Circadian regulation of the transcriptome in a complex polyploid crop

Hannah Rees[1], Rachel Rusholme-Pilcher[1], Paul Bailey[2], Joshua Colmer[1], Benjamen White[1], Connor Reynolds[1], Sabrina Jaye Ward[1], Calum A. Graham[3,4], Luíza Lane de Barros Dantas[3], Antony N. Dodd[3], Anthony Hall[1]

[1] Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ, UK
[2] Royal Botanic Gardens Kew, Richmond, Surrey, TW9 3AE, UK
[3] John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK
[4] School of Biological Sciences, University of Bristol, Bristol, BS8 1TQ, UK.

Abstract

The circadian clock is a finely balanced time-keeping mechanism that coordinates programmes of gene expression. It is currently unknown how the clock regulates expression of homoeologous genes in polyploids. Here, we generate a high-resolution time-course dataset to investigate the circadian balance between sets of three homoeologous genes (triads) from hexaploid bread wheat. We find a large proportion of circadian triads exhibit imbalanced rhythmic expression patterns, with no specific sub-genome favoured. In wheat, period lengths of rhythmic transcripts are found to be longer and have a higher level of variance than in other plant species. Expression of transcripts associated with circadian controlled biological processes are largely conserved between wheat and *Arabidopsis*, however striking differences are seen in agriculturally critical processes such as starch metabolism. Together, this work highlights the ongoing selection for balance versus diversification in circadian homoeologs, and identifies clock-controlled pathways that might provide important targets for future wheat breeding.

Introduction

Circadian clock homologs have been both inadvertently selected during crop domestication and identified as crop improvement targets[1–4]. Understanding circadian regulation of the transcriptome in crops such as bread wheat (*Triticum aestivum*) may provide useful insights for future crop improvement. Wheat also provides an excellent model system to explore how the circadian clock and its outputs are co-ordinated in a recently formed, complex

allopolyploid. In *Arabidopsis*, circadian transcription factors act in a dose-dependent manner, with both knock-out and over-expression mutants resulting in altered function of the circadian oscillator[5–8]. It is not yet understood how rhythmic gene expression is balanced in species with multiple copies of the same gene. *T. aestivum* is a hexaploid (AABBDD) formed through interspecific hybridisation of three diploid ancestors around 10,000 years ago[9,10]. 51.7% of high-confidence wheat genes still exist in triads; sets of three homoeologous genes present on each of the A, B and D genomes[11]. As these homoeologs evolved independently for several million years prior to hybridization, it is plausible that these independent species might have been subject to different selective pressures on their clocks (Fig. 1a).

The circadian network in *Arabidopsis* comprises a series of interlocking negative transcriptional feedback loops connected by key activators[12]. Although monocots such as wheat diverged from their dicot relatives over 140 million years ago[13], many circadian oscillator components seem to have been conserved, particularly those forming the core loop network. Orthologs of *TIMING OF CAB EXPRESSION 1 (TOC1*) and other *PSEUDO-RESPONSE REGULATOR (PRR)* genes have been identified in wheat, rice and barley, and several loci within these genes have been associated with altered flowering times, most notably *(ppd-1*) within *TaPRR3/7*[14–16]. Likewise, mutants of orthologs of *LATE ELONGATED HYPOCOTYL (LHY)*, *GIGANTEA* (*GI*), *EARLY FLOWERING 3 (ELF3),* and *LUX ARRYTHMO (LUX)* have been identified that alter heading dates, pathogen susceptibility, plant height or lower grain yields[17–21].

Circadian control of carbon fixation and starch metabolism are thought to form part of the selective advantage conferred by the clock[22,23]. This is apparent in the *lhy⁻/cca1⁻* short period double mutant, where night-time starch levels reach exhaustion earlier compared to wild-type, triggering early onset starvation responses that reduce plant productivity[23]. Similarly, genes encoding photosynthesis-related proteins are well-established targets of the circadian clock and include the *LIGHT HARVESTING CHLOROPHYLL A/B BINDING PROTEIN* genes (*LHCB* also known as *CAB* genes) and photosystem I and II reaction centre genes[24,25].

Here, we investigate circadian balance within wheat triads to understand how circadian control is co-ordinated in a polyploid crop with three subgenomes. Second, we examine similarities and differences between the circadian transcriptome in wheat and its distant dicot relative *Arabidopsis,* at a global level and at the level of genes encoding key pathways such as primary metabolism and photosynthesis.

65 Results

66 Global analysis of the circadian transcriptome in wheat

67 We generated a circadian RNA-seq time-course and compared it with a recently published
68 dataset from *Arabidopsis*[26] over 24h - 68h following transfer to constant light. Rhythmicity
69 was assessed using Metacycle Benjamini-Hochberg (BH) *q*-values. Of the 86,567 genes
70 expressed in wheat, 33.0% were rhythmically expressed with a BH $q < 0.05$ and 21.5% with
71 a BH $q < 0.01$ (Supplementary Note 1, Supplementary Table 1). This was significantly lower
72 than the proportions of rhythmically expressed genes in the *Arabidopsis* dataset (50.7% BH q
73 < 0.05, 39.1% BH $q < 0.01$) using the same criteria (X^2 (1) = 2727.1, $p < 0.001$, one-tailed,
74 two-proportions z-test). Circadian waveform characteristics of the rhythmically expressed
75 genes (BH $q < 0.01$) in the wheat and *Arabidopsis* datasets were quantified using algorithms
76 in Metacycle (JTK, ARSER, LS and meta2d) and Biodare2 (FFT-NLLS and MESA). Period,
77 phase, and amplitude estimates from FFT-NLLS and meta2d were well-correlated for
78 individual genes (Supplementary Fig. 1). All models reported that mean period length in
79 wheat was approximately 3h longer than in *Arabidopsis* (wheat = 25.9 - 27.5h,
80 *Arabidopsis* = 22.6 - 24.4h; t(36067) = 101.58, $p<0.001$, Welch's two sample t-test;
81 Supplementary Fig. 2). There was no significant difference between mean periods across the
82 three wheat sub-genomes (Fig. 1c, F(2, 28,276) = 0.179, p=0.836, One-way ANOVA).
83 We used meta2d to compare period means and distributions from four previously published
84 circadian datasets, and found that period lengths in wheat were longer and had higher
85 standard deviation than period distributions from *Arabidopsis*, *Brassica rapa*, *Brachypodium*
86 *distachyon* and *Glycine max* (Fig 1b; Supplementary Table 2).
87 We investigated how wheat periods changed over the course of the three-day experiment and
88 found that periods were longer immediately after transfer to constant light (28.61h,
89 SD=3.421h), and progressively shortened over the following days (Supplementary Note 2).
90 One explanation for this initial lengthening of period is that the dusk zeitgeber is an important
91 signaling cue for wheat circadian expression. It is possible that the mean period of expressed
92 transcripts initially lengthens trying to follow this missing dark signal before the free-running
93 endogenous period asserts itself.
94 For both *Arabidopsis* and wheat, we recalculated phases of rhythmic transcripts relative to
95 endogenous period (circadian time; CT). Across all algorithms, most transcripts in
96 *Arabidopsis* peaked during the subjective night (around CT12-24; Supplementary Fig. 3). In

97 wheat, the greatest numbers of rhythmic genes peaked during the subjective day (around
98 CT6-8) with a second, smaller group being expressed in the night (~CT20). When we
99 grouped transcripts into 2h period bins, we found that transcripts with short periods contained
100 proportionally more dawn-peaking transcripts, whereas those with longer periods contained
101 proportionally more dusk-peaking transcripts (Supplementary Note 3, Supplementary Fig. 4).

102 Balance of circadian regulation within triads

103 Although previous studies have examined the relationships between circadian regulated
104 orthologs in different plant species[26–28] and within paralogs in *Brassica rapa*[29], hexaploid
105 wheat provides an opportunity to study the relationships between recently formed circadian
106 regulated homoeologs acting within the same organism. In wheat, over 72% of syntenic triads
107 are estimated to have "balanced" expression, with similar relative abundance of transcripts
108 from each of the three homoeologs[30]. Due to the importance of the clock in coordinating
109 dosage of gene expression, our hypothesis was that many circadian triads would also have
110 balanced circadian regulation. We defined imbalanced circadian regulation as triads
111 harbouring differences in rhythmicity (i.e., BH q-values), period lengths, phases, and relative
112 amplitudes.
113 Of the 16,359 expressed triads in our dataset, 9901 (60.52%) had at least one rhythmic
114 homoeolog, and 3448 (21.08%) had three rhythmically expressed genes (BH $q < 0.05$), with
115 the latter hereafter termed "rhythmic triads" (purple segment, Fig. 1d). 6453 triads lacked
116 rhythmicity in either one or two expressed homoeolog(s) (green and blue segments, Fig. 1d).
117 In both cases, there was no bias for absence of rhythmicity in the A, B or D copy ($\chi^2(2)=$
118 6.8415, $p = 0.40$ where one gene is arrhythmic, $\chi^2(2)= 6.8415$, $p = 0.03$ where two genes are
119 arrhythmic). We found cases where high-confidence rhythmic homoeologs (BH $q < 0.01$)
120 occurred alongside arrhythmic homoeologs (BH $q > 0.05$) represented by light-shaded outer-
121 ring segments in Fig. 1d. In total there were 3450 of these imbalanced-rhythmicity triads
122 (Fig. 1h,i). To explore other forms of circadian imbalance, we assessed whether phase, period
123 and relative amplitude were conserved between homoeologs within the rhythmic triad set
124 (purple segment, Fig. 1d). Differences in phases were quantified by a cross-correlation
125 analysis to assess whether the correlation between homoeologs was improved with a time lag
126 of 4, 8 or 12 hours. We identified 464 triads with imbalanced phases with an optimum lag of
127 >0h between homoeologs (Fig. 1e,j). 1,139 triads had imbalanced periods with more than 2h
128 difference in period between homoeologs (Fig. 1f,k). 701 triads had imbalanced relative
129 amplitudes with more than two-fold difference in relative amplitude (Fig. 1g,l). Within this

last group, the homoeolog with the lowest amplitude was still rhythmic, as observed when data are mean-normalized (Fig. 1l,m). In summary, the largest cause of imbalanced circadian expression within triads was absence of rhythmicity (67.89%) with differences in period (22.41%), relative amplitude (13.79%) and phase (9.13%) occurring more infrequently and with some overlap between categories.

Out of all expressed triads in our dataset, around 11.1% had balanced circadian expression, 31.1% had imbalanced circadian expression, 39.5% were arrhythmic and 18.4% were borderline triads which did not fit into the categories imposed by our cut-offs. There is therefore a ratio of approximately 3:1 imbalanced to balanced circadian triads in wheat. This finding was initially surprising given that Ramírez-González et al. reported 72.5% of wheat triads showed balanced expression. We found that 64.15% of the triads classified as circadian imbalanced in our data would be classified as balanced in the Ramírez-González study (Supplementary Fig. 5). However, if we consider that triads with highly imbalanced circadian regulation can be classified as balanced in their expression at a single timepoint (as demonstrated in Supplementary Fig. 6), and that there are multiple ways in which homeologs can become circadian imbalanced (phase, period, rhythmicity etc.), then it is quite reasonable that only a small proportion of triads are classified as having balanced circadian regulation in this study. This insight highlights the importance of considering temporal dynamics when studying gene expression.

One explanation for imbalanced rhythmicity is that arrhythmic homoeologs are silenced. In support of this, we found that the rhythmic homoeologs in imbalanced triads were expressed at a significantly higher baseline level than their arrhythmic homoeologs (Fig. 1n; $F(16, 35{,}148) = 6.94$, $p<0.001$, Two-level, nested ANOVA on Log10 transformed data). We also found that triads with balanced rhythmicity were expressed at a uniformly higher level than the most highly expressed homoeolog(s) in the imbalanced triads (Fig. 1n; $F(7, 35{,}148) = 570.909$, $p<0.001$). Therefore, in imbalanced rhythmicity triads, the rhythmic homoeolog does not appear to compensate for reduced expression of the other homoeolog(s), so the overall expression across the triad is reduced. This is supported by data from diploid *Brassica rapa*, where circadian regulated paralogs are expressed at a higher level than single copy genes[29].

To investigate whether certain biological processes were associated with circadian balance, we compared GO-slim terms enriched in the 1816 circadian balanced triads, the 5082 differently circadian regulated triads and the 6458 arrhythmic triads, identifying significant terms unique to each group (Table 2). Some terms were enriched only in circadian balanced

164 triads (*p*-value <0.0001, Fisher's exact test e.g., "photosynthesis", "generation of precursor
165 metabolites and energy", "gene-expression" and "translation"). In contrast, GO-slim terms:
166 "developmental process involved in reproduction", "and "system development" were
167 enriched significantly in triads with differently regulated homoeologs (*p*-value <0.0001) but
168 not in triads with circadian balanced homoeologs (*p*-value >0.5). A possible explanation for
169 this enrichment could be that imbalanced circadian triads are more likely to be dynamically
170 expressed over developmental stages or show local dominance of a sub-genome in a
171 particular tissue type. Transcription factor (TF) triads were as likely to be circadian
172 balanced/imbalanced as non-transcription factors (χ^2 (2, N = 13,356) = 3.03, p = 0.08, chi-
173 square test). Previously validated wheat TFs with imbalanced circadian expression included
174 *WCBF2* (aka *TaCBF1*) and *TaPCF5*, both of which regulate abiotic stress responses[31,32]
175 (Supplementary Note 4, Supplementary Fig. 7).
176 When we compared GO-slim terms associated with paralogs in *Brassica rapa*[29], we found
177 that *Brassica* paralogs with similar circadian expression patterns (as characterized in
178 Greenham et al, (2020) using DiPALM[29]) were also associated with "photosynthesis" and
179 "generation of precursor metabolites and energy" (*p*-value <0.001, Fisher's exact test). These
180 terms were not enriched in circadian paralogs with differential expression patterns (*p*-value
181 >0.5; Supplementary Table 3). Examples of genes having similar (balanced) circadian
182 expression within *Brassica rapa* paralogs and in all three homoeologs in wheat triads
183 included orthologs of the PSI light harvesting complex genes *LHCA1*, *LHCA2* and *LHCA3*,
184 and the RNA polymerase sigma factor *SIG5.* It is possible that conservation of circadian
185 expression of these duplicate genes poses an evolutionary benefit, as similar regulation has
186 been retained for both duplicate copies in ancient paralogs of *Brassica rapa* and within more
187 recently formed wheat homoeologs arising through polyploidization.
188
189 <u>Patterns of triad circadian balance across the genome</u>
190 Next, we wanted to determine whether there are genomic regions that incorporate a clustering
191 of circadian balance in rhythmicity. We hypothesized that if we found sequential "runs" of
192 triads with arrhythmic expression on a particular chromosome, this might indicate an
193 overlying region of differential chromatin accessibility or transcriptional suppression. To
194 investigate this, we looked for regions on each set of chromosomes where there were
195 sequential triads having the same number of rhythmic homeologs (i.e. runs of three, two or
196 one rhythmic genes). We also looked for runs of sequential rhythmic triads specific to a
197 particular chromosome (e.g. runs of two rhythmic genes on chromosomes A and B). In both

