Short title: Expression analyses in isolated cowpea reproductive cell-types.

Author for Contact: Anna M.G. Koltunow

Article title:

Gene expression in reproductive cell-types isolated in temporal sequence from meiosis to early seed initiation in cowpea (*Vigna unguiculata* L. Walp)

Author names and affiliations:

Nial Gursanscky[1], Danielle Marzurkiewicz[1], Martina Juranić[1], Susan D. Johnson[1], Gloria León[2], Rocio Escobar-Guzmán[2], Rigel Salinas-Gamboa[2], Itzel Amasende-Morales[2], Matteo Riboni[1], Melanie Hand[1],' Andrew Spriggs[3], Jean-Philippe Vielle-Calzada[2] and Anna M.G. Koltunow[1,4]

1: Commonwealth Scientific and Industrial Research Organisation (CSIRO) Agriculture and Food, Waite Campus, Urrbrae, South Australia 5064, Australia

2: Group of Reproductive Development and Apomixis, UGA Laboratorio Nacional de Genómica para la Biodiversidad, CINVESTAV Irapuato, Guanajuato 36821, México.

3: CSIRO Agriculture and Food, Black Mountain Science and Innovation Park, Canberra, ACT 2601, Australia.

Author Contributions:

D.M. isolated LCM cell types, and obtained transcriptomes; N.G. conducted data analyses identifying expression profiles and developmental trends; M.J. and S.D.J. contributed to cell

cycle and transgenic cowpea analyses; G.L., R.E-G and R.S-G. performed *in situ* hybridizations; M.R verified gene expression by qRT-PCR; M.H. prepared constructs; I.A-M. carried out promoter characterization in *Arabidopsis*; A.S. assembled transcriptomes, carried out global alignments of RNA-seq to genomic sequences ; J-P.V-C. and A.M.G.K. designed experiments and analyzed data. N.G., J-P.V-C. and A.M.G.K wrote the paper. A.M.G.K. agrees to serve as the author responsible for contact and ensures communication.

Funding information: This work was funded by a grant from the Bill and Melinda Gates foundation to the Commonwealth Scientific and Industrial Research Organisation (CSIRO) (OPP1076280).

Present address:

[4.] Queensland Alliance for Agriculture and Food Innovation (QAAFI) Centre for Crop Science. The University of Queensland. St Lucia. Queensland 4072. Australia.

Email address of author for contact: a.koltunow@uq.edu.au

One sentence summary

Transcriptomes from cowpea reproductive cell-types isolated in temporal sequence from meiosis to seed initiation revealed cell-type gene expression profiles enabling comparison with gene expression pathways known to function in cell specification and reproductive development in other plants.

Abstract

Molecular knowledge of pathways regulating seed formation in legumes, remains scarce. Thirteen isolated cell-type transcriptomes were developed, spanning temporal events of male and female gametogenesis and seed initiation, to examine pathways involved in cowpea seed formation. *In situ* hybridization confirmed localization of *in silico* identified cell-specific genes, verifying transcriptome utility. Cowpea and *Arabidopsis* reproductive cells showed some conservation in regulators enabling cell-type expression as some cowpea cell-specific genes promoters and their *Arabidopsis* homologs directed expression to identical reproductive cell-types in transgenic plants. *In silico* analyses revealed gene expression similarities and differences with genes in pathways regulating reproductive events in other plants. Meiosis-related genes were expressed at mitotic stages of gametogenesis and during sporophytic development in cowpea. Plant hormone pathways showing preferential expression at particular reproductive stages were identified. Expression of epigenetic pathways, resembling those found in *Arabidopsis,* including microRNA mediated gene silencing, RNA directed DNA methylation and histone modification were associated with particular stages of male and female gametophyte development, suggesting roles in gametogenic cell specification and elaboration. Analyses of cell-cycle related gene expression in mature cowpea female gametophytes, indicated that the egg and central cell were arrested at the G1/S and G2/M cell cycle phases, respectively, prior to fertilization. Pre-fertilization female gametophyte arrest was characterized by barely detectable auxin biosynthesis gene expression levels, and elevated expression of genes involved in RNA-mediated gene silencing and histone modification. These transcriptomes provide a useful resource for additional interrogation to support functional analyses for development of higher yielding cowpea and syntenic legume crops.

70 **Introduction**

71 Legumes contribute directly and indirectly to world food supply. Higher in protein, they
72 provide forage for animals and increase soil fertility by fixing nitrogen. Agronomic
73 improvement of eight grain legumes, including cowpea (*Vigna unguiculata* L. Walp) has
74 been targeted by the Consultative Group for International Agricultural Research
75 (https://storage.googleapis.com/cgiarorg/2018/11/SC7-B_Breeding-Initiative-1.pdf). We
76 aim to increase seed yield and quality in cowpea, which originated in Africa and remains an
77 important subsistence crop in sub-Saharan Africa (Singh, 2014). Although cowpea is
78 relatively tolerant to drought, it is prone to high levels of flower and pod drop, resulting in
79 decreased seed yields. There is limited knowledge of the molecular pathways supporting
80 successful elaboration of cowpea male and female gametophytes and seeds.

81

82 Current genomic resources for cowpea include a reference genome (Lonardi et al., 2019,
83 Munoz-Amatriain et al., 2017) together with survey genomes and transcriptomes from
84 various plant tissues (Spriggs et al., 2018, Yao et al., 2016). Reproductive events occur only
85 in a few cell-types buried within cowpea floral organs (Salinas-Gamboa et al., 2016).
86 Therefore, laser capture microdissection (LCM) is a useful approach to isolate reproductive
87 cell-types to generate informative transcriptomes. Interrogation of these reproductive cell-
88 type transcriptomes in conjunction with cowpea genome and tissue transcriptome
89 resources could accelerate identification and functional examination of pathways
90 supporting successful cowpea seed formation.

91

92 A cowpea reproductive calendar has been developed linking morphological floral features in
93 cowpea to reproductive events in floral tissues to support collection of tissues for molecular

94 analyses (Salinas-Gamboa et al., 2016). Male gametophyte precursor cells, termed pollen
95 mother cells (PMC) undergo meiosis (sporogenesis) forming tetrads of microspores (mTET)
96 in the anther. Individual microspores (MIC) then undergo mitosis (microgametogenesis) and
97 differentation to form the mature pollen grain (MP-SC) containing a generative and a
98 vegetative cell (Fig. 1A). By contrast, during female gametophyte development, a single
99 megaspore mother cell (MMC) in each ovule undergoes meiosis to form a tetrad of
100 megaspores (fTET). Three megaspores degenerate and the surviving functional megaspore
101 (FM) undergoes three rounds of mitosis to form a syncytium of eight nuclei
102 (megagametogenesis). Cellularization events result in a 7-celled *Polygonum*-type female
103 gametophyte, or embryo sac, and three of these degrade. The mature cowpea embryo sac
104 (MES) contains an egg cell (EC) progenitor of the embryo (Em), flanked by two synergids
105 (Sy), which guide fertilization events and the central cell (CC) containing two fused haploid
106 nuceli, which is the progenitor of the endosperm (En; Fig. 1A; Salinas-Gamboa et al., 2016).
107
108 The events of fertilization initiate with pollen grain germination and emergence of the
109 pollen tube on the stigma. We have observed that mitosis of the generative cell to form the
110 two sperm cells predominantly occurs after pollen tube germination (see Results). Pollen
111 tube growth through maternal tissues into the ovule, synergid puncture and pollen tube
112 burst releases both sperm cells into the embryo sac. One sperm cell fuses with the egg to
113 initiate embryogenesis and the other fuses with with the polar nucleus of the central cell to
114 produce the triploid endosperm. Endosperm tissue is transient in copwpea as it is utilized
115 during embryo development (Salinas-Gamboa et al., 2016); Fig. 1A).
116

117 Studies in *Arabidopsis*, rice, maize, *Boechera* and *Hieracium* have used LCM to isolate cells
118 involved in sexual and asexual (apomictic) reproductive events to examine cell-type specific
119 gene expression (Anderson et al., 2013, Kubo et al., 2013, Zhan et al., 2015, Belmonte et al.,
120 2013, Okada et al., 2013). These analyses have typically focused on the expression in cells
121 involved in specific aspects of the reproductive pathway due to the biological questions
122 being addressed, and potentially, the laborious nature of the method. To our knowledge, a
123 transcriptome series spanning reproductive events in isolated cells from the initiation of
124 meiosis to seed initiation has not been generated in a species. This would enable direct
125 comparisons of gene expression during male and female gametogenesis and seed
126 development within the plant and also with pathways known to functionally regulate cell
127 specification and reproductive development in other plants.
128
129 Here, we used LCM and additional tissue isolation procedures, to develop a suite of 13
130 cowpea reproductive cell-type transcriptomes spanning the temporal events of male and
131 female gametophyte development, and early seed initiation in cowpea. *In vitro* and *in*
132 *planta* analyses were used to verify transcriptome integrity. Following an analysis of global
133 gene expression profiles across the transcriptome set, an *in silico* analysis of the expression
134 of genes involved in epigenetic regulation, hormone biosynthesis and signal transduction,
135 and cell cycle progression was undertaken as these pathways regulate reproductive aspects
136 in other angiosperms. These analyses revealed commonalities between cowpea and other
137 species with respect to some of these pathways, in addition to novel cowpea-related
138 differences identifying genes and pathways as candidates for future functional testing.

139 **Results**

140 **Cowpea IT86D-1010 leaf and reproductive cell-type transcriptomes and PCR verifications**

141 A total of 13 reproductive transcriptomes were generated, in duplicate, using RNA extracted
142 from individual reproductive cells that span male and female gametogenesis, and early seed
143 initiation. Laser capture microdissection was used to generate 12 reproductive cell-type
144 transcriptomes, while pollen and sperm cells were sampled separately (see below, Fig. 1A).
145 Female gametophyte-related cell-type transcriptomes were generated for the megaspore
146 mother cell (MMC) during meiosis I, the tetrad of post-meiotic megaspores (fTET),
147 developing mitotic embryo sacs with two or four nuclei (ES2n, ES4n), and the mature
148 embryo sac (MES) containing the egg cell (EC), the flanking synergids (Sy), and the central
149 cell (CC), at the time of stamen emergence (anthesis). Individual EC and CC transcriptomes
150 were also made. The early embryo (Em) transcriptome comprised tissue from developing
151 zygotes to early globular stages of cowpea embryogenesis (Fig 1A).

152

153 Male gametophyte-related transcriptomes were generated for pollen mother cells at early
154 and later stages of meiosis I (PMC.E, PMC.L), microspore tetrads (mTET), and uninucleate
155 microspores (MIC). A pure sperm cell transcriptome was not generated. Analyses of cleared
156 anthers indicated that 9.7% of mature pollen grains contained two sperm cells (n=500),
157 indicating the final generative cell mitosis giving rise to the sperm cells primarily occurs
158 following pollen tube germination. This is also observed in soybean and 70% of flowering
159 plant species (Brewbaker, 1967, Williams et al., 2014, Haerizadeh et al., 2009,
160 Wojciechowski et al., 2004). Anthers and pistils harvested at anthesis were processed to
161 create a transcriptome referred to as mature pollen-sperm cell (MP-SC; see Methods for

162 details). An additional transcriptome generated from young expanding cowpea leaves
163 served as a control comparison for gene expression in cells not undergoing reproductive
164 events.
165
166 A total of 74,839 genes were predicted in the IT86D-1010 genome (Spriggs et al., 2018)
167 using Augustus gene prediction software when *Arabidopsis* genes were used as the training
168 data set (Supplemental Data 1). Transcriptome sequences were aligned with the predicted
169 genes in the IT86D-1010 cowpea genome to identify common and uniquely expressed genes
170 across all datasets (Supplemental Table 1, Supplemental Data 2). The total number of
171 uniquely expressed genes within each transcriptome ranged from 22,561 in the female
172 megaspore tetrad (fTET) to 60,359 in the MP-SC sample (Fig. 1B). Only 3,769 of the genes
173 expressed in the MP-SC transcriptome were not detected in at least one other cell-type. As
174 a likely consequence of its isolation method, the high transcript diversity in the MP-SC
175 transcriptome contrasts with that found in pollen and sperm cell transcriptomes of
176 *Arabidopsis*, rice and maize, where the number of genes detected is reduced compared to
177 sporophytic cell-types (Rutley and Twell, 2015).
178
179 Gene expression across all of the transcriptomes was normalized using reads per million
180 (RPM) given the absence of a correlation between gene length and read alignment number.
181 Investigation of unique gene expression levels within each transcriptome indicated that the
182 majority of unique genes (68-85%) were expressed at very low to moderate levels (<1 to 5-
183 20 RPM; Fig 1C). Principal component analysis (PCA) using the individual transcriptome
184 replicates showed that most replicates clustered together. The transcriptomes from male
185 and female reproductive cell-types co-segregated forming two distinct male and female

186 clusters. The embryo transcriptomes associated with the female reproductive transcriptome
187 cluster, while the leaf and MP-SC transcriptomes were outliers (Supplemental Fig.2).
188
189 The quality of the generated transcriptomes was evaluated by examining if cell-type specific
190 genes could be identified via *in silico* methods in the transcriptomes, and then by verifying
191 gene expression by qRT-PCR, *in situ* and transgenic approaches. Candidate cell-type specific
192 genes were initially identified using rsgcc software in the MMC, PMC, EN2n, En4n, MES, CC,
193 ES, Em and sperm transcriptomes (Fig. 1A). Candidate genes identified ranged in number
194 from 126 in early pollen mother cells to 25,986 in the MP-SC transcriptome (Supplemental
195 Data 3 and 4). A filtering approach based on identification of genes expressed at least five
196 times higher in the CC, EC or MP-SC transcriptomes relative to all others resulted in 10, 34
197 and 31 candidate specific genes, respectively (Supplemental Table 2).
198
199 Two of the cowpea egg cell-specific candidates identified using the filtering approach were
200 homologues of the *Arabidopsis* egg cell-specific genes *AtEC1.1*, (Sprunck et al., 2012) and
201 *AtRKD1* (Koszegi et al., 2011) (*VuEC1.1_1* and *VuRKD1*; Table 1). Therefore, cowpea
202 homologues of 11 additional *Arabidopsis* reproductive cell-type expressed genes were
203 identified for verification analyses. They included *VuKNU-like1* and *2*, putative cowpea
204 homologs of the *Arabidopsis KNUCKLES* gene expressed during male and female
205 gametogenesis (Payne et al., 2004, Tucker et al., 2012) and genes expressed in specific male
206 and female gametophytic cell-types (Table 1; Supplemental Table 3). Twenty of the 28 (71%)
207 candidate cell-type specific genes tested by qRT-PCR, were confirmed as significantly
208 enriched using cDNA isolated from tissues at comparable developmental stages from which

209 the LCM samples were obtained in addition to other tissue types (Table 1, Supplemental
210 Table 4).
211
212 **Verification of cell-type specific expression using *in situ* hybridization and plant**
213 **transformation**
214 The expression of 11 candidate cell-type specific genes verified using qRT-PCR analyses
215 above was examined by *in situ* hybridization (ISH), using gene-specific digoxygenin-labelled
216 probes (Fig. 2 and Supplemental Fig. 4). ISH was performed on mature pollen grains to
217 examine expression of four candidate sperm-cell specific genes, as we had previously
218 determined that approximately 10 percent of mature pollen grains contain two sperm cells
219 (Table 1). *VuXTH6* mRNA was detected at high levels in the sperm cells and in the vegetative
220 nucleus, with low levels detected in the vegetative cell cytoplasm (Fig. 2A-C). By contrast,
221 mRNA for the other three examined genes was detected in the sperm cells, the vegetative
222 nucleus and at high levels in the vegetative cell cytoplasm (Fig. 2D-F and Supplemental Fig.
223 4A-F).
224
225 A combination of whole-mount and section-based *in situ* hybridization was used in
226 developing cowpea ovules used to examine expression of seven genes during female
227 gametogenesis. Transcripts of *VuKNU-like1*, the candidate cowpea homolog of the
228 *Arabidopsis KNUCKLES* gene were enriched in the megaspore mother cell; however,
229 expression was also observed in nucellar and inner integumentary ovule cells (Fig 2G).
230 Transcripts from the candidate egg cell specific *VuEC1.1_1* and *VuRKD1* genes (Table 1) were
231 highly enriched in the egg cell, in comparison to the rest of the embryo sac and ovule. In
232 both cases, mRNA localization was also evident in the synergids, but not in the central cell

233 (Fig. 2H; Supplemental Fig. 4G-J). Three of the four candidate central cell genes tested by *in*
234 *situ* hybridization, displayed specific mRNA localization in the central cell (*VuSDR1,*
235 *VuGULL06*; Table 1; Supplemental Fig. 4N and O, Fig 2I). *VuNB-ARC,* mRNA was evident in a
236 nuclear dot patterning in the cowpea central cell nucleus (Fig. 2I). Nuclear dots are tightly
237 associated with nascent transcripts of actively expressed genes in the central cell of
238 *Arabidopsis* (Vielle-Calzada et al., 1999), and in embryonic cells of *Drosophila melanogaster*
239 (Shermoen and O'Farrell, 1991). *VuMEE23* mRNA was localized in the central cell and in
240 adjacent ovule cells (Table 1).
241
242 As eleven of the cowpea genes predicted to be expressed in specific reproductive cell- types
243 showed sequence similarity to previously described *Arabidopsis* cell-type expressed genes,
244 cowpea plants were transformed with constructs comprising *Arabidopsis* (*At*) promoters of
245 these genes fused to fluorescent reporters to examine conservation of targeting to specific
246 reproductive cell-types (Table 1; Supplemental Table 3). Six of the tested promoters
247 directed similar spatial and temporal patterns in cowpea as observed in *Arabidopsis*. The
248 remaining five showed either non-specific or no detectable expression in transgenic cowpea
249 (Table 1; Fig. 2 and Supplemental Fig. 4). The *AtKNU* promoter directed expression to the
250 MMC and PMC cells, as published previously (Tucker et al., 2012) but expression was also
251 evident in the chalazal region of the cowpea ovule integument (Supplemental Fig. 5F). *In situ*
252 hybridization using probes to detect native cowpea *VuKNU* expression confirmed expression
253 in the ovule integument (Supplemental Fig.4K). The *AtDD45* promoter directed expression
254 to the egg cell (Fig. 2J) and both *AtDD9* and *DD25* promoters enabled central cell expression
255 (Fig. 2K). The *AtEC1.1* and *RKD2* promoters directed reporter expression to egg cells of

256 transgenic cowpea (Fig. 2K-M). Expression of fluorescent reporters from *AtRKD2* and *DD45*
257 promoters was also observed in early cowpea embryos (Fig. 2N-O).
258
259 Given the low efficiency and lengthy nature of the cowpea transformation protocol used,
260 we also examined if upstream promoter regions of 11 of the cowpea candidate cell-type
261 specific genes identified from *in silico* analyses would direct the same cell-type expression in
262 *Arabidops*is. Genomic fragments of 0.8 to 3.2 kb, corresponding to their upstream putative
263 promoter regions were fused to the *uidA* (GUS) reporter gene and transformed into the
264 *Arabidopsis* Columbia (*Col*-0) ecotype. For each transcriptional fusion, at least 15
265 independent transformants were recovered and histologically analyzed to examine gene
266 expression patterns during male and female gametogenesis. GUS activity was not detected
267 in reproductive organs of *Arabidopsis* in seven of the 11 tested constructs. However, a 1,429
268 bp region upstream of *VuKNU-L1* drove GUS expression in the *Arabidopsis* MMC and
269 adjacent sporophytic ovule cells, in a pattern similar to the *VuKNU-L1* mRNA ISH localization
270 pattern in cowpea (Supplemental Fig. 6A and Supplemental Table 21). The promoter
271 fragments from the *pVuXTH32* and *pVuGIM2* genes drove spatial expression in both pollen
272 sperm and vegetative cells as observed in cowpea ISH experiments (Supplemental Fig. 6F-G).
273 By contrast, the promoter of *pVuRKD1* drove expression in the egg cell, and in both sperm
274 and vegetative cells in 5 of the 15 *Arabidopsis* transformants analyzed (Supplemental Fig.
275 6B-D). In seven of the transformants, expression was seen only in sperm cells (Supplemental
276 Fig. 6C). Finally, the promoter of the cowpea gene *VuGULL06*, a candidate selected for
277 potential central cell expression drove expression in the *Arabidopsis* egg cell and synergids
278 (Supplemental Fig. 6E), suggesting that in some cases cowpea promoters can be influenced
279 by heterologous factors intrinsic to the *Arabidopsis* genome.

280

281 The promoter analyses indicated that there is some conservation in regulators enabling
282 reproductive cell-type specific expression in cowpea and Arabidopsis given the conservation
283 of gene expression directed to common cell types in transgenic studies, in some cases.
284 Taken together, these verification analyses collectively indicted that the developed
285 transcriptomes were a reliable resource for further *in silico* interrogation of gene expression
286 profiles defining specific reproductive cell-types during male and female gametogenesis and
287 early cowpea embryogenesis.

