XAP5 CIRCADIAN TIMEKEEPER regulates RNA splicing and the circadian clock via genetically separable pathways

Authors:

Hongtao Zhang (张弘韬)[1], Roderick W. Kumimoto[1], Shajahan Anver[1,2], Stacey L. Harmer[1*]

Author affiliations:

[1]Department of Plant Biology, College of Biological Sciences, University of California, Davis, CA 95616, USA

[2]Department of Genetics, Evolution and Environment, University College London, London, UK

[*]**Corresponding author:**

slharmer@ucdavis.edu

Running title:

Separable roles for *XCT* in splicing and the clock

One sentence summary:

XCT plays genetically separable roles in controlling the fidelity of 3' splice site selection during pre-mRNA splicing and the pace of circadian clock in Arabidopsis

Abstract

The circadian oscillator allows organisms to synchronize their cellular and physiological activities with diurnal environmental changes. In plants, the circadian clock is primarily composed of multiple transcriptional-translational feedback loops. Regulators of post-transcriptional events, such as pre-mRNA splicing factors, are also involved in controlling the pace of the clock. However, in most cases the underlying mechanisms remain unclear. We have previously identified *XAP5 CIRCADIAN TIMEKEEPER* (*XCT*) as an *Arabidopsis thaliana* circadian clock regulator with uncharacterized molecular functions. Here, we report that XCT physically interacts with components of the spliceosome, including members of the Nineteen Complex (NTC). PacBio Iso-Seq data show that *xct* mutants have transcriptome-wide pre-mRNA splicing defects, predominantly aberrant 3' splice site selection. Expression of a genomic copy of *XCT* fully rescues those splicing defects, demonstrating that functional *XCT* is important for splicing. Dawn-expressed genes are significantly enriched among those aberrantly spliced in *xct* mutants, suggesting that the splicing activity of *XCT* may be circadian regulated. Furthermore, we show that loss of function mutations in *PRP19A* or *PRP19B*, two homologous core NTC components, suppress the short circadian period phenotype of *xct-2*. Interestingly, we do not see rescue of the splicing defects of core clock genes in *prp19 xct* mutants. Similarly, we find that transgenic expression of the fission yeast ortholog of *XCT* in Arabidopsis shortens the circadian period without affecting splicing. Taken together, our results suggest that *XCT* may regulate splicing and the clock function through genetically separable pathways.

Introduction

Most eukaryotes have evolved an endogenous timekeeper known as the circadian clock, which allows them to anticipate the daily fluctuating environmental conditions caused by the earth's rotation (Harmer, 2009). Although the central oscillators of circadian clocks in diverse eukaryotes lack conserved individual components, they share similar general architectures (Nohales and Kay, 2016). In plants, the approximately 24-h periodicity of the clock is maintained by a complex gene regulatory network consisting primarily of repressors and activators of transcription. Those regulators, often referred to as core circadian clock genes, regulate each other's expression and the expression of thousands of output genes (Creux and Harmer, 2019).

32

33 Additionally, post-transcriptional and post-translational mechanisms such as alternative splicing
34 of precursor messenger RNAs (pre-mRNAs) provide critical regulation of clock function. It has
35 been suggested that changes in splicing of core circadian clock genes are important for
36 Arabidopsis in response to environmental stresses (James et al., 2012; Kwon et al., 2014). For
37 example, cold temperatures significantly suppress production of a splice variant of *CIRCADIAN*
38 *CLOCK-ASSOCIATED1* (*CCA1*) in which intron four is retained. This incompletely spliced
39 isoform encodes a truncated protein which competitively inhibits the function of fully spliced
40 *CCA1* (Seo et al., 2012). However, in most cases the effects of alternative splicing of pre-
41 mRNAs on circadian clock function remain unclear.
42
43 The mechanisms underlying pre-mRNA splicing are increasingly well understood (Wilkinson et
44 al., 2020). There are two catalytic transesterification steps, which sequentially remove first the 5'
45 and then the 3' splice sites (5'SS and 3'SS) of introns from their adjacent exons (Shi, 2017). This
46 process is carried out by five small nuclear ribonucleoproteins (snRNPs) and hundreds of non-
47 snRNP splicing factor proteins which assemble on a pre-mRNA to form the spliceosome
48 complex (Wilkinson et al., 2020). One of these spliceosome complex components is the Nineteen
49 Complex (NTC, also known as PRP19 complex), named after PRECURSOR RNA
50 PROCESSING 19 (PRP19), a U-box E3 ubiquitin ligase that forms the core of the NTC
51 (Hatakeyama et al., 2001; Koncz et al., 2012). The NTC is associated with the spliceosome
52 during the two transesterification steps and helps facilitate conformational rearrangements and
53 promote splicing fidelity (Fig. 1; Hogg et al., 2010). The NTC is highly conserved across
54 eukaryotes. In Arabidopsis, multiple orthologs of yeast NTC proteins including two PRP19
55 homologs (PRP19A/MAC3A and PRP19B/MAC3B) have been shown to physically interact
56 with the spliceosome (Monaghan et al., 2009; Deng et al., 2016). Plants mutant for *prp19a*
57 *prp19b* or homologs of other NTC-associated proteins such as *pleiotropic regulatory locus1*
58 (*prl1*) and *snw/ski-interacting protein* (*skip*) have genome-wide intron retention defects (Jia et
59 al., 2017; Wang et al., 2012; Li et al., 2019). These data indicate that NTC components act as
60 evolutionarily conserved splicing factors in Arabidopsis.
61