198 cases, we found no evidence that triads with specific categories of rhythmic homoeologs
199 were grouped together more often than would be expected by chance (Supplementary Note 5,
200 Supplementary Table 4). This suggests that distributions of rhythmic balance appear to be
201 randomly distributed across the genome (Supplementary Fig. 8).
202

203 Clustering of gene expression and GO-term enrichment

204 To establish whether similar phased transcripts in wheat and *Arabidopsis* had similar
205 biological roles, we clustered rhythmic transcripts (BH $q < 0.01$) into 9 expression modules
206 for each species and identified GO-slim terms enriched in each module ($p < 0.01$). Circadian
207 characteristics of module eigengenes are shown in Supplementary Table 6. We compared the
208 correlation and cross-correlation of pairwise modules in the two species to find modules that
209 correlated with a peak lag of 0 (synchronous phase) or with a peak lag of 4, 8, or 12h
210 (asynchronous phase). Overall, modules with synchronous phases in wheat and *Arabidopsis*
211 shared more GO-slim terms than modules with asynchronous phases ($F(3,77) = 4.79$, p
212 <0.01, One-Way ANOVA), indicating that these modules in wheat and *Arabidopsis* contain
213 genes with similar functions (Supplementary Fig. 9). We focused on four of these
214 synchronous module pairs, broadly peaking at dawn, midday, dusk, and night for further
215 analysis (Fig. 2). Eigengenes for dawn peaking modules A9 and W9 were highly correlated
216 ($r>0.9$) and shared 14 overlapping enriched GO-slim terms ($p < 0.01$; Fig. 2a,b,f). These
217 included terms for translation and gene expression as well as terms related to protein, amide,
218 nitrogen and organonitrogen biosynthetic and metabolic processes (full lists in
219 Supplementary Table 7). Co-expressed genes in the dawn-expressed modules included
220 several orthologs involved in light, heat, and biological defence, as well as 45 ribosomal
221 protein orthologs (Supplementary Note 6). Transcripts for ribosomal proteins in mouse liver
222 and *Neurospora crassa* have also been reported to oscillate, suggesting a conserved role for
223 the circadian clock in co-ordinating ribosome biogenesis[33,34].
224 In addition to enriched GO-slim terms, we investigated enrichment for transcription factor
225 (TF) superfamilies and transcription factor binding sites in each wheat module of clock-
226 regulated genes. In late-night/dawn modules W8 and W9, transcripts encoding MYB
227 transcription factors were significantly enriched and included putative TFs involved in leaf
228 morphogenesis, plant growth, regulation of flavonoid biosynthesis, and developmental
229 transition to flowering (Supplementary Note 6, Supplementary Fig. 10).

230 Eigengenes for wheat and *Arabidopsis* modules peaking in the day (W3 and A2) had a
231 relatively low correlation (r=0.491), but peaked with similar CT values (CT 6.34h and
232 6.19h) given the longer circadian period in wheat. 5 out of 15 of the GO-slim terms enriched
233 in the W3 module were also found in the A2 module ($p < 0.01$; yellow triangles in Fig.
234 2a,c,f). These included terms relating to "photosynthesis", "response to radiation" and
235 "generation of precursor metabolites and energy". Co-expressed genes peaking in day-time
236 modules included light-harvesting and light signaling genes as well as *CYP709B3,* which
237 protects the plant from transpiration-triggered salinity stress during the day[35,36].
238 In dusk-peaking modules A5 and W4, 8 significantly enriched GO-slim terms were shared
239 between *Arabidopsis* and wheat ($p < 0.01$, green triangles in Fig. 2a,d,f). Several genes co-
240 expressed in these dusk modules were involved in auxin transport and signalling including
241 the endosomal sorting complex protein *CHMP1A* which ensures proper sorting of auxin
242 carriers (Supplementary Note 6)[37]. There was also a significant enrichment for expression of
243 transcripts encoding AP2-EREBP (ethylene responsive) and ARF (auxin responsive)
244 transcription factor superfamilies within the W4 module (Supplementary Fig. 10).
245 Interestingly, this was followed two hours later by the expression of genes with AP2-EREBP
246 transcription factor binding sites in their promoter region (W5, Supplementary Fig 11).
247 Finally, two evening-phased modules W5 and A6 (r=0.80) were enriched for GO-slim terms
248 concerning several metabolic processes (blue triangles in Fig. 2f). Co-expressed orthologs in
249 these two modules included *SEVEN IN ABSENTIA2* that regulates ABA-mediated stomatal
250 closure and drought tolerance in *Arabidopsis*[38], and *HYDROPEROXIDE LYASE1* that
251 contributes to responses to insect attack and mechanical wounding[39].

252 <u>Components of the core circadian clock in *Arabidopsis* and wheat</u>

253 We next compared the dynamics of circadian oscillator components in wheat and
254 *Arabidopsis*. Clock gene orthologs belonging to large gene families were detected by
255 phylogenetic analysis (Supplementary Fig. 12–16; Supplementary Table 8). Overall, wheat
256 circadian clock genes were expressed rhythmically and with a similar phase to their
257 *Arabidopsis* counterparts (Fig. 3). However, the free-running rhythms of clock transcripts in
258 wheat had a mean circadian period that was approximately 3.49h longer than in *Arabidopsis*
259 (27.23h and 23.74h, respectively).
260 *TaLHY* and *TaTOC1* peaked sharply at dawn and dusk, respectively, during the first cycle in
261 constant light, and maintained an >8h difference in phase throughout the experiment (Fig.
262 3a,b). This is consistent with their phasing in *Arabidopsis.* All three homoeologs for *TaGI*

263 were robustly rhythmic (BH $q < 0.01$) and peaked at CT7 (Fig. 3c). *TaPRR73* transcripts
264 peaked approximately 5h before *TaPRR37* transcripts, consistent with the phase divergence
265 of *Arabidopsis PRR7* and *PRR3* (Fig. 3d,e). However, wheat homoeologs *TaPRR59* and
266 *TaPRR95* had similar expression profiles (R^2=0.68) peaking marginally apart at CT8 and
267 CT10, in between the peak phases of *Arabidopsis PRR9* (CT5) and *PRR5* (CT11) (Fig. 3f,g).
268 Therefore, the *PRR* gene family in wheat peaks in the order of; *TaPRR73*, [*TaPRR37*,
269 *TaPRR95*, *TaPRR59]* in quick succession, and finally *TaTOC1.* This sequential pattern
270 matches the expression of PRR homologs in rice[40].
271 Transcripts encoding evening complex components, *LUX, ELF3* and *ELF4,* are circadian
272 regulated in *Arabidopsis* and peak simultaneously at dusk. Three wheat triads for LUX-like
273 genes were identified; one with higher identity to *LUX/BOA* and two similar to other *LUX*-
274 like *Arabidopsis* genes (Supplementary Fig. 15). Transcripts from all three of these triads
275 accumulated rhythmically and peaked from midday to dusk, *TaLUX-Lb* at CT7, *TaLUX/BOA*
276 at CT10 and *TaLUX-La* at CT12 (Fig. 3h,i; Supplementary Fig. 17). Five wheat transcripts
277 with homology to *Arabidopsis ELF4* and *ELF4-L1-4* accumulated with a mean circadian
278 phase of 12.6h; similar to *ELF4,* but with lower relative amplitudes (Fig. 3j,k; Supplementary
279 Fig. 17). *TaELF3* transcripts were all arrhythmic (BH $q > 0.36$), and the *TaELF3-1D*
280 homoeolog was expressed with a particularly low baseline level of 0.73 TPM in comparison
281 to the other two homoeologs (8.7-9.7 TPM) (Fig. 3l). This is consistent with previous
282 findings linking a deletion in *TaELF3* to the *eps* QTL on chromosome 1D in Cadenza[18].
283 We next assessed the balance of circadian expression between triad homoeologs in the core
284 clock network. *TaLHY, TaGI*, and *TaPRR59* had notably similar expression patterns in terms
285 of phase, period, relative amplitudes, and baseline expression over all timepoints (Fig.
286 3a,c,f,g), suggesting that unbalance in these triads is strongly selected against. Homoeologs
287 of *TaTOC1*, *TaPRR73* and *TaPRR37(Ppd)* had similar phases, but all had reduced expression
288 in the A-genome homoeolog (Fig. 3b,d,e). *TaLUX/BOA-3A*, *TaLUX-La-3B* and *TaLUX-Lb-*
289 *1D* had marginally shorter periods (>2h) and delayed phases (>2h) compared to their
290 respective homoeologs (Supplementary Table 9).
291 The REVEILLE family are CCA1/LHY-like MYB-domain transcription factors that are
292 predominantly activators of evening expressed genes[41,42]. The wheat *RVE* genes could be
293 split into a *LHY* clade (containing the *TaLHY* triad described above), a *RVE6/8*-like clade
294 containing three wheat triads and a *RVE1/2/7*-like clade also containing three triads
295 (Supplementary Fig. 12). All *TaRVE6/8* transcripts peaked at CT0-4 concurrently with
296 *TaLHY* (Supplementary Fig. 17). The *TaRVE2/7* transcripts peaked with distinct phases;

297 *TaRVE27b* in phase with *TaLHY*, *TaRVE27c* 4h before *TaLHY* and *TaRVE27a*
298 approximately 12h before *TaLHY* (Supplementary Fig. 17). Based on their phylogenetic
299 relationships, it is probable that several *RVE2/7* clade paralogs in wheat and *Arabidopsis*
300 arose independently after their evolutionary divergence, and it is therefore interesting that
301 they both show distinct phases of expression, suggesting homoplastic circadian functions.
302 Expression of orthologs for additional transcripts involved in circadian regulation (*FKF1*,
303 *ZTL*, *LKP2*, *LNK1/2*, *CHE* and *LWD*) are reviewed in Supplementary Note 7 and
304 Supplementary Fig. 17.

305 Circadian control of photosystem and light signalling gene expression is largely conserved
306 between *Arabidopsis* and wheat

307 A further GO-slim analysis across all rhythmically-expressed genes in *Arabidopsis* and wheat
308 identified enrichment of similar GO-slim processes including "photosynthesis" (p<1x10-14),
309 "rhythmic process" (p<1x10-6), "response to abiotic stimulus" (p<1x10-13) and "cellular
310 macromolecule biosynthetic process" (p<1x10-5, Fisher's exact test, Supplementary Table
311 10). We used genes associated with some of these GO-slim terms as case-studies to highlight
312 similarities and differences in circadian control between the two species. Expression data and
313 Metacycle statistics for all transcripts in this analysis are in Supplementary Table 11.
314 In considering photosynthesis, we examined specifically nuclear genome-encoded
315 photosystem (PS) proteins. Transcripts encoding the PSI components *LHCA1-6,* the PSI
316 reaction centre subunits *PSAD* and *PSAE* and the *PSII* subunits *LHCB1-7* were rhythmically
317 expressed in both species and had conserved phases (Supplementary Fig. 18). *LHCA1-4*
318 peaked towards the end of the subjective day and *LHCA5* and *6* peaked during the subjective
319 night. *PSAD* and *PSAE* peaked concurrently with *LHCA1-4*. In both species, *LHCB7*
320 transcripts had lower relative amplitudes compared to other LHCB transcripts. PSB27 is a
321 protein associated transiently with the PSII complex involved in adaption to fluctuating light
322 intensities[43]. Transcripts for this protein peaked during the subjective day in *Arabidopsis* and
323 during the subjective night in wheat.
324 In considering the GO-slim term "response to abiotic stimulus", we next investigated
325 expression of transcripts for photoreceptors and light signalling proteins due to their
326 pervasive influence upon development, metabolism, and circadian timing. Although
327 transcripts for the UV-B photoreceptor *UVR8* accumulated with a circadian rhythm in both
328 species, only one PHYTOCHROME ortholog (*PHYA)* and three *CRYPTOCHROME*

329 orthologs *(CRY1)* were rhythmic in wheat out of 18 orthologs identified (Supplementary Fig.
330 18). This contrasts with *Arabidopsis*, where *PHYA-C, CRY1* and *CRY2* accumulated with a
331 circadian rhythm.
332 Downstream light signalling proteins COP1 and SPA form complexes that degrade positive
333 regulators of photomorphogenesis (e.g. *HFR1* and *HY5*) under dark conditions[44]. Transcripts
334 for *COP1*, *SPA4*, *HFR1* and *HY5* accumulated rhythmically and with conserved phases in
335 both species (Supplementary Fig. 19). *COP1/SPA4* peaked synchronously around the end of
336 the subjective night. Surprisingly, given the similar role HFR1 and HY5 proteins have in
337 preventing hypocotyl elongation in low light, *HFR1* and *HY5* transcripts were expressed anti-
338 phase to each other. HY5 and HFR1 act synergistically to coordinate the photomorphogenesis
339 response, although it has been suggested that their activation is regulated through
340 independent pathways[45].
341 Wheat triads with identity to *Arabidopsis PIN1*, *PIN4, PIN5,* and *PIF4/5* were rhythmically
342 expressed, alongside two triads with high similarity to rice *OsPIL11* and *OsPIL13* ([46];
343 Supplementary Fig. 19). Overall, we observe that the arrhythmic accumulation of most of the
344 wheat PHY and CRY transcripts is not reflected in the rhythmic expression of several
345 downstream light signalling transcripts. This supports the notion that regulatory signals from
346 photoreceptors might occur at the level of protein stability and localisation rather than at the
347 level of transcript accumulation, as occurs for *ZTL* or *HY5* in *Arabidopsis*.
348 A set of proteins that link light signalling, circadian regulation and chloroplasts are the sigma
349 factors[47]. These light-responsive nuclear-encoded regulators of chloroplast transcription
350 guide promoter recognition and transcription initiation by plastid encoded RNA-polymerase
351 (PEP) on the chloroplast genome[48–51]. In *Arabidopsis*, *SIG1, 3, 4*, *5* and *6* were rhythmically
352 transcribed (Supplementary Fig. 19). In wheat, all homoeologs in triads orthologous to *SIG1*,
353 *SIG3* and *SIG5* were also rhythmic (BH $q < 0.01$). Whilst the dawn phase of *TaSIG5*
354 transcripts were similar to *AtSIG5*, *TaSIG1* transcripts were expressed over 10h earlier than
355 *AtSIG1* (Supplementary Fig. 19). Previous research has shown that activity of *AtSIG1* can be
356 regulated through redox-dependent phosphorylation[52], and activity of all sigma factors are
357 likely to be subject to multiple layers of regulation in addition to circadian control of
358 transcript expression.
359

360 Similarities and differences in circadian control of primary metabolism genes in *Arabidopsis*
361 and wheat

362 Expression profiles of genes with key roles in primary metabolism were compared in
363 *Arabidopsis* and wheat with a focus on enzymes that regulate trehalose 6 phosphate (Tre6P)
364 and starch metabolism (Fig. 4). Tre6P synthase (TPS) and Tre6P phosphatase (TPP)
365 participate in the synthesis and dephosphorylation of Tre6P, respectively. Tre6P is an
366 important signalling metabolite associated with both sucrose regulation and circadian
367 regulation in *Arabidopsis*[53–55]. Tre6P also affects grain yield and drought resilience in wheat,
368 maize, and rice[56]. Transcripts for *TPS1*, *2*, *6*, *8*, *9*, *10* and *11* and *TPPA*, *E*, *F*, *G* and *H* were
369 expressed rhythmically in *Arabidopsis* (Supplementary Fig. 20). Wheat transcripts for *TPS1*
370 (the most well-characterised of the T6P synthases) were arrhythmic, however rhythmic
371 transcripts were found in triads more closely related to TPS11, 6 and 7 (Supplementary Fig.
372 20). We identified three rhythmic TPP triads in wheat, two of which were orthologous to
373 *Arabidopsis TPPA*, *F* and *G.* The third TPP triad was part of a monocot-specific clade
374 identified by Paul et al. (2018), which also included *Zm00001d032298,* a crop improvement
375 target in maize[56].
376 Ribulose bisphosphate carboxylase (Rubisco) comprises eight small (RbcS) and 8 large
377 (RbcL) subunits, which are encoded by the nuclear and chloroplast genomes, respectively[57].
378 Rubisco requires activation by Rubisco activase (RCA) to release its activity from inhibitory
379 substrates[58]. In our wheat expression data, 22 putative wheat orthologs for the small subunit
380 of Rubisco were rhythmic, peaking during the subjective night, as *RBCS1A*, *RBCS1B*,
381 *RBCS2B* and *RBCS3B* do in *Arabidopsis* (Supplementary Fig. 20). Two triads with identity to
382 Rubisco activase were identified, one of which accumulated rhythmically (peaking at CT0, as
383 with *Arabidopsis RCA*).
384 Circadian regulation has a pervasive influence on starch metabolism in *Arabidopsis*,
385 particularly the nocturnal rate of transitory starch degradation[23,59]. Chloroplast phospho-
386 glucose isomerase 1 (*PGI1*) and chloroplast phosphoglucose mutase (*PGM1*) are essential
387 enzymes that link the Calvin-Benson cycle with starch biosynthetic pathway[60–62]. In
388 *Arabidopsis,* these transcripts accumulated with a circadian rhythm (BH $q < 1\text{x}10^{-4}$); *PGM1*
389 peaked just after dusk (CT14), and *PGI1* slightly later at CT20. In contrast, only one wheat
390 *TaPGI1* homoeolog was rhythmic (BH $q < 0.01$), which had a low relative amplitude (0.16)
391 and a peak phase of CT8. No homoeologs for *TaPGM1* were rhythmically expressed (BH q >
392 0.01, Supplementary Fig. 20).