288 **Enriched biological processes in cowpea reproductive transcriptomes**

289

290 Differentially expressed (DE) genes were identified in each transcriptome type to examine
291 gene expression changes occurring during the temporal events of male and female
292 gametogenesis and early embryogenesis (Supplemental Data 4). Percentages of
293 differentially expressed genes varied widely from 3.4% to 25% amongst the transcriptomes
294 (Supplemental Table 5). Comparisons between the cowpea female reproductive cell-type
295 transcriptomes generated using LCM and those previously generated using LCM in
296 *Arabidopsis* revealed low levels of common genes. Comparisons were made between the
297 cowpea data set and those generated in *Arabidopsis* for the MMC (Schmidt et al., 2011),
298 ovule nucellus, early mitotic female gametophytes (Tucker et al., 2012) and the egg and
299 central cell (Wuest et al., 2010, Steffen et al., 2007). The highest similarity, with a 24%
300 overlap, was observed between the cowpea MMC dataset and the *Arabidopsis* female
301 nucellus plus MMC data set (Tucker et al., 2012); Supplemental Table 6). The low levels of

302 common genes observed may reflect differences in species-specific genes, staging, and
303 methods used to generate transcriptomes and identify expressed genes.
304
305 To obtain an overview of the major biological processes occurring during gametogenesis and
306 seed initiation, each transcriptome was initially surveyed by examining significantly enriched
307 gene ontology (GO) terms in the biological processes' category (Fig. 3A- B). Examination of
308 the control expanding leaf transcriptome showed that it was, unsurprisingly, enriched in GO
309 terms related to light transduction, chlorophyll, photosynthesis, and sugar metabolism (Fig.
310 3). Each cowpea reproductive transcriptome was characterized by a relatively distinct set of
311 significantly enriched biological process terms, supporting the cytologically distinct nature of
312 the cell-types collected (Fig. 3A-B).
313
314 Meiosis and recombination, and cell wall biogenesis terms, which characterize the
315 functional and cytological events of male and female sporogenesis in angiosperms, were
316 found in cowpea meiotic cell transcriptomes. Terms related to a range of pathways involved
317 in epigenetic regulation were also prevalent in transcriptomes related to the elaboration of
318 the female and male gametophytes (Fig. 3A-B). Regulation by small RNAs and global
319 changes in DNA and histone methylation modifications influence determination, selection
320 and development of gametophytic cells (Tucker et al., 2012, Olmedo-Monfil et al., 2010).
321 They prevent transposon activation during male and female gametogenesis (Slotkin et al.,
322 2009, Schoft et al., 2011, Pillot et al., 2010, Ibarra et al., 2012, Martinez et al., 2016). In
323 *Arabidopsis*, Histone H3 methylation (at lysine (K) 27) by the *Polycomb*-Group protein
324 repressive complex 2 (PRC2), results in gene silencing and regulates *Arabidopsis* male
325 gametophyte formation and female gametophyte arrest, pre-fertilization (Mozgova and

326 Hennig, 2015, Grossniklaus and Paro, 2014, Wang and Kohler, 2017, Kawashima and Berger,
327 2014).
328
329 The plant hormone auxin is involved in a range of reproductive events in angiosperms,
330 including initiation of seed formation and embryo patterning. Auxin-related terms were very
331 evident in this global analysis. Interestingly, the GO enrichment analysis suggested potential
332 involvement of other plant hormone pathways at particular cowpea developmental stages.
333 For example, the abscisic acid signalling pathway featured in the female MMC and male late
334 PMC transcriptomes. Brassinosteroid-related terms were evident in dividing embryo sacs
335 with 2 and 4 nuclei, and terms related to jasmonic acid biosynthetic processes and
336 responses to cytokinin were evident in the mature cowpea embryo sac (Fig. 3A and
337 Supplemental Data 6).
338
339 In order to gain greater understanding of reproduction-related gene expression pathways
340 that cowpea may, or may not have in common with other angiosperms, we profiled the
341 transcriptome set for the expression of cowpea homologues of *Arabidopsis* genes involved
342 in meiosis, auxin biosynthesis and signal transduction, epigenetic pathways and cell-cycle
343 progression. The results of these analyses, in the sections below, collate examined gene
344 expression profiles during male and female gametogenesis, pre-fertilization female
345 gametophyte arrest and early embryogenesis.
346
347 **Gene expression during temporal events of cowpea male gametophyte formation**
348 The expression of 18 cowpea meiosis-related gene homologues during male gametogenesis
349 is shown in Fig. 3C. The profiled genes have putative functions in chromosome pairing,

350 homologous recombination, and subsequent meiotic events (see Supplemental Table 7 for
351 functional information and references). A number of these genes were clearly expressed in
352 non-meiotic cell-type transcriptomes and in leaf (Fig. 3C), consistent with expression data
353 from other species (Klepikova et al., 2016).
354
355 Analyses of the expression of 276 auxin biosynthesis and signal transduction pathway genes
356 in male gametogenesis showed high levels of the cowpea *YUCCA6*-like auxin biosynthesis
357 gene in transcriptomes of cowpea late pollen mother cells, tetrads and microspores (Fig. 4A,
358 Supplemental Table 9). This reflects the requirement of *YUCCA6* and auxin production in
359 corresponding cells for viable *Arabidopsis* pollen development (Yao et al., 2018). Cowpea
360 homologs of the *Arabidopsis ARF19*, *AIR12* and *DRB1* auxin-responsive genes were also
361 upregulated during male gametophyte formation (Fig 4B). In *Arabidopsis*, ARF19 enables
362 auxin-activated transcriptional responses to phosphate starvation (Huang et al., 2018).
363 AIR12 is thought to link auxin signal transduction and reactive oxygen signalling (Gibson and
364 Todd, 2015) and DRB1 is involved in cleavage to produce miR399 phosphate homeostasis
365 regulation (Pegler et al., 2019).
366
367 Cowpea *ARGONAUTE* (*AGO*) genes were identified because AGO proteins are central to the
368 small RNA-mediated processes of Post-transcriptional Gene Silencing (PTGS), microRNA
369 (miRNA) mediated gene silencing and RNA-directed DNA methylation (RdDM). The 13
370 cowpea *AGO*-like genes identified showed close phylogenetic relationships with known
371 *AGOs* in other plants, although *AGO8* and *AGO9* were not evident in the cowpea genome
372 (Supplemental Fig. 7).
373

374 Analyses of cowpea homologs of genes required for the 21 nt producing *Arabidopsis*
375 microRNA pathway showed that *AGO1*, *HEN1*, *HYL1* and *HST1* were co-expressed during
376 meiotic and mitotic stages of male gametophyte development. This suggested potential
377 conservation of this pathway with a functional role during cowpea male gametophyte
378 elaboration (Fig. 6A).
379
380 By contrast, the case for a functional *Arabidopsis*-like PTGS pathway operating during male
381 gametogenesis appeared doubtful. The *Arabidopsis* PTGS pathway typically produces 21-24
382 nt small interfering (si) RNAs. It requires the function of AGO1, RDR6, SGS3 and DCL4 as core
383 protein components. However, AGO2, AGO7, and DCL2 can act with partial redundancy with
384 their respective paralogs (Borges and Martienssen, 2015). Homologs of all core *Arabidopsis*
385 PTSG components including two *AGO2*, two *AGO7*, four *AGO10* and three *DCL2*-like genes
386 were identified in the cowpea genome. All of the core components required for *Arabidopsis*-
387 like PTGS were expressed during male gametogenesis except for *RDR6*, which was
388 undetectable or expressed at very low levels in all male cell-transcriptome types and at low
389 levels in the mature pollen-sperm cell sample. By contrast, moderate *RDR6* expression was
390 evident together with similar or higher expression of the PTSG components in the leaf
391 transcriptome (Fig. 6A and Supplemental Table 10 and 11). The typical *Arabidopsis*-like PTSG
392 pathway may not be active during cowpea male gametogenesis.
393
394 The *Arabidopsis* RdDM pathway results in the production of 24nt siRNAs and site-specific *de*
395 *novo* methylation at their target DNA sites by DNA methyltransferases MET1 and DRM2
396 (Borges and Martienssen, 2015). Genes required for RdDM including *AGO4*, *AGO5*, *DCL3*,
397 *RDR2*, *NRPD1* and *NRPE1* (encoding the largest subunits of POL IV and V respectively) were

398 detected at various stages during male gametogenesis. Transcript levels of *NRPD1*, *RDR2*
399 and *DCL3* were relatively low. However, highest transcript levels of DNA methyltransferase
400 *MET1*, *DRM2_1* and *DRM2_2* were evident during male meiosis, decreasing during the
401 mitotic events of gametogenesis (Fig. 6B and Supplemental Table 11). This suggests an
402 *Arabidopsis*-like RdDM pathway is likely to be functional during cowpea male
403 gametogenesis.
404
405 Other DNA methyltransferases including CMT2 and CMT3 DNA can maintain CHG/CHH and
406 CHG methylation, respectively, independently from the RdDM pathway in *Arabidopsis*.
407 Cowpea homologs of both genes were strongly expressed during male meiosis and mitosis
408 (Fig. 5B). Interestingly, high transcript levels from the DNA glycosylase *DME* that mediates
409 DNA demethylation by removing methyl-cytosine and replacing it with unmethylated
410 cytosine in *Arabidopsis* (Penterman et al., 2007) were found during early cowpea PMC
411 meiosis (Fig. 6B and Supplemental Table 13). Collectively, these data suggest multiple
412 pathways inducing dynamic DNA methylation changes may operate during cowpea male
413 gametogenesis.
414
415 Changes in chromatin histone composition are known to occur during pollen development
416 in *Arabidopsis*. A total of 44 cowpea histone genes belonging to the core histone families
417 H2A, H2B, H3 and H4 were identified in cowpea (Talbert et al., 2012); Fig. 6A and
418 Supplemental Table 12), and their expression profiled during male gametogenesis. A
419 prevalence of H3.3 histone transcript variants are evident in mature *Arabidopsis* pollen and
420 H3.1 variants were absent (Borg and Berger, 2015). Analysis of histone gene transcription
421 during cowpea male gametogenesis showed high levels of the core histone H3.3 transcripts

422 throughout male gametogenesis, including the mature pollen-sperm cell sample. H3.1_3
423 histone transcript levels were also detected during male gametogenesis but were
424 undetectable in the mature pollen-sperm cell sample, as was a range of other cowpea
425 histone transcripts (Fig. 6A). These data suggest conservation of transcriptional patterns of
426 H3.3 and 3.1 expression during pollen maturation in cowpea and *Arabidopsis*.
427
428 Epigenetic processes involving the Polycomb-group repressive complex 2 (PRC2), which
429 applies a repressive trimethylation mark on lysine 27 of histone H3 (H3K27me3) regulate a
430 number of events in the plant lifecycle. Proteins forming the PRC2 complex vary in relation
431 to these functional contexts (Fig. 6B). In *Arabidopsis*, the PRC2 VRN complex, containing
432 CLF/SWN, VRN2 and FIE proteins, functions during vegetative development and the floral
433 transition, whereas the EMF complex, containing FIE, MSI1, CLF/SWN and EMF2, functions
434 during inflorescence development and flower organogenesis. The PRC2 FIS complex,
435 containing MEA/SWN, FIS2, FIE and MSI1, functions to silence target genes during early seed
436 development and gametogenesis (Pien and Grossniklaus, 2007). Cowpea homologs of
437 *Arabidopsis FIE, MSI1, EMF2, VRN5, VIN3* and *EMF1* genes were co-expressed during male
438 gametogenesis. (Fig. 6C). High levels of a cowpea *SWINGER*-like (*SWN*-like) transcript were
439 observed during late pollen mother cell meiosis and microspore formation. The expression
440 patterns of the PRC2 complex-like genes during male gametogenesis in cowpea suggest
441 there is a potential for the formation and function of novel, possibly multiple PRC2
442 complexes resembling modified versions of the EMF and VRN PRC2 types during male
443 gametogenesis.
444

445 Jumonji-domain proteins (JMJ) function to demethylate histones, antagonizing the
446 establishment of histone methylation. *JMJ706_1*, *JMJ16* and *JMJ25_1* and *JMJ25_5*-like
447 genes, potentially encoding proteins which demethylate histone methylation marks, other
448 than H3K27 (Fig. 6D) were expressed at elevated levels during cowpea male gametogenesis
449 (Fig. 6D). In rice, JMJ706 regulates transcription factor expression required for spikelet
450 development, floral morphology and organ number (Sun and Zhou, 2008) *Arabidopsis* JMJ16
451 represses leaf senescence (Liu et al., 2019) and JMJ25 functions in poplar to remove
452 H3K9me2 marks and modulate anthocyanin biosynthesis (Fan et al., 2018).
453
454 **Gene expression during temporal events of cowpea female gametophyte formation**
455
456 Meiosis-related genes were expressed at a number of stages during female gametophyte
457 development as also observed during male gametogenesis. Transcript levels of meiosis
458 genes were evident at lower levels throughout female gametogenesis (Fig. 3C). Auxin
459 biosynthesis gene transcripts were detected from *YUCCA* and *TAR*-like genes during cowpea
460 megaspore mother cell meiosis (Fig. 4A). In contrast with male gametophyte development,
461 auxin biosynthesis gene transcripts were generally undetectable in developing cowpea
462 female gametophytes in most of the post-meiotic developmental stages. Exceptions were a
463 *TAR2_3*-like gene expressed at relatively low levels throughout the stages of female
464 gametogenesis, which was also expressed throughout male gametogenesis, during early
465 embryogenesis and in leaf (Fig. 4A). *YUCCA6, 8, 10* and *TAR4_2*-like genes were co-
466 expressed in the EC and CC with *TAR2_2* showing higher relative EC expression (Fig. 4A;
467 Supplemental Table 9).
468

469 In direct contrast to the subdued nature of auxin biosynthesis gene expression in developing
470 female gametophytes, a total of 47 genes annotated generally as "auxin-responsive" were
471 expressed with RPM counts over 10 in cell-types of the developing female gametophyte
472 (Fig. 4B and Supplemental Table 9). These genes included 15 *AUXIN RESPONSE FACTORS*
473 (*ARF*s and five *AUX-IAA* genes, five auxin transporters belong to the *LAX* and *BIG* gene
474 families, two polar auxin transport *PIN* homologs and three *SMALL AUXIN UP RNA* (SAUR)
475 which are induced in response to auxin (Fig. 4B and Supplemental Table 9). The auxin
476 induced *AIR12*-like gene was expressed at high levels in both the 4 nucleate and the mature
477 embryo sacs, which suggests that regulation of reactive oxygen species may be important if
478 the function of *Arabidopsis* AIR12 is conserved in cowpea.
479
480 The expression of gene components required for *Arabidopsis*-like micro-RNA-mediated
481 silencing was evident during female gametophyte development. Low levels of *DCL1*, *HYL* and
482 *HEN* were observed in the mature embryo sac. The relative expression of *DCL1* was
483 approximately five times higher in the EC than CC (Figure 5A and Supplemental Table 11).
484 These observations suggest that miRNA induced gene silencing is ongoing during the events
485 of meiosis and mitosis during female gametogenesis. Potentially, miRNA production may be
486 more active in the egg than the central cell in mature embryo sacs.
487
488 Examination of expression of homologs of genes involved in the *Arabidopsis* PTGS pathway
489 showed non-detectable *RDR6* expression as also observed during in male gametogenesis.
490 Thus either the typical PTGS pathway described for *Arabidopsis* is not active within male and
491 female cowpea reproductive cell types examined, or other cowpea-specific elements
492 substitute for generation of 21-24 nt siRNAs (Fig. 5A).

493

494 Expression of the RdDM pathway components during female gametogenesis suggested
495 functional gene silencing is active during meiosis and mitotic events. High expression levels
496 of *DME* were also observed in the post-meiotic megaspore tetrad and bi-nucleate mitotic
497 embryo sacs, suggesting methylation turnover is required for the initiation of the mitotic
498 embryo sac progression (Fig. 5B). Interestingly, expression of both *AGO4* and *AGO5* were
499 approximately two times and five times higher in the EC than in the CC, respectively. In
500 addition, *MET1* and *DRM2_1* were expressed at three-fold and 1.5-fold higher levels in the
501 EC than CC, respectively. This suggests relatively lower gene silencing activity via the RdDM
502 pathway in the CC than in the EC (Fig. 5B and Supplemental Table 11). Expression of
503 homologs of *CMT2* and *CMT3-B* which maintain methylation independently from RdDM
504 were approximately three times lower in egg than in central cell (Fig. 5B and Supplemental
505 Table 11). These results suggest a hypothesis that possibly *de novo* DNA methylation via the
506 RdDM pathway may be reduced in the cowpea CC compared to the EC, while maintenance
507 of CHG and CHH methylation may be higher in the cowpea CC than the EC. DNA methylation
508 turnover in the EC and CC may be mediated by *DME* and ROS as indicated by moderate
509 levels of expression in these cell-types (Fig. 5B).

510

511 Most core histones including *H3.3-1* were expressed at moderate levels during female
512 development in cowpea. (Fig. 6A and Supplemental Table 14). The MMC showed higher
513 relative levels of histone *H3.3_1*, *H4_1, 2*, and *3.1_1* transcripts. *Histone3.3* variants were
514 expressed throughout gametogenic development, as were a number of *H3.1* variants which
515 contrasts with the trend observed in male gametogenesis (Fig.6A). High transcript levels of
516 the JMJ25-like gene were evident at MMC meiosis, together with moderate levels of the

517 JMJ706 H3K9 demethylase as seen in late pollen mother cell meiosis, suggesting a possible
518 requirement for H3K9 demethylation in both male and female cowpea meiosis. Expression
519 of the JMJ histone demethylase JMJ30, which removes H3K27me2/3 deposited by PRC2
520 complexes was higher in the female gametophyte than the male (except for the MP-SC
521 sample; Fig. 6D; Supplemental Table 14).
522
523 In *Arabidopsis*, the PRC2 FIS complex, containing MEA/SWN, FIS2, FIE and MSI1, functions to
524 silence target genes during embryo sac maturation and early seed development (Mozgova
525 and Hennig, 2015, Grossniklaus and Paro, 2014). Dynamic expression patterns of a number
526 of PRC2-like gene homologs were expressed during cowpea female gametogenesis,
527 including *FIE*, *MSI1*, *EMF1*, *EMF2*, *EMF2*-like, *CLF* and *VRN5* (Fig. 6C). In the mature embryo
528 sac, egg and central cell samples, *FIE*, *MSI1*, *EMF2_2*, *EMF2*-like, and *VRN5_2* were all co-
529 expressed. High levels of *FIE*, *EMF2*-like, and *VRN5_ 2* were evident in the central cell (Fig.
530 6C). Given the multimeric protein composition of PRC2 complexes, it is not possible to
531 predict the composition of one or more potential complexes that might be functioning
532 during cowpea female gametogenesis from these transcriptional analyses.

533 **Cell cycle arrest in egg and central cells**

534
535 Cell cycle genes are relatively well conserved across plant species (Wang et al., 2004,
536 Scofield et al., 2014). Analyses of the expression of cell cycle genes known to block and/or
537 promote passage through various phases of the cell cycle in isolated cell types has been
538 used to predict the stage of cell cycle arrest (Juranic et al., 2018, Sornay et al., 2015, Sukawa
539 and Okamoto, 2018, Velappan et al., 2017, Komaki and Sugimoto, 2012). A core set of cell

540 cycle genes from *Arabidopsis* was used to identify 141 non-redundant cowpea genes
541 encoding proteins with various roles in cell cycle regulation, including 46 cyclins (CYC), 12
542 cyclin-dependent kinases (CDKs) and 5 CDK inhibitors (interactor/inhibitor of cyclin-
543 dependent kinases/Kip-related proteins; KRP/ICKs) in addition to other key regulators
544 (Supplemental Table 15 and 16). The expression of these in the central cell and egg cell
545 transcriptomes was used to predict the stage of arrest in both cell-types in mature cowpea
546 female gametophytes (Fig. 7).
547
548 The transition from G1 to S phase requires activity of CDKA/CYCD complexes to
549 phosphorylate, and inactivate RBR, thereby enabling E2F transcription factors to drive S-
550 phase gene expression (e.g. genes such as *PCNA* and *MCM2-7*; (Dante et al., 2014).
551 Functional activity of the CDKA/CYCD complex at the G1/S transition is also inhibited by
552 KRP/ICKs, which act on additional CDK/CYC complexes to block transition from G2 to M
553 (Dante et al., 2014). The protein kinase, Wee1 can contribute to a block in cell-cycle
554 progression at the G2/M transition (De Schutter et al., 2007). The MYB3R family members,
555 MYB3R1 and MYB3R4 transcriptionally activate many G2/M promoting genes (Haga et al.,
556 2011). MYB3R3 and MYB3R5 negatively regulate transcription of the same genes (Kobayashi
557 et al., 2015). However, MYB3R1 can also negatively regulate transcription, redundantly with
558 MYB3R3 and MYB3R5 at G2/M (Kobayashi et al., 2015).
559
560 The egg cell appeared to be arrested in G1/S because expression of RBR and three KRP-like
561 genes *KRP7/ICK5_1*, *KRP3/ICK6_1* and *KRP3/ICK6_2* was moderate to high in the EC.
562 Expression of the *CYCD*s (*CYCD2*, *CYCD3*, *CYCD5* and *CYCD6*), indicative of G1 phase, was low
563 to moderate in the EC while expression of S-Phase genes was moderate to high (Fig. 4,

564 Supplemental Table 16). The high expression of two repressive *MYB3R5* homologs
565 (*MYB3R5_3* and *MYB3R5_4*) and three *Wee1* homologs (*Wee1_1*, *Wee1_2* and *Wee1_3*) in
566 the EC (Fig. 7, Supplemental Table 16) suggests that it is unlikely the EC can progress
567 through the G2/M transition (Kobayashi et al., 2015).
568
569 The central cell appeared to have progressed past S phase and appeared to be arrested in
570 G2/M because *RBR1* and *KRP7*/*ICK5_1* expression were strongly reduced compared to that
571 in the EC, as was expression of all S phase-related genes (Fig. 5, Supplemental Table 16). The
572 expression of three *CYCB*-like homologs promoting the G2/M transition, (*CYCB2;3_1*,
573 *CYCB2;3_2* and *CYCB1;2.1_2*) showed high expression in CC at almost three times that
574 observed in the egg cell (Fig. 5, Supplemental Table 16). The CC showed only moderate
575 expression of two G2/M promoting *MYB3R1* homologs, expression of repressive *Wee1_1*
576 and *Wee1_2* was moderate and *MYB3R4* and *MYB3R5* expression was low to undetectable
577 (Fig. 5, Supplemental Table 16). Taken together, these data indicate the cowpea CC is
578 preparing to enter M phase.
579
580 **Post-fertilization gene expression in cowpea**
581
582 The GO terms characterizing the early post-fertilization embryo transcriptome comprised of
583 elongating egg cells, zygotes and early embryos, provided signatures of growth with
584 microtubule movement, cell proliferation, transcription, translation carbohydrate processes,
585 and cell wall biogenesis (Fig. 3A). Meiosis transcript expression was restricted to homologs
586 of *MSH4*, *PRD1*, *OSD1* and *SMG7* (Fig. 3C). A range of auxin biosynthesis-like genes were
587 expressed including *YUCCA*-likes, *2*,*3*,*4*, *6*, *8* and *10* and *TAR10*-like (Fig. 4A). By contrast, a

588 lower repertoire of auxin-responsive type gene expression was observed relative to that
589 seen during male and particularly female gametophyte formation (Fig. 4B). Evidence for
590 extensive microRNA-induced silencing was low given the low levels of *HEN* and *DCL1*.
591 Activity of an *Arabidopsis*-like PTGS pathway was unlikely with many components
592 undetectable (Fig. 5A), and significant RdDM activity at the cell-type stages collected was
593 also doubtful. Expression of other DNA methylation components was evident as was
594 expression of the DNA demethylase *DME* (Fig. 5B). A repertoire of histones genes were
595 expressed and expression of a number of PRC2-like genes suppressed. High levels of *JMJ25*-
596 like gene expression were noticeable suggesting H3K9 demethylation may be supporting
597 early cowpea embryogenesis (Fig. 6D).