62 Mutation of splicing factors can lead to disruption of circadian clock function (Shakhmantsir and
63 Sehgal, 2019). For example, loss-of-function alleles of NTC components, including *PRP19*,
64 *PRL1* and *SKIP*, cause lengthening of circadian period (Wang et al., 2012; Feke et al., 2019; Li
65 et al., 2019). Aberrantly spliced mRNA variants of core circadian clock genes have been
66 detected in these splicing factor mutants (Sanchez et al., 2010; Wang et al., 2012; Jones et al.,
67 2012; Schlaen et al., 2015; Marshall et al., 2016; Li et al., 2019; Feke et al., 2019), suggesting
68 that changes in the pace of the clock might be caused by aberrant splicing of core clock genes. In
69 some cases, epistasis analysis suggests that this may be true (Marshall et al., 2016; Schlaen et al.,
70 2015). However, in other cases, genetic analysis has either not been performed or has revealed
71 additive interactions between the splicing factor and clock gene mutants. Besides, the levels of
72 mis-spliced mRNA variants of clock genes are usually only a small fraction of total transcripts
73 (Jones et al., 2012; Perez-Santángelo et al., 2014; Feke et al., 2019; Sanchez et al., 2010; Wang
74 et al., 2012). Thus, it is unclear whether splicing factors affect clock function solely by
75 controlling the splicing of core clock genes.
76
77 RNA splicing factors often function in multiple biological pathways: plants deficient for
78 components of the NTC have defects in immunity, microRNA biogenesis, DNA damage
79 response and transcriptional elongation (Koncz et al., 2012; Monaghan et al., 2009; Zhang et al.,
80 2013; Jia et al., 2017). Some splicing factors are known to carry out roles in nuclear processes
81 biochemically separable from their roles in splicing. In addition to its structural role in the
82 spliceosome, PRP19 also senses DNA damage and promotes DNA repair via its ubiquitin ligase
83 activity (Maréchal et al., 2014). Additionally, SKIP acts in two distinct complexes to regulate
84 splicing and the transcription of genes involved in flowering time control (Li et al., 2019).
85 Nonetheless, whether splicing factors might control circadian clock function via splicing-
86 independent activities has not been investigated.
87
88 We previously identified *XAP5 CIRCADIAN TIMEKEEPER* (*XCT*) as a novel regulator of the
89 Arabidopsis circadian clock (Martin-Tryon and Harmer, 2008). Like NTC components, XCT is
90 also well-conserved across eukaryotes. Homologs of XCT (also known as XAP5 proteins) share
91 a highly conserved C-terminal protein domain and are nuclear-localized (Martin-Tryon and
92 Harmer, 2008; Anver et al., 2014; Li et al., 2018; Lee et al., 2020). Previously, we found that

93 transgenic expression of Arabidopsis *XCT* fully rescued the slow-growth phenotype of fission
94 yeast mutant for *xap5* (Anver et al., 2014). Taken together, these data suggest that XCT
95 homologs might share similar molecular and cellular functions across kingdoms.
96
97 It has been reported that FAM50A, one of the two XCT orthologs in humans, physically
98 associates with the spliceosomal C complex and its mutants have defects in RNA splicing
99 (Bessonov et al., 2008; Lee et al., 2020). Similarly, a recent study in Arabidopsis also reported
100 the association of XCT with the spliceosome (Liu et al., 2022), implying that XAP5 proteins may
101 be evolutionarily conserved splicing factors. However, evidence suggests that XAP5 proteins
102 may also participate in fundamental biological processes other than splicing. Fission yeast and
103 *Chlamydomonas* XAP5 proteins associate with chromatin and directly regulate transcription
104 (Anver et al., 2014; Li et al., 2018). In addition to its role in the clock, Arabidopsis *XCT* has been
105 implicated in diverse processes including small RNA production, immune signaling, and DNA
106 damage responses (Fang et al., 2015; Xu et al., 2017; Kumimoto et al., 2021). Notably, NTC
107 components have also been reported to function in all these pathways (Jia et al., 2017; Maréchal
108 et al., 2014; Chanarat and Sträßer, 2013). Although pleiotropic defects are seen in *xct* mutants,
109 interconnections between these phenotypes and the molecular function of *XCT* have not been
110 extensively studied.
111
112 In the current work, we report that XCT physically interacts with NTC and other spliceosome-
113 associated proteins in Arabidopsis. We use long-read RNA sequencing to reveal that *XCT*
114 controls the fidelity of 3' splice site selection for hundreds of pre-mRNA splicing events,
115 probably by rejecting downstream suboptimal 3' splice sites. Intriguingly, circadian-regulated
116 genes that are also aberrantly spliced in *xct* mutants are significantly enriched for peak
117 expression at subjective dawn. This implies that the splicing-related activity of *XCT* may be
118 circadian regulated. We also demonstrate that *PRP19A* and *PRP19B* display epistatic genetic
119 interactions with *XCT* in control of the circadian clock period length but not of the splicing of
120 core clock genes. Finally, our data show that expression of *SpXAP5*, the fission yeast ortholog of
121 *XCT*, shortens the circadian period without impacting splicing of clock genes. Collectively, our
122 results suggest that XCT works close to NTC and may regulate pre-mRNA splicing and the
123 circadian clock function via distinct pathways.

124

125 **Results**

126 **XCT physically interacts with the Nineteen Complex and other spliceosome-associated**
127 **proteins**

128

129 To uncover the molecular functions of XCT, we carried out mass spectrometry (MS)
130 experiments to identify its possible interactors. We expressed epitope-tagged XCT protein under
131 the control of the native *XCT* promoter in *xct-2* mutant background (*xct-2 XCT*). This transgene
132 largely rescued the morphological and circadian clock defects of *xct-2* (Supplemental Fig. 1).
133 Next, we affinity purified tagged XCT from plant extracts and analyzed co-purifying proteins via
134 MS. In total, 26 proteins were significantly enriched in the XCT immunoprecipitation compared
135 to control immunoprecipitations (Welch's two sample *t*-test, $P < 0.05$; Supplemental Dataset 1).
136 Among those XCT-associated proteins, 15 (57.7%) were annotated as mRNA splicing-related,
137 including 11 core NTC or NTC-associated proteins (Table 1), consistent with the MS and co-
138 immunoprecipitation data from a recent report (Liu et al., 2022). Notably, studies in human and
139 yeast have revealed that NTC is physically associated with the spliceosomal complex throughout
140 the two catalytic transesterification steps (Fig. 1; Hogg et al., 2010). Taken together, those results
141 imply that *XCT* may act close to NTC and function in the catalytic steps during pre-mRNA
142 splicing. Interestingly, we also observed 8 (30.8%) chloroplast proteins enriched in the XCT
143 immunoprecipitation (Supplemental Dataset 1), which may be related to the delayed leaf
144 greening phenotypes observed in *xct-2* mutant plants (Martin-Tryon and Harmer, 2008).