393 ADP-glucose pyrophosphorylase (AGPase) mediates the first irreversible and rate-limiting
394 step in starch biosynthesis through the formation of ADP-Glc. In *Arabidopsis,* transcripts
395 encoding the small and large subunits of AGPase (*APL1*, *APL2*, *APL3*, *APS1*) were rhythmic,
396 peaking at night around CT20. In comparison, in wheat only two of the eleven transcripts
397 with homology to *APL1*, *APL2* and *APS1* were rhythmic (BH $q < 0.01$), with the remaining
398 transcripts lacking a discernible rhythm (BH $q > 0.05$) (Supplementary Fig. 20).
399 Starch synthases (SS) represent another group of metabolically important enzymes that use
400 the glucose from ADP-Glc to elongate glucan chains. In *Arabidopsis*, there are five types:
401 SSI, SSII, SSIII, SSIV and granule bound GBSSI. SSI-IV are responsible for synthesis of
402 amylopectin, with SSIII and IV determining starch granule number and morphology[63]. *GBSSI*
403 is a known dawn-expressed gene, regulated directly by CCA1/LHY, specialised for amylose
404 synthesis[64]. In wheat, *GBSSI* orthologs are called *TaWaxy* and cultivars with three null alleles
405 produce amylose-free starch in their grain[65]. Comparison of starch synthase expression in
406 *Arabidopsis* and wheat revealed several differences between the phases and relative
407 amplitudes of these transcripts (Supplementary Fig. 20). In *Arabidopsis, GBSSI* transcripts
408 had by far the greatest relative amplitude (1.26) with peak expression at dawn. The next
409 greatest amplitudes were of *SSIV* transcripts, which peaked at CT17. *SSII* and *SSIII* peaked
410 together at CT21 and *SS1* peaked at CT8 with a much smaller amplitude (0.12). In contrast,
411 in wheat, an *SSIII* triad (*TaSSIIIb*) had the largest relative amplitude rhythms of the wheat
412 starch synthases identified (0.64 - 0.73). Wheat transcripts for *SSI* and *SSIV* also peaked in
413 the morning, whereas wheat *SSII* peaked instead in the subjective night (~CT15). In our data,
414 TaWaxy (GBSSI) transcripts were present at a very low baseline level (<0.01 TPM) and
415 without any circadian oscillation. However, another wheat triad, *TaGBSSII,* shared >62%
416 identity with *Arabidopsis GBSSI,* and the B and D homoeologs had rhythmic expression
417 which peaked at dawn. TaWaxy and *TaGBSSII* are specific to endosperm and leaf tissues,
418 respectively[66], which might explain the distribution of transcript accumulation seen here. We
419 can conclude that the circadian clock regulates the expression of SS transcripts in both
420 *Arabidopsis* and wheat, although there might be an emphasis on different types of SS in each
421 species.
422 The *Arabidopsis* circadian clock regulates the rate of starch degradation so that starch
423 reserves are depleted precisely at subjective dawn[23]. Many transcripts encoding starch-
424 degrading enzymes in *Arabidopsis* had synchronized dusk peaks: Isoamylase-type starch
425 debranching enzyme *ISA3*; alpha-amylase *AMY3*; plastidial phosphorylase *PHS1-2*;
426 disproportionating enzymes *DPE1-2*; glucan, water dikinases *GWD1* and *PWD* and glucan

427 phosphatase *SEX4*. *Arabidopsis* transcripts for *BAM3*, *BAM5* and *PU1* also oscillated with a
428 circadian rhythm, peaking later in the subjective night. Strikingly, wheat orthologs for several
429 of these genes were not rhythmic, including *AMY3*, *DPE1*, *PWD, PHS1, PU1* and *BAM1*.
430 Wheat orthologs for *ISA3* and *DPE2* were expressed rhythmically, but peaked approximately
431 8-12 h ahead of their *Arabidopsis* counterparts. Some starch degradation enzymes had
432 conserved circadian expression patterns in the two species, such as *SEX4*, *GWD1, BAM3* and
433 *BAM5* transcripts. GWD catalyses glucan phosphorylation and *SEX4* encodes a
434 phosphoglucan phosphatase, both of which facilitate hydrolytic attack by β-amylases (BAM)
435 in the early steps of starch degradation[59,67].
436

437 Discussion

438 Conservation of circadian regulation between homoeologous genes

439 We identified a large proportion of imbalanced circadian triads in our dataset. It was our
440 initial expectation that there would be strict balance between the majority of circadian
441 regulated homoeologs due to the critical and finely balanced role the clock has in regulating
442 the transcriptome, and due to the reported high levels of balance reported from single-
443 timepoint transcriptomic analysis in wheat[30]. Instead, we find three times as many triads with
444 imbalanced circadian rhythms as triads with balanced circadian rhythms. This is likely to be
445 partly due to our multiple classification of circadian imbalance as any triad with different
446 rhythmicity, period, phase or relative amplitudes between homeologs. Another factor which
447 distinguishes ratios of transcriptional balance (as defined by Ramírez-González et al. (2018))
448 and circadian balance is that transcriptional balance within circadian triads is often dynamic
449 across a time-course, shifting between balanced, dominant and supressed relationships over
450 time (Supplementary Fig 6).
451 Most of these imbalanced circadian triads were imbalanced due to arrhythmicity in one or
452 two homoeologs expressed at a lower mean level than the rhythmic homeologs. The
453 reduction of expression could be due to constitutive epigenetic silencing or changes to
454 promoter regions, allowing differential binding of transcription factors[68–70]. These are likely
455 to be triads where one or two homeologs take responsibility for performing the biological
456 function of the triad, and the other homoeolog has reduced functionality. We found additional
457 circadian unbalance in the form of altered phase, period, and relative amplitudes. It is
458 possible that some of these differences are due to retention of circadian regulation from the

459 ancestral genome of each homeolog (Fig. 1a), although it is likely that other differences
460 reflect more recent diversification in expression as a step towards neo-functionalisation. It
461 has been previously suggested that functional divergence is a likely fate for duplicated genes
462 in a sufficiently large population[71]. In *B. rapa,* 42% of circadian controlled paralogs had
463 differential expression patterns[29], however these paralogs arose through whole genome
464 duplication events around 13-43 million years ago, so have been exposed to longer periods of
465 time during which selection could act upon these duplicate genes[72]. In comparison,
466 specialisation of circadian homeologs in wheat could be comparatively lower due to the
467 relative infancy of its polyploidisation around 10,000 years ago.
468

469 Differences between periods of rhythmic transcripts in *Arabidopsis* and wheat

470 The mean period of circadian regulated genes in wheat was over three hours longer than in
471 *Arabidopsis*. Period length is affected by a range of exogenous conditions (e.g. light and
472 temperature), and varies between tissues and plant age[73]. There is also evidence that longer
473 periods have been selected for during cultivation of crops at higher latitudes[1,2,74], potentially
474 due to enhanced seasonal tracking capability enabling precision timing of growth and
475 flowering[75]. Compared to other plant circadian transcriptome data sets, rhythmic wheat
476 transcripts also had higher period variance (Fig 1b). The broad period distribution in wheat
477 might be due to inclusion of all aerial material in our sampling strategy. Variation in free-
478 running periods could occur at the organ-, tissue- or cellular-level, and transcripts which are
479 highly expressed in those regions may reflect those period differences[76,77]. An alternative
480 possibility is that period variation is due to uncoupling of multiple circadian oscillators within
481 the same cell which control expression of subsets of transcripts[78–80]. Future research could
482 examine the relationship between period distributions of circadian transcriptomes and the
483 effects of domestication, latitudinal adaption, monocot-dicot divergence, or polyploidy.
484

485 Similarities and differences in circadian regulation between wheat and *Arabidopsis*
486 Our analysis revealed extensive conservation of time-of-day specific GO-slim processes and
487 co-expressed genes between *Arabidopsis* and wheat. These included genes involved in
488 photosynthesis (e.g., photosystem proteins), light signalling (e.g., *HFR1, HY5, PINs* etc*),*
489 translation (e.g., ribosome proteins) and auxin and ethylene responsive transcription factors.
490 The striking conservation of photosynthesis related genes was also reflected by the
491 enrichment of these genes in both balanced wheat triads and similarly expressed circadian

492 paralogs in *Brassica napus*. Photosynthetic outputs have also been reported to be governed
493 by the circadian clock in the liverwort species *Marchantia polymorpha*, in diazotrophic
494 cyanobacterium *Cyanothece* sp. and in alga *Aegagropila linnaei*, perhaps suggesting a
495 widespread control mechanism with an ancient evolutionary origin[81–83].
496 We also identified several interesting differences between *Arabidopsis* and wheat, including
497 absence of rhythmicity in wheat *PHY* and *CRY* transcripts and antiphase expression of the
498 wheat sigma factor *SIG1.* Furthermore, we found differences in rhythmic expression of many
499 transcripts involved in regulating Tre6P and starch metabolism.
500 In our data, putative wheat homeologs of *TPS1* were arrhythmic. Instead, rhythmic *TPS*
501 transcripts in wheat had similarity to *Arabidopsis TPS11*, *6* and *7* (Supplementary Fig. 20). In
502 *Arabidopsis*, TPS1 is the most catalytically active and best characterised TPS, and feeds back
503 into the entrainment of the circadian clock[54,84]. If the lack of rhythmicity in wheat *TPS1*
504 transcripts is reflected at the level of protein activity, it may indicate that Tre6P synthesis is
505 not regulated as tightly by the circadian clock in wheat as in *Arabidopsis*. On the other hand,
506 circadian control of other TPS triads may have implications for biotic or abiotic defence in
507 wheat. TPS5-11 have been previously implicated in control of stomatal aperture[85],
508 thermotolerance[86], and defence against fungal, bacterial and aphid attack[87,88]. In rice, *OsTPS8*
509 influences drought resistance through suberin deposition[89], and wheat *TaTPS11* participates
510 in a cold stress response[90].
511 In wheat, transcripts for starch degradation enzymes (*PHS1, DPE1, BAM1, PU1, AMY3,*
512 *PWD*) and starch biosynthesis enzymes (*PGI1*, *PGM1*, *ISA1* and *ATPase*) had either
513 arrhythmic expression or low relative amplitudes compared with the robust rhythms of many
514 of these transcripts in *Arabidopsis.* Additionally, *ISA2*, *ISA3* and the starch synthases (*SSI-IV*)
515 had differing circadian phases between the two species. While it is possible that rhythmic
516 expression of a reduced number of genes (e.g.: *SEX4*, *GWD1, BAM3* and *BAM5)* is sufficient
517 to mediate circadian control of starch degradation in wheat, these data suggest that the
518 circadian clock has a less pervasive influence upon transcriptional control of starch
519 metabolism in wheat compared to *Arabidopsis* (Supplementary Note 8).
520
521 <u>Conclusions</u>
522 Our data reveal the influence of circadian regulation on the wheat transcriptome and highlight
523 several intriguing differences between rhythmically expressed transcripts in *Arabidopsis* and
524 wheat. It explores the added complexity of co-ordinating circadian expression across multiple
525 sub-genomes in a hexaploid species. Given the circadian clock has been under selection

526 during domestication and presents multiple targets for crop improvement, it is likely that this
527 new insight into the clock in wheat will be important in the development of new sustainable
528 and resilient cultivars. It is our hope that these data provide a resource for identifying target
529 genes regulated by the circadian clock, allowing the relationships between chronotype, yield
530 and resilience to be explored in future studies.
531

532 Methods

533 Plant materials and growth conditions

534 Wheat: Wheat seeds of the spring wheat cv. Cadenza were imbibed for three days on damp
535 filter paper on a Petri dish at 4°C. Plates were moved at dawn (06.00 = ZT0), to a growth
536 cabinet set to 22°C under 12:12 light: dark cycles (approximately 200 µmol m−2 s−1). After
537 two days, only seedlings with fully emerged radicles were sown, 3 cm deep in Petersfield
538 cereal mix in 9cm pots. Plants were not vernalized. Seedlings were grown under
539 12hlight:12hdark conditions for 14 days. After 14 days, at dawn (ZT0) seedlings were
540 transferred to constant light conditions, tissue was sampled every 4h for 3 days (18 samples
541 in total). At each timepoint, we sampled the entire aerial tissue from 3 replicate plants, which
542 was frozen immediately in liquid nitrogen before storage at -80°C. Total RNA was extracted
543 using Qiagen RNeasy plant mini kits (cat. no. 74904) with on-column DNAse treatment
544 (RNAse-Free DNase Set (cat. no. 79254). RNA concentration and integrity were quantified
545 using a Nanodrop Spectrophotometer and Perkin Elmer LabChip GX Nucleic acid analyser
546 before sequencing.
547 Details of growth conditions for Arabidopsis[26], Brassica rapa[29], Brachypodium distachyon[27]
548 and Glycine max[28] datasets can be viewed in their source manuscripts. Briefly, all circadian
549 data were measured under constant light and temperature following 12h:12h light:dark
550 entrainment other than Glycine max[28] which was entrained under 16h:8h light:dark cycles.