598 **Discussion**

599 **Cowpea reproductive cell-type transcriptomes: a legume resource**

600

601 The generated cowpea transcriptomes spanning the temporal sequence of events during
602 male and female gametophyte and early seed development enabled detection of enriched,
603 specifically expressed genes including cell-type specific genes expressed at low levels.
604 Verification levels of 71% of cell-type specific expression by qRT-PCR and spatial
605 confirmation *in situ* hybridization confirmed the utility of these transcriptomes for further
606 dissection of regulatory networks in cowpea and syntenic legumes. Comparative analyses of
607 cowpea cell-type specific genes and those specifically expressed in *Arabidopsis*, led to the
608 identification of *Arabidopsis*-derived promoters directing gene expression to meiotic cells,
609 the egg and central cells and early embryos of cowpea. This supports the conservation of

610 some regulatory transcriptional networks during reprodutive cell differentiation in both
611 cowpea and *Arabidopsis.*
612
613 The cowpea meiosis genes identified show strong conservation with those found in other
614 angiosperms. Cowpea meiosis genes were, however, found to be expressed at moderate
615 levels in non-meiotic cell-types. *Arabidopsis* tissue RNAseq experiments have also found
616 that meiosis genes are present in many tissues throughout development (Klepikova et al.,
617 2016). This is consistent with their roles in DNA repair, cell-cycle control and recombination,
618 and highlights the likely role of cell-specific protein-protein interactions in determining their
619 molecular function in either meiosis or mitosis (Cromer et al., 2012).
620
621 With respect to the involvement of epigenetic pathways during cowpea reproductive
622 development, genes encoding cowpea AGO proteins were identified, and transcriptome
623 analysis suggests an *Arabidopsis*-like RdDM pathway is likely to regulate a number of
624 reproductive stages. *AGO5*, *AGO4* and *AGO6*-like are expressed throughout female
625 gametogenesis. Previous work has suggested that *Arabidopsis AGO5* is expressed in
626 sporophytic tissue and plays an important role in defining the MMC prior to meiosis, but is
627 absent from female gametophyte tissue during subsequent development (Tucker et al.,
628 2012). The roles of RdDM in female gametophyte development may differ between
629 *Arabidopsis* and cowpea, and this can be functionally tested.
630
631 By contrast, an *Arabidopsis*-like PTGS pathway does not appear to be functional in the
632 stages analyzed because RDR6 expression, while detectable in leaf, was not expressed in
633 reproductive cell-types. An alternative component may substitute for RDR6 in cowpea

634 reproductive cells. The cowpea cell-type transcriptomes also suggested potential for PRC2-
635 like complex formation and function during male and female gametophyte formation. The
636 data revealed the possibility for a number of novel combinations of proteins in such
637 potential cowpea-PRC2 multimeric complexes. The identified gene candidates could be
638 tested in protein interaction studies, and mutagenic approaches to confirm complex
639 components and their functional relevance. Additional insights into pathways evident during
640 cowpea reproductive development which have been uncovered from *in silico* analyses are
641 discussed further below.
642
643 **Plant hormone pathways involved in cowpea reproductive development and patterning**
644
645 Plant hormones contribute significantly to plant differentiation and growth. In both the
646 megaspore mother cell and late pollen mother cells, GO terms for "abscisic acid (ABA)
647 signalling pathway" and response to abscisic acid were enriched. Underpinning genes were
648 predicted to encode homologs of calcium-dependent protein kinases involved in ABA signal
649 transduction that influence panicle development and seed formation in rice (Ray et al.,
650 2007). Homologues of genes involved in the biosynthesis of brassinosteroids, a HERK
651 receptor-like kinase and a SCARECROW-like transcription factor, involved in perception and
652 positive regulation of brassinosteroid-mediated gene expression, were expressed during
653 female embryo sac mitosis. In addition, homologs of genes functioning in reproductive
654 development in other species, and involved in synthesis and perception of jasmonic acid and
655 cytokinin were upregulated in mature embryo sacs (Li et al., 2004, Bartrina et al., 2011, Kim
656 et al., 2005). Functional roles of these genes in cowpea remain to be determined.
657

658 Auxin biosynthesis and auxin-responsive genes were upregulated during early cowpea
659 pollen development, reflecting the functional requirement for auxin in early *Arabidopsis*
660 pollen development (Yao et al., 2018). Auxin biosynthesis is also considered important
661 during *Arabidopsis* female gametophyte development. For example, the auxin biosynthesis
662 genes *YUCCA1* and *YUCCA2* are considered to be required early for cell specification. Later
663 expression of *YUCC8*, *TAA1* and *TAR2* biosynthesis genes are required for nuclear
664 proliferation, vacuole formation and embryo sac growth (Panoli et al., 2015). Studies have
665 suggested an auxin gradient serves as the morphogen driving female gametophyte cell
666 specification in *Arabidopsis* (Pagnussat et al., 2009). However, auxin activity has not been
667 experimentally determined in the developing *Arabidopsis,* maize and *Hieracium pilosella*
668 female gametophytes during mitotic divisions (Lituiev et al., 2013).
669
670 While *YUCCA* and *TAR*-like auxin biosynthesis genes were expressed in the cowpea female
671 megaspore mother cell, the expression of auxin biosynthesis genes was barely detectable in
672 subsequent stages of megagametogenesis except for low levels of *TAR2*. Expression of
673 *YUCCA6*, *8* and *10* homologs, and *TAR2* and *TAR4* homologs was detected at low levels in
674 the egg and central cell of the mature embryo sac. The auxin response machinery including
675 multiple *ARF* transcription factors, *AUX-IAA* genes, *SAUR* genes and auxin transporters was
676 detected at moderate to high levels. Highest expression levels were evident in the egg cell
677 and central cell. In the presence of low levels of auxin, AUX-IAA proteins interact with the
678 ARF transcription factors and inhibit transcription of auxin responsive genes. At higher auxin
679 levels, AUX-IAA proteins are degraded by the ubiquitination pathway, freeing the ARF
680 transcription factors to promote expression of their target genes (Wang and Estelle, 2014).
681

682 Given that expression of auxin biosynthesis genes was very low in the developing cowpea
683 female gametophytes, this may indicate that some auxin-related processes are suppressed.
684 The contrasting pattern of very low levels of auxin biosynthesis gene expression and higher
685 relative abundance of auxin responsive genes and transporters suggests that if auxin is
686 involved in cowpea female gametophyte patterning it may be transported in from the
687 sporophytic tissues post-meiosis. Experiments using auxin sensors in transgenic cowpea may
688 resolve this possibility. Alternatively, auxin may have effects in cowpea sporophytic tissue
689 that indirectly affect cell fate decisions in the female gametophyte as has been suggested
690 for *Arabidopsis* and maize (Lituiev et al., 2013).
691

692 **Cell cycle arrest in male and female cowpea gametophytes: involvement of epigenetic**
693 **regulation**

694 Unlike most other eukaryotes, where the gametes remain in the G1 cell cycle phase through
695 nuclear fusion at fertilization (karyogamy), higher plants form gametes that undergo
696 karyogamy in G1, S or G2 phases of the cell cycle (Friedman, 1999). We observed that
697 cowpea pollen is predominantly bi-cellular, containing a vegetative and a generative cell at
698 maturity, as approximately 10% of pollen grains contained a vegetative cell and two sperm
699 cells. Therefore, the events of pollen mitosis II in cowpea continue after pollen-tube
700 germination. *Arabidopsis* pollen grains are tri-cellular and sperm cells progress through S-
701 phase after pollen germination to arrive at G2 immediately prior to fertilization (Friedman,
702 1999).
703

704 Although the cell cycle stage of cowpea sperm cells at fertilization remain to be determined,
705 similarities in histone transcription were observed between *Arabidopsis* and cowpea during

706 pollen maturation. Histone *H3.3* and *H3.1* transcripts were detected during cowpea male
707 gametophyte formation, but histone *H3.1* transcripts were not detected in the mature
708 cowpea pollen-enriched sperm cell sample, as is also observed in *Arabidopsis*. H3.3 variants
709 correlate with gene expression in *Arabidopsis,* while H3.1 variants are incorporated during
710 the S-phase of the cell cycle (Borg and Berger, 2015), and references therein). In rice, EMF2
711 function is essential for pollen development, and it is thought to interact with a rice MSI1
712 homolog, PTC1 (Deng et al., 2017). Genes encoding homologs of a number of these PRC2
713 components were also expressed in cowpea pollen development and may have comparable
714 roles.
715
716 The central cell and egg cell are arrested at different stages of the cell-cycle prior to
717 fertilization. Studies in maize, rice, barley and *Arabidopsis* suggest that the egg is arrested at
718 G1 (Chen et al., 2017, Mogensen and Holm, 1995, Sukawa and Okamoto, 2018), while, in
719 *Arabidopsis,* it has been proposed that the central cell is most likely to be at G2 (Berger et
720 al., 2008). Analysis of cell cycle gene expression in cowpea egg and central cells indicated
721 their likely arrest in G1/S and G2/M cell cycle phases, respectively, consistent with that for
722 *Arabidopsis.*
723
724 Epigenetic pathways involving RdDM and DNA methylation repress transposon activity and
725 regulate female gametophyte arrest. DNA methylation is very low in the central cell and
726 reduced in the egg cell (Kawashima and Berger, 2014). RdDM and DNA methylation are
727 decreased in the central cell, and this allows expression of transposons, which are normally
728 repressed. Subsequent production of siRNAs that then direct methylation within the egg
729 cell, prevent transposon activation (Ibarra et al., 2012). In cowpea, expression of *AGO4*,

730 *AGO5* and *MET1* homologs was several times lower in the central cell than in the EC,
731 suggesting that *de novo* RdDM may also be less active in the central cell. Furthermore, the
732 miRNA pathway is likely more active in the cowpea egg cell than in the central cell as the key
733 components, *AGO1* and *DCL1* are expressed around five times higher in the egg cell. This
734 corresponds with observations of Wuest et al. (2010), who reported high egg cell expression
735 levels of *DCL1* and *AGO1* in *Arabidopsis.*
736
737 The transcriptome analyses also suggested that one or more novel PRC2-like complexes may
738 function to repress female gametophyte initiation in the absence of fertilization via histone
739 methylation. The conservation of a post-fertilization role of auxin in facilitating embryo
740 patterning in cowpea is supported by the observation of increased transcription of auxin
741 biosynthesis genes in the embryo as found in other species (Hands et al., 2016).
742

743 **Conclusions**

744 In this study, we have isolated and analysed cell-specific transcriptomes spanning the events
745 of male and female gametogenesis and early seed formation in cowpea. Cell-specific genes
746 were identified and validated using a combination of techniques including qRT-PCR and *in*
747 *situ* hybridization. The transcriptomes have identified gene expression pathways in common
748 with reproduction in other angiosperms, in addition to novel differences that can be
749 functionally examined. These cowpea reproductive cell-type transcriptomes provide a useful
750 resource for the further isolation of cell-type specific promoters, and uncovering
751 informative pathways which may aid in improving reproductive resilience and yield in
752 cowpea and related legumes.

753 **Materials and Methods**

754 **Plant material, tissue collection and generation of RNA sequences**

755 Cowpea plants (*Vigna unguiculata* IT-86D-1010) were grown in the glasshouse under
756 previously described conditions (Salinas-Gamboa et al., 2016). To collect RNA samples,
757 tissues at appropriate developmental stages were fixed and processed for LCM according to
758 previously published protocols (Okada et al., 2013). To collect sperm cells, anthers and
759 stigmas were collected from flowers at anthesis and subjected to osmotic shock in
760 Brewbaker and Kwack medium (Brewbaker and Kwack, 1963), pH 6.5, supplemented with
761 12.5% (w/v) sucrose in 15 ml centrifuge tubes. The homogenate was centrifuged at 130
762 rpm, for 30 minutes at room temperature, then filtered through 150 µm and then 30 µm
763 CellTrics nylon sieves (Partec GmbH) and collected in 2 ml microfuge tubes. The
764 homogenate was centrifuged at maximum speed for 2 minutes to pellet cells, and 50 µl was
765 layered on 0.5ml of 10% Percoll in 2ml microfuge tubes that were centrifuged at 900xg for 2
766 minutes. 1.5 ml of a solution of 0.52 M (10%) mannitol, 10 mM MOPS buffer (pH 7.5) was
767 added, and the tubes were centrifuged for 2 minutes at maximum speed. 20µl of the
768 resulting pellet was added to 30ul of ARCTURUS PicoPure RNA isolation buffer and snap
769 frozen in liquid nitrogen before storing at -80C. Total RNA was extracted from the LCM
770 samples and sperm cell enriched samples using the ARCTURUS® PicoPure® RNA Isolation Kit
771 (Applied Biosystems®) according to the manufacturer's protocol. RNA was subjected to two
772 rounds of amplification (three rounds for fTET, ES2n and ES4n) using the MessageAmp II
773 RNA amplification kit (Ambion). Total RNA was extracted from leaf as described (Spriggs et
774 al., 2018).

775

776 **Sequencing and read alignment**

777 Illumina sequencing libraries were prepared from the RNA samples by the Australian
778 Genome Research Facility (AGRF) and run on the Illumina HiSeq2500 sequencing platform
779 multiplexed in 3 lanes. At least 36 million reads were produced per sample (approximately
780 1.6 billion reads, total) of which ~86% could be uniquely aligned to the IT86D-1010 genome
781 (Spriggs et al., 2018) using the BioKanga software package (Supplemental Table 1,
782 Supplemental Data 8). A total of 74,839 genes were predicted in the cowpea IT86D-1010
783 genome (Spriggs et al., 2018) using Augustus v 3.1.0 (Stanke and Waack, 2003) and
784 *Arabidopsis thaliana* (TAIR 10) as the gene training set. As no correlation between gene
785 length and number of reads aligned was observed (Pearson´s correlation coefficient = -
786 0.010), the Reads-Per-Million (RPM) normalization factor was used in all cases reporting
787 normalized read counts. The RPM normalized counts between replicates of a cell-type
788 sample were strongly positively correlated, with a Pearson's correlation coefficient (PCC)
789 over 0.8 for most samples. Exceptions were replicates in the megaspore mother cell (MMC;
790 PCC = 0.72), two-nucleate female gametophyte (ES2n; PCC= 0.766) and the pollen
791 microspore (MIC; PCC = 0.62; Supplemental Fig. 1).
792
793 **R versions and rsgcc analyses**
794 Analyses made using R software packages used R Studio version 1.0.143, with R version
795 3.4.3 running under Windows 10. The rsgcc software of Ma and Wang (2012) was used as
796 described by Zhan et al. (2015). The mean RPM normalized counts from each cell-type was
797 used together with the getsgene() function of the rsgcc R package (rsgcc version 1.0.6) with
798 the following parameters "getsgene(data, Log = FALSE, tsThreshold = 0.50, MeanOrMax =
799 "Max")".
800

801 **Gene Ontology annotations**

802 Annotations were made by aligning the predicted IT86D-1010 coding sequences
803 (Supplemental Data 9) against the NCBI nr protein database (version December 23rd 2017)
804 using the DIAMOND aligner (Buchfink et al., 2015), version 0.9.13 returning at most ten
805 alignments with a maximum e-value of 1e-5. The generated XML file was supplied to
806 Blast2GO (version 5.0.13 running on windows 10 with Java version 1.0.8 152), which was run
807 with default parameters, to assign Gene Ontology (GO) annotations (Ashburner et al., 2000,
808 The Gene Ontology, 2017) to each predicted gene. This approach assigned putative
809 functional annotations to 54,223 predicted genes, with 39,653 predicted genes assigned GO
810 terms (Supplemental Data 1).

811

812 **Filtering to identify cell-specific genes in the egg cell, central cell and sperm cell**
813 **transcriptomes**

814 Genes specifically expressed in central cell, egg cell and sperm were selected using the filter
815 function in Microsoft Excel to choose genes with a mean RPM at least five times higher in
816 the cell-type of interest than any other cell-type (see full criteria in Supplemental Table 2).
817 No filter was applied to other cell-types in the female gametophyte when selecting egg cell
818 and central cell genes where the contrast between those two cells was the key criteria.

819

820 **Differential expression analysis**

821 Differential expression analysis was carried out using the DESeq2 (version 1.18.1) software
822 package in R (Love et al., 2014). Analyses were carried out as described except that
823 differentially expressed genes in each cell-type were identified using the contrast argument
824 of the results function with parameters set to contrast the expression of each gene in one

825 cell-type with the average expression of all other cell-types "results(dds, contrast =
826 list(c(cell_X), c(all_other_cells)), listValues = c(1, -1/13), alpha = .05)". Genes with adjusted p
827 value < 0.05 and log2 fold change > 2 were selected as significant (Supplemental Data 5).
828
829 **Principal Component Analysis**
830 Principal component analysis was carried out on each replicate of our developed
831 transcriptomes using the base R prcomp() function on read counts transformed by the
832 variance stabilising transformation, vsd(), function of DESeq2. Principal components were
833 plotted in 3D using the plot3d() function of the rgl package. Colours suitable for all vision
834 types were chosen from the colour alphabet previously described (Green-Armytage, 2010).
835 Plots were exported as .svg files and figures were assembled in Adobe Illustrator.
836
837 **Comparisons between cowpea and Arabidopsis female gametophyte LCM data sets**
838 The best *Arabidopsis* homolog for each predicted cowpea gene was identified by using the
839 DIAMOND aligner (Buchfink et al., 2015) (version 0.9.13) to align cowpea genes against the
840 TAIR 10 gene database, and selecting the best match for each cowpea gene (Supplemental
841 Data 7). *Arabidopsis* gene sets from the MMC (Schmidt et al., 2011), ovule nucellus, early
842 mitotic female gametophytes (Tucker et al., 2012) and the egg and central cell (Wuest et al.,
843 2010, Steffen et al., 2007) were compared with the *Arabidopsis* annotations of each cowpea
844 gene in the differentially expressed genes (log2 fold-change >2) in corresponding cell-types
845 from the cowpea dataset.
846

847 **Identification of core meiosis, auxin biosynthesis, auxin-responsive genes and cowpea cell-**
848 **cycle genes**

849 A set of 18 core cowpea meiosis genes were identified by BlastP search of key genes in the
850 literature (Supplemental Table 7 and references therein) against the amino acid sequence of
851 the cowpea IT86D-1010 predicted genes. In each case the best match based on bit-score
852 was chosen for further analysis. To identify cell-cycle related cowpea genes, the amino acid
853 sequences of a set of cell-cycle genes from *Arabidopsis* (Juranic et al., 2018) were used in
854 BlastP searches against the predicted cowpea protein sequences. BlastP results with an e
855 value < 1e-20 were chosen, and the annotation of cowpea sequences matching multiple
856 *Arabidopsis* genes was chosen based on the highest bit score. The *KRP* gene family and
857 *cdc*25 homologs were selected with an e value < 1e-10. Where multiple cowpea genes
858 matched a single *Arabidopsis* gene, these were notated with an underscore and number
859 after the gene name starting with 1 for the best match and increasing (e.g. VuGene_1).
860 Auxin-related genes were identified in the laser captured cell-type transcriptomes, based on
861 their GO annotations (Supplemental Table 9).

862

863 **Identification of cowpea RNA silencing, RdDM and histone genes**

864 A total of 37 key genes in RNA silencing and RdDM pathways were identified by searching
865 the annotations produced by the Blast2GO software for genes described in key literature
866 (Supplemental Table 10). Fragments not encoding full length protein homologs were not
867 included in further analysis.

868 A total of 49 putative histone encoding genes and 20 JMJ homologs were identified in
869 annotations produced by Blast2GO as described above. The predicted histone variant

870 encoded by each gene was assigned based on annotations, and on homology of the
871 predicted protein to *Arabidopsis* histone amino acid sequence (Supplemental Fig. 7).
872
873 **Phylogenetic tree construction**
874 Phylogenetic trees were made using Geneious version 10.2 (Biomatters,
875 https://www.geneious.com). Trees were built using the following parameters: ClustalW
876 multiple-sequence alignment with BLOSUM cost matrix, Gap open cost 10, Gap extend cost
877 0.1, Jukes-Cantor Neighbour-Joining consensus tree, no outgroup, Bootstrap resampling
878 with 1000 replicates, 50% support threshold.
879
880 **Validation by qPCR**
881 Whole tissues were collected from the following tissues, emerging leaves, developing stem,
882 petiole, shoot apical meristems and roots (Supplemental Table 3), frozen in liquid nitrogen
883 and RNA was extracted with a QIAGEN RNeasy Plant mini kit according to the
884 manufacturer's protocol, including on-column DNAse treatment. cDNA was synthesised
885 using 1ug of RNA with the SuperScript™ III First-Strand Synthesis System (Thermo Fisher
886 Scientific). Central cell and sperm genes were selected for testing by qPCR and *in situ*
887 hybridisation based on specificity in expression and putative function of the predicted
888 genes. Egg cell genes were chosen for testing based on the specificity of their expression
889 and homology to *Arabidopsis* genes known to be specifically expressed in egg cell. Genes in
890 other cell-types tested were selected based on their specificity score in the rsgcc analysis
891 and on our ability to design primer pairs (Supplemental Table 17) that specifically amplified
892 the predicted region, including intron structures, from genomic and cDNA. qRT-PCR was
893 carried out on a Roche Light Cycler ® 480 thermal cycler using SYBR Green I master mix

894 according to the manufacturer's protocols. Quantification was done using LinRegPCR
895 software as described (Ruijter et al., 2009, Ramakers et al., 2003).
896
897 **In situ hybridization.**
898 ISH was performed as previously described (Jackson, 1992) with some modifications for
899 improved resolution in reproductive organs of cowpea. For ISH performed on sectioned
900 specimens, mature unpollinated ovules were fixed in 4% paraformaldehyde and embedded
901 in Paraplast. Sections at 12 µg thickness were attached to ProbeOnPlus slides (Fisher
902 Biotech) and processed as previously described (Jackson, 1992). For whole-mount ISH,
903 developing ovule primordia, unpollinated mature ovules, and mature anthers were fixed in
904 paraformaldehyde (4% paraformaldehyde, 2% Triton, and 1× PBS in diethylpyrocarbonate
905 DEPC-treated water) for 2 h at room temperature with gentle agitation, washed three times
906 in 1× PBS-DEPC water, and embedded in 15% acrylamide:bisacrylamide (29:1) using pre-
907 charged slides (Fisher Probe-On) treated with poly-l-Lys as described (Bass et al., 1997).
908 Fixed pollen grains were stripped from the anthers before embedding, and incubated with
909 an enzymatic solution containing 1% driselase, 0.5% cellulose, and 1% pectolyase, for 30
910 minutes (min) in a humid chamber at 37ºC. All samples were subsequently incubated with
911 0.2M HCl solution for 30 min at room temperature (RT) and in 1ug/ul proteinase K for 30
912 min at 37ºC, before being processed as previously described (García-Aguilar et al., 2005).
913 Slides containing paraffin sections were mounted on Cytoseal, whereas slides containing
914 whole specimens were mounted in either 50% (ovules) or 20% (pollen grains) glycerol, and
915 analyzed with a Leica DMRB microscope under Nomarski illumination.
916
917 **Plasmid construction and cowpea transformation**

918 Gateway entry vectors containing cell-type specific promoter elements together with genes
919 encoding fluorescent proteins were gifts from Shai Lawit and Marc Albertsen (Lawit et al.,
920 2013) as follows: *AtDD1$_{pro}$:ZsYellow1*, *AtDD9$_{pro}$:DsRed-Express*, *AtDD25$_{pro}$:DsRed-Express*,
921 *AtDD31$_{pro}$:AcGFP1*, *AtDD45$_{pro}$:DsRed-Express*, *AtDD65$_{pro}$:AmCyan1*, *AtLat52$_{pro}$:AcGFP1*,
922 *AtRKD2$_{pro}$:DsRed-Express*. The promoter element of *AtEC1.1* was amplified from *Arabidopsis*
923 Col-0 DNA by PCR with primers including adapter elements, and cloned into gateway entry
924 vectors containing AmCyan1. The *AtDUO1* promoter was synthesized and cloned into pMK-
925 RQ by GeneArt (Thermo Fisher Scientific) with SfiI and BglII sites introduced for subsequent
926 cloning into gateway entry vector containing AcGFP1.
927
928 A gateway-compatible binary vector pOREOSAr4r3 was created by insertion of a Gateway
929 recombinational cassette amplified from pDESTr4r3 (Invitrogen) into ClaI and KpnI sites of
930 pOREOSA vector backbone, which was a gift from Thomas J. Higgins, CSIRO. Reporter entry
931 clones were introduced into the pOREOSAgw expression vector via multisite Gateway
932 recombination. Final constructs contained either triple or quadruple fluorescent labels
933 (Supplemental Table 20). AtKNU:YFP:3'KNU cassette was previously described (Tucker et al.,
934 2012) and was inserted into the pOREOSA via HindIII digestion. Gateway cloning reactions
935 were performed with LR Clonase II Enzyme mix (ThermoFisher Scientific). All plasmid vectors
936 were verified by sequencing and final constructs electroporated into *Agrobacterium*
937 *tumefaciens* strain AGL1 for use in cowpea stable transformation (see below). Details of the
938 primer sequences and constructs transformed into cowpea are described in Supplemental
939 Tables 19 and 20. Cowpea transformation was carried out using *Agrobacterium*-mediated
940 transformation of cotyledonary nodes (Popelka et al., 2006). Transformed explants were
941 first selected on a medium containing 100 mg/L kanamycin and then on 20 mg/L geneticin

942 (G-418) for up to 3-6 months. Shoots developing healthy roots were transferred into 90mm
943 small pots containing sterilized soil mixture (Van Schaik's Bio-Gro Pty Ltd, Australia),
944 acclimatized in the growth room at 22°C with 16h photoperiod for up to 4 weeks, and then
945 transferred to the glasshouse in larger 4.5L pots. PCR was performed to confirm the
946 presence of the *nptII* and reporter genes with the primers listed in Supplemental Table 19.
947 The number of independent cowpea transgenic lines per construct varied from three to 15.
948

949 **Fluorescent microscopy and clearing procedures**

950 Samples were examined using a Zeiss fluorescence microscope (Axio Imager M2; Carl Zeiss)
951 with an HXP 120 light source. Digital images were captured using an AxioCam HRc camera
952 and ZEN 2.6 software (Carl Zeiss). Signals from dsRED-Express were observed with the Zeiss
953 filter set 43, AcGFP1 signals were observed with the Zeiss filter set 13, AmCyan1 signals
954 were observed with Zeiss filter set 47, while YFP and ZsYellow1 signals were observed with
955 Zeiss filter set 46. The exposure time was adjusted to as appropriate for each sample.
956 Images were processed in ZEN2.6 software and assembled in Adobe Photoshop. The
957 method previously described in (Salinas-Gamboa et al., 2016) was used when tissue clearing
958 was required. For light microscopy, the Zeiss Axioskop 2 microscope equipped with
959 Nomarski optics, Spot Flex colour camera and Spot 5.1 software was used to capture images
960 (Diagnostic Instruments, Inc).