145

146 ***XCT* is required for the fidelity of 3' splice site selection during pre-mRNA splicing**

147

148 To investigate a possible role for *XCT* in RNA splicing, we performed PacBio Isoform
149 Sequencing (Iso-Seq) on wild-type Col-0, reduction-of-function allele *xct-1*, null allele *xct-2*, and
150 the complemented line *xct-2 XCT*. Additionally, we also sequenced *prl1-2*, which contains a T-
151 DNA insertion mutation *in PRL1*, a NTC member that has been demonstrated to control genome-
152 wide splicing efficiency (Jia et al., 2017). Since both transcript levels and RNA splicing of a
153 large proportion of the Arabidopsis transcriptome are circadian regulated (Romanowski et al.,
154 2020; Yang et al., 2020), the time of day at which plants are harvested has significant effects on

155 gene expression and splicing analysis (Hsu and Harmer, 2012). Therefore, to minimize any
156 differences in subjective time of day between wild-type plants, the long-circadian-period *prl1-2*,
157 and the short-period *xct* mutants (Supplemental Fig. 2), we grew plants in constant
158 environmental conditions for ten days and then pooled seedlings harvested at 2-hour intervals
159 across a twenty-four hour period (Fig. 2A). For each genotype, we sequenced three biological
160 replicates and acquired an average of 784,832 full-length transcript reads per genotype per
161 replicate (Supplemental Dataset 2). With no alignment needed, those reads were directly mapped
162 to the TAIR 10 Arabidopsis reference genome and then subjected to differential splicing analysis
163 using the JunctionSeq R package (Hartley and Mullikin, 2016).
164
165 Previous studies have demonstrated that intron retention is the most prevalent type of alternative
166 splicing event in Arabidopsis (Wang and Brendel, 2006; Filichkin et al., 2010). Indeed, plants
167 mutant for multiple NTC components have been reported to have global intron retention defects
168 (Jia et al., 2017; Meng et al., 2022). In our analysis of *prl1-2*, we found 5730 out of 44,496
169 detected splicing events were significantly differentially enriched from Col-0 (Supplemental
170 Dataset 3). Notably, 99% of these events were decreases in known junctions (Supplemental Fig.
171 3), consistent with previous studies suggesting that loss of *PRL1* function mainly causes intron
172 retention without affecting splice site selection (Jia et al., 2017).
173
174 We next examined transcript composition in *xct* mutants. In the null allele *xct-2*, we detected
175 75,073 splicing events, of which 875 were significantly down-regulated and 532 were up-
176 regulated compared with Col-0 (FDR<0.05, Fig. 2B-C). Meanwhile, 209 and 337 splicing events
177 were significantly decreased or increased, respectively, in the partial loss-of-function mutant *xct-*
178 *1*. Comparing mis-regulated splicing between the two *xct* mutant alleles, 196 (93.8%) of the
179 down-regulated and 329 (97.6%) of the up-regulated events were shared. Thus, the two *xct*
180 alleles have similar splicing phenotypes but the severity of the defects depends on the degree of
181 XCT function lost. Next, we analyzed the splicing events in the complemented line *xct-2 XCT*.
182 Only 16 out of 84,965 events were significantly differentially enriched compared with Col-0,
183 indicating that the splicing defects of *xct-2* were almost completely rescued by the restored *XCT*
184 expression. Taken together, these results demonstrate that *XCT* is important for transcriptome-
185 wide pre-mRNA splicing.

186

187 To further characterize the major types of splicing defects caused by loss of *XCT* function, we
188 analyzed the splice sites for the differentially spliced junctions in *xct* mutants. Specifically, we
189 categorized *xct*-induced mis-splicing events into four classes based on whether the 5' and 3'
190 splice sites used were previously annotated (known) or not (novel). As expected, most down-
191 regulated splicing events displayed decreases in usage of junctions with known 5' and 3' splice
192 sites (Fig. 2D). Interestingly, 307 (90.8%) in *xct-1* and 462 (86.5%) in *xct-2* of up-regulated
193 splicing events involved increased usage of junctions with a known 5' splice site but a novel 3'
194 splice site (Fig. 2E). The decreases in abundance of junctions with known splice sites hence
195 reflects both intron retention events and novel 3' splice sites usage. Therefore, our results
196 demonstrate that in addition to controlling the efficient removal of introns, *XCT* is also
197 responsible for the fidelity of 3' splice site selection.
198
199 To investigate how *XCT* contributes to the fidelity of 3' splice site selection, we compared the
200 significantly up- or down-regulated 3' splice sites in *xct* mutants with all detected 3' splice sites
201 in Col-0. Intriguingly, most of the up-regulated novel 3' splice sites in *xct-1* and *xct-2* were
202 located less than 50 nucleotides downstream of the wild-type 3' splice sites (Fig. 3B;
203 Supplemental Fig. 4A). Studies in human and yeast have demonstrated that the sequences
204 preceding a 3' splice site, including the polypyrimidine tract (PPT) and the pyrimidine at the -3
205 position (Fig. 3A), are important for the strength of the 3' splice site (Horowitz, 2012).
206 Therefore, we further examined the 23-mer sequences flanking the 3' splice sites (from -20 to +3
207 position). We found that the frequency of the canonical AG sequence at the 3' splice sites was
208 not altered in either *xct-1* or *xct-2* (Fig. 3C; Supplemental Fig. 4B-C). However, the average
209 predicted strength of the 875 down-regulated 3' splice sites in *xct-2* was significantly weaker
210 than the average of total detected 3' splice sites in Col-0 (Fig. 3D). Specifically, the percentage
211 of pyrimidine residues at the -3 position was significantly reduced in down-regulated 3' splice
212 sites compared to that found in all detected 3' splice sites (Fisher's exact test, $P < 0.001$, Fig.
213 3E), suggesting that functional *XCT* is important for the removal of 3' splice sites with a
214 suboptimal sequence at the -3 position. Moreover, the 532 up-regulated 3' splice sites had an
215 even lower 3' splice site score than the down-regulated sites (Fig. 3D). The frequency of
216 pyrimidines both at the -3 position and throughout the PPT region was significantly lower in the

217 up-regulated junctions (Fig. 3E-F), showing that *xct-2* is less able to discriminate between 3'
218 splice sites during splicing. Taken together, our Iso-Seq data demonstrates that *XCT* controls the
219 accuracy of 3' splice site selection, possibly by helping to recognize weaker 3' splice sites and
220 reject sub-optimal downstream sites.
221
222 **Genes expressed at subjective dawn are enriched among those aberrantly spliced in *xct***
223 **mutants**
224
225 To explore the possible link between the circadian clock and splicing phenotypes in *xct* mutants,
226 we looked for enrichment of circadian clock regulation within genes aberrantly spliced in *xct.*
227 Among the 5602 previously reported clock-regulated genes (Romanowski et al., 2020) that are
228 detected in our Iso-Seq data, we identified 182 and 402 genes that are mis-spliced in *xct-1* and
229 *xct-2*, respectively (Fig. 4A; Supplemental Dataset 4). However, there is no significant over-
230 representation in either mutant (one-tailed Fisher's exact test, $P = 0.87$ for *xct-1* and 0.99 for *xct-*
231 *2*), suggesting that *XCT* does not preferentially affect the splicing of clock-regulated genes.
232 Previous studies have identified transcripts that are differentially spliced at various time of a day
233 (Romanowski et al., 2020; Yang et al., 2020). This suggest that the splicing activity may be
234 circadian clock regulated. Therefore, to test whether the circadian regulator *XCT* preferentially
235 affects pre-mRNA splicing at certain time of a day, we examined the distribution of estimated
236 peaked expression times of genes aberrantly spliced in *xct* mutants. We grouped all 5602 clock-
237 regulated genes into twelve 2-hour intervals based on their estimated peaked expression time
238 (Fig. 4B). Genes that are aberrantly spliced in *xct-2* are significantly enriched for peaked
239 expression between ZT22 and ZT2 compared with the other intervals of the day (Fig. 4C-D,
240 Fisher's exact test, $P < 1e\text{-}4$). Similarly, clock-regulated genes that are aberrantly spliced in *xct-1*
241 also showed a dawn-enriched expression pattern (Supplemental Fig. 5). Collectively, those
242 results indicate that *XCT* activity within the spliceosome may be regulated by the circadian
243 clock.
244
245 Alternative splicing of core clock genes is important for Arabidopsis responses to environmental
246 stresses such as shifts in ambient temperature (James et al., 2012; Kwon et al., 2014). Recent
247 studies showed that aberrant splicing of core clock genes might fully or partially contribute to the