551 Wheat mRNA sequencing, read alignment and quantification

552 Library preparation was carried out following the Illumina TruSeq protocol and reads were
553 sequenced on a NovaSeq S2 flow cell at the Earlham Institute. 150bp paired-end reads were
554 generated from each library to an average depth of 84M reads per replicate. Reads were
555 filtered for quality and any remaining adaptor sequence was trimmed with Trimmomatic[91].
556 Surviving reads were aligned to the Chinese Spring RefSeq v1.1 wheat genome[11] using

557 HISAT2[92] with default parameters. Uniquely mapping reads were then quantified using
558 StringTie[93] and TPM values were extracted for each gene per sample.

559 Processing and quantification of previously published datasets

560 Raw reads from previously published circadian datasets were downloaded for *Arabidopsis*[26],
561 *Brassica rapa*[29], and *Brachypodium distachyon*[27]. These reads were filtered for quality, and
562 any remaining adaptor sequence trimmed with Trimmomatic[91]. Surviving reads were aligned
563 using HISAT2[92] to *A. thaliana* genome (TAIR 10), *B. rapa* genome (v1.0) and the *B.*
564 *distachyon genome* (v3.0) respectively. For the *Arabidopsis* alignment, maximum intron
565 length was set to 5000nt consistent with pre-processing in[26,94]. StringTie[93] was used to
566 quantify uniquely mapping reads before TPM value extraction at gene level. For *Glycine*
567 *max*[28], FPKM normalised reads were downloaded from the *Glycine max* RNA-seq Database[95]
568 (accession GSE94228) and were converted from FPKM to TPM prior to analysis.
569

570 Homolog identification of circadian clock and circadian controlled genes

571 Wheat homologs of *Arabidopsis* core circadian clock genes were identified in the wheat
572 genome by detecting similarity to the following conserved protein family domains that are
573 present in the proteins encoded by these genes: MYB1R, a subtype of MYB domain that
574 contains a distinctive SHAQKY sequence motif (present in the CCA1, LHY and RVE[1-8])
575 or a distinctive SHLQKY sequence motif (present in LUX), PAS (present in ZTL), PRR
576 (present in TOC1 and PRR[3579]) and ELF4 (present in ELF4). A hidden Markov model
577 (HMM) for each domain was used in HMMER 3.2.1 HMMSEARCH[96] to search for
578 members of the domain family in the following proteome datasets: Araport11 (*Arabidopsis*
579 *thaliana*), RGAP7 (*Oryza sativa*), JGI Phytozome version 12 (*Brachypodium distachyon*),
580 IBSC (*Hordeum vulgare*), SpudDB PGSC v4.03 (*Solanum tuberosum*) and IWGSC Refseq
581 v1.1 (*Triticum aestivum*). The HMMs provided by Pfam (https://pfam.xfam.org/) were used
582 for the PAS domain (PAS_9, PF13426), the PRR domain (Response_reg, PF00072) and the
583 ELF4 domain (PF07011). For the MYB domain, an HMM was built for the MYB1R
584 subfamily using HMMER3 HMMBUILD[96] with an alignment of protein sequences from
585 *Arabidopsis* and rice, previously established as being members of this subfamily. The
586 sequences found from these genomes were re-aligned to the original alignment using
587 HMMER 3.2.1 HMMALIGN[96]. Amino acids with non-match states in the HMM were
588 removed from the alignment and alignment columns with <70% occupancy were also

589 removed. The longest splice variant of each protein was selected to estimate a phylogenetic
590 tree with bootstrap support using RAxML 8.2.12[97] with the following method parameters set:
591 -f a, -x 12345, -p 12345, -# 100, -m PROTCATJTT. The trees were mid-point rooted and
592 images created using the Interactive Tree of Life (iToL) tool[98]. For the larger MYB and PRR
593 families, proteins from the tree clades containing known clock gene(s) were re-aligned across
594 their full-length and a "nested" phylogenetic tree was re-estimated with RAxML as described
595 above. The tree was visualised in the Interactive Tree Of Life (iTOL) website alongside the
596 corresponding alignment. This view provided increased detail about the relationships within
597 the clade and enabled orthologous sequences to be inferred. Wheat homologues for *ELF3*,
598 *GI*, *LWD1/2*, *CHE,* and *LNK1/2* were identified by BLASTP searches using previously
599 identified wheat and *Brachypodium* predicted proteins confirmed by reciprocal BLAST
600 searches against *Arabidopsis*. IDs and source references can be viewed in Supplementary
601 Table 8.
602 Putative wheat orthologs for *Arabidopsis* circadian controlled pathway genes involved in
603 photosynthesis, light-signalling and primary metabolism were first extracted using Biomart
604 v0.7[99] available from Ensembl Plants and taken forward if they had >40% identity in the
605 DNA sequence. Orthologs were then verified using BLASTP using *Arabidopsis* protein
606 sequences as a query against the wheat protein database to confirm the wheat gene IDs.
607 Complete lists of wheat gene IDs used in the pathway analysis can be viewed in
608 Supplementary Table 11.
609

610 <u>Circadian quantification using Metacycle and Biodare2</u>

611 To estimate proportions of rhythmic genes expressed in *Arabidopsis* and wheat, we removed
612 only genes with 0 TPM at all timepoints. This approach has been used in several previous
613 studies[26,100,101] and allows detection of low-expression rhythmic transcripts. An analysis of
614 how filtering for low-expression genes affects the estimates of proportions of rhythmically
615 expressed genes is discussed in Supplementary Note 1 and Supplementary Table 1.
616 The R package MetaCycle[102] was used to identify rhythmically expressed transcripts
617 (Benjamini-Hochberg *q*-values) and to quantify period lengths (hours), absolute phase
618 (hours), baseline expression (TPM), amplitudes (TPM) and relative amplitudes of circadian
619 waveforms. Relative amplitude is the ratio between amplitude and baseline TPM if the
620 baseline is greater than 1. Metacycle integrates results from three independent algorithms
621 (ARSER, JTK_CYCLE and Lomb-Scargle) to produce summary "meta2d" statistics that

622 combine the outcome from these algorithms. Metacycle was run using the following
623 parameters; minper = 12, maxper = 35, adjustPhase = "predictedPer". Transcripts were
624 defined as rhythmic if they had *q*-values < 0.05 and high confidence rhythmic transcripts if
625 they have *q*-values < 0.01. To calculate circadian phase (CT; relative to period length=24),
626 meta2d phase estimates were multiplied by 24 and then divided by the period estimates for
627 each transcript. Circular phase means were calculated using the package 'circular'
628 implemented in R[103].
629 There are many different algorithms available for quantification of rhythmicity within time-
630 series data, some of which perform better on datasets with higher levels of noise, non-24h
631 periods or various sampling frequencies. To validate the meta2d results we also used the
632 FFT-NLLS and MESA algorithms implemented in Biodare2 to verify our observations about
633 period, phase and rhythmicity[104]. FFT-NLLS also provides relative amplitude error (RAE)
634 statistics which represent a useful metric for assessing rhythmic robustness. FFT-NLLS and
635 MESA were run using the BH *q* < 0.01 filtered transcripts categorized in Metacycle, and with
636 the following parameters: no dtr, min-max, p(12.0-35.0).
637 To enable as close a comparison with the *Arabidopsis* dataset as possible, the wheat time-
638 course was cropped to a data window of 24-68h for approximation of period, phase and
639 relative amplitude unless specified otherwise. This data-window also ensures that
640 measurements are being made under circadian conditions following transfer to constant light.
641 For the triad analysis, meta2d estimates were measured over the full time-course (0-68h) as
642 differentiation of homeolog behaviour was the main interest, including the response to
643 transfer to L:L.
644

645 <u>Clustering of rhythmic genes into expression modules</u>

646 Gene co-expression analysis was carried out using the R package WGCNA (Langfelder and
647 Horvath, 2008; R version 3.6.0.).
648 *Arabidopsis*: The 10,317 genes identified by MetaCycle as significantly rhythmic (q-value <
649 0.01) were filtered and genes with greater than 0.5 TPM average expression at more than
650 three timepoints were retained for further analysis. The average expression at each timepoint
651 for the remaining 10,129 genes was used to construct signed hybrid networks on a replicate
652 basis using the blockwiseModules() function. The soft power threshold was calculated as 18,
653 and the following parameters were used; minModuleSize = 30, corType = bicor,
654 maxPOutliers = 0.05, mergeCutHeight = 0.15. Highly connected hub genes were identified

655 for each of the 9 co-expression modules using the function chooseTopHubInEachModule()
656 and eigengenes were identified for each module using the moduleEigengenes() function.
657 Wheat: The 18,633 genes identified by MetaCycle as significantly rhythmic across 12
658 timepoints ZT24 - ZT68 (q-value < 0.01) were filtered and genes with greater than 0.5 TPM
659 average expression at more than three timepoints were retained for further analysis. The
660 average expression at each timepoint for the remaining 16,327 genes was used to construct
661 signed hybrid networks using the blockwiseModules() function. A soft power threshold of 18
662 was used, together with the following parameters; minModuleSize = 30, corType = bicor,
663 maxPOutliers = 0.05, mergeCutHeight = 0.15. Eigengenes were identified for each module
664 using the moduleEigengenes() function. Modules with closely correlated eigengenes were
665 merged using the mergeCloseModules() function, with the parameters; cutHeight = 0.25,
666 iterate = F) and new module eigengenes were calculated for the resulting 9 modules.
667

668 Cross-correlation analysis

669 A cross-correlation analysis was used to find the shift in time (lag) which produced the
670 highest (peak) correlation between two rhythms. This approach was used to identify modules
671 which peaked synchronously (had a peak lag of 0h) or asynchronously (had a peak lag of 4, 8
672 or 12h) by correlating eigengenes for each module (Supplementary Fig. 9). We also used
673 cross-correlation to identify imbalanced phases within rhythmic triads (Fig. 1E). Before
674 calculating the cross-correlation between two expression rhythms, we first scaled both
675 expression patterns using their means and standard deviations, so the output reflects a time-
676 dependent Pearson correlation coefficient ranging between -1 and 1:

677 $$Z_A = \frac{X_A - \bar{X}_A}{S_A}, Z_B = \frac{X_B - \bar{X}_B}{S_R}$$

678 Where Z_i, X_i, $\bar{X}_i$ and S_i represent the standardised expression level, tpm expression level,
679 mean expression level, and standard deviation of gene A and B respectively. Once both
680 expression patterns have been scaled, the discrete cross-correlation between the two
681 expression patterns is calculated using the np.correlate function and is divided by the number
682 of time points in the expression signal returning the Pearson correlation coefficient at
683 different lags. The index of the array with the largest Pearson correlation coefficient score
684 corresponds to the lag that maximises the phase similarity between the two temporal
685 expression patterns.

686 Mean-normalised data for oscillation plots

687 Oscillation plots in Supplementary Fig. 18-20 were mean normalised to aid visualisation of
688 period and phase differences between transcripts. Data was adjusted by dividing the TPM
689 values at each timepoint by the mean across all timepoints for each gene so that the baseline
690 expression was equal to 1.

691 Gene ontology term enrichment

692 Functional enrichment of differentially expressed genes for biological processes within each
693 module was performed using the gene ontology enrichment analysis package, topGO[105] in R
694 (version 3.6.0, with the following parameters: nodeSize = 10, algorithm = "parentchild",
695 classicFisher test $p < 0.05$). Enrichment of terms in all rhythmic genes in *Arabidopsis* and
696 wheat was compared against a background 'gene universe' of all expressed genes in each
697 dataset (26,392 genes for *Arabidopsis* and 86,567 for wheat). This gene universe was also
698 used in the GO-slim analysis for enrichment in circadian balanced versus imbalanced triads.
699 Enrichment of terms in expression modules was compared against a background of all
700 rhythmically expressed genes (BH $q < 0.01$) which clustered into modules in each dataset
701 (10,129 genes for *Arabidopsis* and 16,327 for wheat). GO-slim terms refer to ontology terms
702 for biological processes unless otherwise specified and were obtained from Ensembl Plants
703 51[11], using the BioMart tool. The bubble plot was plotted using ggplot in R adapting code
704 from De Vega et al., 2021[106].
705 Enrichment of GO-slim terms in *B. rapa* circadian paralogs with similar and differential
706 expression patterns was conducted using previously published DiPALM results for pattern
707 change in a LDHC circadian time course (Supplementary File 4, Greenham et al. 2020 [29]).
708 Paralog pairs with a pattern change *p*-value of <0.001 were termed differential patterns and
709 pairs with a pattern change *p*-value of >0.1 were considered to be similar patterns. In this
710 analysis we ignored differences in expression change for consistency with our wheat triads.
711 Data was first filtered for rhythmicity using Metacycle q-values <0.01. Only paralog pairs
712 with two significantly rhythmic paralogs were retained for the GO-slim analysis. Enrichment
713 of terms in similarly expressed circadian paralogs (1562 genes) or differentially expressed
714 paralogs (1438 genes) in B. napus was compared against a background of 4646 genes
715 expressed in the Greenham dataset in paralogs and which had GO-slim annotation available.
716
717

718 Enrichment analysis of transcription factor superfamilies in wheat co-expression modules

719 Genes annotated as members of transcription factor superfamilies[30] were identified in each
720 co-expression module and the frequency of each TF superfamily compared to the frequency
721 observed in the 16,327 genes submitted to WGCNA. TF families were classed as either
722 significantly under or overrepresented in each module using Fisher's exact test ($p <= 0.05$).
723

724 Enrichment analysis of transcription factor binding sites in wheat co-expression modules

725 1.5 kb of sequence upstream of the transcription start site (TSS) was extracted for each of the
726 16,327 genes submitted to WGCNA. FIMO, from the MEME tool suite (v 4.11.1) was used
727 to predict TFBS in these regions based on similarity with previously DAP-seq validated
728 TFBS identified in *Arabidopsis*[107]. FIMO was run as reported in Ramírez-González et al.,
729 2018 (p-value threshold of <1e-04 (default), --motifpseudo set to 1e-08 as recommended for
730 use with PWMs and a --max-stored-scores of 1,000,000). The background model was
731 generated from the 16,327 promoter sequences using MEME fasta-get-markov. As the
732 significance of multiple matches of a single TFBS family in the putative promoter region for
733 each gene is unknown, we derived a non-redundant (nr) list of matched TFBS motifs for each
734 gene within each of the nine modules and for the complete set of 16,327 genes, where
735 multiple occurrences of a TFBS superfamily in a single promoter sequence were only
736 counted once. The frequency of these nrTFBS motifs for each co-expression module was
737 compared to the frequency of nrTFBS seen across all 16,327 genes and families significantly
738 under or overrepresented in each module were identified using Fisher's exact test ($p <= 0.05$).
739

740 Loom plots

741 Genome position data for the plots are based on annotations from Chinese spring
742 (Triticum_aestivum.IWGSC.52.gtf) downloaded from Ensembl Plants. Code for creating
743 Loom plots (Supplementary Fig 8) is implemented in R and a code package with
744 accompanying R markdown notebooks is available from our groups GitHub repository
745 (https://github.com/AHallLab/triad.expression).

