 2015), and *vrs2* in barley, which was also found
613 to be consistently differentially expressed across the spike (Youssef et al., 2016). *VRS2* has
614 been shown to maintain a basal to apical expression pattern across three, post awn initiation
615 developmental timepoints in barley (Youssef et al., 2016). The study of *vrs2* mutants revealed
616 that *VRS2* is furthermore engaged with the basal-apical patterns of auxin, cytokinin, and
617 gibberellin across the spike. While hormonal gradients across the spike in early development
618 have not been studied in great detail in wheat, they have been shown to play crucial roles in
619 floral induction and development in *Arabidopsis* (Reinhardt et al., 2000). Their patterns across
620 the spike should be investigated in future studies addressing developmental differences across
621 the spike.

622

623 **A model for the regulation of leaf and spikelet ridge outgrowth in the base of the spike**
624 Recently, Meir et al. (2021) proposed that in shoot apical meristems of tomato, similar to
625 processes during embryonic development, transient programmes are required to inhibit a
626 preceding setup (i.e. vegetative growth), before a new developmental program (flowering) can
627 be initiated. We propose that the altered gene expression and development of the basal spikelets
628 could be a consequence of their initiation during the transient phase between vegetative and
629 floral network shifts and thus being exposed to mixed signals of development. Upon floral
630 transition, the lower (leaf) ridge is supressed, while the growth of spikelet ridges from the
631 previously supressed axillary meristems is activated. Development of lower ridges subtending
632 all branching events is suppressed in grass inflorescences upon flowering transition (Whipple
633 et al., 2010). Li et al. (2019) noticed that this suppression was disrupted in the double *vrn1ful2*
634 and triple *vrn1ful2ful3*-null mutants, which fail to down-regulate *SVP* genes. In these mutants,
635 the upper spikelet meristems generate vegetative structures resembling tillers that are
636 subtended by bracts or leaves originating from the lower leaf ridge.

637 We observed that genes that were highly expressed in the basal section of the inflorescence
638 (cluster 4) have previously been shown to be expressed specifically in the lower/bract ridge
639 and before or at vegetative to floral transition. This is also supported by the GO-term
640 enrichment of photosynthesis related terms in cluster 4. Our tissue sections do not allow us to
641 distinguish lower and upper ridge tissues, however, the two ridges have been separately

642 collected and sequenced via LCM in barley (Thiel et al., 2021). In this barley dataset, we found
643 a higher expression of *HvVRT2 (HORVU7Hr1G036130)* and *FLOWERING LOCUS C*
644 *(HvFLC; HORVU7Hr1G054320)* in the lower ridge compared to the upper ridge, whereas
645 *HvSVP1* (*HORVU6Hr1G077300*) was also marginally more highly expressed (Supplemental
646 Figure S5). Furthermore, the barley *MND1* gene (*HORVU7Hr1G113480*) has recently been
647 shown to be expressed in leaf primordia and during the Double Ridge stage in the basal region
648 of the spike in barley (Walla et al., 2021), while it is most highly expressed in the vegetative
649 meristem and lower/leaf ridge in the LCM data (Thiel et al 2021). We observed that in our data,
650 the wheat *MND1* orthologs (*TraesCS7A02G506400*, *TraesCS7B02G413900*,
651 *TraesCS7D02G494500*) were significantly more highly expressed in the basal section than the
652 apical section at both DR and GP stage (Supplemental Table S2). The suppressed leaf ridge (or
653 bract) has been proposed to act as a signalling centre, regulating the fate of the upper spikelet
654 meristem ridge (Whipple, 2017). Insufficient bract suppression during the formation of the
655 basal spikelets might therefore negatively affect initiation and development of spikelets.