248 altered circadian clock period phenotype in some splicing factor mutants (Sanchez et al., 2010;
249 Wang et al., 2012; Jones et al., 2012; Schlaen et al., 2015; Marshall et al., 2016; Li et al., 2019;
250 Feke et al., 2019). Similarly, in the *xct-1* and *xct-2* mutants, we detected shared aberrant splicing
251 events in five core circadian clock genes, including *LHY*, *LNK2*, *PRR7*, *TOC1* and *TIC* (Fig. 4A;
252 Fig. 5A-D), one or more of which might potentially cause the short clock period phenotype. Four
253 of these genes are transcriptionally regulated by the circadian clock, with peak *LHY* and *LNK2*
254 expression at dawn and peak *PRR7* and *TOC1* expression late in the afternoon (Creux and
255 Harmer, 2019). Thus, to further investigate whether *XCT* preferentially affects splicing of
256 transcripts produced at dawn, we monitored the abundance of the splice variants of these four
257 circadian-clock-regulated core clock genes together with one non-clock-regulated gene via
258 quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) over a 24-hour
259 light/dark period (LD) followed by 48 hours in constant light (LL).
260
261 As we expected, total transcript levels of clock-regulated genes displayed rhythmic abundance
262 with anticipated peak phases of expression (Supplemental Fig. 6). Notably, both the peak and
263 average abundance of the aberrantly spliced isoforms were significantly different in *xct-2*
264 compared with Col-0 (Fig. 5F-J). This serves as validation of the pooling strategy we used to
265 generate our Iso-Seq libraries, where the daily abundance of transcripts was averaged (Fig. 2A;
266 Fig. 5A-E). However, instead of preferentially accumulating at dawn, abundance of all splice
267 variants in *xct-2* was synchronized with total transcript levels (Fig. 5F-I; Supplemental Fig. 6).
268 Likewise, the mis-spliced isoform of the non-clock-regulated gene *SPPA* did not show a
269 significant morning peak (Fig. 5J). Meanwhile, the genomic *XCT* complemented line almost
270 completely rescued the splicing defects across the whole experimental period. Taken together,
271 those results suggest that although morning-expressed genes are preferentially enriched among
272 *XCT* targets, *XCT* is important for the accuracy of pre-mRNA splicing throughout the day under
273 both LD and free-running LL conditions.
274
275 **Reduction of *PRP19* function rescues the circadian clock but not the splicing defects in *xct-***
276 ***2***
277

278 Previous studies revealed that NTC components participate in regulation of circadian clock
279 function in Arabidopsis (Feke et al., 2019; Li et al., 2019). Here we examined circadian period in
280 mutants for several NTC components. We found that mutation in either of the two Arabidopsis
281 *PRP19* homologs, *PRP19A* (*MAC3A*) or *PRP19B* (*MAC3B*), caused minor lengthening of
282 circadian period. Moreover, *prp19a-1 prp19b-1* double mutants had a significantly slower
283 circadian clock than Col-0 (Supplemental Fig. 7A), consistent with previous reports (Feke et al.,
284 2019). Similarly, loss of function of other NTC members, including *PRL1*, *CDC5* and *SKIP*, also
285 lengthened the clock period by 1 to 3 hours (Supplemental Fig. 2B; Supplemental Fig. 7B-C),
286 similar to previous reports for *prl1-9* and *skip-1* mutants (Li et al., 2019).
287
288 Since *PRP19* co-purifies with *XCT* (Table 1) and they both control the pace of the circadian
289 clock (Supplemental Fig. 7A), we hypothesized they might function in the same pathway to
290 regulate the clock. To test this hypothesis, we introduced *prp19a-1* and *prp19b-1* mutant alleles,
291 which express greatly decreased and near-null levels of *PRP19A* and *PRP19B*, respectively
292 (Supplemental Fig. 8), into the *xct-2* background and assayed their circadian clock phenotypes.
293 Interestingly, we found that the circadian periods of both *prp19a-1 xct-2* and *prp19b-1 xct-2*
294 double mutants were not significantly different from Col-0 (Fig. 6A; Supplemental Fig. 9),
295 indicating that *prp19a-1* and *prp19b-1* both fully suppress the short period phenotype of *xct-2*.
296 Thus, our data suggest that functional *PRP19A* and *PRP19B* are both necessary for *XCT* in
297 regulation of circadian period.
298
299 To further investigate the functional association between *XCT* and *PRP19*, we next asked
300 whether *prp19* mutants suppress other developmental and physiological defects observed in *xct-*
301 *2*. In contrast to the circadian period, the reduced rosette size in *xct-2* was not significantly
302 restored in either *prp19a-1 xct-2* or *prp19b-1 xct-2* (Supplemental Fig. 10), suggesting that the
303 regulation of *XCT* on rosette development and the circadian clock period are genetically
304 separable. This is consistent with our observation that *xct-1* and *xct-2* have similar short
305 circadian period phenotypes but only *xct-2* is morphologically different from Col-0
306 (Supplemental Fig. 1).
307