746 Statistical analysis

747 Statistical tests including Welch's two sample t-test, Two-proportions z-test, One-way
748 ANOVA, Two-level, nested ANOVA and Chi-square tests of independence were all
749 conducted in the R 'stats' package (version 4.0.0) with default parameters.
750

751 Data availability

752 Fastq data from the RNA-seq circadian time course are available to view from the Grassroots
753 Data Repository: https://opendata.earlham.ac.uk/opendata/data/wheat_circadian_Rees_2021.
754 *[Data will be uploaded to the European Nucleotide Archive (ENA) during the review*
755 *process]* A summary csv table with expression of wheat genes (TPM), Metacycle estimates,
756 gene annotations and triad balance classification can be viewed in Supplementary Table 12
757 also available here:
758 https://opendata.earlham.ac.uk/opendata/data/wheat_circadian_Rees_2021.
759

760 References

761 1. Greenham, K. *et al.* Geographic variation of plant circadian clock function in natural and agricultural
762 settings. *J. Biol. Rhythms* **32**, 26–34 (2017).
763 2. Müller, N. A. *et al.* Domestication selected for deceleration of the circadian clock in cultivated tomato.
764 *Nat. Genet.* **48**, 89–93 (2016).
765 3. McClung, C. R. Circadian Clock Components Offer Targets for Crop Domestication and Improvement.
766 *Genes (Basel).* **12**, 374 (2021).
767 4. Steed, G., Ramirez, D. C., Hannah, M. A. & Webb, A. A. R. Chronoculture, harnessing the circadian
768 clock to improve crop yield and sustainability. *Science (80-.).* **372**, (2021).
769 5. Hazen, S. P. *et al.* LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms.
770 *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10387–10392 (2005).
771 6. Nakamichi, N., Kita, M., Ito, S., Yamashino, T. & Mizuno, T. PSEUDO-RESPONSE REGULATORS,
772 PRR9, PRR7 and PRR5, Together play essential roles close to the circadian clock of Arabidopsis
773 thaliana. *Plant Cell Physiol.* **46**, 686–698 (2005).
774 7. Somers, D. E., Webb, A. A. R., Pearson, M. & Kay, S. A. The short-period mutant, toc1-1, alters
775 circadian clock regulation of multiple outputs throughout development in Arabidopsis thaliana.
776 *Development* **125**, 485–494 (1998).
777 8. Wang, Z. Y. & Tobin, E. M. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1
778 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* **93**, 1207–1217 (1998).
779 9. Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R. & Martin, W. Genetics and geography of wild
780 cereal domestication in the near east. *Nature Reviews Genetics* vol. 3 429–441 (2002).
781 10. Araus, J. L., Slafer, G. A., Romagosa, I. & Molist, M. FOCUS: Estimated wheat yields during the

782 emergence of agriculture based on the carbon isotope discrimination of grains: Evidence from a 10th
783 millenuim BP site on the euphrates. *J. Archaeol. Sci.* **28**, 341–350 (2001).
784 11. IWGSC *et al.* Shifting the limits in wheat research and breeding using a fully annotated reference
785 genome. *Science (80-.).* **361**, (2018).
786 12. Hernando, C. E., Romanowski, A. & Yanovsky, M. J. Transcriptional and post-transcriptional control of
787 the plant circadian gene regulatory network. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1860**, 84–94
788 (2017).
789 13. Chaw, S. M., Chang, C. C., Chen, H. L. & Li, W. H. Dating the monocot-dicot divergence and the
790 origin of core eudicots using whole chloroplast genomes. *J. Mol. Evol.* **58**, 424–441 (2004).
791 14. Turner, A., Beales, J., Faure, S., Dunford, R. P. & Laurie, D. A. The pseudo-response regulator Ppd-H1
792 provides adaptation to photoperiod in barley. *Science (80-.).* **310**, 1031–1034 (2005).
793 15. Shaw, L. M., Turner, A. S. & Laurie, D. A. The impact of photoperiod insensitive Ppd-1a mutations on
794 the photoperiod pathway across the three genomes of hexaploid wheat (Triticum aestivum). *Plant J.* **71**,
795 71–84 (2012).
796 16. Laurie, D. A. Comparative genetics of flowering time. *Plant Mol. Biol.* **35**, 167–177 (1997).
797 17. Xiang, Y. Z., Mao, S. L., Jia, R. L., Chun, M. G. & Xian, S. Z. The wheat TaGI1, involved in
798 photoperiodic flowering, encodes an Arabidopsis GI ortholog. *Plant Mol. Biol.* **58**, 53–64 (2005).
799 18. Zikhali, M., Wingen, L. U. & Griffiths, S. Delimitation of the Earliness per se D1 (Eps-D1) flowering
800 gene to a subtelomeric chromosomal deletion in bread wheat (Triticum aestivum). *J. Exp. Bot.* **67**, 287–
801 299 (2016).
802 19. Wang, J. *et al.* TaELF3-1DL, a homolog of ELF3, is associated with heading date in bread wheat. *Mol.*
803 *Breed.* **36**, 1–9 (2016).
804 20. Mizuno, N., Nitta, M., Sato, K. & Nasuda, S. A wheat homologue of PHYTOCLOCK 1 is a candidate
805 gene conferring the early heading phenotype to einkorn wheat. *Genes Genet. Syst.* **87**, 357–367 (2012).
806 21. Zhang, W. *et al.* Functional studies of heading date-related gene TaPRR73, a paralog of ppd1 in
807 common wheat. *Front. Plant Sci.* **7**, 772 (2016).
808 22. Dodd, A. N. *et al.* Plant circadian clocks increase photosynthesis, growth, survival, and competitive
809 advantage. *Science (80-.).* **309**, 630–633 (2005).
810 23. Graf, A., Schlereth, A., Stitt, M. & Smith, A. M. Circadian control of carbohydrate availability for
811 growth in Arabidopsis plants at night. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9458–9463 (2010).
812 24. Harmer, S. L. *et al.* Orchestrated transcription of key pathways in Arabidopsis by the circadian clock.
813 *Science (80-.).* **290**, 2110–3 (2000).
814 25. Millar, A. J. & Kay, S. A. Circadian Control of cab Gene Transcription and mRNA Accumulation in
815 Arabidopsis. *Plant Cell* **3**, 541–550 (1991).
816 26. Romanowski, A., Schlaen, R. G., Perez-Santangelo, S., Mancini, E. & Yanovsky, M. J. Global
817 transcriptome analysis reveals circadian control of splicing events in Arabidopsis thaliana. *Plant J.* **103**,
818 889–902 (2020).
819 27. MacKinnon, K. J. M. *et al.* Changes in ambient temperature are the prevailing cue in determining
820 Brachypodium distachyon diurnal gene regulation. *New Phytol.* **227**, 1709–1724 (2020).
821 28. Li, M. *et al.* Comprehensive mapping of abiotic stress inputs into the soybean circadian clock. *Proc.*

822 *Natl. Acad. Sci.* **116**, 23840–23849 (2019).
823 29. Greenham, K. *et al.* Expansion of the circadian transcriptome in Brassica rapa and genome-wide
824 diversification of paralog expression patterns. *Elife* **9**, 1–68 (2020).
825 30. Ramírez-González, R. H. *et al.* The transcriptional landscape of polyploid wheat. *Science (80-.).* **361**,
826 (2018).
827 31. Zhao, J. *et al.* Genome-Wide Identification and Expression Profiling of the TCP Family Genes in Spike
828 and Grain Development of Wheat (Triticum aestivum L.). *Front. Plant Sci.* **9**, (2018).
829 32. Takumi, S., Shimamura, C. & Kobayashi, F. Increased freezing tolerance through up-regulation of
830 downstream genes via the wheat CBF gene in transgenic tobacco. *Plant Physiol. Biochem.* **46**, 205–211
831 (2008).
832 33. Jouffe, C. *et al.* The Circadian Clock Coordinates Ribosome Biogenesis. *PLoS Biol.* **11**, 1001455
833 (2013).
834 34. Correa, A. *et al.* Multiple oscillators regulate circadian gene expression in Neurospora. *Proc. Natl.*
835 *Acad. Sci.* **100**, 13597–13602 (2003).
836 35. Mao, G., Seebeck, T., Schrenker, D. & Yu, O. CYP709B3, a cytochrome P450 monooxygenase gene
837 involved in salt tolerance in Arabidopsis thaliana. *BMC Plant Biol.* **13**, 1–13 (2013).
838 36. Pan, Y. *et al.* Cytochrome P450 monooxygenases as reporters for circadian-regulated pathways. *Plant*
839 *Physiol.* **150**, 858–878 (2009).
840 37. Spitzer, C. *et al.* The ESCRT-Related CHMP1A and B Proteins Mediate Multivesicular Body Sorting of
841 Auxin Carriers in Arabidopsis and Are Required for Plant Development. *Plant Cell* **21**, 749–766 (2009).
842 38. Bao, Y. *et al.* The tumor necrosis factor receptor-associated factor (TRAF)-like family protein SEVEN
843 IN ABSENTIA 2 (SINA2) promotes drought tolerance in an ABA-dependent manner in Arabidopsis.
844 *New Phytol.* **202**, 174–187 (2014).
845 39. Noordermeer, M. A., Veldink, G. A. & Vliegenthart, J. F. G. Fatty acid hydroperoxide lyase: A plant
846 cytochrome P450 enzyme involved in wound healing and pest resistance. *ChemBioChem* **2**, 494–504
847 (2001).
848 40. Murakami, M., Ashikari, M., Miura, K., Yamashino, T. & Mizuno, T. The Evolutionarily Conserved
849 OsPRR Quintet: Rice Pseudo-Response Regulators Implicated in Circadian Rhythm. *Plant Cell Physiol.*
850 **44**, 1229–1236 (2003).
851 41. Shalit-Kaneh, A., Kumimoto, R. W., Filkov, V. & Harmer, S. L. Multiple feedback loops of the
852 Arabidopsis circadian clock provide rhythmic robustness across environmental conditions. *Proc. Natl.*
853 *Acad. Sci. U. S. A.* **115**, 7147–7152 (2018).
854 42. Hsu, P. Y., Devisetty, U. K. & Harmer, S. L. Accurate timekeeping is controlled by a cycling activator
855 in Arabidopsis. *Elife* **2013**, 473 (2013).
856 43. Hou, X., Fu, A., Garcia, V. J., Buchanan, B. B. & Luan, S. PSB27: A thylakoid protein enabling
857 Arabidopsis to adapt to changing light intensity. *Proc. Natl. Acad. Sci.* **112**, 1613–1618 (2015).
858 44. Ponnu, J. & Hoecker, U. Illuminating the COP1/SPA Ubiquitin Ligase: Fresh Insights Into Its Structure
859 and Functions During Plant Photomorphogenesis. *Front. Plant Sci.* **0**, 486 (2021).
860 45. Jang, I.-C., Henriques, R. & Chua, N.-H. Three Transcription Factors, HFR1, LAF1 and HY5, Regulate
861 Largely Independent Signaling Pathways Downstream of Phytochrome A. *Plant Cell Physiol.* **54**, 907–

862 916 (2013).
863 46. Nakamura, Y., Kato, T., Yamashino, T., Murakami, M. & Mizuno, T. Characterization of a set of
864 phytochrome-interacting factor-like bHLH proteins in Oryza sativa. *Biosci. Biotechnol. Biochem.* **71**,
865 1183–1191 (2007).
866 47. Noordally, Z. B. *et al.* Circadian Control of Chloroplast Transcription by a Nuclear-Encoded Timing
867 Signal. *Science (80-.).* **339**, 1316–1319 (2013).
868 48. Tanaka, K. *et al.* Characterization of three cDNA species encoding plastid RNA polymerase sigma
869 factors in Arabidopsis thaliana: Evidence for the sigma factor heterogeneity in higher plant plastids.
870 *FEBS Lett.* **413**, 309–313 (1997).
871 49. Kanamaru, K. *et al.* Plastidic RNA polymerase σ factors in Arabidopsis. *Plant Cell Physiol.* **40**, 832–
872 842 (1999).
873 50. Allison, L. A. The role of sigma factors in plastid transcription. *Biochimie* vol. 82 537–548 (2000).
874 51. Kanamaru, K. & Tanaka, K. Roles of chloroplast RNA polymerase sigma factors in chloroplast
875 development and stress response in higher plants. *Biosci. Biotechnol. Biochem.* **68**, 2215–2223 (2004).
876 52. Shimizu, M. *et al.* Sigma factor phosphorylation in the photosynthetic control of photosystem
877 stoichiometry. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10760–10764 (2010).
878 53. Figueroa, C. M. & Lunn, J. E. A tale of two sugars: Trehalose 6-phosphate and sucrose. *Plant*
879 *Physiology* vol. 172 7–27 (2016).
880 54. Frank, A. *et al.* Circadian Entrainment in Arabidopsis by the Sugar-Responsive Transcription Factor
881 bZIP63. *Curr. Biol.* **28**, 2597-2606.e6 (2018).
882 55. Fichtner, F. & Lunn, J. E. The Role of Trehalose 6-Phosphate (Tre6P) in Plant Metabolism and
883 Development. *Annu. Rev. Plant Biol.* **72**, 737–760 (2021).
884 56. Paul, M. J., Gonzalez-Uriarte, A., Griffiths, C. A. & Hassani-Pak, K. The Role of Trehalose 6-
885 Phosphate in Crop Yield and Resilience. *Plant Physiol.* **177**, 12–23 (2018).
886 57. Gruber, A. V. & Feiz, L. Rubisco assembly in the chloroplast. *Frontiers in Molecular Biosciences* vol. 5
887 24 (2018).
888 58. Bhat, J. Y., Thieulin-Pardo, G., Hartl, F. U. & Hayer-Hartl, M. Rubisco activases: AAA+ chaperones
889 adapted to enzyme repair. *Frontiers in Molecular Biosciences* vol. 4 20 (2017).
890 59. Smith, S. M. *et al.* Diurnal changes in the transcriptome encoding enzymes of starch metabolism
891 provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in
892 arabidopsis leaves. *Plant Physiol.* **136**, 2687–2699 (2004).
893 60. Periappuram, C. *et al.* The Plastidic Phosphoglucomutase from Arabidopsis. A Reversible Enzyme
894 Reaction with an Important Role in Metabolic Control. *Plant Physiol.* **122**, 1193 (2000).
895 61. Bahaji, A. *et al.* Plastidic phosphoglucose isomerase is an important determinant of starch accumulation
896 in mesophyll cells, growth, photosynthetic capacity, and biosynthesis of plastidic cytokinins in
897 Arabidopsis. *PLoS One* **10**, e0119641 (2015).
898 62. Lee, S. K. *et al.* Plastidic phosphoglucomutase and ADP-glucose pyrophosphorylase mutants impair
899 starch synthesis in rice pollen grains and cause male sterility. *J. Exp. Bot.* **67**, 5557–5569 (2016).
900 63. Nielsen, M. M. *et al.* Crystal structures of the catalytic domain of Arabidopsis thaliana starch synthase
901 IV, of granule bound starch synthase from CLg1 and of granule bound starch synthase I of Cyanophora