656 At DR, the widest point of the spike is indeed as expected the base and not the central section
657 (Figure 1A). The lower ridge is however much less developed in the central section and can be
658 hardly seen in the apical ridges. Interestingly, mutants failing to repress the lower ridge growth,
659 such as *third outer glume1* (*trd1*), the barley ortholog of maize *tsh1*, develop large bracts from
660 the lower ridge in basal spikelets, unlike apical spikelets, which do not develop bracts from
661 their lower ridges regardless of the absence of *TRD*. This is reminiscent of the gradient in the
662 strength of the phenotypic effects observed from the top to the base of the inflorescence in
663 multiple *Poaceae* mutants (discussed above). We therefore hypothesise that the basal
664 meristems develop into smaller spikelets and larger bract primordia due to a slow suppression
665 of "vegetative growth signals" (e.g., *SVPs*) and a concomitant slow upregulation of "floral
666 growth signals" (e.g., *SEPs*) upon floral transition. To investigate how a change from
667 vegetative to floral signalling might affect the development of individual meristems, we
668 modelled the genetic interaction of *SVPs* and *SEPs,* as proposed by Li et al. (2021), in the
669 spatial context of a growing spike (Supplemental File S1). Under the assumptions that SVP
670 supresses *SEP* expression, *SVP* expression is downregulated upon flowering, and that SEP
671 promotes spikelet outgrowth, the model could recapitulate (a) the observed opposing gradients
672 in expression of *SVPs* and *SEPs* along the spike, and (b) the formation of a lanceolate shaped
673 wheat spike with reduced spikelet elongation and stronger bract growth in the most basal

674 spikelets. Thus, whilst this hypothesis will require further investigation and testing, modelling
675 supports its plausibility.

676

677 Materials and Methods

678

679 Plant materials

680 Hexaploid wheat (*Triticum aestivum*) germplasm used in this study includes wildtype
681 hexaploid wheat cultivar Paragon and *P1/VRT2 germplasm* described in Adamski et al. (2021)
682 including *P1* NILs, recombinants, and T_1 transgenic lines carrying the *T. polonicum VRT-A2b*
683 copy under the native promoter. *T. polonicum* accession T1100002 was obtained from the John
684 Innes Centre Germplasm Resources Unit (https://www.seedstor.ac.uk/search-
685 infoaccession.php?idPlant=27422). For field experiments, we used between two to four sibling
686 BC_4/ BC_6 NILs differing for the *VRT-A2b* allele.

687

688 Low input RNA sequencing

689 Paragon seedlings were grown in a single batch in a controlled environment growth chambers
690 in 24-cell seed trays under long-day (16 h light/8 h dark) photoperiods at 300 µmol m^{-2} s^{-1},
691 with a day temperature of 20 °C and a night temperature of 15 °C. Inflorescences for Double
692 Ridge (DR) stage were collected 18 days after sowing, while inflorescences for Glume
693 Primordia (GP) stage were collected 22 days after sowing. All plants were grown in "John
694 Innes Cereal Mix" (40% Medium Grade Peat, 40% Sterilized Soil, 20% Horticultural Grit, 1.3
695 kg·m^{-3} PG Mix 14-16-18 + Te Base Fertiliser, 1 kg·m^{-3} Osmocote Mini 16-8-11 2 mg + Te
696 0.02% B, Wetting Agent, 3 kg·m^{-3} Maglime, 300 g·m^{-3} Exemptor).

697 Four individual spikes per developmental stage (DR and GP) were dissected into apical,
698 central, and basal sections (1:1:1 ratio) using a stereo microscope (Leica MZ16). Sections were
699 immediately placed into 96-well plates (on ice) containing 10 µL of RLT plus (Qiagen, Hilden,
700 Germany). All instruments and surfaces were cleaned with 80% ethanol, RNAse-free water
701 and lastly RNAse-out solution after each sample to reduce cross-contamination and RNA
702 degradation. Samples were stored at -80 °C until cDNA preparation, using the G&T-seq
703 method as previously described (Macaulay et al., 2015). cDNA was normalised to 0.2 ng/ µL
704 before Nextera (Illumina, San Diego, CA, USA) library preparation using a Mosquito HV

705 liquid handler (STP, Royston, UK) in a total reaction volume of 4 µL as described in Mora-
706 Castilla et al. (2016). Libraries were pooled by volume and sequenced on a single lane of a
707 NovaSeq 6000 (NVS200S2 flow cell, 100 bp paired-end reads).