308 Next, to determine whether suppression of the short clock period phenotype in *xct-2 prp19*
309 mutants is due to reversal of the *xct-2* splicing defects, we conducted qRT-PCR experiments to
310 examine the abundance of *xct*-induced aberrantly-spliced mRNA isoforms of core clock genes in
311 *prp19a-1 xct-2* and *prp19b-1 xct-2*. Surprisingly, none of the aberrantly-spliced isoforms tested
312 showed significantly different abundance in *prp19a-1 xct-2* and *prp19b-1 xct-2* relative to *xct-2*
313 (Fig. 6B-F). We also checked the expression levels of fully spliced or total mRNA isoforms of
314 the five core clock genes that were mis-spliced in *xct-2*. In all cases, addition of *prp19a-1* or
315 *prp19b-1* to *xct-2* did not alter the abundance of functionally spliced isoforms of the clock genes
316 (Supplemental Fig. 11). Thus, similar to the developmental defects, the splicing defects of core
317 clock genes are genetically separable from the circadian clock phenotype of *xct-2*.
318
319 **Fission yeast *XCT* ortholog *SpXAP5* accelerates the Arabidopsis circadian clock without**
320 **affecting splicing**
321
322 XAP5 proteins are highly conserved across eukaryotes (Martin-Tryon and Harmer, 2008).
323 Previous studies showed that transgenic expression of Arabidopsis *XCT* in the fission yeast
324 *Schizosaccharomyces pombe xap5* mutant can fully rescue its slow-growing phenotype (Anver et
325 al., 2014), indicating that there are shared molecular functions between these two homologous
326 genes. Therefore, to further investigate the molecular function of *XCT* that regulates the clock in
327 Arabidopsis, we examined whether a *SpXAP5* transgene could complement the short circadian
328 period phenotype of Arabidopsis *xct* mutants. To our surprise, we found that expression of
329 *SpXAP5* under the control of the native Arabidopsis *XCT* promoter not only failed to rescue the
330 short circadian period phenotype in *xct-1* mutant, but actually further accelerated the clock in
331 both Col-0 and *xct-1* backgrounds (Fig. 7A; Supplemental Fig. 12), suggesting a gain-of-function
332 effect of *SpXAP5* on controlling the pace of the Arabidopsis circadian clock.
333
334 To determine whether *SpXAP5* accelerates the Arabidopsis clock by exacerbating the splicing
335 defects of *xct* mutants, we performed PacBio Iso-Seq to examine changes in the transcriptome-
336 wide splicing pattern caused by expressing the *SpXAP5* transgene in the *xct-1* background.
337 Overall, only 3 out of the 90,114 detected splicing events were significantly, albeit only slightly,
338 enriched in *xct-1 SpXAP5* compared to *xct-1* (FDR<0.05, Fig. 7B). Furthermore, we conducted

339 qRT-PCR experiment to determine whether *SpXAP5* alters the splicing patterns of the core clock
340 genes targeted by *XCT*. Consistent with the Iso-Seq results (Fig. 5A-E), ratios of aberrantly
341 spliced isoforms to fully spliced isoforms in *xct-1* and *xct-2* were significantly different from
342 Col-0, whereas those ratios in *SpXAP5* and *xct-1 SpXAP5* showed no significant difference
343 compared to Col-0 and *xct-1*, respectively (Fig. 7C-F). Taken together, these results indicate that
344 the *SpXAP5* transgene shortens the Arabidopsis circadian clock period length in a gain-of-
345 function manner, but not by affecting splicing of the same set of core clock genes targeted by
346 *XCT*.
347
348 **Discussion**
349 The accurate removal of introns from pre-mRNAs is an essential step of gene expression in all
350 eukaryotes. In this work, we report that *XCT* is a global regulator of pre-mRNA splicing in
351 Arabidopsis. Multiple lines of evidence suggest that orthologs of XCT share conserved
352 molecular and cellular functions across kingdoms. Previous mass spectrometry data showed that
353 FAM50A and FAM50B, two human homologs of XCT, physically associate with affinity-
354 purified spliceosomal C complex (Bessonov et al., 2008; Bessonov et al., 2010). Loss of
355 *FAM50A* function induces transcriptome-wide pre-mRNA splicing defects in both human and
356 zebrafish (Lee et al., 2020). Similarly, a recent study in Arabidopsis also found that XCT
357 associates with spliceosomal proteins and regulates splicing (Liu et al., 2022). In this paper, we
358 show that XCT physically co-purifies with PRP19 and other NTC-associated proteins (Table 1).
359 Furthermore, our PacBio Iso-Seq data indicate that the efficiency and fidelity of pre-mRNA
360 splicing are negatively impacted in both the partial loss-of-function mutant *xct-1* and the null
361 mutant *xct-2* but are largely restored in *xct-2 XCT* (Fig. 2). Therefore, XCT orthologs likely play
362 evolutionarily conserved roles in pre-mRNA splicing. This is supported by our prior finding that
363 Arabidopsis *XCT* rescues the growth defects in yeast *xap5* mutant (Anver et al., 2014). However,
364 we now report that expression of fission yeast *SpXAP5* fails to rescue splicing defects in *xct-1*
365 mutant (Fig. 7B-F), indicating that any conserved biochemical functions between yeast SpXAP5
366 and XCT are not sufficient for appropriate splicing regulation in Arabidopsis. It would be
367 interesting to apply structural biology approaches to further investigate what conserved
368 functional domains among XCT homologs contribute to their splicing activities.
369

370 Studies have shown that intron retention is the most prevalent type of alternative splicing event
371 in Arabidopsis (Wang and Brendel, 2006; Filichkin et al., 2010). Correspondingly, mutations in
372 many Arabidopsis splicing factors mainly cause intron retention (Supplemental Fig. 3; Schlaen et
373 al., 2015; Li et al., 2019). Here we identify *XCT* as an unusual splicing regulator that specifically
374 controls the fidelity of 3' splice site selection (Fig. 2D-E; Fig. 3). Biochemical studies of splicing
375 in yeast revealed that several DEAH-box ATPases are important for the selection of 3' splice
376 sites (Horowitz, 2012). For example, PRP22, a DEAH-box ATPases that co-purifies with XCT
377 (Table 1; Liu et al., 2022), represses usage of aberrant 3' splice sites and promotes spliceosome
378 scanning for downstream alternative 3' splice sites in yeast (Mayas et al., 2006; Semlow et al.,
379 2016). Studies in human showed that hPRP22, FAM50A and FAM50B, are all abundant in the
380 spliceosomal C complex but nearly absent in B and B* complexes (Bessonov et al., 2008;
381 Bessonov et al., 2010; Zhan et al., 2022). This implies that PRP22 and XCT may have
382 evolutionarily conserved interactions during the second transesterification reaction of splicing,
383 when 5' and 3' splice sites are joined (Fig. 1). Intriguingly, we detected decreased fidelity of 3'
384 splice site selection in both *xct-1* and *xct-2* mutants (Fig. 3; Supplementary Fig. 4), which is also
385 observed in yeast *prp22* mutants (Semlow et al., 2016). Therefore, our results imply that XCT
386 may work with PRP22 to control the fidelity of 3' splice site selection. Future research on the
387 biochemical functions of XCT using *in vitro* and *in vivo* systems could help reveal how it
388 controls splicing fidelity.
389
390 Previous microarray and RNAseq studies demonstrated that transcription and splicing of a large
391 proportion of Arabidopsis genes are under circadian clock regulation (Covington and Harmer,
392 2007; Hsu and Harmer, 2012; Romanowski et al., 2020; Yang et al., 2020). Consequently, timing
393 of sample collection is an important consideration in gene expression and splicing analysis,
394 especially when comparing genotypes with different circadian periods. Here we employed a
395 more efficient sampling method by harvesting plants at twelve evenly distributed time points of a
396 day and pooling them before analysis. Our pooling strategy enabled us to detect transcripts of
397 genes with the distribution of peak phases of expression mirroring that of all clock-regulated
398 genes (Fig. 4B; Romanowski et al., 2020). Overall, this strategy allowed us to identify more
399 clock-regulated differentially spliced transcripts in *xct-2* mutants than in a recent study in which
400 plants were collected at a single time point (Liu et al., 2022). For example, we report novel