961

962 **Generation of transgenic *Arabidopsis* with cowpea promoter constructs**

963 The regulatory region of eleven cowpea genes was transcriptionally fused to the *uidA* (*GUS*)
964 reporter gene by amplifying a fragment corresponding to either the complete intergenic

965 region (*VuNB-RCpro; VuSDR1pro*) or an arbitrary fragment (*VuBAS1pro; VuGULL06pro;*
966 *VuRKD1pro; VuEC1.1pro; VuGIM2pro; VuXTH6pro; VuXTH32pro; VuAt3g09950pro; VuKNU-*
967 *L1pro*) located upstream of its coding sequence but excluding the 5 untranslated region
968 (see Suppl Table 21 for primers). Amplicons were cloned into pCR8 TOPO TA (Invitrogen)
969 and used as donors in LR recombination (LR Clonase II; Invitrogen) with pMDC162 (Curtis
970 and Grossniklaus, 2003), producing a binary vector that contains the *uidA* reporter gene.
971 Transgenic Col-0 plants were obtained by floral dipping as previously described (Clough and
972 Bent, 1998). At least 15 T1 individuals were obtained and analyzed to quantify the
973 frequency of GUS expression at different stages of male and female reproductive
974 development. GUS staining assays for stages before fertilization were conducted as
975 described (Vielle-Calzada et al., 2000).
976

977 **Accession numbers**

978 **Large datasets**

979 .bam files of sequences from
980 TO BE CONFIRMED
981
982 Fasta.gz archives of raw reads for MMC, EM and Sp
983

984 **Supplemental Material**

985 **Supplemental Tables**

986 **Supplemental Table 1. Align stats**

987 Number and proportion of Illumina reads aligned against the IT86D genome for each sample

988 **Supplemental Table 2. Filtering**

989 Criteria and results of selecting genes based on filtering in Excel

990 **Supplemental Table 3. *Arabidopsis* Reproductive genes and promoters**

991 Summary of *Arabidopsis* reproductive cell-specific gene expression of their Cowpea

992 homologs in cowpea and activity of their promoters driving transgenes in *Arabidopsis*

993 **Supplemental Table 4. qPCR**

994 Table of tissues and results of qRT-PCR validation of cell-specific gene expression

995 **Supplemental Table 5. CS_DE_gene_overlap**

996 Table summarising the number of common genes between cell-specific and differentially

997 expressed datasets.

998 **Supplemental Table 6. Compare datasets**

999 Comparison of published cell-specific datasets with our transcriptomes

1000 **Supplemental Table 7. Meiosis genes**

1001 Names and references for the cowpea homologs of genes involved in meiosis

1002 **Supplemental Table 8. Meiosis RPM**

1003 Expression of meiosis genes as RPM counts

1004 **Supplemental Table 9. Auxin genes**

1005 Expression of annotated auxin related genes as RPM counts

1006 **Supplemental Table 10. RNAi, RdDM refs**

1007 Names, functions and references for the cowpea homologs of genes involved in RNA
1008 silencing and RdDM

1009 **Supplemental Table 11. RNAi, RdDM RPM counts**

1010 Expression of RNAi and RdDM genes as RPM counts

1011 **Supplemental Table 12. Histone RPM counts**

1012 Expression of histone genes as RPM counts

1013 **Supplemental Table 13. PRC RPM counts**

1014 Expression of PRC genes as RPM counts

1015 **Supplemental Table 14. JMJ RPM counts**

1016 Expression of JMJ genes as RPM counts

1017 **Supplemental Table 15. Cell-cycle genes**

1018 Names and references for the cowpea homologs of genes involved in cell-cycle progression

1019 **Supplemental Table 16. Cell-cycle RPM**
1020 Expression of cell-cycle genes as RPM counts

1021 **Supplemental Table 17. Oligonucleotide sequences for qRT-PCR**

1022 Sequences of oligonucleotide primer sequences used for qRT-PCR in this study

1023 **Supplemental Table 18. Probe sequences**

1024 Sequences used in RNA *in situ* probes

1025 **Supplemental Table 19. Cloning Primers**

1026 Sequences of primers used in creating plasmids used in this study

1027 **Supplementary Table 20. Plasmid constructs**

1028 List of plasmid constructs used in this study

1029 **Supplemental Table 21. Activity of Cowpea promoters**

1030 Activity of cowpea promoters in *Arabidopsis thaliana*
1031

1032 **Supplemental Data**

1033 **Supplemental Data 1: IT86D-1010 genes and GO annotation**

1034 Large table with GO annotations and descriptions for all cowpea predicted genes

1035 **Supplemental Data 2: Mean RPM counts**

1036 Large table with the average RPM counts for all genes in all cell-types

1037 **Supplemental Data 3: Cell-specificity scores and RPM counts**

1038 Large set of tables with all genes in each of the cell-specific gene sets

1039 **Supplemental Data 4: Cell-specific gene annotations**

1040 Large set of tables with all cell-specific genes for each cell-type and their GO annotations

1041 **Supplemental Data 5: log2 FC2 DE genes**

1042 Large set of tables with all genes DE at log2 fold-change 2 in each of the cell-types

1043 **Supplemental Data 6: DE enriched GO terms**

1044 Large set of tables with the enriched biological process GO terms from each of DE gene sets

1045 **Supplemental Data 7: BlastP comparing Vu and At**

1046 Large table containing the best *Arabidopsis* protein match by homology to the translated

1047 coding sequence of each predicted cowpea gene

1048 **Supplemental Data 8: Raw aligned read counts**

1049 A csv file with the raw counts of aligned to each gene in each replicate.

1050 **Supplemental Data 9: Predicted gene coding sequences**

1051 A fasta file with the coding sequences of all IT86D-1010 predicted genes.

1052 **Acknowledgments**

1053 We are grateful to Tracy How and Dilrukshi Nagahatenna for their assistance with cowpea

1054 transformation and growth, to Natalia Bazanova for assistance with cloning, and to Steven

Table 1. Genes and promoters tested for cell-type specificity by qPCR, *in situ* hybridization and transformation in cowpea.

Cowpea gene or reporter	Arabidopsis homolog	Arabidopsis annotations / predicted localisation	Cell-type localization	Tissue localization by qPCR	Cell type localization by *in situ*[c]	Promoter directed gene

1055 Henderson for suggestions concerning the *in silico* identification of cowpea cell-specific

1056 genes.

construct			(*in silico* RNA-Seq)			expression in transgenic cowpea
VuMAP3K17-like	At2g32510	mitogen-activated protein kinase kinase kinase 17	MMC	MMC	-	-
VuCPK1	At5g04870	calcium dependent protein kinase 1	MMC	MMC, MIC	-	-
VuLBD12	At2g30130	Lateral organ boundaries domain family protein	fTET	fTET	-	-
Vu.g6173	At4g35670	Pectin lyase-like superfamily protein	fTET	MIC, SC	-	-
VuLHL3	At1g64625	Plant-specific basic helix-loop-helix (bHLH) protein	ES4n	ES2n/4n	-	-
Vu.g43722	At1g68040	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	ES2n	ES2n/4n	-	-
Vu.g2461	??	unknown	ES2n	ES2n/4n	-	-
VuGULL06	At2g46670	FAD-dependent oxidoreductase	CC	CC/EC	CC	-
VuNB-ARC	At3g14470	NB-ARC domain disease resistance protein	CC	fTET, ES2n/4n	CC[d]	-
VuSDR1	At3g61220	NADP-binding Rossman-fold protein	CC	NS	CC	-
VuMEE23	At2g34790	Berberine Bridge Enzyme-like protein	CC	CC/EC	CC, NC	-
VuBAS1	At2g26710	Cytochrome P450 superfamily protein	CC	CC/EC	-	-
VuEC1.1_1	At1g76750	cytosine-rich secreted protein	EC	CC/EC	EC, Sy	-
VuEC1.1_2	At1g76750	cytosine-rich secreted protein	EC	CC/EC	EC, Sy	-
VuRKD1	At1g18790	RWP-RK domain transcription factor	EC	CC/EC	EC, Sy	-
VuCAP	At3g19690	CAP superfamily protein	Em	Em	-	-
VuZFP	At5g01860	C2HC zinc fingers superfamily protein	Em	MMC, fTET, mTET, Em	-	-
VuVRN1	At3g18990	AP2/B3-like transcriptional factor family protein	Em	Roots, Petiole, Leaf, EC,Em	-	-
VuLBD13	At2g30340	LOB domain-containing protein 13	Em	Em	-	-
Vu.g56782	At4g24130	Protein of unknown function, DUF538	mTET	mTET	-	-
VuPPCK	At3g04530	phosphoenolpyruvate carboxylase kinase 2	mTET	mTET	-	-
VuXND1	At5g64530	xylem NAC domain 1	mTET	mTET	-	-

VuSRO2	At1g23550	similar to RCD one	mTET	Seedling, roots	-	-
VuXTH6	At5g65730	xyloglucan endotransglucosylase hydrolase	SC	SC	SC, VN	-
VuXTH32	At2g36870	xyloglucan endotransglucosylase hydrolase	SC	SC	SC, VC	-
VuGIM2	At2g36690	2-oxoglutarate-dependent dioxygenase	SC	SC	SC, VC	-
VuAt3g09950	At3g09950	unknown	SC	Roots, seedlings	SC, VC	-
VuKNU-like1	At5g14010	C2H2 zinc finger protein	MMC, PMC, mTET, SC	MMC, mTET, SC	MMC+NC+II	-
VuKNU-like2	At5g14010	C2H2 zinc finger protein	MMC, Em	MMC, MIC	-	-
AtRKD2 [e]	At1g74480	EC	-	-	-	EC, Em
AtEC1.1 [e]	At1g76750	EC	-	-	-	EC
AtDD45 [e]	At2g21740	EC	-	-	-	EC
AtDD1 [e]	At1g36340	AC	-	-	-	non-specific
AtDD9 [e]	At1g26795	CC	-	-	-	CC
AtDD25 [e]	At3g04540	CC	-	-	-	CC
AtDD65 [e]	At3g10890	CC	-	-	-	ND
AtDD31 [e]	At1g47470	SC	-	-	-	ND
AtDUO1 [e]	At3g60460	SC	-	-	-	ND
AtLAT52 [e]	At5g45880	VC	-	-	-	ND
AtKNU [e]	At5g14010	MMC + PMC	-	-	-	MMC, PMC, II

Footnotes

[a]Based on *Arabidopsis thaliana* genome annotation TAIR 10.

[b]Based on laser-capture cell transcriptomes

[c]Based on *in situ* hybridization

[d]Shows nuclear dots corresponding to active transcription in the central cell nuclei

[e]Fluorescent reporter gene constructs

Abbreviations: CC, central cell; EC, egg cell; II, inner ovule integument; MMC, megaspore mother cell;
NC, nucellar cells; SC, sperm cell; Sy, synergid; VC, vegetative cell; VN, vegetative nucleus.

1057 **Table 2.** The number of genes upregulated in each cell-type compared to the average
1058 expression in all cell-types.

Cell type	Number of genes differentially upregulated
MMC	5878
fTET	2316
ES2n	2627
ES4n	1423
CC	1897
EC	1550
MES	3376
Em	5051
PMC E	1282
PMC L	2653
mTET	3743
MIC	3101
sperm	38691
Leaf	9140

1059

1060 **Figure Legends**

1061

1062 **Figure 1**. Stages of reproductive development in cowpea and gene expression in developed
1063 transcriptomes.

1064 A, Cartoon of temporal stages of male and female gametophyte development showing cell-
1065 type stages collected by LCM for RNAseq. For detail refer to text. B, The number of genes
1066 expressed in each transcriptome set (RPM >0). C, The proportion of genes expressed at very
1067 low (<1RPM), low (1-5RPM), moderate (5-20RPM) or high (>20RPM) levels in RNAseq sets.
1068 Abbreviations: MMC, megaspore mother cell; fTET, female tetrads; FM, functional
1069 megaspore selection stage (not collected); ES2n, mitotic embryo sac with 2 nuclei; ES4n,
1070 embryo sac with 4 nuclei; MES, mature embryo sac at anthesis, where antipodals have
1071 degenerated, containing the egg cell (EC), flanking synergids (Sy, not collected), and central
1072 cell (CC); Em, embryo; En, endosperm (not collected); PMC.E, early pollen mother cell;
1073 PMC.L, late pollen mother cell; mTET, male tetrads; MIC, uninucleate microspore; Bi,
1074 bicellular pollen stage; Tri, tricellular pollen stage, two sperm cells (SC) in the emerging
1075 pollen tube; MP-SC, mature pollen-sperm cell.

1076

1077 **Figure 2.** Gene expression in cowpea reproductive cell-types**.**

1078 A – J, Whole-mount *in situ* hybridization (WISH) in pollen grains of cowpea using four
1079 different antisense probes. A-C, *VuXTH6* mRNA in sperm cells A and vegetative nucleus B,
1080 with no signal in the sense control C. D and E, *VuXTH32* mRNA was detected in sperm cell
1081 and vegetative cell cytoplasm. F, *VuXTH32* sense controls. G, Whole-mount ovule undergoing
1082 megasporogenesis showing *VuKNU1* mRNA localized in the MMC, the nucellus and the inner
1083 integument. H, Paraffin section through a differentiated female gametophyte showing

1084 *VuRKD1* mRNA localization in the egg cell and synergids. I, Paraffin section through a
1085 differentiated female gametophyte showing *VuNB-ARC* mRNA localization in the central cell.
1086 Inset shows fused polar nuclei showing active transcription of *VuNB-ARC*. J,
1087 *AtRKD2*$_{pro}$*:dsRED-Express* (red) in EC and *AtDD9*$_{pro}$*:AmCyan1* (cyan) in the CC one day before
1088 anthesis. K, *AtEC1.1*$_{pro}$*:AmCyan1* (cyan) in EC and *AtDD25*$_{pro}$*:dsRED-Express* (red) in CC. L-M,
1089 *AtDD45*$_{pro}$*:dsRED-Express* (red) in an isolated egg cell, N, two-celled proembryo and O, early
1090 embryo.

1091 Abbreviations: EC, egg cell; CC, central cell; Sy, synergids; SC, sperm cells; VN, vegetative
1092 nucleus, MMC, megaspore mother cell.
1093 Scale bars: A-G, L-M = 10 µm, H = 8 µm, I = 14 µm, J-K = 50 µm, N-O = 20 µm.
1094

1095 **Figure 3.** Differentially enriched Biological Process GO terms and expression of meiosis genes
1096 during gametogenesis.

1097 A, Enriched GO terms primarily in female gametophyte and early embryo transcriptomes. B,
1098 Enriched GO terms primarily in the male gametophyte and leaf transcriptomes. The
1099 statistical significance of the GO enrichment is represented according to the scale of
1100 weighted Fisher p value shown at the bottom right. C, Expression of genes involved in
1101 meiosis in the generated cowpea transcriptomes (See also Supplemental Tables 7 and 8).
1102 Expression levels relate to the RPM scale.
1103

1104 **Figure 4.** Expression of auxin biosynthesis and auxin responsive genes during meiosis and
1105 gametophyte development.

1106 A, Expression of genes belonging to the YUCCA and TAR auxin biosynthesis families. B,
1107 Expression of 60 auxin responsive genes sorted in decreasing order of expression level in the

1108 2-nucleate embryo sac. Expression levels are related to the RPM scale with each respective
1109 panel. (See also Supplemental Table 9).
1110
1111 **Figure 5.** Heatmaps representing the expression of selected genes involved in RNA-mediated
1112 gene silencing and DNA methylation.
1113 A, Expression of genes involved in the miRNA-mediated gene silencing pathway and in PTGS.
1114 B, Expression of genes that interact with RNA and are involved in RdDM, DNA
1115 methyltransferases that mediate RdDM and DNA glycosylases that mediate demethylation of
1116 DNA. Expression level is according to the RPM scales for each panel. (See also Supplemental
1117 Tables 12 and 13).
1118
1119
1120 **Figure 6.** Histone gene expression and the expression of selected genes involved in histone
1121 modification.
1122 A, Expression of genes belonging to the core histone family (See Supplementary Table 18).
1123 B, Expression of selected PcG genes involved in histone methylation resulting in gene
1124 silencing (See Supplementary Table 19). C, Known protein interactions forming Polycomb-
1125 group repressive 2 complex types. Accessory proteins (X) are often associated with PRC2
1126 complex function. Asterisks indicate *Arabidopsis* specific proteins. MEA and *Arabidopsis*
1127 counterparts, SWN and CLF are SET domain proteins and homologs of *Drosophilia* Enhancer
1128 of Zeste. FIS2 and others listed are VEFS-box proteins. FIE and MSI1 are WD40 repeat
1129 protein containing proteins (Mozgova and Henning, 2015). D, Expression of selected JMJ
1130 family histone demethylase genes and demethylation targets, at left (See also Supplemental
1131 Tables 14, 15, 16). Expression levels relate to the RPM scales associated with each panel.

1132

1133 **Figure 7.** Expression of selected cell cycle genes in egg, central cell, mature pollen-sperm

1134 and leaf transcriptomes.

1135 Cell cycle genes are grouped according to their function relating to roles in promoting or

1136 blocking progression through various cell cycle phases. Relative expression level is indicated

1137 according to the RPM scale shown at the bottom of each panel. (See also Supplemental

1138 Tables 10 and 11).

1139

1140 **Supplemental Figure. 1.** Correlation between RNAseq replicates.

1141 Scatter plot of log-transformed mean RPM counts of replicates for each tissue. Replicate 1

1142 and replicate 2 for each cell-type are shown on the Y and X axes, respectively.

1143

1144 **Supplemental Figure. 2.** Principal component analysis (PCA).

1145 PCA plot showing relationships between all transcriptomes.

1146 Abbreviations: MMC, megaspore mother cell; fTET, female tetrads; ES2n, mitotic embryo

1147 sac with 2 nuclei; ES4n, embryo sac with 4 nuclei; MES, mature embryo sac at anthesis,

1148 containing the egg cell (EC) and central cell (CC); Em, embryo; PMC.E, early pollen mother

1149 cell; PMC.L, late pollen mother cell; mTET, male tetrads; MIC, uninucleate microspore; MP-

1150 SC, mature pollen-sperm cell.

1151

1152 **Supplemental Figure. 3.** Cell-specific genes expression

1153 Heat-map of scaled RPM expression values of the cell-specific genes identified using the

1154 rsgcc software package

1155

1156 **Supplemental Figure. 4.** *In situ* hybridization of cell-specific cowpea genes in gametophytic
1157 tissue and fluorescent reporters expressed from *Arabidopsis* cell-specific promoters in
1158 cowpea.
1159 A-B, *VuXTH32* mRNA localized in the sperm cells and vegetative cell cytoplasm. C, *VuXTH32*
1160 sense control. D-E, *VuAt3g09950* mRNA abundantly localized in the sperm and vegetative
1161 cells. F, sense *VuAt3g09950* control. G, Paraffin section through a fully differentiated female
1162 gametophyte showing *VuEC1.1* mRNA localization in the egg cell. H, Consecutive section to
1163 (G) showing *VuEC1.1* mRNA localization in the synergids. I, Detail of G showing mRNA
1164 localization in the egg cell. J, Detail of H showing mRNA localization in the synergids. K,
1165 Whole-mounted ovule showing *VuMEE23* mRNA localization in the central cell and adjacent
1166 nucellar cells. L-M, $AtDUO1_{pro}$*:AcGFP1* no expression. N, Whole-mounted ovule showing
1167 *VuSDR1* mRNA localization in the central cell. O, Whole-mounted ovule showing *VuGULL06*
1168 mRNA localization confined to the central cell. P, No expression of $AtLAT52_{pro}$*:AcGFP1* in
1169 germinating pollen. Q, Mature female gametophyte with egg cell, polar nuclei and central
1170 cell in cleared ovule. R, $AtDD45_{pro}$*:dsRED-Express* (red) in EC and general pattern of the
1171 $AtDD1_{pro}$*:ZsYellow1* (yellow) in cells external to the embryo sac and in CC. S,
1172 $AtEC1.1_{pro}$*:AmCyan1* (cyan) in zygote and $AtDD25_{pro}$*:dsRED-Express* (red) in early endosperm.
1173 T, $AtRKD2_{pro}$*:dsRED-Express* (red) in zygote and $AtDD9_{pro}$*:AmCyan1* (cyan) in early
1174 endosperm. U, $AtRKD2_{pro}$*:dsRED-Express* in early embryo. *AtDD9* is not expressed at this
1175 stage. V, Cleared ovule showing early embryo corresponding to the stage in U.
1176 Scale bars: A-F, I-J = 10 µm; G-H = 18 µm; K, N-O = 35 µm; L-M, P-T = 50 µm; U-V = 100 µm
1177 Abbreviations: EC, egg cell; CC, central cell; Sy, synergids; SC, sperm cells; VN, vegetative
1178 nucleus; PN, polar nuclei; Em, embryo; En, endosperm; Zy, zygote.
1179

1180 **Supplemental Figure. 5.** Expression of the KNU genes in cowpea.

1181 A, Whole-mount pre-meiotic ovule showing *VuKNU1* mRNA localized in the MMC and the

1182 developing nucellus at the onset of integument initiation. B-D, Micrographs of cowpea

1183 ovules expressing a *AtKNU$_{pro}$:nlsYFP* reporter construct marking MMC fate. E-F, At the end

1184 of meiosis the KNU promoter turns off in female functional megaspore, but the expression

1185 persists in somatic ovule cells at the chalazal end. G-H, Expression of

1186 *AtKNU$_{pro}$:nlsYFP* reporter construct during male meiosis. I, Male tetrad showing no

1187 expression of reporter construct. J, After the first pollen mitosis, the expression of reporter

1188 construct is evident in both vegetative and generative nucleus. K, Mature pollen grain does

1189 not show any expression.

1190 Abbreviations: MMC, megaspore mother cell; GC, generative cell; VN, vegetative nucleus.

1191 Scale bars: A = 10 µm; B = 100 µm; C-H = 50 µm; I-K = 20 µm.

1192

1193 **Supplemental Figure. 6.** Specific activity of cowpea promoters in reproductive cells of

1194 *Arabidopsis thaliana*. A, *VuKNU-L1$_{pro}$:GUS* expression in the developing ovule of *Arabidopsis*.