708 **Bioinformatic analysis**

709 For the RNA-seq analysis, we used the RefSeqv1.0 genome assembly and the RefSeqv1.1 gene
710 annotation (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Assemblies/v1.0/;
711 IWGSC et al. (2018)). Reads were trimmed and adapters were removed using trim-galore
712 v.0.4.2 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with settings: "--
713 paired --fastqc --a GGTATCAACGCAGAGT --clip_R1 20 --clip_R2 20 --trim-n". Minimum
714 length of reads retained was set to 50 bp. Reads were aligned to the RefSeqv1.0 genome
715 assembly using HISAT2 v. 2.1.0 (https://daehwankimlab.github.io/hisat2/; Kim et al. (2019))
716 with the following parameters: "--pen-noncansplice 20 --mp 1,0 --rna-strandness RF".
717 Alignment files were converted to BAM format, sorted, indexed, filtered, and purged of all
718 none-primary alignments (0x100 flag) using samtools (v. 1.9; Li et al. (2009)). HTSeq v.0.6.1
719 (https://htseq.readthedocs.io/en/master/; Anders et al. (2015)) was used to count the read
720 numbers mapped to the RefSeqv1.1 gene models.

721 HT-read count normalization and differential expression analyses were performed using the
722 DESeq2 v.1.28.1 R packages
723 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html; Love et al., 2014; RStudio
724 1.2.5001). Genes with an average expression below 10 HT-count, and which were not
725 expressed (i.e. ≤ 10 HT counts) in at least three libraries, were removed from the analysis.
726 Correlation between expressed genes and Waddington stage and/or section was tested by
727 ANOVA. Raw read data from Li et al. (2018) and Feng et al. (2017) were pseudo-aligned
728 using Kallisto Sleuth pipeline (https://scilifelab.github.io/courses/rnaseq/labs/kallisto) and the
729 coefficient of variation was calculated for each gene (by condition) using R (RStudio 1.2.5001)
730 ddply (plyr 1.8.6). Differentially expressed genes (DEGs) between the two Waddington stages
731 (DR and GP) were calculated with the design "~plant + section", while DEGs between the
732 three sections (apical, central, basal) were determined using the design "~section +
733 waddington:plant + waddington". DEGs among the sections within each Waddington stage
734 were determined with the design "~plant + section". For each gene, an adjusted *P*-value was
735 computed by DESeq2 (using the using the Benjamini and Hochberg method (Benjamini and
736 Hochberg, 1995)), and those with an adjusted *P*-value of ≤ 0.05 were considered differentially
737 expressed. DeSeq2 also computed Log2FoldChanges as well as the associated uncertainty

738 (lfcSE, see Love et al. (2014) for further detail). The "contrast" function was used to determine
739 pairwise comparison *P*-values. The full set of expression data and comparisons is presented in
740 Supplemental Data Set S1. Enrichment of GO-terms was performed using the online tool
741 "PLAZA" (https://bioinformatics.psb.ugent.be/plaza; Van Bel et al. (2017)) using the
742 recommended settings, and all enriched GO-terms of Biological function (BF) and Cellular
743 Compartment (CC) were retained. In brief, PLAZA determines the overrepresentation of a
744 certain GO-term in a gene set compared to the genome-wide background frequency (= all
745 expressed genes in this experiment; submitted manually). The significance of over- or
746 underrepresentation is determined using the hypergeometric distribution and the Bonferroni
747 method is applied to correct for multiple testing. Note that enrichment folds are reported in
748 log2 fold scale. Enrichment of TF families (Genes that were annotated as TFs were obtained
749 from https://opendata.earlham.ac.uk/wheat/under_license/toronto/Ramirez-
750 Gonzalez_etal_2018-06025-Transcriptome-Landscape/data/data_tables/ (Ramirez-Gonzalez
751 et al., 2018)) and MADS-box TFs (based on Schilling et al. (2020)) was performed in R using
752 the phyper() function from stats package v.4.0.1 to test for Hypergeometric Distribution. All
753 DEGs were scaled and centred using R-base function "scale". All cluster analysis was
754 performed on scaled data using R (stats) functions kmeans and hclust, followed by visualisation
755 through pheatmap v.1.0.12 (https://cran.rstudio.com/web/packages/pheatmap/index.html).
756 Correlation to centroid cluster shape of each gene expression pattern was calculated using the
757 "cor" function from R stats.