differential splicing events in afternoon-peaked genes including *PRR7* and *TOC1* (Fig. 4A; Supplemental Fig. 6C-D). In fact, our time-course qPCR data suggest that abundance of the aberrantly spliced isoforms of these clock-regulated genes fluctuates by over 99% depending on time of sample collection (Fig. 5F-I). Thus, our results demonstrate the advantages of pooling samples when conducting transcriptome-wide splicing analysis of clock-regulated genes.

Using this strategy, we were able to investigate whether *XCT* preferentially affects splicing at certain time of the day. Indeed, we found that dawn-expressed genes are significantly more likely to be mis-spliced in *xct* mutants than those with other peak phases (Fig. 4C-D; Supplemental Fig. 5). This suggests that *XCT* activity in or its association with the spliceosome may be clock-regulated. An alternative explanation for the overrepresented mis-splicing of dawn-expressed genes could be that *cis*-regulatory elements of those genes might preferentially recruit XCT to control their splicing. This possibility is supported by previous studies showing that XCT orthologs in fission yeast and *Chlamydomonas* are chromatin-associated and in the latter case recruit RNA polymerase II (Pol II) to promoter regions (Anver et al., 2014; Li et al., 2018).

Alternative splicing regulates various biological processes, including the function of the circadian clock (Hsu and Harmer, 2014; Nolte and Staiger, 2015). In Arabidopsis, such regulation is supported by identification of splicing factor mutants that alter circadian clock period length (Shakhmantsir and Sehgal, 2019). Some splicing factors, such as *GEMIN2* and *SICKLE*, interact epistatically with one or more alternatively spliced clock genes in the control of period length (Schlaen et al., 2015; Marshall et al., 2016), suggesting changes in the pace of the clock are due to altered splicing of these clock genes. Whereas in other cases, the mis-regulated circadian period can only be partially attributed to changes in levels of splicing variants of clock genes (Sanchez et al., 2010; Wang et al., 2012). Additionally, in most splicing factor mutants, only small fractions of total transcripts are aberrantly processed (Jones et al., 2012; Perez-Santángelo et al., 2014; Feke et al., 2019; Li et al., 2019). Therefore, whether these small changes in levels of aberrantly spliced isoforms could lead to significantly decreased levels of functional proteins and thereby cause the observed circadian phenotypes remains unclear.

431 Here, we demonstrate that in *xct* mutants the circadian clock phenotypes are genetically
432 separable from altered levels of aberrantly spliced mRNA variants of core clock gene. We found
433 that loss of *XCT* function accelerates the clock and causes aberrant splicing of five core clock
434 genes (Supplemental Fig. 1; Fig. 4A). Surprisingly, the clock phenotype but not the splicing
435 defects of the five clock genes in *xct-2* is suppressed by reduction of *PRP19A* or *PRP19B*
436 function (Fig. 6). Similarly, the clock in *xct-1* runs slightly faster than *xct-2* but the splicing
437 defects, including the aberrant splicing of *TOC1* and *TIC*, are more severe in the latter (Fig. 2;
438 Fig. 7D, F). Furthermore, we show that expression of *SpXAP5* in Col-0 or *xct-1* shortens the
439 circadian period without impacting splicing of *XCT*-targeted clock genes (Fig. 7). Collectively,
440 these data demonstrate that the short period phenotypes and altered levels of clock gene splicing
441 variants are genetically separable in *xct* mutants.
442
443 How *XCT* regulates the pace of the circadian clock is still an outstanding question. Although we
444 argue that changes in levels of aberrantly spliced clock mRNAs are not responsible for the
445 accelerated clock in *xct* mutants, one possibility is that alterations in the overall kinetics of pre-
446 mRNA splicing may cause circadian period phenotypes. Indeed, pharmacological perturbations
447 of global transcription and translation efficiency can both lengthen the circadian period (de Melo
448 et al., 2021; Uehara et al., 2022). Those results suggest that changes in the kinetics of RNA
449 processing might cause the period phenotypes observed in plants mutated for some splicing
450 factors.
451
452 Alternatively, *XCT* may control the circadian clock function independent of its role in pre-
453 mRNA splicing. In Chlamydomonas, the XCT homolog XAP5 co-immunoprecipitates with
454 RNA Pol II (Li et al., 2018), suggesting that XCT orthologs may possess transcriptional
455 regulatory activities. In this paper, we show that NTC components physically and genetically
456 interact with XCT to regulate circadian clock period (Table 1; Fig. 6A). Besides splicing, another
457 well-characterized role of the NTC in gene expression is regulation of transcriptional elongation
458 (Chanarat and Sträßer, 2013). Studies in both yeast and Arabidopsis have revealed that multiple
459 NTC members, including PRP19, PRL1 and CDC5, physically associate with RNA Pol II and
460 participate in transcription elongation (Chanarat et al., 2011; Zhang et al., 2013; Zhang et al.,
461 2014). In addition, the NTC component SKIP interacts with Polymerase-Associated Factor 1