902 paradoxa illustrate substrate recognition in starch synthases. *Front. Plant Sci.* **9**, 1138 (2018).
903 64. Tenorio, G., Orea, A., Romero, J. M. & Mérida, Á. Oscillation of mRNA level and activity of granule-
904 bound starch synthase I in Arabidopsis leaves during the day/night cycle. *Plant Mol. Biol.* **51**, 949–958
905 (2003).
906 65. Yamamori, M., Nakamura, T., Endo, T. R. & Nagamine, T. Waxy protein deficiency and chromosomal
907 location of coding genes in common wheat. *Theor. Appl. Genet.* **89**, 179–184 (1994).
908 66. Vrinten, P. L. & Nakamura, T. Wheat granule-bound starch synthase I and II are encoded by separate
909 genes that are expressed in different tissues. *Plant Physiol.* **122**, 255–263 (2000).
910 67. Kaplan, F., Sung, D. Y. & Guy, C. L. Roles of β-amylase and starch breakdown during temperatures
911 stress. *Physiol. Plant.* **126**, 120–128 (2006).
912 68. Otto, S. P. The Evolutionary Consequences of Polyploidy. *Cell* **131**, 452–462 (2007).
913 69. Gardiner, L. J. *et al.* A genome-wide survey of DNA methylation in hexaploid wheat. *Genome Biol.* **16**,
914 1–15 (2015).
915 70. Chen, Z. J. & Mas, P. Interactive roles of chromatin regulation and circadian clock function in plants.
916 *Genome Biol.* **20**, 1–12 (2019).
917 71. Walsh, J. B. How often do duplicated genes evolve new functions? *Genetics* **139**, 421–428 (1995).
918 72. Lou, P. *et al.* Preferential retention of circadian clock genes during diploidization following whole
919 genome triplication in Brassica rapa. *Plant Cell* **24**, 2415–2426 (2012).
920 73. Rees, H. *et al.* A high-throughput delayed fluorescence method reveals underlying differences in the
921 control of circadian rhythms in Triticum aestivum and Brassica napus. *Plant Methods* **15**, 51 (2019).
922 74. Yarkhunova, Y. *et al.* Selection during crop diversification involves correlated evolution of the
923 circadian clock and ecophysiological traits in Brassica rapa. *New Phytol.* **210**, 133–144 (2016).
924 75. Michael, T. P. *et al.* Enhanced Fitness Conferred by Naturally Occurring Variation in the Circadian
925 Clock. *Science (80-.).* **302**, 1049–1053 (2003).
926 76. Gould, P. D. *et al.* Coordination of robust single cell rhythms in the Arabidopsis circadian clock via
927 spatial waves of gene expression. *Elife* **7**, (2018).
928 77. Takahashi, N., Hirata, Y., Aihara, K. & Mas, P. A Hierarchical Multi-oscillator Network Orchestrates
929 the Arabidopsis Circadian System. *Cell* **163**, 148–159 (2015).
930 78. Locke, J. C. W. *et al.* Experimental validation of a predicted feedback loop in the multi-oscillator clock
931 of Arabidopsis thaliana. *Mol. Syst. Biol.* **2**, 59 (2006).
932 79. Hall, A., Kozma-Bognár, L., Bastow, R. M., Nagy, F. & Millar, A. J. Distinct regulation of CAB and
933 PHYB gene expression by similar circadian clocks. *Plant J.* **32**, 529–537 (2002).
934 80. Hotta, C. T. *et al.* Are there multiple circadian clocks in plants? *Plant Signal. Behav.* **3**, 342 (2008).
935 81. Cuitun-Coronado, D., Rees, H., Hall, A., de Barros Dantas, L. L. & Dodd, A. N. Circadian and diel
936 regulation of photosynthesis in the bryophyte Marchantia polymorpha. *bioRxiv* (2022)
937 doi:10.1101/2022.01.11.475783.
938 82. Cano-Ramirez, D., Fraine, T., Griffiths, O. & Dodd, A. Photosynthesis and circadian rhythms regulate
939 the buoyancy of marimo lake balls. *Curr. Biol.* **28**, R869–R870 (2018).
940 83. Schneegurt, M. A., Tucker, D. L., Ondr, J. K., Sherman, D. M. & Sherman, L. A. Metabolic rhythms of
941 a diazotrophic cyanobacterium, cyanothece sp. strain atcc 51142, heterotrophically grown in continuous

942 dark. *J. Phycol.* **36**, 107–117 (2000).
943 84. Delorge, I., Figueroa, C. M., Feil, R., Lunn, J. E. & Van Dijck, P. Trehalose-6-phosphate synthase 1 is
944 not the only active TPS in Arabidopsis thaliana. *Biochem. J.* **466**, 283–290 (2015).
945 85. Tian, L. *et al.* The trehalose-6-phosphate synthase TPS5 negatively regulates ABA signaling in
946 Arabidopsis thaliana. *Plant Cell Rep.* **38**, 869–882 (2019).
947 86. Suzuki, N., Bajad, S., Shuman, J., Shulaev, V. & Mittler, R. The Transcriptional Co-activator MBF1c Is
948 a Key Regulator of Thermotolerance in Arabidopsis thaliana. *J. Biol. Chem.* **283**, 9269–9275 (2008).
949 87. Wang, X., Du, Y. & Yu, D. Trehalose phosphate synthase 5-dependent trehalose metabolism modulates
950 basal defense responses in Arabidopsis thaliana. *J. Integr. Plant Biol.* **61**, 509–527 (2019).
951 88. V, S. *et al.* TREHALOSE PHOSPHATE SYNTHASE11-dependent trehalose metabolism promotes
952 Arabidopsis thaliana defense against the phloem-feeding insect Myzus persicae. *Plant J.* **67**, 94–104
953 (2011).
954 89. Vishal, B., Krishnamurthy, P., Ramamoorthy, R. & Kumar, P. P. OsTPS8 controls yield-related traits
955 and confers salt stress tolerance in rice by enhancing suberin deposition. *New Phytol.* **221**, 1369–1386
956 (2019).
957 90. Liu, X. *et al.* Overexpression of the wheat trehalose 6-phosphate synthase 11 gene enhances cold
958 tolerance in Arabidopsis thaliana. *Gene* **710**, 210–217 (2019).
959 91. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data.
960 *Bioinformatics* **30**, (2014).
961 92. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and
962 genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, (2019).
963 93. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of
964 RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, (2016).
965 94. Gardiner, L. J. *et al.* Interpreting machine learning models to investigate circadian regulation and
966 facilitate exploration of clock function. *Proc. Natl. Acad. Sci. U. S. A.* **118**, 2021 (2021).
967 95. Zhang, H. *et al.* A Comprehensive Online Database for Exploring ∼20,000 Public
968 Arabidopsis RNA-Seq Libraries. *Mol. Plant* **13**, 1231–1233 (2020).
969 96. Eddy, S. R. Accelerated Profile HMM Searches. *PLOS Comput. Biol.* **7**, e1002195 (2011).
970 97. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
971 phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
972 98. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of
973 phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
974 99. Smedley, D. *et al.* BioMart – biological queries made easy. *BMC Genomics 2009 101* **10**, 1–12 (2009).
975 100. Kim, J. A. *et al.* Transcriptome analysis of diurnal gene expression in Chinese cabbage. *Genes (Basel).*
976 **10**, (2019).
977 101. Yang, Y., Li, Y., Sancar, A. & Oztas, O. The circadian clock shapes the Arabidopsis transcriptome by
978 regulating alternative splicing and alternative polyadenylation. *J. Biol. Chem.* **295**, 7608–7619 (2020).
979 102. Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K. & Hogenesch, J. B. MetaCycle: an integrated R
980 package to evaluate periodicity in large scale data. *Bioinformatics* **32**, 3351–3353 (2016).
981 103. Agostinelli, C. & Lund, U. R package 'circular': Circular Statistics. (2017).

104. Zielinski, T., Moore, A. M., Troup, E., Halliday, K. J. & Millar, A. J. Strengths and Limitations of Period Estimation Methods for Circadian Data. *PLoS One* **9**, e96462 (2014).

105. Alexa, A. & Rahnenfuhrer, J. topGO: Enrichment Analysis for Gene Ontology. R package version 2.44.0. *Alexa* (2021).

106. De Vega, J. J. *et al.* Differential expression of starch and sucrose metabolic genes linked to varying biomass yield in Miscanthus hybrids. *Biotechnol. Biofuels* **14**, 98 (2021).

107. O'Malley, R. C. *et al.* Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**, (2016).

108. Marcussen, T. *et al.* Ancient hybridizations among the ancestral genomes of bread wheat. *Science (80-.).* **345**, 6194 (2014).

109. Zeeman, S. C., Smith, S. M. & Smith, A. M. The diurnal metabolism of leaf starch. *Biochem. J.* **401**, 13–28 (2007).

110. López-González, C., Juárez-Colunga, S., Morales-Elías, N. C. & Tiessen, A. Exploring regulatory networks in plants: Transcription factors of starch metabolism. *PeerJ* **2019**, e6841 (2019).

111. Skryhan, K., Gurrieri, L., Sparla, F., Trost, P. & Blennow, A. Redox regulation of starch metabolism. *Front. Plant Sci.* **9**, 1344 (2018).

Acknowledgements

H.R., R.R.P. and A.H. were funded by the BBSRC Core Strategic Programme Grant (Genomes to Food Security) BB/CSP1720/1 and its constituent work package, BBS/E/T/000PR9819 (WP2 Regulatory interactions and Complex Phenotypes). B.W., R.R.P. and A.H. was supported by the BBSRC Designing Future Wheat grant BB/P016855/1; BBS/E/T/000PR9783 (DFW WP4 Data Access and Analysis). C.R. by a BBSRC grant BB/V509267/1 and Wave 1 of The UKRI Strategic Priorities Fund under the EPSRC Grant EP/T001569/1, particularly the "AI for Science" theme within that grant & The Alan Turing Institute. J.C. by the BBSRC funded Norwich Research Park Biosciences Doctoral Training Partnership grant BB/M011216/1. We would also like to acknowledge BBS/E/T/000PR9816 (NC1 - Supporting EI's ISPs and the UK Community with Genomics and Single Cell Analysis) for data generation and BB/CCG1720/1 for the physical HPC infrastructure and data centre delivered via the NBI Computing infrastructure for Science (CiS) group. P.B. was supported by a BBSRC TRDF grant BB/N023145/1. A.N.D., L.L.B.D. and C.A.G. are funded by BBSRC ISP Genes in the Environment (BB/P013511/1). C.A.G. was also funded by UK BBSRC SWBIO DTP (BB/M009122/1). We thank Dr. Susan Duncan for her contribution to preliminary experiments prior to the start of this project.

Author contributions

H.R. and A.H. were involved with project concept and designed experiments. H.R. performed time-course experiments. R.R.P. processed data, quantified read counts and conducted clustering of rhythmic transcripts. P.B. conducted the phylogenetic analysis of core circadian protein families. J.C. conducted the cross-correlation analysis. C.R. processed previously published circadian datasets. S.J.W produced loom plots for circadian triad balance. H.R., L.L.B.D., C.A.G., B.W., R.R.P. A.H., and A.N.D. analysed and interpreted the RNA-seq data. H.R. wrote the initial manuscript and all authors contributed to subsequent drafts.

Corresponding author

Correspondence to Anthony Hall.

Ethics declarations

Competing interests: The authors declare no competing interests.

Main Figures and Tables

[Figures are provided as separate files]

1037 **Fig. 1. Circadian regulation of homoeolog expression of wheat triads.**

1038 **a,** Schematic of the origins of hexaploid wheat, showing circadian clocks evolving
1039 independently in the ancestors of the A, B and D subgenomes following divergence from a
1040 common ancestor approximately 6.5 million years ago**.** Colours of clock icons represent
1041 theoretical differences in clock regulation integrated in the tetraploid and hexaploid hybrids
1042 either through circadian balance or through dominance of a particular homoeolog copy.
1043 Speciation and hybridisation event dates are based on estimates from[108]. **b,** Density plot
1044 showing the distribution of period lengths across rhythmic transcripts (BH $q < 0.01$) in
1045 *Arabidopsis*, *Brassica rapa*, *Brachypodium distachyon*, *Glycine max* (Soybean) and wheat
1046 based on meta2d estimates on 24-68h data following transfer to constant light. **c,** Histogram
1047 showing distribution of period lengths in wheat split between the A, B and D subgenomes.
1048 Dotted line indicates the mean period for the A, B and D subgenomes. **d,** Proportions of triads
1049 with either zero (red segment), one (green segment), two (blue segment) or three (purple
1050 segment) rhythmic gene(s) out of the 16,359 expressed triads in this dataset. Lighter shading
1051 in the outer segments represents cases where one/two homeolog(s) have high confidence
1052 rhythmicity (BH $q < 0.01$) alongside an arrhythmic homeolog (BH $q > 0.05$). We term these
1053 genes "imbalanced rhythmicity" triads. Of the 3448 triads with three rhythmic genes
1054 (represented by the purple segment in d), we also looked for triads with circadian imbalance
1055 in: phase (**e**), period (**f**) or relative amplitude (**g**). 464 triads had homoeologs which peaked
1056 with an optimum lag of 4, 8 or 12h following cross-correlation analysis. 1,139 triads had
1057 homoeologs with period differences of more than 2h. 701 triads had homoeologs with a more
1058 than two-fold difference in relative amplitude. **h,i,** Example triads for imbalanced rhythmicity,
1059 where either one or two homoeologs are rhythmic respectively. **j,** Example triad where the D
1060 genome homeolog lags by 8h. **k,** Example of a triad where the A genome homoeolog has a
1061 period estimate 4h longer than the D genome homoeolog. **l,** Example triad where the relative
1062 amplitude of the D-genome homoeolog is more than four times that of the A-genome
1063 homoeolog. **m,** The rhythmicity of all three homoeologs in **l,** is evident when the expression is
1064 mean normalized. **n,** Mean expression of transcripts across all timepoints in the A, B and D
1065 subgenomes within imbalanced rhythmicity triads compared with circadian balanced and
1066 arrhythmic triads. Error bars represent standard error.

1067 Circadian statistics are meta2d estimates from data 0-68h after transfer to L:L. Data represent
1068 the mean of three biological replicates with transcript expression collapsed to gene level. Genes
1069 in example triads are: [Triad 1664: TraesCS3A02G177600, TraesCS3B02G207400,
1070 TraesCS3D02G183200], [Triad 408: TraesCS3A02G533700, TraesCS3B02G610500,
1071 TraesCS3D02G539000], [Triad 13405: TraesCS6A02G269100, TraesCS6B02G296400,
1072 TraesCS6D02G245800], [Triad 10854: TraesCS6A02G166500, TraesCS6B02G194000,
1073 TraesCS6D02G155100] and [Triad 2454: TraesCS2A02G333000, TraesCS2B02G348800,
1074 TraesCS2D02G329900].