758

759 **Quantitative real-time PCR analysis**

760 *P1* NILs were grown in controlled growth chambers in 24-cell seed trays under the same
761 conditions as used in the low input RNA-seq experiment (see above). For each biological
762 replicate, we pooled 30 inflorescences for DR stage, 15 for GP stage, and nine for Terminal
763 Spikelet stage (n = 4 biological replicates per stage). Inflorescences from NILs were dissected
764 using a stereo microscope (Leica MZ16). Inflorescences were dissected into apical, central,
765 basal and peduncle sections (1:1:1:1 ratio). At Double Ridge stage, inflorescences were only
766 dissected into apical, basal and peduncle section as the inflorescences were too small to be
767 accurately dissected into four sections for all 30 plants per biological replicate. Each section
768 was immediately placed into 1.5-mL tubes on dry ice and tubes were snap frozen in liquid
769 nitrogen as soon as all plants for the sample were collected. Samples were stored at −80°C until
770 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/hlalab/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 qRT-PCR 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 sec at 95 °C, 15 sec at 62 °C, 30 sec 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 *TaActin* as the reference gene (Uauy et al., 2006). Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) with a common calibrator so that values are comparable across genes, tissues, and developmental stages. All primers used in qRT-PCR came from Adamski et al., 2021 can be found in Supplemental Table S10.

All qRT-PCR data was normalised 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 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 Supplemental Table S6.

801 **Field experiments and phenotyping**

802 *VRT-A2* NILs were evaluated in four field experiments. Three trials were located at The Morley
803 Agricultural Foundation trials site, Morley St Botolph, UK (52°33'15.1"N 1°01'59.2"E) in
804 2017, 2018 and 2020 and one trial was sown in 2020 at the John Innes Experimental trials site
805 in Norwich, UK (52°37'50.7"N 1°10'39.7"E). In Morley (2017) we analysed two BC_4 lines of
806 *VRT-A2a* and three BC_4 lines of *VRT-A2b*. In Morley (2019) we analysed two BC_6 and one BC_4
807 line per *VRT-A2* allele and in Morley and Church Farm 2020 we analysed two BC_6 and two
808 BC_4 lines for each *VRT-A2* allele. All experiments were drilled as yield-scale plots (6 m x 1.2
809 m) and sown by grain number for comparable plant densities aiming for 275 seeds m^{-2}. The
810 trials were arranged in a randomised complete block design (RCBD) with five replicates per
811 sibling line per location. Developmental and plant architecture traits were evaluated throughout
812 the growing period. A 10-ear grab sample was collected from each plot pre-harvest for the
813 assessment of rudimentary basal spikelet (RBS) numbers and other phenotypes (recorded in
814 Adamski et al. 2021). RBS were defined as spikelets carrying no grain at maturity and counted
815 for each spike individually. To determine the differences between the $P1^{POL}$ and $P1^{WT}$ NILs,
816 we performed analysis of variance (ANOVA) on the multiple field trials phenotypic data. For
817 the analysis of individual trials, we used a two-way ANOVA including Genotype + Block
818 performed in R ('car' package version 3.0-10; RStudio 1.2.5001).

819

820 **Glasshouse phenotyping**

821 We evaluated the BC_4 NILs and BC_6F_3 recombinant lines, as well as *T. polonicum* accession
822 T1100002, under standard glasshouse conditions. 18-20 plants per genotype were grown in 1
823 L pots containing John Innes Cereal Mix under long day conditions (16 h light, 8 h dark). The
824 genotypes of all plants were confirmed using KASP marker *SP1Pol* (Adamski et al., 2021).We
825 counted the number of rudimentary basal spikelets (RBS) for all tillers of all biological
826 replicates at maturity. To evaluate the differences in RBS between genotypes, we performed a
827 two-way ANOVA analysis and post-hoc multi-pairwise comparisons Sidak test ('car' package
828 version 3.0-10; RStudio 1.2.5001).

829

830 **Phenotyping of transgenic lines**

831 T_1 lines from Adamski et al. (2021) differing for the copy number of the *VRT-A2b* transgenic
832 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).

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 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),

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

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 Supplemental Figure S6.

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

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

Data Availability

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

Acknowledgments

We thank Phil Robinson from the Scientific Photography at JIC for taking high-quality images of spikes and spikelets and Tobin Florio (http://flozbox-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.

Competing Interest

The authors declare no competing interest.

Additional Files

Supplemental Figure S1: Principal Component analysis (PCA) of RNA-seq libraries.