462 complex to regulate transcription in a splicing-independent manner (Li et al., 2019).
463 Interestingly, a recent study showed that inhibition of transcriptional elongation by decreasing
464 phosphorylation of the RNA Pol II C-terminal domain lengthens circadian period in Arabidopsis
465 (Uehara et al., 2022). It is therefore possible that *XCT* affects circadian period by altering
466 transcriptional elongation.
467
468 Yet another possibility is that *XCT* affects a cellular process independently of its role in gene
469 expression. For example, PRP19 is known to facilitate DNA repair by its E3 ubiquitin ligase
470 activity independent of its involvement in the spliceosomal complex (Maréchal et al., 2014;
471 Idrissou and Maréchal, 2022). Indeed, there is increasing evidence that many RNA binding
472 proteins directly participate in DNA double-strand break responses (Klaric et al., 2021). A recent
473 study showed that the NTC-associated protein MOS4-ASSOCIATED COMPLEX SUBUNIT
474 5A (MAC5A) physically interacts with the 26S proteasome and regulates its activities in
475 response to DNA damage (Meng et al., 2022). Intriguingly, we recently reported that loss of *XCT*
476 function also disturbs the DNA damage response pathway (Kumimoto et al., 2021). Thus, it is
477 possible that *XCT* works with *PRP19* in this pathway separately from its role in RNA processing.
478 Future research is required to understand the nature of the relationship of between XCT and the
479 NTC in the control of circadian clock function and other essential biological processes.
480
481 **Materials and Methods**
482 **Plant materials and growth conditions**
483
484 All *Arabidopsis thaliana* plants used in this study are Columbia-0 (Col-0) ecotype and contain a
485 *CCR2*::*LUC* reporter for circadian clock assays. The *xct-1*, *xct-2*, *xct-2* p*XCT*::g*XCT-YFP-HA*,
486 *prp19a-1*, *prp19b-1*, *prl1-2*, *cdc5-1* and *skip-1* genotypes have been previously described
487 (Martin-Tryon and Harmer, 2008; Monaghan et al., 2009; Zhang et al., 2014; Wang et al., 2012).
488 All the double mutants in this study were produced by crossing. The fission yeast transgene was
489 generated by cloning *SpXAP5* cDNA sequence driven by an Arabidopsis *XCT* promoter into the
490 pCR8/GW/TOPO gateway entry vector (Invitrogen) and then transferring it into the promoter-
491 less pEarleyGate destination vector pEG301. The construct was transformed into *xct-1* mutant
492 plants using *Agrobacterium*-mediated floral dip method (Clough and Bent, 1998) and introduced

493 into the Col-0 background by crossing. Unless otherwise specified, seeds were surface sterilized
494 with chlorine gas for 3 hours (50 ml 100% bleach + 3ml 1M HCl) and then plated on 1x
495 Murashige and Skoog (MS) growth media containing 0.7% agar, pH 5.7. After 3d stratification
496 in dark at 4˚C, plates were transferred to 12-h light (cool white fluorescent bulbs, 55 µmol m^{-2} s^{-}
497 ^{1}) / 12-h dark cycles at 22˚C for a variable number of days depending on the experiment.
498
499 **Immunoprecipitations and mass spectrometry**
500
501 Plants were grown in 12-h light / 12-h dark cycles and harvested on day 10 at ZT3 or ZT17.
502 Approximately 7.5g of vegetative tissue was flash frozen in liquid nitrogen and ground into fine
503 powder with a mortar and pestle. Ground tissue was resuspended in nuclei enrichment buffer (50
504 mM Tris pH7.5, 400 mM sucrose, 2.5 mM $MgCl_2$, 1 mM DTT, 1 mM PMSF, cOmplete Protease
505 Inhibitor cocktail (Roche)), the nuclear pellet was collected, resuspended in IP buffer (100mM
506 Tris pH7.5, 1mM EDTA, 75mM NaCl, 10% glycerol, 0.3% Triton X-100, 0.05% SDS, 10 mM
507 MG-132, 1 mM PMSF, , cOmplete Protease Inhibitor cocktail (Roche)) and pelleted again.
508 Nuclei were resuspended in IP buffer, disrupted via sonication, and adjusted to 1 mg/mL in IP
509 buffer. 1 mg of extract was incubated with µMACS MicroBeads conjugated to an anti-HA
510 monoclonal antibody (Miltenyi Biotec) and beads were then captured on µMACS M-colums.
511 Beads were washed 3x with ice cold IP buffer and 1x with ice cold TE buffer (10 mM Tris-HCl
512 pH 8.0, 0.1 mM EDTA). Proteins were eluted off the beads with elution buffer (Miltenyi Biotec).
513 Proteins were resolved using SDS polyacrylamide gel electrophoresis and gels were sent to the
514 Rutgers Biological Mass Spectrometry Facility, Robert Wood Johnson Medical School. Proteins
515 were eluted and subjected to mass spectrometry as previously described (Wei et al., 2020).
516
517 **PacBio Iso-Seq and bioinformatic analysis**
518
519 Arabidopsis seeds were surface sterilized in 70% ethanol prepared in 0.1% (v/v) Triton X-100
520 (Sigma) for 5 minutes and then in 100% Ethanol for 20 minutes. After air drying, seeds were
521 plated on 1x MS growth media containing 0.7% agar. After cold stratification for 3 days, plates
522 were transferred to constant light (55 µmol m^{-2} s^{-1}) and temperature (22˚C). Approximately 60
523 mg of whole seedlings of each genotype were collected every 2h over 24 hours on day 11 and

524 pooled as indicated in Fig. 2A. Three biological replicates were collected for each genotype.
525 Samples were flash-frozen in liquid nitrogen, and ground into fine powder using a beadbeater.
526 Total RNA was extracted using an RNeasy Plant Mini Kit (Qiagen) followed by purification
527 using the RNA Clean & Concentrator Kit (Zymo Research). Purified RNAs were quantified
528 using a Nanodrop (ThermoFisher) and quality control was performed by an Agilent Bioanalyzer
529 2100. All samples had a 260/230 nm ratio higher than 1.9, 260/280nm ratio between 2 and 2.15,
530 and RIN score higher than 8. PacBio Sequel II library preparation and RNA-sequencing (Iso-
531 Seq) was performed by UC Davis DNA Technologies & Expression Analysis Core
532 (https://dnatech.genomecenter.ucdavis.edu/pacbio-library-prep-sequencing/).
533
534 Raw reads generated by the PacBio Sequel II sequencer were imported into PacBio SMRT Link
535 v 8.0 for Circular Consensus Sequence (CCS) calling and demultiplexing. Next, poly(A) tails
536 and concatemers were removed using the refine command from isoseq3 package via Bioconda (v
537 3.10) with the option '--require-polya'. Then the fastq files containing full-length non-
538 concatemer reads were mapped to Arabidopsis TAIR 10 genome assembly using minimap2 (v
539 2.17) with the parameter '-ax splice -t 30 -uf --secondary=no -C5'. For differential splicing
540 analysis, counts of exonic regions and known/novel splice junctions were generated using
541 QoRTs software package (v 1.3.6) (Hartley and Mullikin, 2015) with the parameter '--stranded --
542 singleEnded --stranded_fr_secondstrand --keepMultiMapped'. To adapt to the long-read data,
543 the 'maxPhredScore' was set to 93 and the 'maxReadLength' was adjusted manually for each
544 library. To increase the power of detecting novel splice junctions, the '--minCount' threshold was
545 set to 5. The raw count data was then loaded to JunctionSeq R package (v 1.5.4) (Hartley and
546 Mullikin, 2016) to determine differential usage of exons or splice junctions relative to the overall
547 expression of the corresponding gene with a false discovery rate (FDR) of 0.05. Genes with at
548 least one differentially spliced exon/junction were compared to the total circadian-clock-
549 regulated genes (Romanowski et al., 2020) as described in Fig. 4.
550
551 **RNA extraction and RT-PCR**
552
553 For time-course qRT-PCR analysis, Arabidopsis seedlings were entrained in 12-h light / 12-h
554 dark cycles for 9 days before transferred to constant light and temperature (55 µmol m^{-2} s^{-1},