1195 B, *VuRKD1$_{pro}$:GUS* expression in the egg cell. C, *VuRKD1$_{pro}$:GUS* expression in a sperm cell. D,

1196 *VuRKD1$_{pro}$:GUS* expression in the vegetative cell. E, *VuGULL06$_{pro}$:GUS* in the egg apparatus.

1197 F, *VuXTH32$_{pro}$:GUS* expression in the vegetative cell. G, *VuGIM2$_{pro}$:GUS* expression in the

1198 vegetative cell. Scale Bars: A, F, and G = 10 µm; B and D = 20 µm; C and D = 8 µm.

1199

1200 **Supplemental Figure. 7.** Phylogeny of the AGO family in cowpea and *Arabidopsis*

1201 A Jukes-Cantor Neighbour-Joining consensus tree showing the relationship between the

1202 AGO genes identified in cowpea (Vu) and those known in *Arabidopsis* (At).

1203 Parameters for tree building were ClustalW MSA with BLOSUM cost matrix, Gap open cost
1204 10, Gap extend cost 0.1. Jukes-Cantor Neighbour-Joining consensus tree, no outgroup,
1205 Bootstrap resampling with 1000 replicates, 50% support threshold.
1206
1207 **Supplemental Figure. 8.** Phylogeny of *Arabidopsis* and cowpea histone proteins
1208 Jukes-Cantor Neighbour-Joining consensus tree showing the relationship between the
1209 *Arabidopsis* and cowpea histones is shown. Histone1 (orange), Histone2 (blue), Histone3
1210 (green), Histone4 (pink), CenH3 (yellow). A Parameters for tree building were ClustalW MSA
1211 with BLOSUM cost matrix, Gap open cost 10, Gap extend cost 0.1. Jukes-Cantor Neighbour-
1212 Joining consensus tree, no outgroup, Bootstrap resampling with 1000 replicates, 50%
1213 support threshold.
1214
1215

1216 **Literature Cited**

1217 ANDERSON, S. N., JOHNSON, C. S., JONES, D. S., CONRAD, L. J., GOU, X., RUSSELL, S. D. &
1218 SUNDARESAN, V. 2013. Transcriptomes of isolated Oryza sativa gametes
1219 characterized by deep sequencing: evidence for distinct sex-dependent chromatin
1220 and epigenetic states before fertilization. *Plant J,* 76**,** 729-41.
1221 ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS,
1222 A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-
1223 TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD,
1224 M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene ontology: tool for the unification of
1225 biology. The Gene Ontology Consortium. *Nat Genet,* 25**,** 25-9.
1226 BARTRINA, I., OTTO, E., STRNAD, M., WERNER, T. & SCHMULLING, T. 2011. Cytokinin
1227 regulates the activity of reproductive meristems, flower organ size, ovule formation,
1228 and thus seed yield in Arabidopsis thaliana. *Plant Cell,* 23**,** 69-80.
1229 BASS, H. W., MARSHALL, W. F., SEDAT, J. W., AGARD, D. A. & CANDE, W. Z. 1997. Telomeres
1230 cluster de novo before the initiation of synapsis: a three-dimensional spatial analysis
1231 of telomere positions before and during meiotic prophase. *J Cell Biol,* 137**,** 5-18.
1232 BELMONTE, M. F., KIRKBRIDE, R. C., STONE, S. L., PELLETIER, J. M., BUI, A. Q., YEUNG, E. C.,
1233 HASHIMOTO, M., FEI, J., HARADA, M., MUNOZ, M. D., LE, B. H., DREWS, G. N.,
1234 BRADY, S. M., GOLDBERG, R. B. & HARADA, J. J. 2013. Comprehensive developmental

1235 profiles of gene activity in regions and subregions of the Arabidopsis seed.
1236 *Proceedings of the National Academy of Sciences of the United States of America,*
1237 110, E435-E444.
1238 BERGER, F., HAMAMURA, Y., INGOUFF, M. & HIGASHIYAMA, T. 2008. Double fertilization -
1239 caught in the act. *Trends Plant Sci,* 13, 437-43.
1240 BORG, M. & BERGER, F. 2015. Chromatin remodelling during male gametophyte
1241 development. *Plant J,* 83, 177-88.
1242 BORGES, F. & MARTIENSSEN, R. A. 2015. The expanding world of small RNAs in plants.
1243 *Nature Reviews Molecular Cell Biology,* 16, 727-741.
1244 BREWBAKER, J. L. 1967. THE DISTRIBUTION AND PHYLOGENETIC SIGNIFICANCE OF
1245 BINUCLEATE AND TRINUCLEATE POLLEN GRAINS IN THE ANGIOSPERMS. *American*
1246 *Journal of Botany,* 54, 1069-1083.
1247 BREWBAKER, J. L. & KWACK, B. H. 1963. THE ESSENTIAL ROLE OF CALCIUM ION IN POLLEN
1248 GERMINATION AND POLLEN TUBE GROWTH. *American Journal of Botany,* 50, 859-
1249 865.
1250 BUCHFINK, B., XIE, C. & HUSON, D. H. 2015. Fast and sensitive protein alignment using
1251 DIAMOND. *Nat Methods,* 12, 59-60.
1252 CHEN, J., STRIEDER, N., KROHN, N. G., CYPRYS, P., SPRUNCK, S., ENGELMANN, J. C. &
1253 DRESSELHAUS, T. 2017. Zygotic Genome Activation Occurs Shortly after Fertilization
1254 in Maize. *Plant Cell,* 29, 2106-2125.
1255 CLOUGH, S. J. & BENT, A. F. 1998. Floral dip: a simplified method for Agrobacterium-
1256 mediated transformation of Arabidopsis thaliana. *Plant J,* 16, 735-43.
1257 CROMER, L., HEYMAN, J., TOUATI, S., HARASHIMA, H., ARAOU, E., GIRARD, C., HORLOW, C.,
1258 WASSMANN, K., SCHNITTGER, A., DE VEYLDER, L. & MERCIER, R. 2012. OSD1
1259 promotes meiotic progression via APC/C inhibition and forms a regulatory network
1260 with TDM and CYCA1;2/TAM. *PLoS Genet,* 8, e1002865.
1261 CURTIS, M. D. & GROSSNIKLAUS, U. 2003. A gateway cloning vector set for high-throughput
1262 functional analysis of genes in planta. *Plant Physiol,* 133, 462-9.
1263 DANTE, R. A., LARKINS, B. A. & SABELLI, P. A. 2014. Cell cycle control and seed development.
1264 *Front Plant Sci,* 5, 493.
1265 DE SCHUTTER, K., JOUBES, J., COOLS, T., VERKEST, A., CORELLOU, F., BABIYCHUK, E., VAN
1266 DER SCHUEREN, E., BEECKMAN, T., KUSHNIR, S., INZE, D. & DE VEYLDER, L. 2007.
1267 Arabidopsis WEE1 kinase controls cell cycle arrest in response to activation of the
1268 DNA integrity checkpoint. *Plant Cell,* 19, 211-25.
1269 DENG, L., ZHANG, S., WANG, G., FAN, S., LI, M., CHEN, W., TU, B., TAN, J., WANG, Y., MA, B.,
1270 LI, S. & QIN, P. 2017. Down-Regulation of OsEMF2b Caused Semi-sterility Due to
1271 Anther and Pollen Development Defects in Rice. *Front Plant Sci,* 8, 1998.
1272 FAN, D., WANG, X., TANG, X., YE, X., REN, S., WANG, D. & LUO, K. 2018. Histone H3K9
1273 demethylase JMJ25 epigenetically modulates anthocyanin biosynthesis in poplar.
1274 *Plant J,* 96, 1121-1136.
1275 FRIEDMAN, W. E. 1999. Expression of the cell cycle in sperm of Arabidopsis: implications for
1276 understanding patterns of gametogenesis and fertilization in plants and other
1277 eukaryotes. *Development,* 126, 1065-1075.
1278 GARCÍA-AGUILAR, M., DORANTES-ACOSTA, A., PÉREZ-ESPAÑA, V. & VIELLE-CALZADA, J.-P.
1279 2005. Whole-mount in situ mRNA localization in developing ovules and seeds
1280 ofArabidopsis. *Plant Molecular Biology Reporter,* 23, 279-289.

1281 GIBSON, S. W. & TODD, C. D. 2015. Arabidopsis AIR12 influences root development. *Physiol*
1282 *Mol Biol Plants,* 21**,** 479-89.
1283 GREEN-ARMYTAGE, P. 2010. A Colour Alphabet and the Limits of Colour Coding. *Colour:*
1284 *Design & Creativity,* 5**,** 1-23.
1285 GROSSNIKLAUS, U. & PARO, R. 2014. Transcriptional silencing by polycomb-group proteins.
1286 *Cold Spring Harb Perspect Biol,* 6**,** a019331.
1287 HAERIZADEH, F., WONG, C. E., BHALLA, P. L., GRESSHOFF, P. M. & SINGH, M. B. 2009.
1288 Genomic expression profiling of mature soybean (Glycine max) pollen. *BMC Plant*
1289 *Biol,* 9**,** 25.
1290 HAGA, N., KOBAYASHI, K., SUZUKI, T., MAEO, K., KUBO, M., OHTANI, M., MITSUDA, N.,
1291 DEMURA, T., NAKAMURA, K., JURGENS, G. & ITO, M. 2011. Mutations in MYB3R1
1292 and MYB3R4 cause pleiotropic developmental defects and preferential down-
1293 regulation of multiple G2/M-specific genes in Arabidopsis. *Plant Physiol,* 157**,** 706-17.
1294 HANDS, P., RABIGER, D. S. & KOLTUNOW, A. 2016. Mechanisms of endosperm initiation.
1295 *Plant Reprod,* 29**,** 215-25.
1296 HUANG, K. L., MA, G. J., ZHANG, M. L., XIONG, H., WU, H., ZHAO, C. Z., LIU, C. S., JIA, H. X.,
1297 CHEN, L., KJORVEN, J. O., LI, X. B. & REN, F. 2018. The ARF7 and ARF19 Transcription
1298 Factors Positively Regulate PHOSPHATE STARVATION RESPONSE1 in Arabidopsis
1299 Roots. *Plant Physiol,* 178**,** 413-427.
1300 IBARRA, C. A., FENG, X., SCHOFT, V. K., HSIEH, T. F., UZAWA, R., RODRIGUES, J. A., ZEMACH,
1301 A., CHUMAK, N., MACHLICOVA, A., NISHIMURA, T., ROJAS, D., FISCHER, R. L.,
1302 TAMARU, H. & ZILBERMAN, D. 2012. Active DNA demethylation in plant companion
1303 cells reinforces transposon methylation in gametes. *Science,* 337**,** 1360-1364.
1304 JACKSON, D. 1992. In-Situ Hybridization in Plants. *In:* GURR, S. J., MCPHERSON, M. J. &
1305 BOWLES, D. J. (eds.) *Molecular Plant Pathology: A Practical Approach* Oxford
1306 University Press.
1307 JURANIC, M., TUCKER, M. R., SCHULTZ, C. J., SHIRLEY, N. J., TAYLOR, J. M., SPRIGGS, A.,
1308 JOHNSON, S. D., BULONE, V. & KOLTUNOW, A. M. 2018. Asexual Female
1309 Gametogenesis Involves Contact with a Sexually-Fated Megaspore in Apomictic
1310 Hieracium. *Plant Physiol,* 177**,** 1027-1049.
1311 KAWASHIMA, T. & BERGER, F. 2014. Epigenetic reprogramming in plant sexual reproduction.
1312 *Nat Rev Genet,* 15**,** 613-24.
1313 KIM, T. W., HWANG, J. Y., KIM, Y. S., JOO, S. H., CHANG, S. C., LEE, J. S., TAKATSUTO, S. &
1314 KIM, S. K. 2005. Arabidopsis CYP85A2, a cytochrome P450, mediates the Baeyer-
1315 Villiger oxidation of castasterone to brassinolide in brassinosteroid biosynthesis.
1316 *Plant Cell,* 17**,** 2397-412.
1317 KLEPIKOVA, A. V., KASIANOV, A. S., GERASIMOV, E. S., LOGACHEVA, M. D. & PENIN, A. A.
1318 2016. A high resolution map of the Arabidopsis thaliana developmental
1319 transcriptome based on RNA-seq profiling. *Plant J,* 88**,** 1058-1070.
1320 KOBAYASHI, K., SUZUKI, T., IWATA, E., MAGYAR, Z., BOGRE, L. & ITO, M. 2015. MYB3Rs,
1321 plant homologs of Myb oncoproteins, control cell cycle-regulated transcription and
1322 form DREAM-like complexes. *Transcription,* 6**,** 106-11.
1323 KOMAKI, S. & SUGIMOTO, K. 2012. Control of the plant cell cycle by developmental and
1324 environmental cues. *Plant Cell Physiol,* 53**,** 953-64.
1325 KOSZEGI, D., JOHNSTON, A. J., RUTTEN, T., CZIHAL, A., ALTSCHMIED, L., KUMLEHN, J., WUST,
1326 S. E., KIRIOUKHOVA, O., GHEYSELINCK, J., GROSSNIKLAUS, U. & BAUMLEIN, H. 2011.

1327 Members of the RKD transcription factor family induce an egg cell-like gene
1328 expression program. *Plant J,* 67**,** 280-91.
1329 KUBO, T., FUJITA, M., TAKAHASHI, H., NAKAZONO, M., TSUTSUMI, N. & KURATA, N. 2013.
1330 Transcriptome analysis of developing ovules in rice isolated by laser microdissection.
1331 *Plant Cell Physiol,* 54**,** 750-65.
1332 LAWIT, S. J., CHAMBERLIN, M. A., AGEE, A., CASWELL, E. S. & ALBERTSEN, M. C. 2013.
1333 Transgenic manipulation of plant embryo sacs tracked through cell-type-specific
1334 fluorescent markers: cell labeling, cell ablation, and adventitious embryos. *Plant*
1335 *Reproduction,* 26**,** 125-137.
1336 LI, L., ZHAO, Y. F., MCCAIG, B. C., WINGERD, B. A., WANG, J. H., WHALON, M. E., PICHERSKY,
1337 E. & HOWE, G. A. 2004. The tomato homolog of CORONATINE-INSENSITIVE1 is
1338 required for the maternal control of seed maturation, jasmonate-signaled defense
1339 responses, and glandular trichome development (vol 16, pg 126, 2004). *Plant Cell,*
1340 16**,** 783-783.
1341 LITUIEV, D. S., KROHN, N. G., MULLER, B., JACKSON, D., HELLRIEGEL, B., DRESSELHAUS, T. &
1342 GROSSNIKLAUS, U. 2013. Theoretical and experimental evidence indicates that there
1343 is no detectable auxin gradient in the angiosperm female gametophyte.
1344 *Development,* 140**,** 4544-4553.
1345 LIU, P., ZHANG, S., ZHOU, B., LUO, X., ZHOU, X. F., CAI, B., JIN, Y. H., NIU, LIN, J., CAO, X. &
1346 JIN, J. B. 2019. The Histone H3K4 Demethylase JMJ16 Represses Leaf Senescence in
1347 Arabidopsis. *Plant Cell,* 31**,** 430-443.
1348 LONARDI, S., MUNOZ-AMATRIAIN, M., LIANG, Q., SHU, S., WANAMAKER, S. I., LO, S.,
1349 TANSKANEN, J., SCHULMAN, A. H., ZHU, T., LUO, M. C., ALHAKAMI, H., OUNIT, R.,
1350 HASAN, A. M., VERDIER, J., ROBERTS, P. A., SANTOS, J. R. P., NDEVE, A., DOLEZEL, J.,
1351 VRANA, J., HOKIN, S. A., FARMER, A. D., CANNON, S. B. & CLOSE, T. J. 2019. The
1352 genome of cowpea (Vigna unguiculata [L.] Walp.). *Plant J,* 98**,** 767-782.
1353 LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and
1354 dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.
1355 MA, C. & WANG, X. 2012. Application of the Gini correlation coefficient to infer regulatory
1356 relationships in transcriptome analysis. *Plant Physiol,* 160**,** 192-203.
1357 MARTINEZ, G., PANDA, K., KOHLER, C. & SLOTKIN, R. K. 2016. Silencing in sperm cells is
1358 directed by RNA movement from the surrounding nurse cell. *Nat Plants,* 2**,** 16030.
1359 MOGENSEN, H. L. & HOLM, P. B. 1995. Dynamics of Nuclear-DNA Quantities during Zygote
1360 Development in Barley. *Plant Cell,* 7**,** 487-494.
1361 MOZGOVA, I. & HENNIG, L. 2015. The polycomb group protein regulatory network. *Annu*
1362 *Rev Plant Biol,* 66**,** 269-96.
1363 MUNOZ-AMATRIAIN, M., MIREBRAHIM, H., XU, P., WANAMAKER, S. I., LUO, M., ALHAKAMI,
1364 H., ALPERT, M., ATOKPLE, I., BATIENO, B. J., BOUKAR, O., BOZDAG, S., CISSE, N.,
1365 DRABO, I., EHLERS, J. D., FARMER, A., FATOKUN, C., GU, Y. Q., GUO, Y. N., HUYNH, B.
1366 L., JACKSON, S. A., KUSI, F., LAWLEY, C. T., LUCAS, M. R., MA, Y., TIMKO, M. P., WU, J.,
1367 YOU, F., BARKLEY, N. A., ROBERTS, P. A., LONARDI, S. & CLOSE, T. J. 2017. Genome
1368 resources for climate-resilient cowpea, an essential crop for food security. *Plant J,*
1369 89**,** 1042-1054.
1370 OKADA, T., HU, Y. K., TUCKER, M. R., TAYLOR, J. M., JOHNSON, S. D., SPRIGGS, A., TSUCHIYA,
1371 T., OELKERS, K., RODRIGUES, J. C. M. & KOLTUNOW, A. M. G. 2013. Enlarging Cells
1372 Initiating Apomixis in Hieracium praealtum Transition to an Embryo Sac Program
1373 prior to Entering Mitosis. *Plant Physiology,* 163**,** 216-231.

1374 OLMEDO-MONFIL, V., DURAN-FIGUEROA, N., ARTEAGA-VAZQUEZ, M., DEMESA-AREVALO,
1375 E., AUTRAN, D., GRIMANELLI, D., SLOTKIN, R. K., MARTIENSSEN, R. A. & VIELLE-
1376 CALZADA, J. P. 2010. Control of female gamete formation by a small RNA pathway in
1377 Arabidopsis. *Nature,* 464**,** 628-32.
1378 PAGNUSSAT, G. C., ALANDETE-SAEZ, M., BOWMAN, J. L. & SUNDARESAN, V. 2009. Auxin-
1379 Dependent Patterning and Gamete Specification in the Arabidopsis Female
1380 Gametophyte. *Science,* 324**,** 1684-1689.
1381 PANOLI, A., MARTIN, M. V., ALANDETE-SAEZ, M., SIMON, M., NEFF, C., SWARUP, R.,
1382 BELLIDO, A., YUAN, L., PAGNUSSAT, G. C. & SUNDARESAN, V. 2015. Auxin Import and
1383 Local Auxin Biosynthesis Are Required for Mitotic Divisions, Cell Expansion and Cell
1384 Specification during Female Gametophyte Development in Arabidopsis thaliana. *Plos*
1385 *One,* 10.
1386 PAYNE, T., JOHNSON, S. D. & KOLTUNOW, A. M. 2004. KNUCKLES (KNU) encodes a C2H2
1387 zinc-finger protein that regulates development of basal pattern elements of the
1388 Arabidopsis gynoecium. *Development,* 131**,** 3737-49.
1389 PEGLER, J. L., OULTRAM, J. M. J., GROF, C. P. L. & EAMENS, A. L. 2019. Profiling the Abiotic
1390 Stress Responsive microRNA Landscape of Arabidopsis thaliana. *Plants (Basel),* 8.
1391 PENTERMAN, J., ZILBERMAN, D., HUH, J. H., BALLINGER, T., HENIKOFF, S. & FISCHER, R. L.
1392 2007. DNA demethylation in the Arabidopsis genome. *Proceedings of the National*
1393 *Academy of Sciences of the United States of America,* 104**,** 6752-6757.
1394 PIEN, S. & GROSSNIKLAUS, U. 2007. Polycomb group and trithorax group proteins in
1395 Arabidopsis. *Biochim Biophys Acta,* 1769**,** 375-82.
1396 PILLOT, M., AUTRAN, D., LEBLANC, O. & GRIMANELLI, D. 2010. A role for
1397 CHROMOMETHYLASE3 in mediating transposon and euchromatin silencing during
1398 egg cell reprogramming in Arabidopsis. *Plant Signal Behav,* 5**,** 1167-70.
1399 POPELKA, J. C., GOLLASCH, S., MOORE, A., MOLVIG, L. & HIGGINS, T. J. V. 2006. Genetic
1400 transformation of cowpea (Vigna unguiculata L.) and stable transmission of the
1401 transgenes to progeny. *Plant Cell Reports,* 25**,** 304-312.
1402 RAMAKERS, C., RUIJTER, J. M., DEPREZ, R. H. L. & MOORMAN, A. F. M. 2003. Assumption-
1403 free analysis of quantitative real-time polymerase chain reaction (PCR) data.
1404 *Neuroscience Letters,* 339**,** 62-66.
1405 RAY, S., AGARWAL, P., ARORA, R., KAPOOR, S. & TYAGI, A. K. 2007. Expression analysis of
1406 calcium-dependent protein kinase gene family during reproductive development and
1407 abiotic stress conditions in rice (Oryza sativa L. ssp. indica). *Mol Genet Genomics,*
1408 278**,** 493-505.
1409 RUIJTER, J. M., RAMAKERS, C., HOOGAARS, W. M. H., KARLEN, Y., BAKKER, O., VAN DEN
1410 HOFF, M. J. B. & MOORMAN, A. F. M. 2009. Amplification efficiency: linking baseline
1411 and bias in the analysis of quantitative PCR data. *Nucleic Acids Research,* 37.
1412 RUTLEY, N. & TWELL, D. 2015. A decade of pollen transcriptomics. *Plant Reproduction,* 28**,**
1413 73-89.
1414 SALINAS-GAMBOA, R., JOHNSON, S. D., SANCHEZ-LEON, N., KOLTUNOW, A. M. G. & VIELLE-
1415 CALZADA, J. P. 2016. New observations on gametogenic development and
1416 reproductive experimental tools to support seed yield improvement in cowpea
1417 [Vigna unguiculata (L.) Walp.]. *Plant Reproduction,* 29**,** 165-177.
1418 SCHMIDT, A., WUEST, S. E., VIJVERBERG, K., BAROUX, C., KLEEN, D. & GROSSNIKLAUS, U.
1419 2011. Transcriptome analysis of the Arabidopsis megaspore mother cell uncovers the