Supplemental Figure S2: Comparison of Coefficient of variation (CV) of gene expression in wheat RNA-seq data sets.

Supplemental Figure S3: Expression patterns of al DEGs clusters (n=7) as identified by k-means clustering.

Supplemental Figure S4: Dissected Floret 1 and 2 of basal and central spikelets of BC_6 NILs before anthesis.

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

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

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

892 **Supplemental Table S2:** Summary of normalised gene expression, statistical analyses and k-
893 means clustering for all differentially expressed genes (*Padj* ≤ 0.05).

894 **Supplemental Table S3:** Enrichment of Gene Ontology (GO)-terms in the seven identified
895 clusters of DEGs.

896 **Supplemental Table S4:** Enrichment of Transcription factor families and MADS-box
897 transcription factor genes in the seven identified clusters of DEGs.

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

900 **Supplemental Table S6:** Statistical analysis of qRT-PCR data from *VRT-A2* NIL spike
901 sections (Supplemental Table S5).

902 **Supplemental Table S7**: Field evaluations for rudimentary basal spikelets (RBS) in *VRT-A2*
903 NILs.

904 **Supplemental Table S8**: Graphical genotype from Adamski et al (2021) and RBS phenotype
905 of BC_6 recombinant inbred lines (RILs) for *VRT-A2*.

906 **Supplemental Table S9**: Rudimentary basal spikelet phenotypic data from *VRT-A2* transgenic
907 lines.

908 **Supplemental Table S10:** List of primers used in qRT-PCR in this study.

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

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

912

913

914 ## References

915 **Adamski NM, Simmonds J, Brinton JF, Backhaus AE, Chen Y, Smedley M, Hayta S, Florio T,**
916 **Crane P, Scott P, Pieri A, Hall O, Barclay JE, Clayton M, Doonan JH, Nibau C, Uauy C** (2021)
917 Ectopic expression of *Triticum polonicum VRT-A2* underlies elongated glumes and grains in hexaploid
918 wheat in a dosage-dependent manner. The Plant Cell **33**:2296-2319

919 **Alon U** (2007) An introduction to systems biology: design principles of biological circuits. Chapman
920 & Hall/CRC, London

921 **Anders S, Pyl PT, Huber W** (2015) HTSeq--a python framework to work with high-throughput
922 sequencing data. Bioinformatics **31:**166-169

923 **Becker A, Theißen G** (2003) The major clades of MADS-box genes and their role in the development
924 and evolution of flowering plants. Molecular Phylogenetics and Evolution **29:** 464-489

925 **Benjamini Y, Hochberg Y** (1995) Controlling the false discovery rate: a practical and powerful
926 approach to multiple testing. J. R. Stat. Soc. Ser. B Stat. Method. **57**: 289–300

927 **Bi X, van Esse W, Mulki MA, Kirschner G, Zhong J, Simon R, von Korff M** (2019)
928 *CENTRORADIALIS* interacts with *FLOWERING LOCUS T-like* genes to control floret development
929 and grain number. Plant Physiology **180:** 1013-1030

930 **Bommert P, Whipple C** (2018) Grass inflorescence architecture and meristem determinacy. Seminars
931 in Cell & Developmental Biology **79:** 37-47

932 **Bonnett OT** (1966) Inflorescences of maize, wheat, rye, barley, and oats: their initiation and
933 development. Bulletin (University of Illinois (Urbana-Champaign campus) Agricultural Experiment
934 Station) **721**

935 **Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S** (2006) *ramosa2* Encodes
936 a LATERAL ORGAN BOUNDARY Domain Protein That Determines the Fate of Stem Cells in Branch
937 Meristems of Maize. The Plant Cell **18:** 574-585

938 **Debernardi JM, Lin H, Chuck G, Faris JD, Dubcovsky J** (2017) microRNA172 plays a crucial role
939 in wheat spike morphogenesis and grain threshability. Development **144:** 1966-1975

940 **Eveland AL, Goldshmidt A, Pautler M, Morohashi K, Liseron-Monfils C, Lewis MW, Kumari S,**
941 **Hiraga S, Yang F, Unger-Wallace E** (2014) Regulatory modules controlling maize inflorescence
942 architecture. Genome Research **24**: 431-443