1075

1076

1077

1078 **Fig. 2. Overlapping GO-slim terms shared between *Arabidopsis* and wheat modules**
1079 **expressed at similar times in the day**

1080 **a,** Pearson correlation coefficient (r) between eigengenes for wheat and *Arabidopsis* expression
1081 modules ordered by circadian phase. Coloured triangles and axes labels correspond to module
1082 expression profiles and columns in bubble-plot. **b-e,** Expression profiles of *Arabidopsis* and
1083 wheat modules compared in the main text normalised to their mean. Solid and dashed black
1084 lines represent the module eigengene for wheat and *Arabidopsis* modules respectively. **f,** GO-
1085 slim terms associated with *Arabidopsis* and Wheat modules. Modules are ordered by predicted
1086 CT phase for each species. Only terms with -Log10p >3 are shown. Wheat W6 and *Arabidopsis*
1087 A4 contained no terms above the significance cut-off and so are not shown. Bubble color
1088 indicates the -Log10p-value significance from Fisher's exact test and size indicates the
1089 frequency of the GO-slim term in the underlying EBI Gene Ontology Annotation database
1090 (larger bubbles indicate more general terms).
1091
1092

1093 **Fig. 3. Free-running expression of core circadian clock genes in wheat and their homologs**
1094 **in *Arabidopsis.* a-l,** Wheat circadian clock genes were identified through alignment of
1095 phylogenetic protein family trees or BLASTP to known clock gene homologs. Gene IDs for
1096 each gene set are in Supplementary Table 8. Wheat homoeologs are coloured according to their
1097 identity to either the A genome (orange), B genome (yellow) or D genome (blue) and grey and
1098 white blocks indicate subjective dark and light time periods under constant conditions. Data
1099 represent the mean of three biological replicates and transcript expression is collapsed to gene
1100 level. Expression profiles for additional core circadian clock genes are in Supplementary Fig.
1101 17. **m,** phases of core clock genes in *Arabidopsis* and wheat (meta2d estimates from data 24-
1102 68h after transfer to L:L). Genes were not plotted if B.H q-values were > 0.01. Wheat values
1103 represent circular mean circadian phases (CT) across homoeologs calculated in Supplementary
1104 Table 9.
1105

1106 **Fig. 4. Similarities and differences in circadian control of transcript accumulation in key**
1107 **genes involved in primary metabolism and signalling.** Circles represent metabolites
1108 involved in the breakdown and biosynthesis of starch. Starch synthesis occurs during the day
1109 and breakdown occurs at night as indicated by the yellow to grey shading gradient. The dotted
1110 line encloses processes which take place in the chloroplast. Abbreviations: HP: Hexose-
1111 phosphate, T6P: Trehalose-6-phosphate, TP: Triose phosphate, 3-PGA: Glycerate 3-
1112 phosphate, Fru6P: Fructose-6-phosphate, Glc6P: Glucose-6-phosphate, Glc1P: Glucose-1-
1113 phosphate, ATP: Adensoine tri phosphate, ADP-Glc: ADP-glucose, TPS: Trehalose phosphate
1114 synthase, TPP: Trehalose phosphate phosphatase, PGK1: Phosphoglycerate kinase 1, PGI1:
1115 Glucose-6-phosphate isomerase, PGM1: Phosphoglucomutase-1, PHS1 and 2: ALPHA-
1116 GLUCAN PHOSPHORYLASE 1 and 2, AGPase: ADP-Glc pyrophosphorylase, BAM1,3,5:
1117 β-amylase 1,3,5, ISA1,2,3: Isoamylase 1,2,3, DPE1,2: Disproportionating enzyme1 and 2,
1118 SBEI,II: Starch branching enzyme I, II, PU1: Pullulanase 1, PWD: Phosphoglucan, water
1119 dikinase, GWD: α-glucan, water dikinase, SEX4: starch excess 4, AMY3: α-amylase, GBSS:
1120 Granule bound Starch synthase, SSI-IV: Starch synthase I-IV. Pathway references:[109–111].
1121

1122 **Table 1: Numbers of rhythmic genes at (BH $q < 0.05$ or BH $q < 0.01$) in *Arabidopsis* and**
1123 **wheat identified using Metacycle Benjamini Hochberg *q*-values.** Periods, relative
1124 amplitudes, and *q*-values are estimates from meta2d. Data windows reflect hours relative to
1125 transfer to constant light from entrained 12:12h light conditions. A repeat of this table with
1126 pre-filtering to remove low-expression genes is provided in Supplementary Figure 1, and the
1127 effects on proportions of rhythmic genes are discussed in Supplementary Note 1.
1128

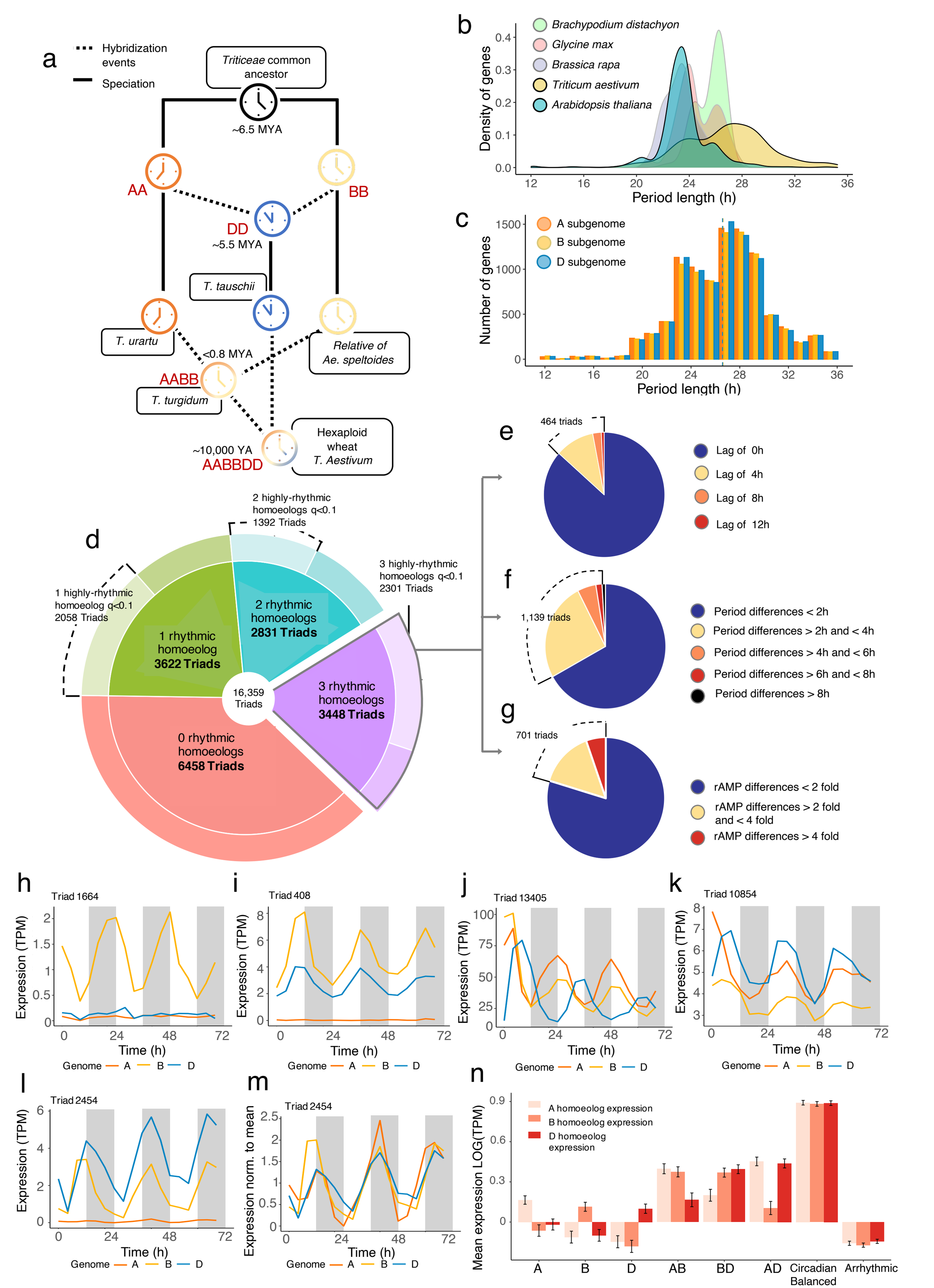
	Wheat data from this study		*Arabidopsis* data from Romanowski et al.
	24-68 data window	0-68 data window	24-68 data window
Total number of expressed genes	86,567	86,567	26,392
Total rhythmic genes (BH $q < 0.05$)	28,594	28,530	13,392
Total rhythmic genes (BH $q < 0.01$)	18,633	21,059	10,317
Mean Period (h) (BH $q < 0.05$)	26.60h (SD 3.62)	26.75h (SD 2.82)	23.50 (SD 2.52)
Mean Period (h) (BH $q < 0.01$)	26.82h (SD 3.21)	26.83h (SD 2.42)	23.62 (SD 2.04)
Mean relative Amplitude (BH $q < 0.05$)	0.24 (SD 0.19)	0.26 (SD 0.20)	0.28 (SD 0.20)
Mean relative Amplitude (BH $q < 0.01$)	0.27 (SD 0.19)	0.29 (SD 0.21)	0.30 (SD 0.20)

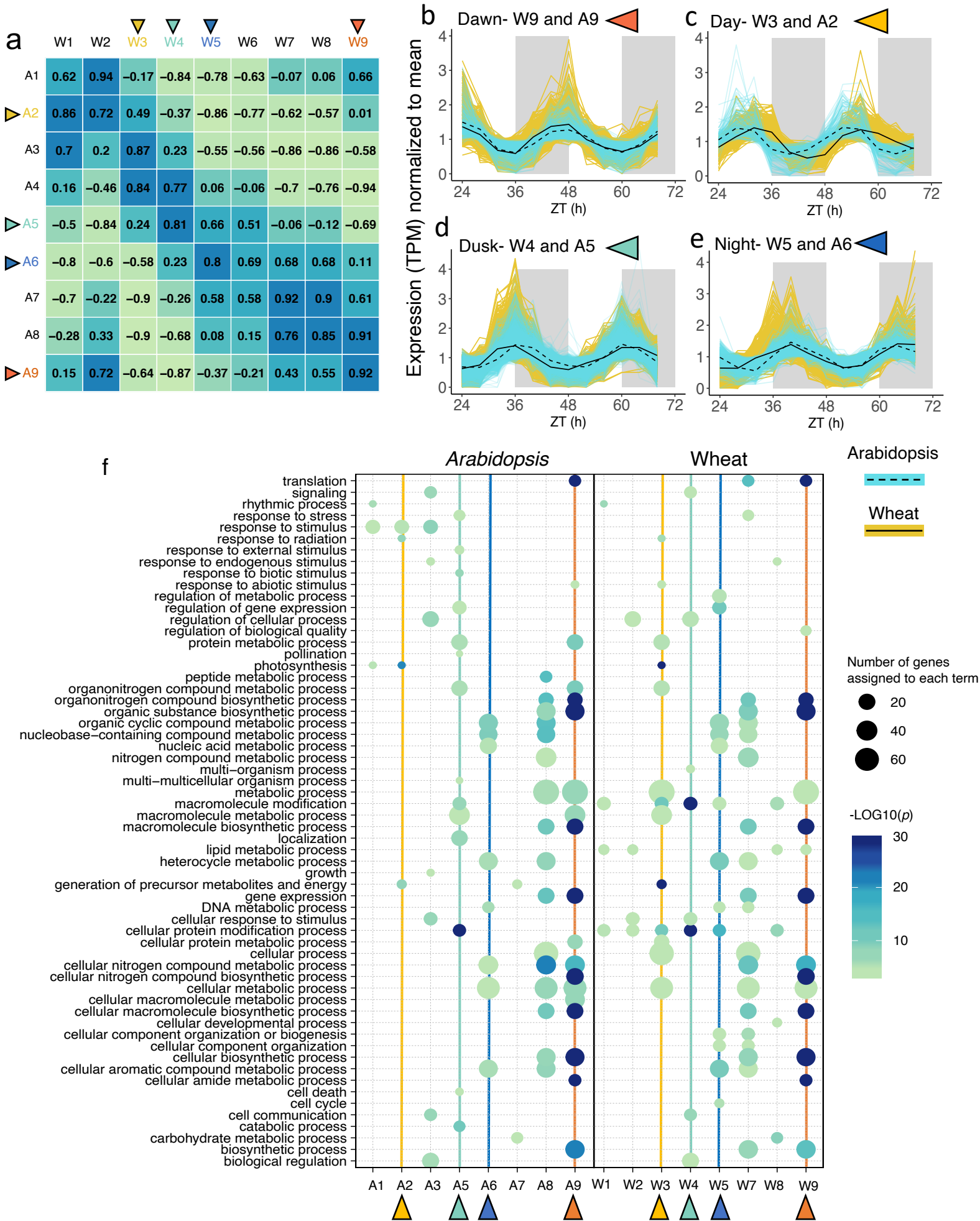
1129 **Table 2: GO-slim terms for biological processes associated with circadian balanced,**
1130 **circadian imbalanced, and arrhythmic wheat triads.** Only enriched terms which were
1131 highly enriched (Fisher's exact test $p<0.01$) in one category and non-significantly expressed
1132 ($p>0.05$) in other categories is displayed.