1420 importance of RNA helicases for plant germline development. *PLoS Biol,* 9**,**
1421 e1001155.
1422 SCHOFT, V. K., CHUMAK, N., CHOI, Y., HANNON, M., GARCIA-AGUILAR, M., MACHLICOVA, A.,
1423 SLUSARZ, L., MOSIOLEK, M., PARK, J. S., PARK, G. T., FISCHER, R. L. & TAMARU, H.
1424 2011. Function of the DEMETER DNA glycosylase in the Arabidopsis thaliana male
1425 gametophyte. *Proc Natl Acad Sci U S A,* 108**,** 8042-7.
1426 SCOFIELD, S., JONES, A. & MURRAY, J. A. 2014. The plant cell cycle in context. *J Exp Bot,* 65**,**
1427 2557-62.
1428 SHERMOEN, A. W. & O'FARRELL, P. H. 1991. Progression of the Cell-Cycle through Mitosis
1429 Leads to Abortion of Nascent Transcripts. *Cell,* 67**,** 303-310.
1430 SINGH, B. B. 2014. Cowpea: The Food Legume of the 21st Century. *Cowpea: The Food*
1431 *Legume of the 21st Century***,** 1-170.
1432 SLOTKIN, R. K., VAUGHN, M., BORGES, F., TANURDZIC, M., BECKER, J. D., FEIJO, J. A. &
1433 MARTIENSSEN, R. A. 2009. Epigenetic reprogramming and small RNA silencing of
1434 transposable elements in pollen. *Cell,* 136**,** 461-72.
1435 SORNAY, E., FORZANI, C., FORERO-VARGAS, M., DEWITTE, W. & MURRAY, J. A. 2015.
1436 Activation of CYCD7;1 in the central cell and early endosperm overcomes cell-cycle
1437 arrest in the Arabidopsis female gametophyte, and promotes early endosperm and
1438 embryo development. *Plant J,* 84**,** 41-55.
1439 SPRIGGS, A., HENDERSON, S. T., HAND, M. L., JOHNSON, S. D., TAYLOR, J. M. & KOLTUNOW,
1440 A. 2018. Assembled genomic and tissue-specific transcriptomic data resources for
1441 two genetically distinct lines of Cowpea (Vigna unguiculata (L.) Walp). *Gates Open*
1442 *Res,* 2**,** 7.
1443 SPRUNCK, S., RADEMACHER, S., VOGLER, F., GHEYSELINCK, J., GROSSNIKLAUS, U. &
1444 DRESSELHAUS, T. 2012. Egg cell-secreted EC1 triggers sperm cell activation during
1445 double fertilization. *Science,* 338**,** 1093-7.
1446 STANKE, M. & WAACK, S. 2003. Gene prediction with a hidden Markov model and a new
1447 intron submodel. *Bioinformatics,* 19 Suppl 2**,** ii215-25.
1448 STEFFEN, J. G., KANG, I. H., MACFARLANE, J. & DREWS, G. N. 2007. Identification of genes
1449 expressed in the Arabidopsis female gametophyte. *Plant J,* 51**,** 281-92.
1450 SUKAWA, Y. & OKAMOTO, T. 2018. Cell cycle in egg cell and its progression during zygotic
1451 development in rice. *Plant Reprod,* 31**,** 107-116.
1452 SUN, Q. & ZHOU, D. X. 2008. Rice jmjC domain-containing gene JMJ706 encodes H3K9
1453 demethylase required for floral organ development. *Proc Natl Acad Sci U S A,* 105**,**
1454 13679-84.
1455 TALBERT, P. B., AHMAD, K., ALMOUZNI, G., AUSIO, J., BERGER, F., BHALLA, P. L., BONNER, W.
1456 M., CANDE, W. Z., CHADWICK, B. P., CHAN, S. W. L., CROSS, G. A. M., CUI, L. W.,
1457 DIMITROV, S. I., DOENECKE, D., EIRIN-LOPEZ, J. M., GOROVSKY, M. A., HAKE, S. B.,
1458 HAMKALO, B. A., HOLEC, S., JACOBSEN, S. E., KAMIENIARZ, K., KHOCHBIN, S.,
1459 LADURNER, A. G., LANDSMAN, D., LATHAM, J. A., LOPPIN, B., MALIK, H. S.,
1460 MARZLUFF, W. F., PEHRSON, J. R., POSTBERG, J., SCHNEIDER, R., SINGH, M. B.,
1461 SMITH, M. M., THOMPSON, E., TORRES-PADILLA, M. E., TREMETHICK, D. J., TURNER,
1462 B. M., WATERBORG, J. H., WOLLMANN, H., YELAGANDULA, R., ZHU, B. & HENIKOFF,
1463 S. 2012. A unified phylogeny-based nomenclature for histone variants. *Epigenetics &*
1464 *Chromatin,* 5.
1465 THE GENE ONTOLOGY, C. 2017. Expansion of the Gene Ontology knowledgebase and
1466 resources. *Nucleic Acids Res,* 45**,** D331-D338.

1467 TUCKER, M. R., OKADA, T., HU, Y., SCHOLEFIELD, A., TAYLOR, J. M. & KOLTUNOW, A. M.
1468 2012. Somatic small RNA pathways promote the mitotic events of
1469 megagametogenesis during female reproductive development in Arabidopsis.
1470 *Development,* 139**,** 1399-404.
1471 VELAPPAN, Y., SIGNORELLI, S. & CONSIDINE, M. J. 2017. Cell cycle arrest in plants: what
1472 distinguishes quiescence, dormancy and differentiated G1? *Ann Bot,* 120**,** 495-509.
1473 VIELLE-CALZADA, J. P., BASKAR, R. & GROSSNIKLAUS, U. 2000. Delayed activation of the
1474 paternal genome during seed development. *Nature,* 404**,** 91-4.
1475 VIELLE-CALZADA, J. P., THOMAS, J., SPILLANE, C., COLUCCIO, A., HOEPPNER, M. A. &
1476 GROSSNIKLAUS, U. 1999. Maintenance of genomic imprinting at the Arabidopsis
1477 medea locus requires zygotic DDM1 activity. *Genes & Development,* 13**,** 2971-2982.
1478 WANG, G. & KOHLER, C. 2017. Epigenetic processes in flowering plant reproduction. *J Exp*
1479 *Bot,* 68**,** 797-807.
1480 WANG, G., KONG, H., SUN, Y., ZHANG, X., ZHANG, W., ALTMAN, N., DEPAMPHILIS, C. W. &
1481 MA, H. 2004. Genome-wide analysis of the cyclin family in Arabidopsis and
1482 comparative phylogenetic analysis of plant cyclin-like proteins. *Plant Physiol,* 135**,**
1483 1084-99.
1484 WANG, R. & ESTELLE, M. 2014. Diversity and specificity: auxin perception and signaling
1485 through the TIR1/AFB pathway. *Curr Opin Plant Biol,* 21**,** 51-58.
1486 WILLIAMS, J. H., TAYLOR, M. L. & O'MEARA, B. C. 2014. Repeated evolution of tricellular
1487 (and bicellular) pollen. *Am J Bot,* 101**,** 559-71.
1488 WOJCIECHOWSKI, M. F., LAVIN, M. & SANDERSON, M. J. 2004. A phylogeny of legumes
1489 (Leguminosae) based on analysis of the plastid matK gene resolves many well-
1490 supported subclades within the family. *Am J Bot,* 91**,** 1846-62.
1491 WUEST, S. E., VIJVERBERG, K., SCHMIDT, A., WEISS, M., GHEYSELINCK, J., LOHR, M.,
1492 WELLMER, F., RAHNENFUHRER, J., VON MERING, C. & GROSSNIKLAUS, U. 2010.
1493 Arabidopsis female gametophyte gene expression map reveals similarities between
1494 plant and animal gametes. *Curr Biol,* 20**,** 506-12.
1495 YAO, S., JIANG, C., HUANG, Z., TORRES-JEREZ, I., CHANG, J., ZHANG, H., UDVARDI, M., LIU, R.
1496 & VERDIER, J. 2016. The Vigna unguiculata Gene Expression Atlas (VuGEA) from de
1497 novo assembly and quantification of RNA-seq data provides insights into seed
1498 maturation mechanisms. *Plant J,* 88**,** 318-327.
1499 YAO, X., TIAN, L., YANG, J., ZHAO, Y. N., ZHU, Y. X., DAI, X., ZHAO, Y. & YANG, Z. N. 2018.
1500 Auxin production in diploid microsporocytes is necessary and sufficient for early
1501 stages of pollen development. *PLoS Genet,* 14**,** e1007397.
1502 ZHAN, J. P., THAKARE, D., MA, C., LLOYD, A., NIXON, N. M., ARAKAKI, A. M., BURNETT, W. J.,
1503 LOGAN, K. O., WANG, D. F., WANG, X. F., DREWS, G. N. & YADEGARIA, R. 2015. RNA
1504 Sequencing of Laser-Capture Microdissected Compartments of the Maize Kernel
1505 Identifies Regulatory Modules Associated with Endosperm Cell Differentiation. *Plant*
1506 *Cell,* 27**,** 513-531.
1507

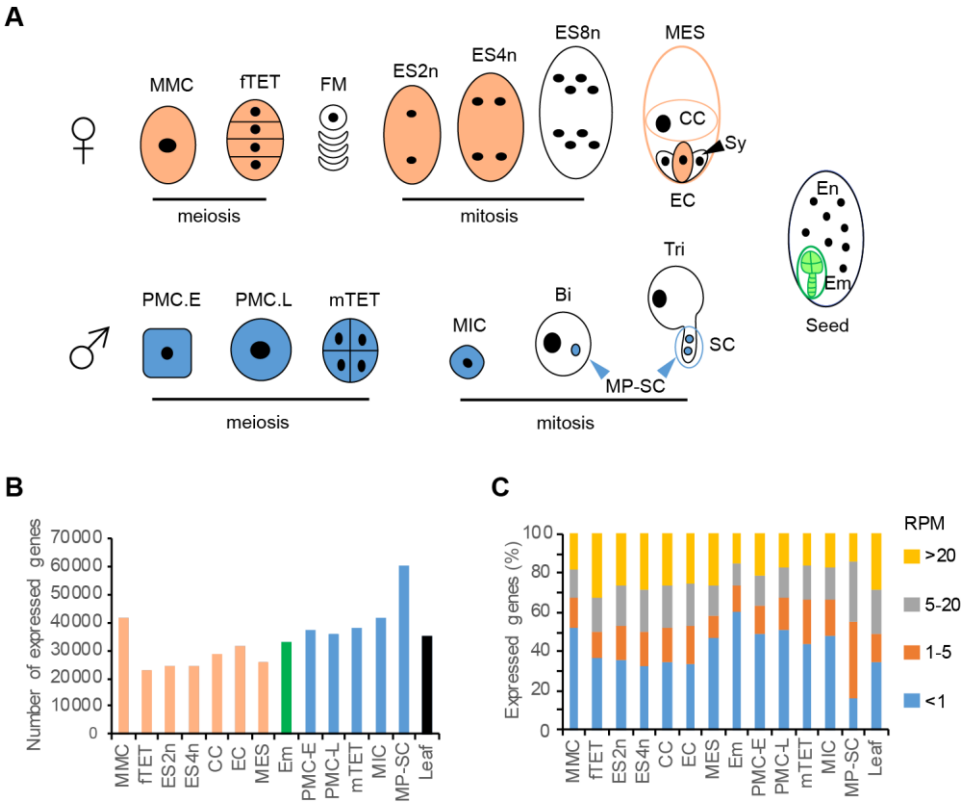


Figure 1. Stages of reproductive development in cowpea and gene expression in developed transcriptomes.

A, Cartoon of temporal stages of male and female gametophyte development showing cell-type stages collected by LCM for RNAseq. For detail refer to text. B, The number of genes expressed in each transcriptome set (RPM >0). C, The proportion of genes expressed at very low (<1RPM), low (1-5RPM), moderate (5-20RPM) or high (>20RPM) levels in RNAseq sets.

Abbreviations: MMC, megaspore mother cell; fTET, female tetrads; FM, functional megaspore selection stage (not collected); ES2n, mitotic embryo sac with 2 nuclei; ES4n, embryo sac with 4 nuclei; MES, mature embryo sac at anthesis, where antipodals have degenerated, containing the egg cell (EC), flanking synergids (Sy, not collected), and central cell (CC); Em, embryo; En, endosperm (not collected); PMC.E, early pollen mother cell; PMC.L, late pollen mother cell; mTET, male tetrads; MIC, uninucleate microspore; Bi, bicellular pollen stage; Tri, tricellular pollen stage, two sperm cells (SC) in the emerging pollen tube; MP-SC, mature pollen-sperm cell.

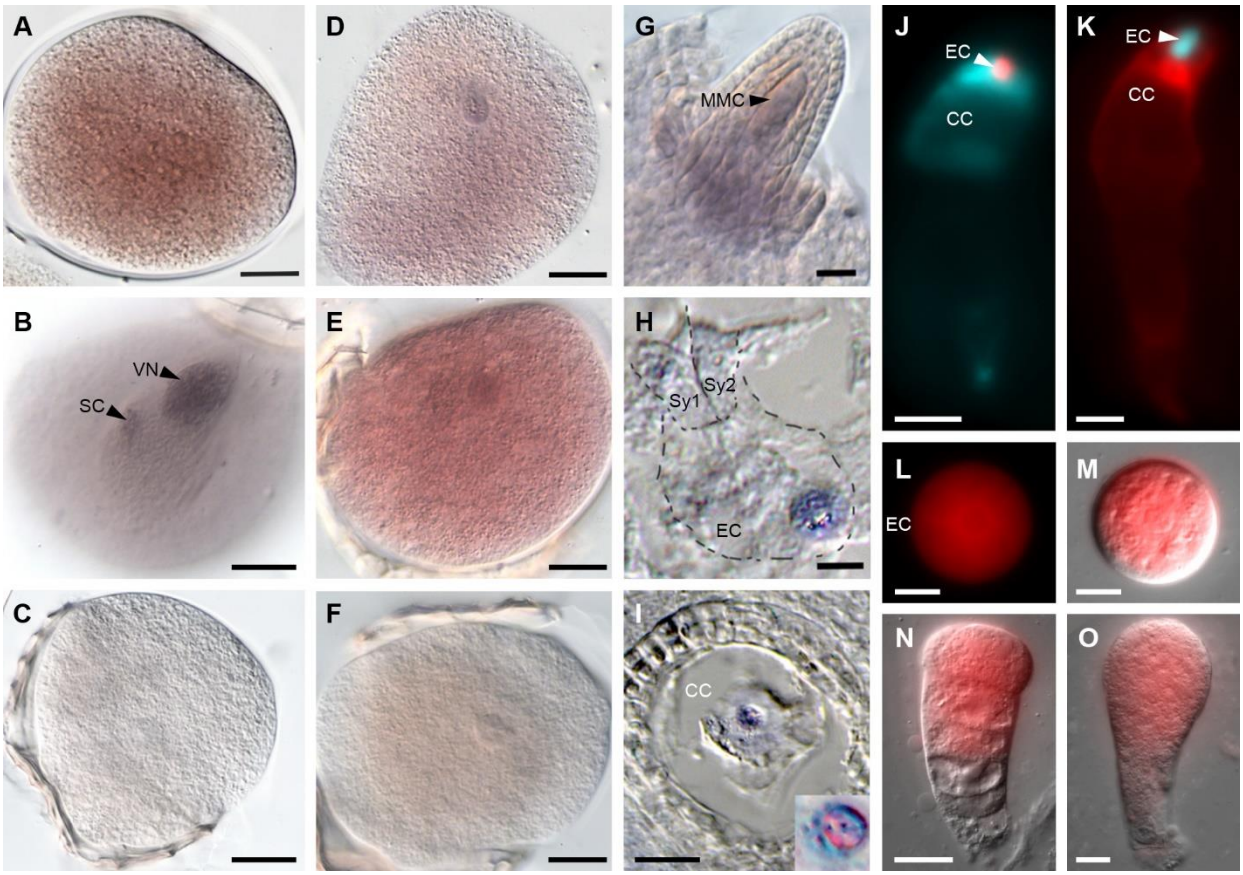


Figure 2. Gene expression in cowpea reproductive cell-types.
A – J, Whole-mount *in situ* hybridization (WISH) in pollen grains of cowpea using four different antisense probes. A-C, *VuXTH6* mRNA in sperm cells A and vegetative nucleus B, with no signal in the sense control C. D and E, *VuXTH32* was detected in sperm cell and vegetative cell cytoplasm. F, *VuXTH32* sense controls. G, Whole-mount ovule undergoing megasporogenesis showing *VuKNU1* mRNA localized in the MMC, the nucellus and the initial inner integument. H, Paraffin section through a differentiated female gametophyte showing *VuRKD1* mRNA localization in the egg cell and synergids. I, Paraffin section through a differentiated female gametophyte showing *VuNB-ARC* mRNA localization in the central cell. Inset shows fused polar nuclei showing active transcription of *VuNB-ARC*. J, $AtRKD2_{pro}$:*dsRED-Express* (red) in EC and $AtDD9_{pro}$:*AmCyan1* (cyan) in CC one day before anthesis. K, $AtEC1.1_{pro}$:*AmCyan1* (cyan) in EC and $AtDD25_{pro}$:*dsRED-Express* (red) in CC. L-M, $AtDD45_{pro}$:*dsRED-Express* (red) in isolated egg cell, N, two-celled proembryo and O, early embryo.
Abbreviations: EC, egg cell; CC, central cell; Sy, synergids; SC, sperm cells; VN, vegetative nucleus, MMC, megaspore mother cell.
Scale bars: A-G, L-M = 10 µm, H = 8 µm, I = 14 µm, J-K = 50 µm, N-O = 20 µm.

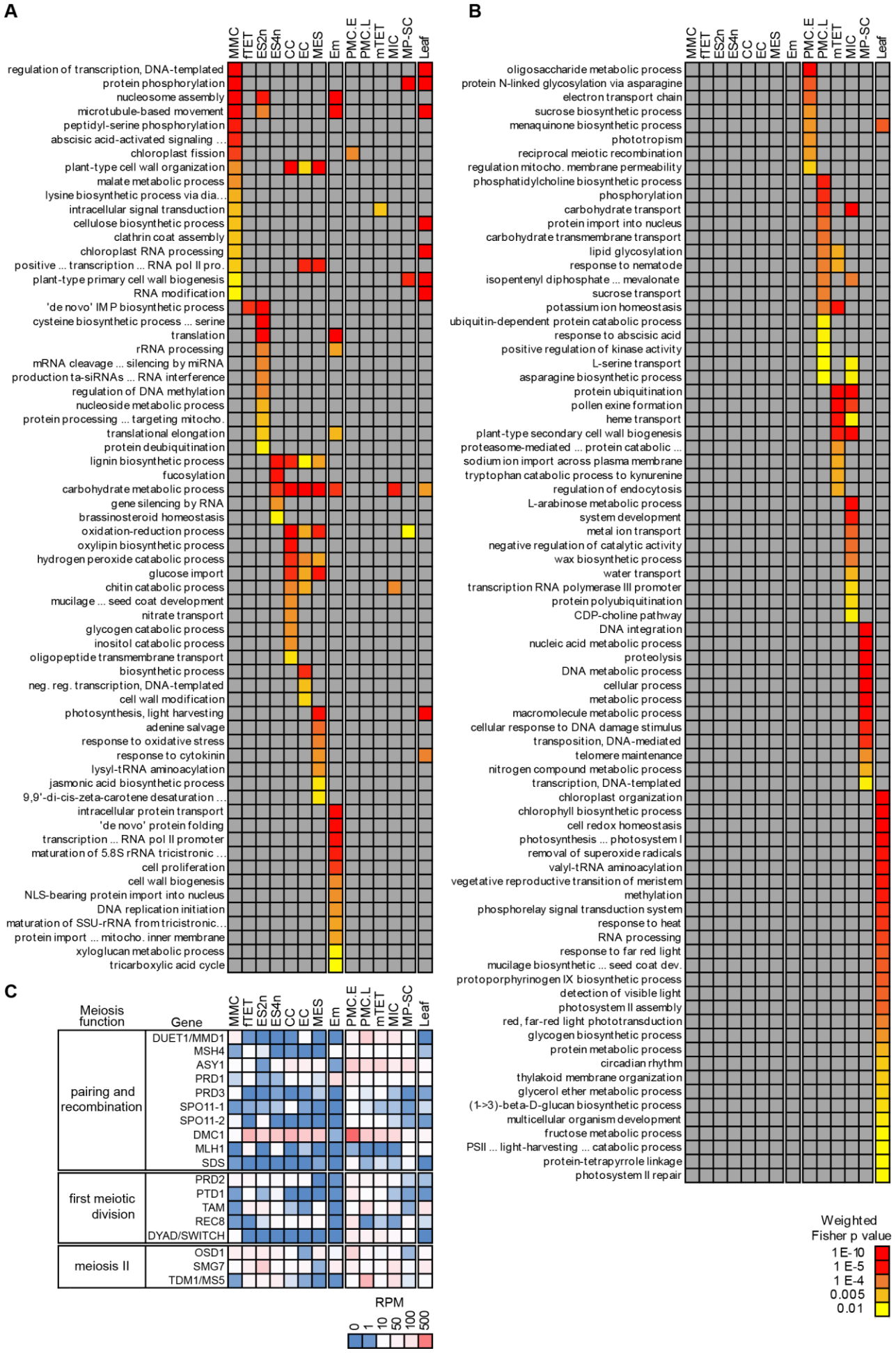

A
MMC
fTET
ES2n
ES4n
CC
EC
MES
Em
PMC.E
PMC.L
mTET
MIC
MP-SC
Leaf
regulation of transcription, DNA-templated
protein phosphorylation
nucleosome assembly
microtubule-based movement
peptidyl-serine phosphorylation
abscisic acid-activated signaling ...
chloroplast fission
plant-type cell wall organization
malate metabolic process
lysine biosynthetic process via dia...
intracellular signal transduction
cellulose biosynthetic process
clathrin coat assembly
chloroplast RNA processing
positive ... transcription ... RNA pol II pro.
plant-type primary cell wall biogenesis
RNA modification
'de novo' IMP biosynthetic process
cysteine biosynthetic process ... serine
translation
rRNA processing
mRNA cleavage ... silencing by miRNA
production ta-siRNAs ... RNA interference
regulation of DNA methylation
nucleoside metabolic process
protein processing ... targeting mitocho.
translational elongation
protein deubiquitination
lignin biosynthetic process
fucosylation
carbohydrate metabolic process
gene silencing by RNA
brassinosteroid homeostasis
oxidation-reduction process
oxylipin biosynthetic process
hydrogen peroxide catabolic process
glucose import
chitin catabolic process
mucilage ... seed coat development
nitrate transport
glycogen catabolic process
inositol catabolic process
oligopeptide transmembrane transport
biosynthetic process
neg. reg. transcription, DNA-templated
cell wall modification
photosynthesis, light harvesting
adenine salvage
response to oxidative stress
response to cytokinin
lysyl-tRNA aminoacylation
jasmonic acid biosynthetic process
9,9'-di-cis-zeta-carotene desaturation ...
intracellular protein transport
'de novo' protein folding
transcription ... RNA pol II promoter
maturation of 5.8S rRNA tricistronic ...
cell proliferation
cell wall biogenesis
NLS-bearing protein import into nucleus
DNA replication initiation
maturation of SSU-rRNA from tricistronic...
protein import ... mitocho. inner membrane
xyloglucan metabolic process
tricarboxylic acid cycle
B
MMC
fTET
ES2n
ES4n
CC
EC
MES
Em
PMC.E
PMC.L
mTET
MIC
MP-SC
Leaf
oligosaccharide metabolic process
protein N-linked glycosylation via asparagine
electron transport chain
sucrose biosynthetic process
menaquinone biosynthetic process
phototropism
reciprocal meiotic recombination
regulation mitocho. membrane permeability
phosphatidylcholine biosynthetic process
phosphorylation
carbohydrate transport
protein import into nucleus
carbohydrate transmembrane transport
lipid glycosylation
response to nematode
isopentenyl diphosphate ... mevalonate
sucrose transport
potassium ion homeostasis
ubiquitin-dependent protein catabolic process
response to abscisic acid
positive regulation of kinase activity
L-serine transport
asparagine biosynthetic process
protein ubiquitination
pollen exine formation
heme transport
plant-type secondary cell wall biogenesis
proteasome-mediated ... protein catabolic ...
sodium ion import across plasma membrane
tryptophan catabolic process to kynurenine
regulation of endocytosis
L-arabinose metabolic process
system development
metal ion transport
negative regulation of catalytic activity
wax biosynthetic process
water transport
transcription RNA polymerase III promoter
protein polyubiquitination
CDP-choline pathway
DNA integration
nucleic acid metabolic process
proteolysis
DNA metabolic process
cellular process
metabolic process
macromolecule metabolic process
cellular response to DNA damage stimulus
transposition, DNA-mediated
telomere maintenance
nitrogen compound metabolic process
transcription, DNA-templated
chloroplast organization
chlorophyll biosynthetic process
cell redox homeostasis
photosynthesis ... photosystem I
removal of superoxide radicals
valyl-tRNA aminoacylation
vegetative reproductive transition of meristem
methylation
phosphorelay signal transduction system
response to heat
RNA processing
response to far red light
mucilage biosynthetic ... seed coat dev.
protoporphyrinogen IX biosynthetic process
detection of visible light
photosystem II assembly
red, far-red light phototransduction
glycogen biosynthetic process
protein metabolic process
circadian rhythm
thylakoid membrane organization
glycerol ether metabolic process
(1->3)-beta-D-glucan biosynthetic process
multicellular organism development
fructose metabolic process
PSII ... light-harvesting ... catabolic process
protein-tetrapyrrole linkage
photosystem II repair
Weighted Fisher p value
1 E-10
1 E-5
1 E-4
0.005
0.01
C
Meiosis function
Gene
MMC
fTET
ES2n
ES4n
CC
EC
MES
Em
PMC.E
PMC.L
mTET
MIC
MP-SC
Leaf
pairing and recombination
DUET1/MMD1
MSH4
ASY1
PRD1
PRD3
SPO11-1
SPO11-2
DMC1
MLH1
SDS
first meiotic division
PRD2
PTD1
TAM
REC8
DYAD/SWITCH
meiosis II
OSD1
SMG7
TDM1/MS5
RPM
0
1
10
50
100
500

Figure 3. Differentially enriched Biological Process GO terms and expression of meiosis genes during gametogenesis.