943 **Feng N, Song G, Guan J, Chen K, Jia M, Huang D, Wu J, Zhang L, Kong X, Geng S, Liu J, Li A,**
944 **Mao L** (2017) Transcriptome Profiling of Wheat Inflorescence Development from Spikelet Initiation
945 to Floral Patterning Identified Stage-Specific Regulatory Genes. Plant Physiology **174:** 1779-1794

946 **Giacomello S, Salmén F, Terebieniec BK, Vickovic S, Navarro JF, Alexeyenko A, Reimegård J,**
947 **McKee LS, Mannapperuma C, Bulone V** (2017) Spatially resolved transcriptome profiling in model
948 plant species. Nature Plants **3**: 17061

949 **Giolai M, Verweij W, Lister A, Heavens D, Macaulay I, Clark MD** (2019) Spatially resolved
950 transcriptomics reveals plant host responses to pathogens. Plant Methods **15:** 114

951 **Harrop TWR., Ud Din I, Gregis V, Osnato M, Jouannic S, Adam H, Kater MM** (2016) Gene
952 expression profiling of reproductive meristem types in early rice inflorescences by laser
953 microdissection. Plant Journal **86**: 75-88

954 **IWGSC, Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, Pozniak CJ, Choulet F,**
955 **et_al.** (2018) Shifting the limits in wheat research and breeding using a fully annotated reference
956 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). Frontiers in Plant Science **4:** 250

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

Koppolu R, Schnurbusch T (2019) Developmental pathways for shaping spike inflorescence architecture in barley and wheat. Journal of Integrative Plant Biology **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 Journal **54:** 93-105

Leiboff S, Hake S (2019) Reconstructing the transcriptional ontogeny of maize and sorghum supports an inverse hourglass model of inflorescence development. Current Biology **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 GPDP (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. 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 **14**: 1472-1488

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. Nature Methods **12**: 519-522

McFaline-Figueroa JL, Trapnell C, Cuperus JT (2020) The promise of single-cell genomics in plants. Current Opinion in Plant Biology **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, Harel TH, Tanay A, Eshed Y (2021) Dissection of floral transition by single-meristem transcriptomes at high temporal resolution. Nature 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. Journal of laboratory automation **21:** 557-567

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

1002 **Prusinkiewicz P, Erasmus Y, Lane B, Harder LD, Coen E** (2007) Evolution and Development of
1003 Inflorescence Architectures. Science **316:** 1452-1456

1004 **Ramirez-Gonzalez RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M,**
1005 **Jacobs J, van Ex F, Pasha A, Khedikar Y, Robinson SJ, Cory AT, Florio T, Concia L, Juery C,**
1006 **Schoonbeek H, Steuernagel B, Xiang D, Ridout CJ, Chalhoub B, Mayer KFX, Benhamed M,**
1007 **Latrasse D, Bendahmane A, Wulff BBH, Appels R, Tiwari V, Datla R, Choulet F, Pozniak CJ,**
1008 **Provart NJ, Sharpe AG, Paux E, Spannagl M, Brautigam A, Uauy C** (2018) The transcriptional
1009 landscape of polyploid wheat. Science **361**: eaar6089

1010 **Reinhardt D, Mandel T, Kuhlemeier C** (2000) Auxin Regulates the Initiation and Radial Position of
1011 Plant Lateral Organs. The Plant Cell **12:** 507-518

1012 **Ren D, Rao Y, Leng Y, Li Z, Xu Q, Wu L, Qiu Z, Xue D, Zeng D, Hu J** (2016) Regulatory role of
1013 OsMADS34 in the determination of glumes fate, grain yield, and quality in rice. Frontiers in plant
1014 science **7:** 1853

1015 **Rich-Griffin C, Stechemesser A, Finch J, Lucas E, Ott S, Schäfer P** (2020) Single-cell
1016 transcriptomics: a high-resolution avenue for plant functional genomics. Trends in Plant Science **25:**
1017 186-197

1018 **Schilling S, Kennedy A, Pan S, Jermiin LS, Melzer R** (2020) Genome-wide analysis of MIKC-type
1019 MADS-box genes in wheat: pervasive duplications, functional conservation and putative
1020 neofunctionalization. New Phytologist **225:** 511-529