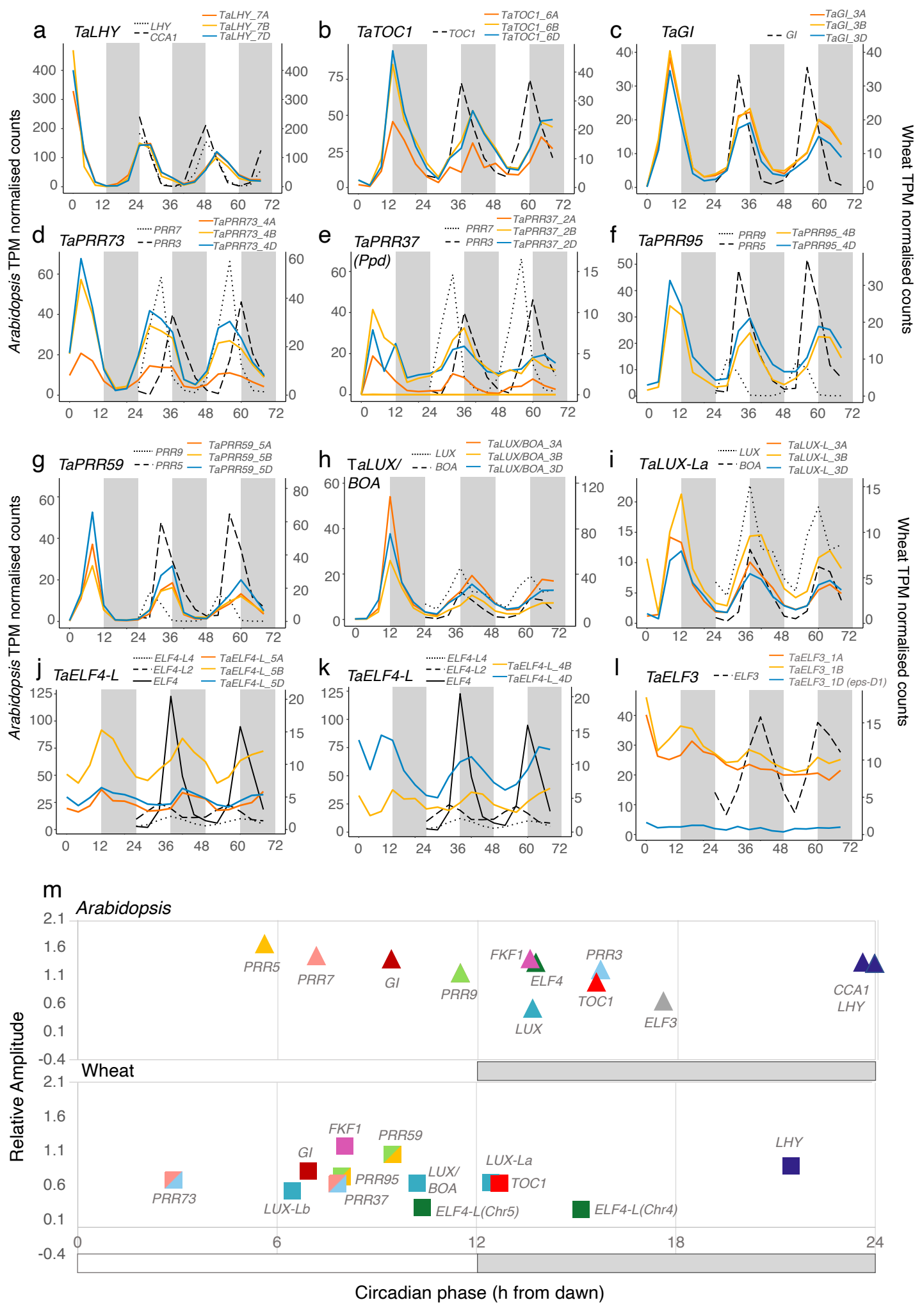
	GO ID	Terms	*p*-value in circadian balanced triads	*p*-value in circadian imbalanced triads	*p*-value in non-rhythmic triads
CIRCADIAN BALANCED	GO:0009628	response to abiotic stimulus	0.00	0.22	0.58
	GO:0015979	photosynthesis	0.00	1.00	1.00
	GO:0006091	generation of precursor metabolites and energy	0.00	1.00	1.00
	GO:0006518	peptide metabolic process	0.00	1.00	0.92
	GO:1901566	organonitrogen compound biosynthetic process	0.00	1.00	0.92
	GO:0009059	macromolecule biosynthetic process	0.00	1.00	0.99
	GO:0006412	translation	0.00	1.00	0.98
	GO:0034645	cellular macromolecule biosynthetic process	0.00	1.00	1.00
	GO:0010467	gene expression	0.00	1.00	0.90
	GO:0019725	cellular homeostasis	0.00	0.54	0.96
	GO:0065008	regulation of biological quality	0.00	0.71	1.00
CIRCADIAN IMBALANCED	GO:0003006	developmental process involved in reproduction	0.85	0.00	0.87
	GO:0090567	reproductive shoot system development	0.92	0.01	1.00
	GO:0009719	response to endogenous stimulus	0.97	0.01	0.10
	GO:0048731	system development	0.98	0.00	0.97
	GO:0048608	reproductive structure development	0.98	0.00	0.97
	GO:0043412	macromolecule modification	1.00	0.00	0.28
	GO:0022414	reproductive process	1.00	0.00	0.47
NON-RHYTHMIC	GO:0044237	cellular metabolic process	0.08	1.00	0.00
	GO:0009605	response to external stimulus	0.41	0.39	0.00
	GO:0009607	response to biotic stimulus	0.58	0.60	0.01
	GO:0060255	regulation of macromolecule metabolic process	0.60	0.76	0.00
	GO:0009056	catabolic process	0.65	0.93	0.00

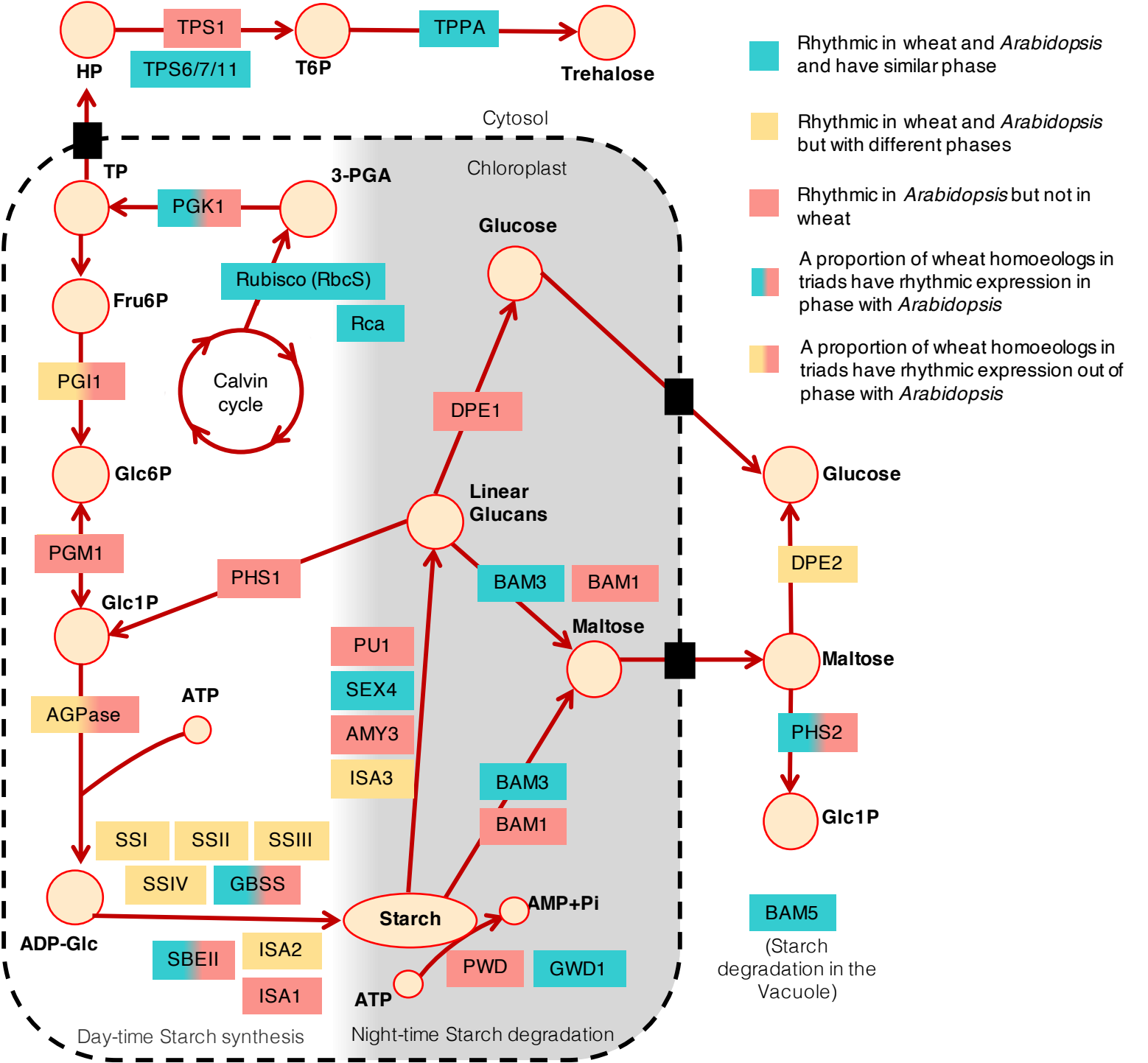
	GO:0048869	cellular developmental process	0.71	0.19	0.00
	GO:0019222	regulation of metabolic process	0.78	0.85	0.00
	GO:0044238	primary metabolic process	0.81	0.24	0.00
	GO:0071704	organic substance metabolic process	0.81	0.26	0.00
	GO:0008219	cell death	0.82	0.94	0.00
	GO:0010468	regulation of gene expression	0.88	0.15	0.00
	GO:0009790	embryo development	0.95	0.98	0.00
	GO:0007049	cell cycle	0.97	1.00	0.00
	GO:0065009	regulation of molecular function	1.00	1.00	0.00
	GO:0006807	nitrogen compound metabolic process	1.00	1.00	0.00

1133


a
Hybridization events
Speciation
Triticeae common ancestor
~6.5 MYA
AA
BB
DD
~5.5 MYA
T. tauschii
T. urartu
<0.8 MYA
Relative of Ae. speltoides
AABB
T. turgidum
~10,000 YA
AABBDD
Hexaploid wheat T. Aestivum
b
Brachypodium distachyon
Glycine max
Brassica rapa
Triticum aestivum
Arabidopsis thaliana
Density of genes
Period length (h)
c
A subgenome
B subgenome
D subgenome
Number of genes
Period length (h)
d
2 highly-rhythmic homoeologs q<0.1 1392 Triads
3 highly-rhythmic homoeologs q<0.1 2301 Triads
1 highly-rhythmic homoeolog q<0.1 2058 Triads
2 rhythmic homoeologs 2831 Triads
1 rhythmic homoeolog 3622 Triads
16,359 Triads
3 rhythmic homoeologs 3448 Triads
0 rhythmic homoeologs 6458 Triads
e
464 triads
Lag of 0h
Lag of 4h
Lag of 8h
Lag of 12h
f
1,139 triads
Period differences < 2h
Period differences > 2h and < 4h
Period differences > 4h and < 6h
Period differences > 6h and < 8h
Period differences > 8h
g
701 triads
rAMP differences < 2 fold
rAMP differences > 2 fold and < 4 fold
rAMP differences > 4 fold
h
Triad 1664
i
Triad 408
j
Triad 13405
k
Triad 10854
Expression (TPM)
Time (h)
Genome A B D
l
Triad 2454
m
Triad 2454
Expression norm. to mean
n
A homoeolog expression
B homoeolog expression
D homoeolog expression
Mean expression LOG(TPM)
A
B
D
AB
BD
AD
Circadian Balanced
Arrhythmic


a
W1 W2 W3 W4 W5 W6 W7 W8 W9
A1 0.62 0.94 −0.17 −0.84 −0.78 −0.63 −0.07 0.06 0.66
A2 0.86 0.72 0.49 −0.37 −0.86 −0.77 −0.62 −0.57 0.01
A3 0.7 0.2 0.87 0.23 −0.55 −0.56 −0.86 −0.86 −0.58
A4 0.16 −0.46 0.84 0.77 0.06 −0.06 −0.7 −0.76 −0.94
A5 −0.5 −0.84 0.24 0.81 0.66 0.51 −0.06 −0.12 −0.69
A6 −0.8 −0.6 −0.58 0.23 0.8 0.69 0.68 0.68 0.11
A7 −0.7 −0.22 −0.9 −0.26 0.58 0.58 0.92 0.9 0.61
A8 −0.28 0.33 −0.9 −0.68 0.08 0.15 0.76 0.85 0.91
A9 0.15 0.72 −0.64 −0.87 −0.37 −0.21 0.43 0.55 0.92
b Dawn- W9 and A9
c Day- W3 and A2
d Dusk- W4 and A5
e Night- W5 and A6
Expression (TPM) normalized to mean
ZT (h)
f
Arabidopsis
Wheat
translation
signaling
rhythmic process
response to stress
response to stimulus
response to radiation
response to external stimulus
response to endogenous stimulus
response to biotic stimulus
response to abiotic stimulus
regulation of metabolic process
regulation of gene expression
regulation of cellular process
regulation of biological quality
protein metabolic process
pollination
photosynthesis
peptide metabolic process
organonitrogen compound metabolic process
organonitrogen compound biosynthetic process
organic substance biosynthetic process
organic cyclic compound metabolic process
nucleobase−containing compound metabolic process
nucleic acid metabolic process
nitrogen compound metabolic process
multi−organism process
multi−multicellular organism process
metabolic process
macromolecule modification
macromolecule metabolic process
macromolecule biosynthetic process
localization
lipid metabolic process
heterocycle metabolic process
growth
generation of precursor metabolites and energy
gene expression
DNA metabolic process
cellular response to stimulus
cellular protein modification process
cellular protein metabolic process
cellular process
cellular nitrogen compound metabolic process
cellular nitrogen compound biosynthetic process
cellular metabolic process
cellular macromolecule metabolic process
cellular macromolecule biosynthetic process
cellular developmental process
cellular component organization or biogenesis
cellular component organization
cellular biosynthetic process
cellular aromatic compound metabolic process
cellular amide metabolic process
cell death
cell cycle
cell communication
catabolic process
carbohydrate metabolic process
biosynthetic process
biological regulation
A1 A2 A3 A5 A6 A7 A8 A9 W1 W2 W3 W4 W5 W7 W8 W9
Arabidopsis
Wheat
Number of genes assigned to each term
20
40
60
-LOG10(p)
30
20
10


a
TaLHY
LHY
CCA1
TaLHY_7A
TaLHY_7B
TaLHY_7D
b
TaTOC1
TOC1
TaTOC1_6A
TaTOC1_6B
TaTOC1_6D
c
TaGI
GI
TaGI_3A
TaGI_3B
TaGI_3D
d
TaPRR73
PRR7
PRR3
TaPRR73_4A
TaPRR73_4B
TaPRR73_4D
e
TaPRR37
(Ppd)
PRR7
PRR3
TaPRR37_2A
TaPRR37_2B
TaPRR37_2D
f
TaPRR95
PRR9
PRR5
TaPRR95_4B
TaPRR95_4D
g
TaPRR59
PRR9
PRR5
TaPRR59_5A
TaPRR59_5B
TaPRR59_5D
h
TaLUX/
BOA
LUX
BOA
TaLUX/BOA_3A
TaLUX/BOA_3B
TaLUX/BOA_3D
i
TaLUX-La
LUX
BOA
TaLUX-L_3A
TaLUX-L_3B
TaLUX-L_3D
j
TaELF4-L
ELF4-L4
ELF4-L2
ELF4
TaELF4-L_5A
TaELF4-L_5B
TaELF4-L_5D
k
TaELF4-L
ELF4-L4
ELF4-L2
ELF4
TaELF4-L_4B
TaELF4-L_4D
l
TaELF3
ELF3
TaELF3_1A
TaELF3_1B
TaELF3_1D (eps-D1)
Arabidopsis TPM normalised counts
Wheat TPM normalised counts
m
Arabidopsis
PRR5
PRR7
GI
PRR9
FKF1
ELF4
LUX
PRR3
TOC1
ELF3
CCA1
LHY
Wheat
PRR73
LUX-Lb
GI
FKF1
PRR95
PRR37
PRR59
LUX/
BOA
ELF4-L(Chr5)
LUX-La
TOC1
ELF4-L(Chr4)
LHY
Relative Amplitude
Circadian phase (h from dawn)


TPS1
TPS6/7/11
TPPA
HP
T6P
Trehalose
Cytosol
Chloroplast
TP
3-PGA
PGK1
Glucose
Rubisco (RbcS)
Rca
Fru6P
Calvin cycle
PGI1
DPE1
Glc6P
Linear Glucans
Glucose
PGM1
PHS1
BAM3
BAM1
DPE2
Glc1P
Maltose
PU1
SEX4
AMY3
ISA3
Maltose
AGPase
ATP
PHS2
BAM3
BAM1
Glc1P
SSI
SSII
SSIII
SSIV
GBSS
ADP-Glc
Starch
AMP+Pi
SBEII
ISA2
ISA1
PWD
GWD1
ATP
Day-time Starch synthesis
Night-time Starch degradation
Rhythmic in wheat and Arabidopsis and have similar phase
Rhythmic in wheat and Arabidopsis but with different phases
Rhythmic in Arabidopsis but not in wheat
A proportion of wheat homoeologs in triads have rhythmic expression in phase with Arabidopsis
A proportion of wheat homoeologs in triads have rhythmic expression out of phase with Arabidopsis
BAM5
(Starch degradation in the Vacuole)