A, Enriched GO terms primarily in female gametophyte and early embryo transcriptomes. B, Enriched GO terms primarily in the male gametophyte and leaf transcriptomes. The statistical significance of the GO enrichment is represented according to the scale of weighted Fisher p value shown at the bottom right. C, Expression of genes involved in meiosis in the generated cowpea transcriptomes (see also Supplemental Tables 7 and 8). Expression levels relate to the RPM scale.

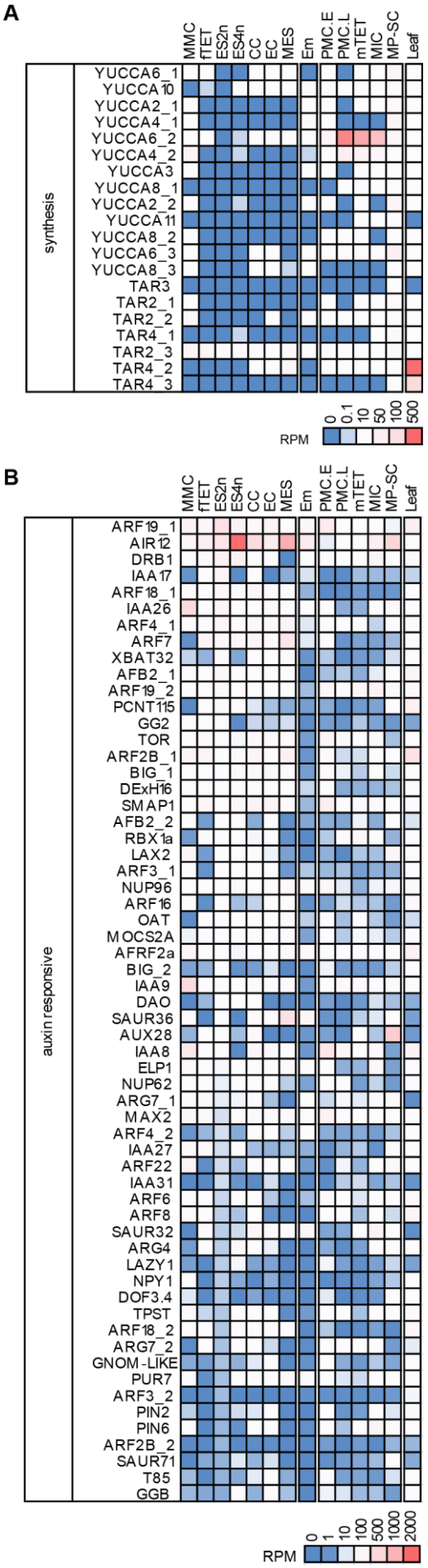


Figure 4. Expression of auxin biosynthesis and auxin responsive genes during meiosis and gametophyte development.
A, Expression of genes belonging to the YUCCA and TAR auxin biosynthesis gene families. B, Expression of 60 auxin responsive genes sorted in decreasing order of expression level in the 2-nucleate embryo sac. Expression levels are related to the RPM scale with each respective panel (see also Supplemental Table 9).

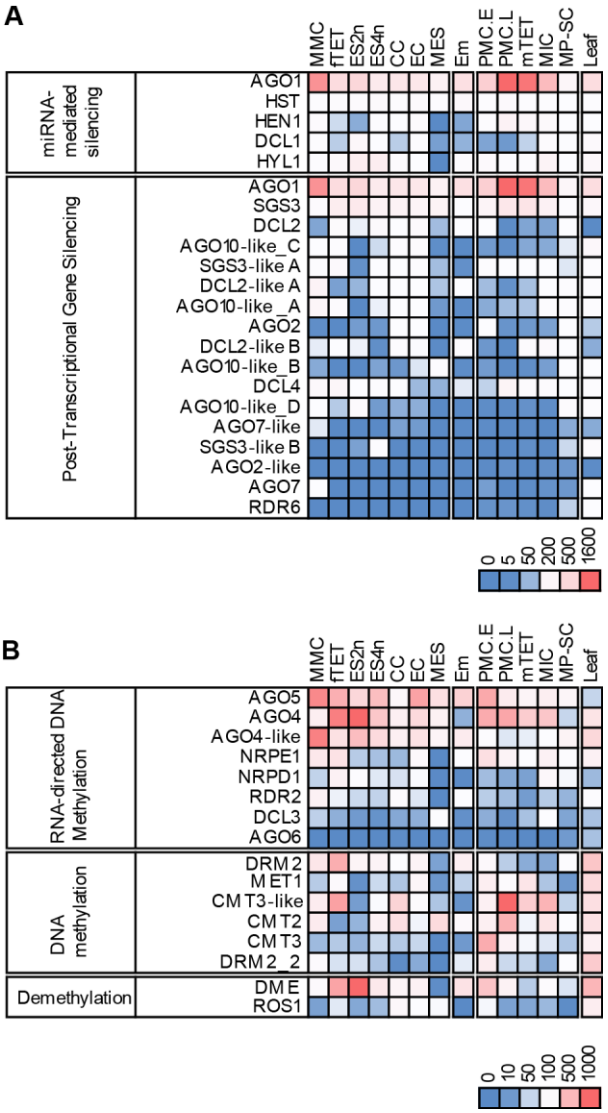


Figure 5. Heatmaps representing the expression of selected genes involved in RNA-mediated gene silencing and DNA methylation.

A, Expression of genes involved in the miRNA-mediated gene silencing pathway and involved in PTGS. B, Expression of genes that interact with RNA and are involved in RdDM, DNA methyltransferases that mediate RdDM and DNA glycosylases that mediate demethylation of DNA. Expression level is according to the RPM scales with each panel (see also Supplemental Tables 12 and 13).

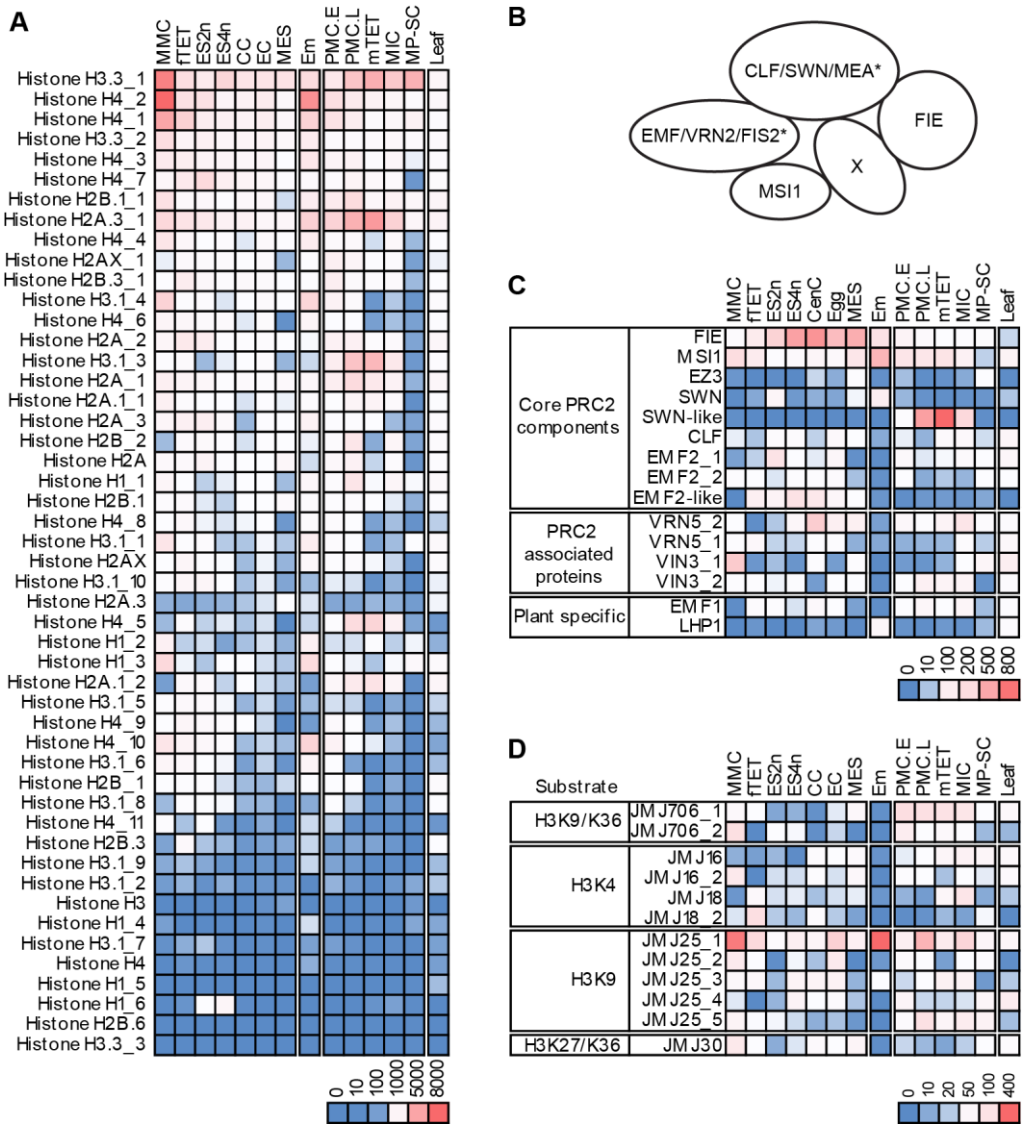


Figure 6. Histone gene expression and the expression of selected genes involved in histone modification.

A, Expression of genes belonging to the core histone family (see Supplemental Table 18). B, Expression of selected PcG genes involved in histone methylation resulting in gene silencing (see Supplemental Table 19). C, Known protein interactions forming Polycomb-group repressive 2 complex types. Accessory proteins (X) are often associated with PRC2 complex function. Asterisks indicate *Arabidopsis* specific proteins. MEA and *Arabidopsis* counterparts, SWN and CLF are SET domain protein and homologs of *Drosophilia* Enhancer of Zeste. FIS2 and others listed are VEFS-box proteins. FIE and MSI1 are WD40 repeat protein containing proteins (Mozgova and Henning, 2015). D, Expression of selected JMJ family histone demethylase genes and demethylation targets, at left (see also Supplemental Tables 14, 15, 16). Expression levels relate to the RPM scales associated with each panel.

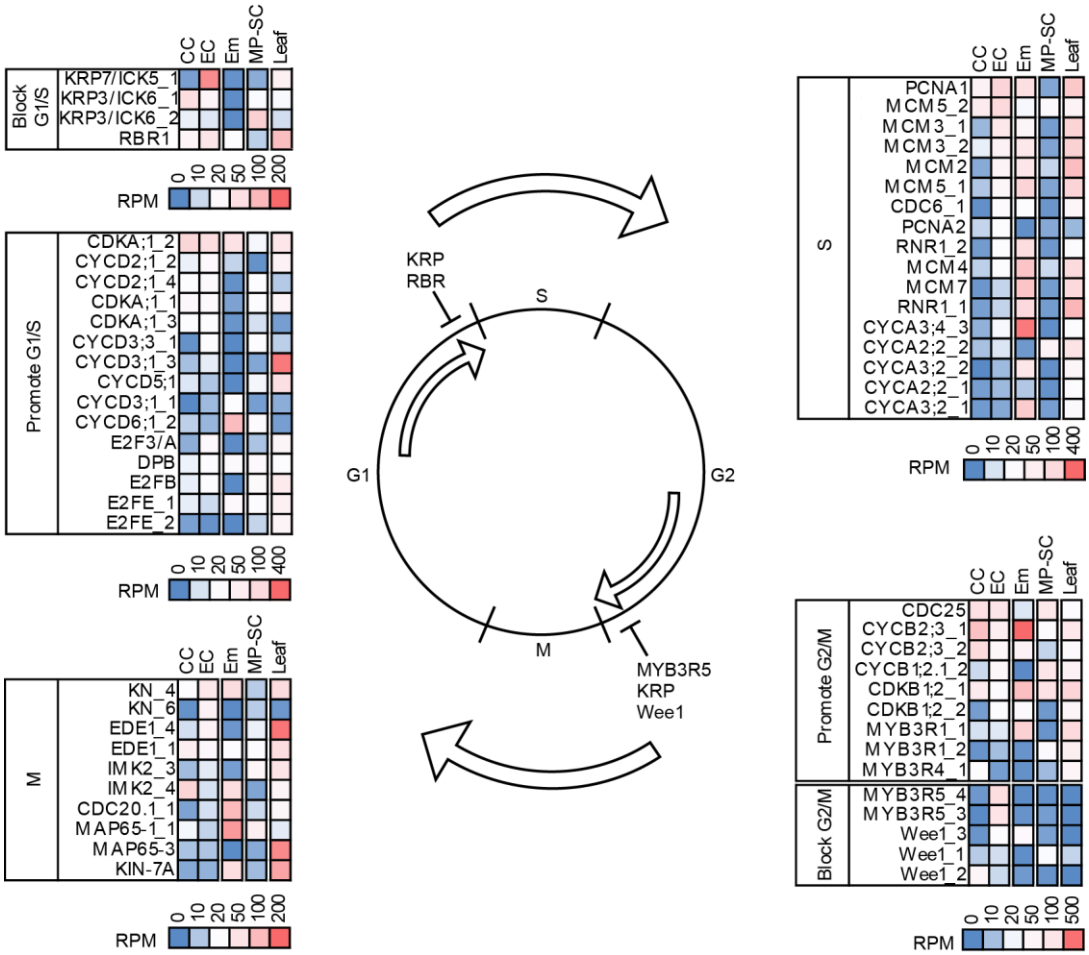
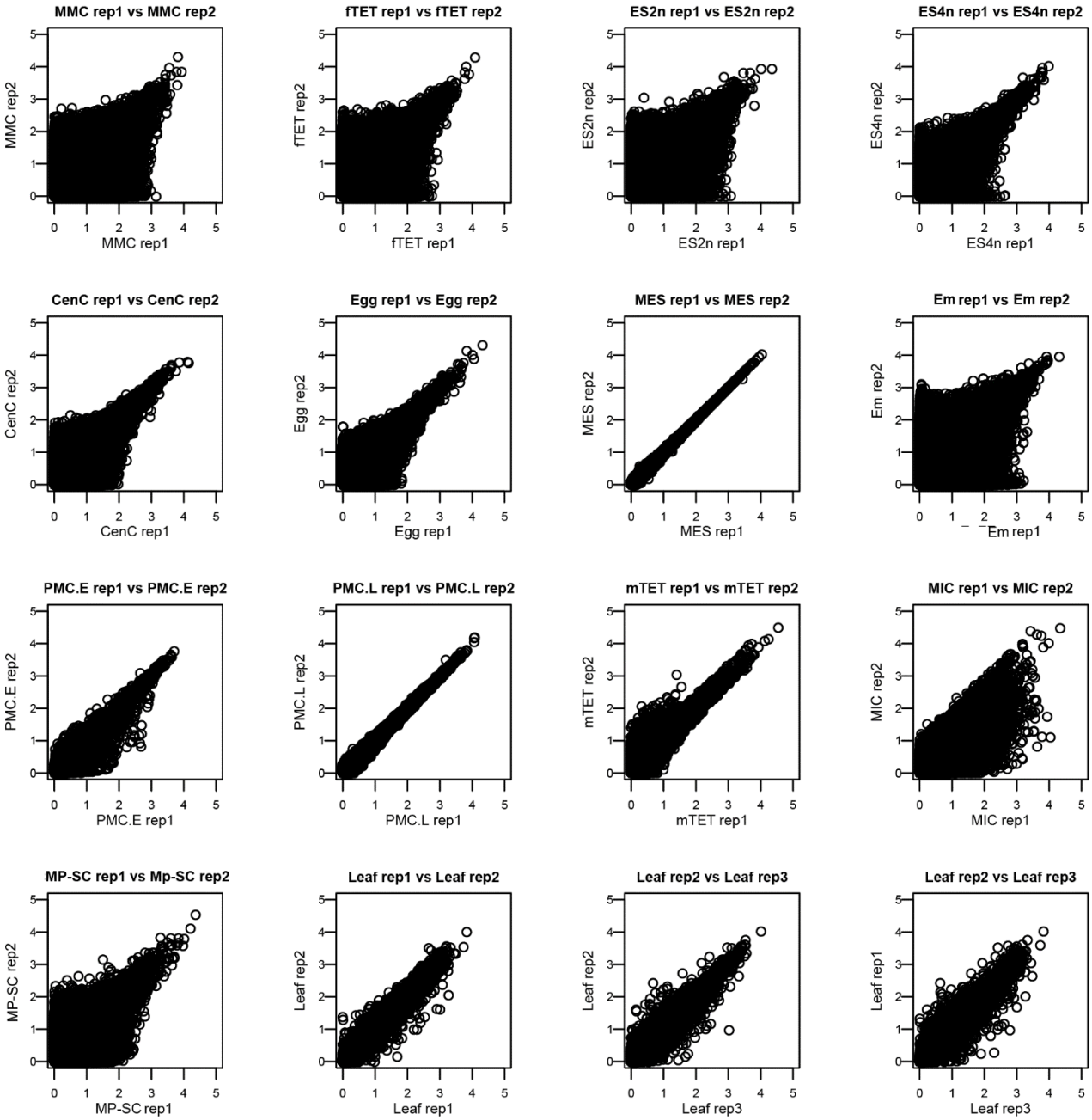
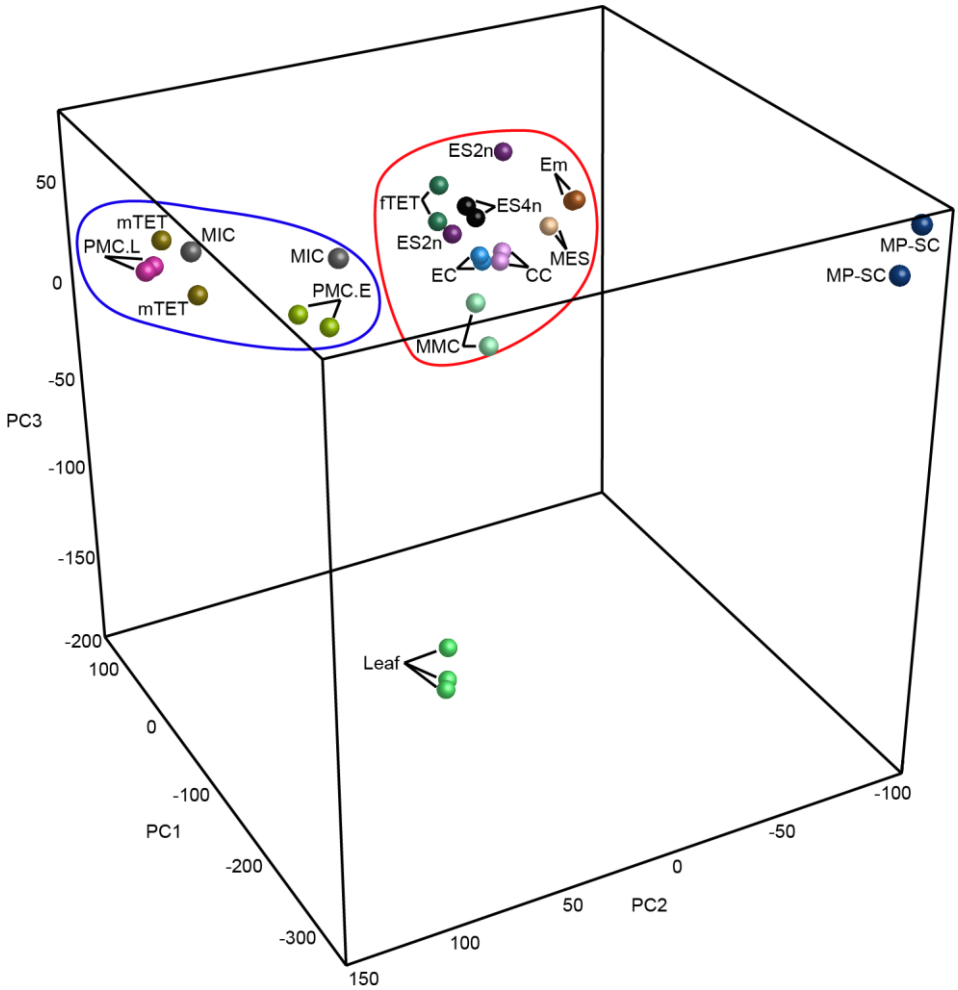


Figure 7. Expression of selected cell cycle genes in egg, central cell, mature pollen-sperm and leaf transcriptomes.

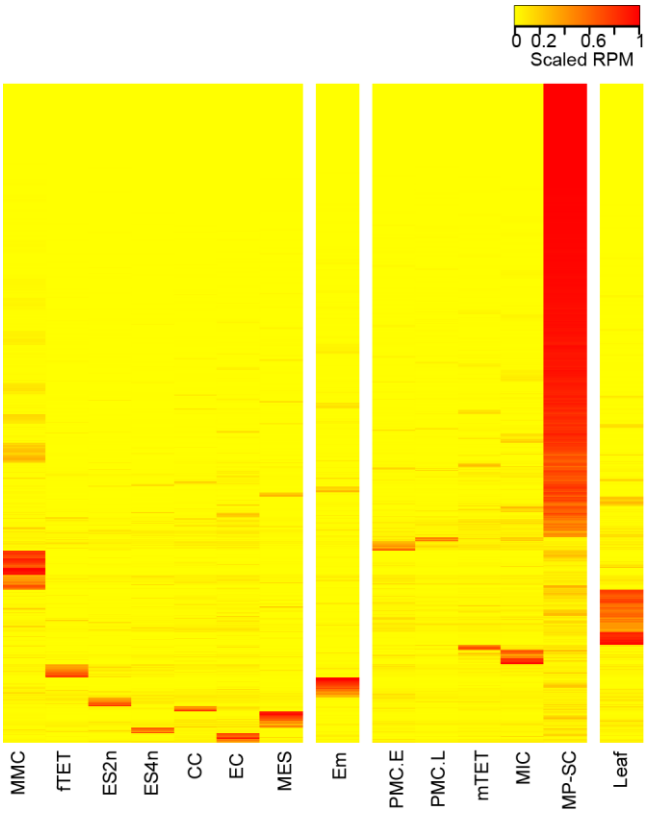
Cell cycle genes are grouped according to their function in the cell cycle relating to roles in promoting or blocking progression through various cell cycle phases. Relative expression level is indicated according to the RPM scale shown at the bottom of each panel (see also Supplemental Tables 10 and 11).