1021 **Schurch NJ, Schofield P, Gierliński M, Cole C, Sherstnev A, Singh V, Wrobel N, Gharbi K,**
1022 **Simpson GG, Owen-Hughes T** (2016) How many biological replicates are needed in an RNA-seq
1023 experiment and which differential expression tool should you use? RNA **22:** 839-851

1024 **Sentoku N, Kato H, Kitano H, Imai R** (2005) *OsMADS22*, an STMADS11-like MADS-box gene of
1025 rice, is expressed in non-vegetative tissues and its ectopic expression induces spikelet meristem
1026 indeterminacy. Molecular Genetics and Genomics **273:** 1-9

1027 **Thiel J, Koppolu R, Trautewig C, Hertig C, Kale S, Erbe S, Mascher M, Himmelbach A, Rutten**
1028 **T, Esteban E** (2021) Transcriptional landscapes of floral meristems in barley. Science Advances **7**:
1029 eabf0832

1030 **Trevaskis B, Tadege M, Hemming MN, Peacock WJ, Dennis ES, Sheldon C** (2007) *Short*
1031 *Vegetative Phase-Like* MADS-Box Genes Inhibit Floral Meristem Identity in Barley. Plant Physiology
1032 **143:** 225-235

1033 **Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J** (2006) A NAC gene regulating senescence
1034 improves grain protein, zinc, and iron content in wheat. Science **314:** 1298-1301

1035 **Van Bel M, Diels T, Vancaester E, Kreft L, Botzki A, Van de Peer Y, Coppens F, Vandepoele K**
1036 (2017) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant
1037 genomics. Nucleic Acids Research **46:** D1190-D1196

1038 **Waddington S, Cartwright P, Wall P** (1983) A quantitative scale of spike initial and pistil
1039 development in barley and wheat. Annals of Botany **51:** 119-130

1040 **Walla A, Wilma van Esse G, Kirschner GK, Guo G, Brünje A, Finkemeier I, Simon R, von Korff**
1041 **M** (2020) An Acyl-CoA *N*-Acyltransferase Regulates Meristem Phase Change and Plant Architecture
1042 in Barley1. Plant Physiology **183:** 1088-1109

1043 **Wang Y, Yu H, Tian C, Sajjad M, Gao C, Tong Y, Wang X, Jiao Y** (2017) Transcriptome
1044 Association Identifies Regulators of Wheat Spike Architecture. Plant Physiology **175:** 746-757

1045 **Whipple CJ** (2017) Grass inflorescence architecture and evolution: the origin of novel signaling
1046 centers. New Phytologist **216:** 367-372

1047 **Whipple CJ, Hall DH, DeBlasio S, Taguchi-Shiobara F, Schmidt RJ, Jackson DP** (2010) A
1048 Conserved Mechanism of Bract Suppression in the Grass Family The Plant Cell **22:** 565-578

1049 **Wu D, Liang W, Zhu W, Chen M, Ferrándiz C, Burton RA, Dreni L, Zhang D** (2018) Loss of
1050 LOFSEP transcription factor function converts spikelet to leaf-like structures in rice. Plant Physiology
1051 **176:** 1646-1664

1052 **Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu Z, Wang L, Fox N, Wang**
1053 **X, Drenkow J, Luo A, Char SN, Yang B, Sylvester AW, Gingeras TR, Schmitz RJ, Ware D, Lipka**
1054 **AE, Gillis J, Jackson D** (2021) Single-cell RNA sequencing of developing maize ears facilitates
1055 functional analysis and trait candidate gene discovery. Developmental Cell **56**: 557-568.e556

1056 **Youssef HM, Eggert K, Koppolu R, Alqudah AM, Poursarebani N, Fazeli A, Sakuma S, Tagiri**
1057 **A, Rutten T, Govind G, Lundqvist U, Graner A, Komatsuda T, Sreenivasulu N, Schnurbusch T**
1058 (2017) VRS2 regulates hormone-mediated inflorescence patterning in barley. Nature Genetics **49**: 157-
1059 161

1060 **Zhao T, Ni Z, Dai Y, Yao Y, Nie X, Sun Q** (2006) Characterization and expression of 42 MADS-box
1061 genes in wheat (*Triticum aestivum* L.). Molecular Genetics and Genomics **276:** 334