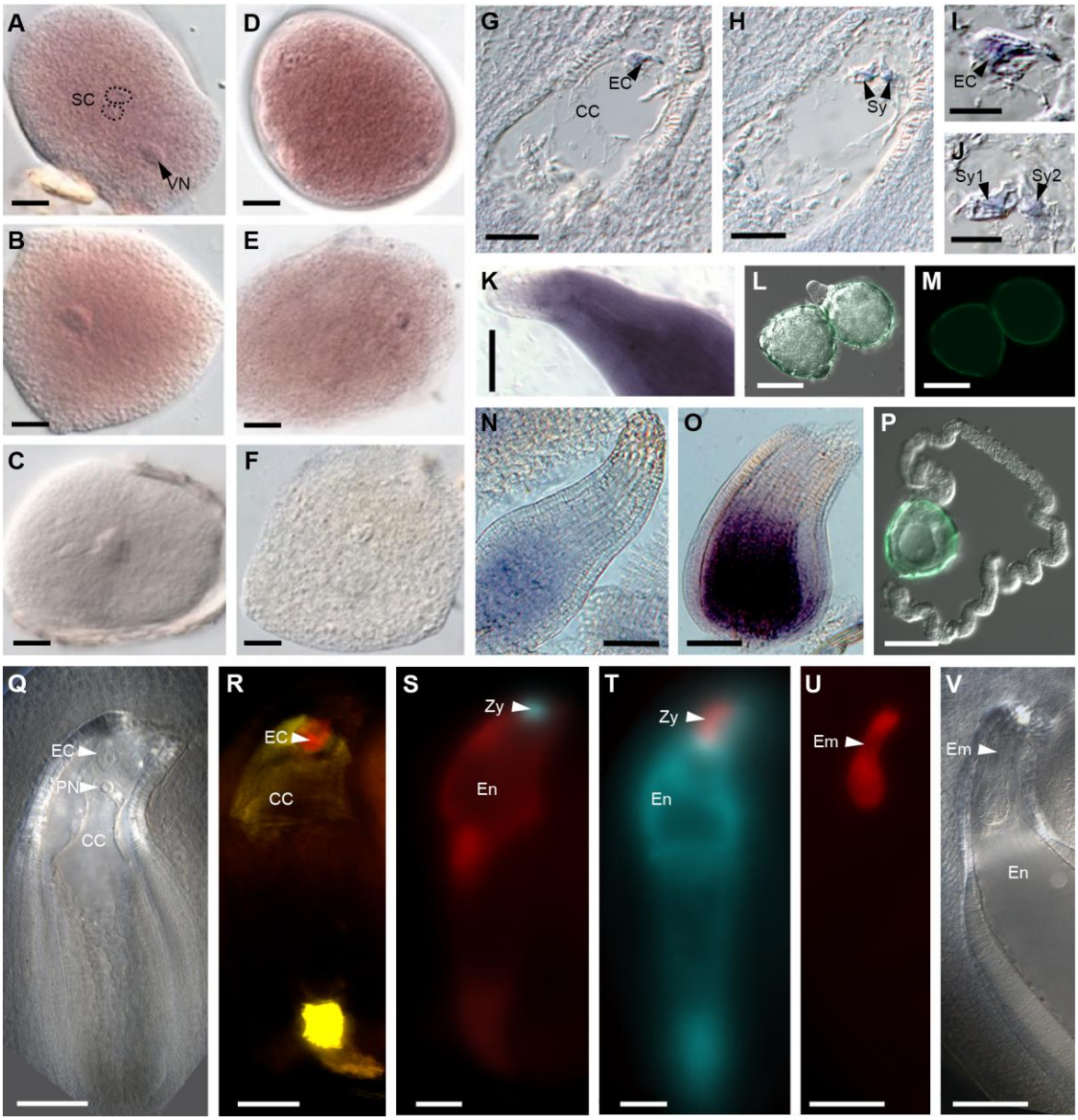
Supplemental Figure 1. Correlation between RNAseq replicates.
Scatter plot of log-transformed mean RPM counts of replicates for each tissue. Replicate 1 and replicate 2 for each cell-type are shown on the Y and X axes, respectively.



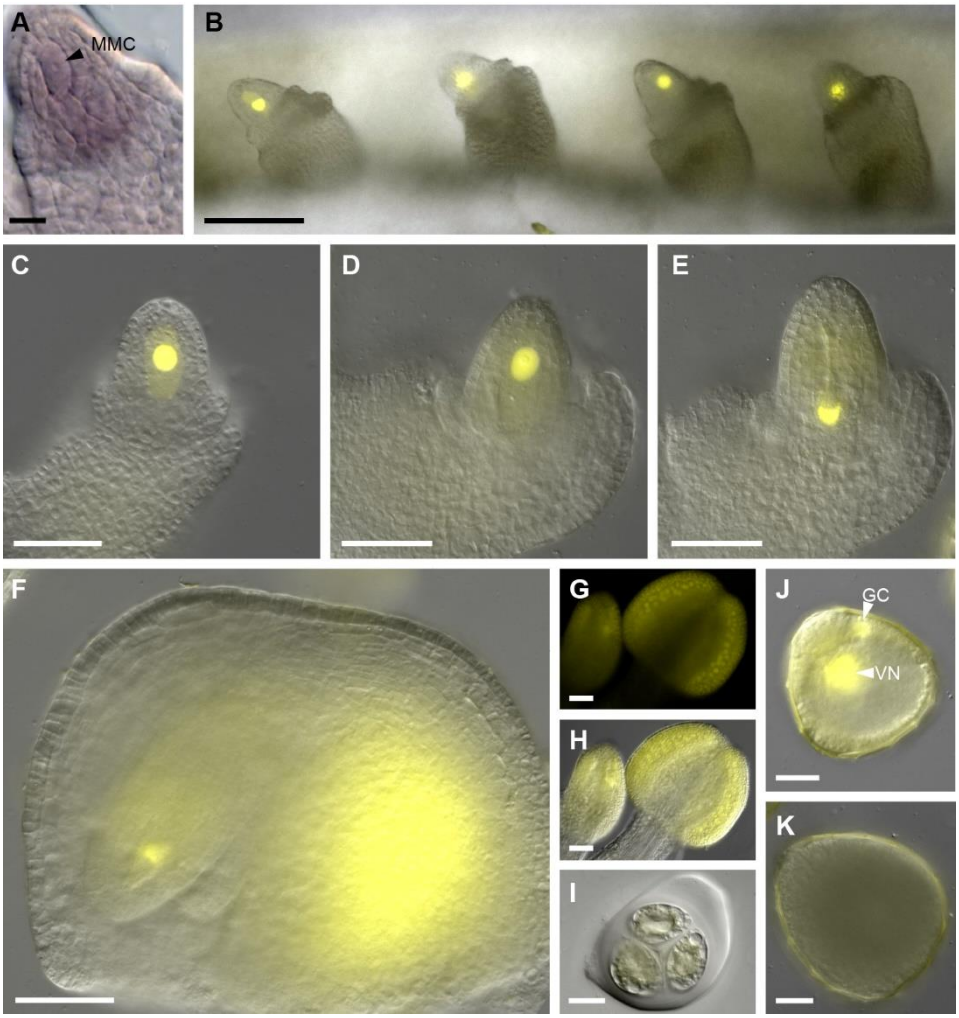
Supplemental Figure 2. Principal component analysis (PCA).
PCA plot showing relationships between all transcriptomes.
Abbreviations: MMC, megaspore mother cell; fTET, female tetrads; ES2n, mitotic embryo sac with 2 nuclei; ES4n, embryo sac with 4 nuclei; MES, mature embryo sac at anthesis, containing the egg cell (EC) and central cell (CC); Em, embryo; PMC.E, early pollen mother cell; PMC.L, late pollen mother cell; mTET, male tetrads; MIC, uninucleate microspore; MP-SC, mature pollen-sperm cell.



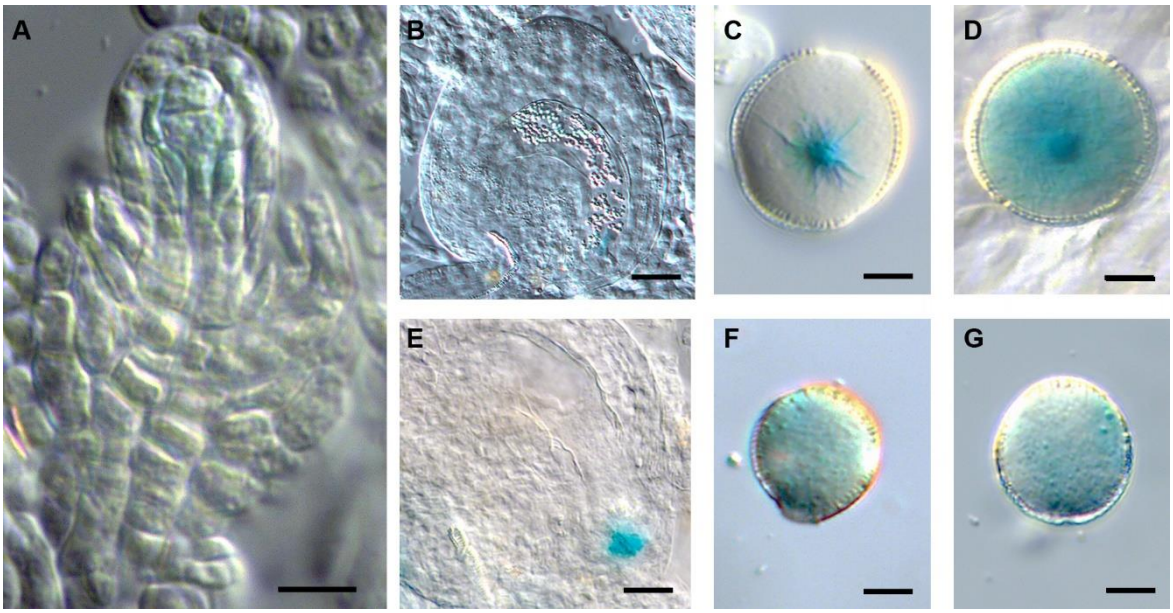
Supplemental Figure 3. Cell-specific genes expression.
Heat-map of scaled RPM expression values of the cell-specific genes identified by using the rsgcc software package.



Supplemental Figure 4. *In situ* hybridization of cell-specific cowpea genes in gametophytic tissue and fluorescent reporters expressed from *Arabidopsis* cell-specific promoters in cowpea.
A-B, *VuXTH32* mRNA localized in the sperm cells and vegetative cell cytoplasm. C, *VuXTH32* sense control. D-E, *VuAt3g09950* mRNA abundantly localized in the sperm and vegetative cells. F, sense *VuAt3g09950* control. G, Paraffin section through a fully differentiated female gametophyte showing *VuEC1.1* mRNA localization in the egg cell. H, Consecutive section to (G) showing *VuEC1.1* mRNA localization in the synergids. I, Detail of G showing mRNA localization in the egg cell. J, Detail of H showing mRNA localization in the synergids. K, Whole-mounted ovule showing *VuMEE23* mRNA localization in the central cell and adjacent nucellar cells. L-M, $AtDUO1_{pro}$*:AcGFP1* no expression. N, Whole-mounted ovule showing *VuSDR1* mRNA localization in the central cell. O, Whole-mounted ovule showing *VuGULL06* mRNA localization confined to the central cell. P, No expression of $AtLAT52_{pro}$*:AcGFP1* in germinating pollen. Q, Mature female gametophyte with egg cell, polar nuclei and central cell in cleared ovule. R, $AtDD45_{pro}$*:dsRED-Express* (red) in EC and general pattern of the $AtDD1_{pro}$*:ZsYellow1* (yellow) in cells external to the embryo sac and in CC. S, $AtEC1.1_{pro}$*:AmCyan1* (cyan) in zygote and $AtDD25_{pro}$*:dsRED-Express* (red) in early endosperm. T, $AtRKD2_{pro}$*:dsRED-Express* (red) in zygote and $AtDD9_{pro}$*:AmCyan1* (cyan) in early endosperm. U, $AtRKD2_{pro}$*:dsRED-Express* in early embryo. *AtDD9* is not expressed at this stage. V, Cleared ovule showing early embryo corresponding to the stage in U.
Scale bars: A-F, I-J = 10 µm; G-H = 18 µm; K, N-O = 35 µm; L-M, P-T = 50 µm; U-V = 100 µm
Abbreviations: EC, egg cell; CC, central cell; Sy, synergids; SC, sperm cells; VN, vegetative nucleus; PN, polar nuclei; Em, embryo; En, endosperm; Zy, zygote.

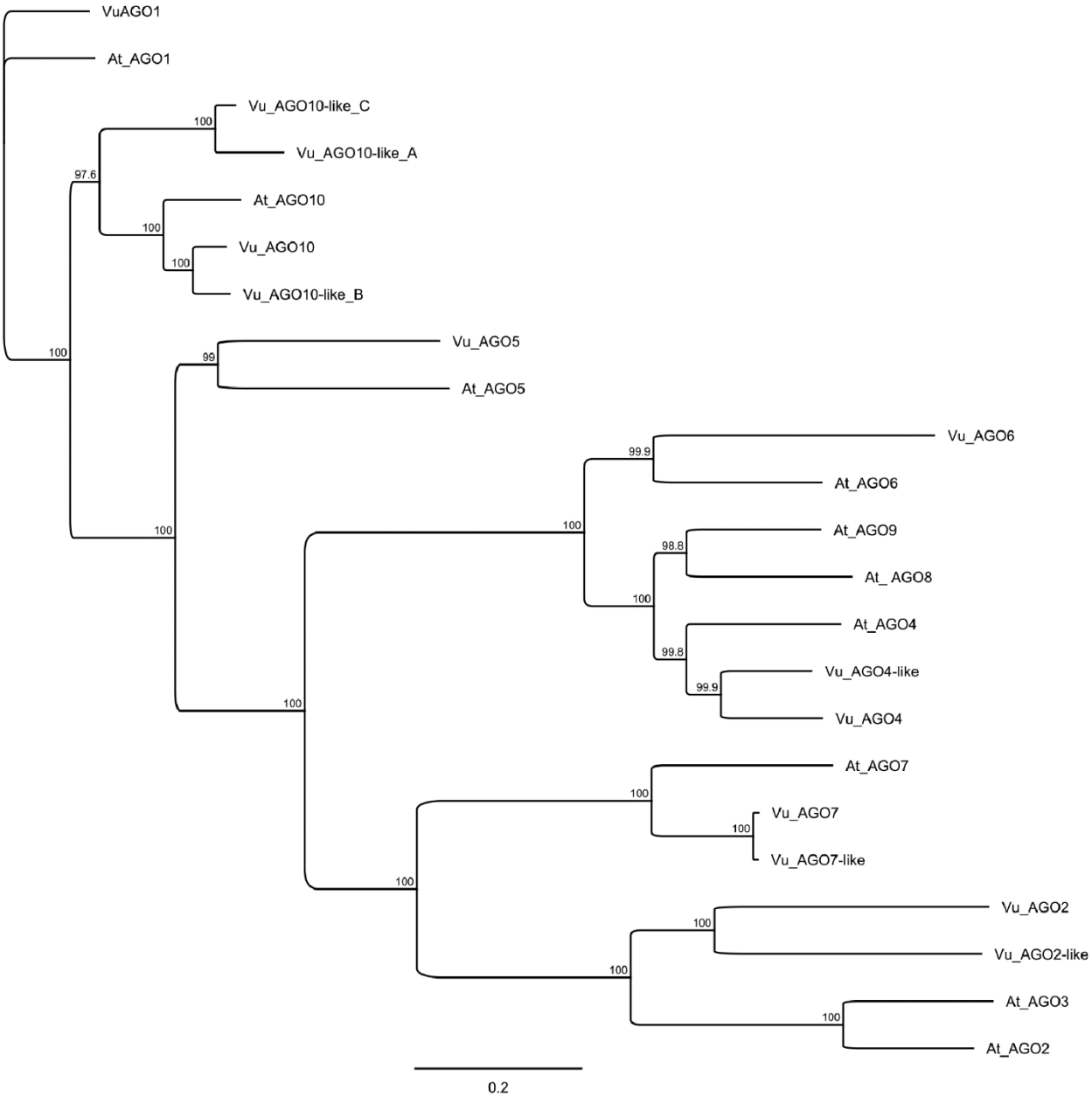


Supplemental Figure 5. Expression of the *KNU* genes in cowpea.
A, Whole-mount pre-meiotic ovule showing *VuKNU1* mRNA localized in the MMC and the developing nucellus at the onset of integument initiation. B-D, Micrographs of cowpea ovules expressing a *AtKNU$_{pro}$:nlsYFP* reporter construct marking MMC fate. E-F, At the end of meiosis *KNU* promoter turns off in female functional megaspore, but the expression persists in somatic ovule cells at the chalazal end. G-H, Expression of *AtKNU$_{pro}$:nlsYFP* reporter construct during male meiosis. I, Male tetrad showing no expression of reporter construct. J, After the first pollen mitosis, the expression of reporter construct is evident in both vegetative and generative nuclei. K, Mature pollen grain does not show any expression.
Abbreviations: MMC, megaspore mother cell; GC, generative cell; VN, vegetative nucleus. Scale bars: A = 10 µm; B = 100 µm; C-H = 50 µm; I-K = 20 µm.

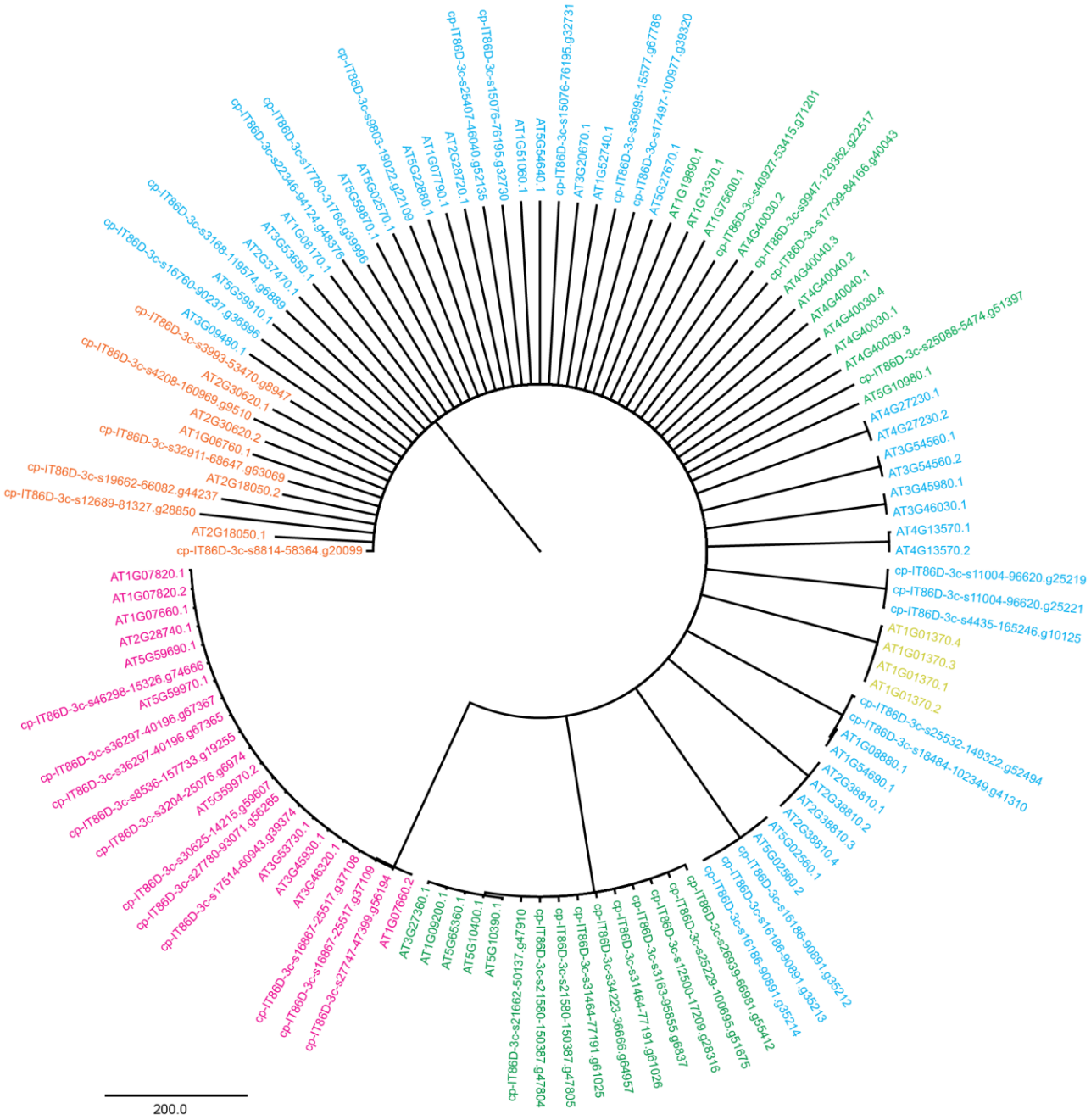


Supplemental Figure 6. Specific activity of cowpea promoters in reproductive cells of *Arabidopsis thaliana*.

A, *VuKNU-L1*$_{pro}$*:GUS* expression in the developing ovule of *Arabidopsis*. B, *VuRKD1*$_{pro}$*:GUS* expression in the egg cell. C, *VuRKD1*$_{pro}$*:GUS* expression in a sperm cell. D, *VuRKD1*$_{pro}$*:GUS* expression in the vegetative cell. E, *VuGULL06*$_{pro}$*:GUS* in the egg apparatus. F, *VuXTH32*$_{pro}$*:GUS* expression in the vegetative cell. G, *VuGIM2*$_{pro}$*:GUS* expression in the vegetative cell. Scale Bars: A, F and G = 10 µm; B and E = 20 µm; C and D = 8 µm.



Supplemental Figure 7. Phylogeny of the AGO family in cowpea and *Arabidopsis.*
A Jukes-Cantor Neighbour-Joining consensus tree showing the relationship between the AGO genes identified in cowpea (Vu) and those known in *Arabidopsis* (At).
Parameters for tree building were ClustalW MSA with BLOSUM cost matrix, Gap open cost 10, Gap extend cost 0.1. Jukes-Cantor Neighbour-Joining consensus tree, no outgroup, Bootstrap resampling with 1000 replicates, 50% support threshold.



Supplemental Figure 8. Phylogeny of *Arabidopsis* and cowpea histone proteins.
Jukes-Cantor Neighbour-Joining consensus tree showing the relationship between the *Arabidopsis* and cowpea histones is shown. Histone1 (orange), Histone2 (blue), Histone3 (green), Histone4 (pink), CenH3 (yellow). A Parameters for tree building were ClustalW MSA with BLOSUM cost matrix, Gap open cost 10, Gap extend cost 0.1. Jukes-Cantor Neighbour-Joining consensus tree, no outgroup, Bootstrap resampling with 1000 replicates, 50% support threshold.