555 22˚C). Starting from day 9, approximately 15-20 whole seedlings of each genotype were
556 harvested every 4h over the 72-h period. For single time point gene expression and splicing
557 analysis, 10-d-old seedlings grown in 12-h light / 12-h dark cycles at 22˚C were harvested at
558 estimated peak time of expression for each studied gene. Collected tissue was immediately
559 frozen in liquid nitrogen and ground into fine powder using a beadbeater. Total RNA was
560 isolated using TRIzol reagent (Invitrogen) and quantified with a Nanodrop (ThermoFisher). 200
561 ng of total RNA was used for cDNA synthesis with an oligo(dT)18 primer and SuperScriptIII
562 (Invitrogen) reverse transcriptase. Diluted cDNAs were used as templates for qRT-PCR and
563 semi-qRT-PCR reactions. The qRT-PCR was carried out as previously described (Martin-Tryon
564 et al., 2007) using a Bio-Rad CFX96 thermocycler. Primers amplifying the aberrantly spliced or
565 total transcripts were tested by standard curve and melt curve assays. Relative expression (ΔΔCq)
566 values were normalized to the geometric mean of *PROTEIN PHOSPHATASE 2a* (*PP2A*) and
567 *ISOPENTENYL DIPHOSPHATE ISOMERASE 2* (*IPP2*) expression levels. For semi-qRT-PCR,
568 splice-junction specific primer pairs were designed to amplify regions flanking the aberrantly
569 spliced introns of interest. The size and abundance of the resultant PCR products were then
570 analyzed and quantified by LabChip GX bioanalyzer (PerkinElmer). The expression analyses
571 represent two to three biological replicates. All primers used in this study are described in
572 Supplemental Dataset 5.
573
574 **Circadian period analysis**
575
576 After growing on MS plates under 12-h light / 12-h dark cycles for 6 days, Arabidopsis seedlings
577 were sprayed with 3 mM D-luciferin (Biosynth) prepared in 0.01% (v/v) Triton X-100 (Sigma)
578 and then transferred to a growth chamber with a constant 22˚C and constant light provided by red
579 and blue LED SnapLites (Quantum Devices, 35 µmol m^{-2} s^{-1} each) for imaging. Luciferase
580 activity was detected using a cooled CCD camera (DU434-BV [Andor Technology]).
581 Bioluminescence signals from the images were quantified using MetaMorph 7.7.1.0 software
582 (Molecular Devices). Subsequent circadian period estimation and rhythmicity analysis were
583 performed using Biological Rhythms Analysis Software System 3.0 (BRASS) by fitting the
584 bioluminescence data to a cosine wave through Fourier Fast Transform-Non-Linear Least
585 Squares (Plautz et al., 1997).

586

587 **Rosette size measurement**

588

589 Seeds on MS plates were germinated in a growth chamber with 12-h light / 12-h dark cycles at a
590 constant 22˚C. 15-day-old seedlings were transferred to soil and grown in 16-h light (cool-white
591 fluorescent bulbs, 75 µmol m^{-2} s^{-1}) / 8-h dark long day condition at a constant 22˚C. Rosette sizes
592 of 39-day-old plants were determined by the greatest distance between rosette leaves using a
593 digital caliper.

594

595 **Accession Numbers**

596

597 All the *Arabidopsis thaliana* and *Schizosaccharomyces pombe* genes studied in this paper can be
598 found under the following accession numbers:
599 *XCT*, AT2G21150
600 *PRP19A*, AT1G04510
601 *PRP19B*, AT2G33340
602 *PRL1*, AT4G15900
603 *SKIP*, AT1G77180
604 *CDC5*, AT1G09770
605 *LHY*, AT1G01060
606 *LNK2*, AT3G54500
607 *TIC*, AT3G22380
608 *PRR7*, AT5G02810
609 *TOC1*, AT5G61380
610 *SPPA*, AT1G73990
611 *IPP2*, AT3G02780
612 *PP2A*, AT1G69960
613 *SpXAP5*, NP_587947

614

615 **Data Availability**

616

All our MS data were deposited and available at Center for Computational Mass Spectrometry (CCMS, Dataset: MSV000090830). Our PacBio Iso-Seq data were deposited and available in NCBI Gene Expression Omnibus (GEO, Accession number: GSE220902).

Fundings

This work was supported by awards from the National Institutes of Health (R01 GM069418 to SLH) and the United States Department of Agriculture-National Institute of Food and Agriculture (CA-D-PLB-2259-H to SLH). HZ was supported by a fellowship from China Scholarship Council (CSC #201806010204 to HZ).

Acknowledgments

We thank the Arabidopsis Biological Resource Center for providing seeds. We thank Haiyan Zheng and the Biological Mass Spectrometry Facility of Robert Wood Johnson Medical School and Rutgers for mass spectrometry analysis. They are supported by NIH Shared Instrumentation Grant S10OD01640. We thank the DNA Technologies and Expression Analysis Core at the UC Davis Genome Center for performing PacBio Iso-Seq experiment. They are supported by NIH Shared Instrumentation Grant 1S10OD010786-01. We thank Julin Maloof for statistical advice, as well as members of the Harmer, Maloof, and Shabek labs for helpful discussions.

Author Contributions

HZ, RWK and SLH designed the research; HZ, RWK, and SA performed the research; HZ, RWK, and SLH analyzed data; HZ and SLH wrote the paper.

Figure Legends

Figure 1. Simplified overview of pre-mRNA splicing reactions. A schematic diagram highlighting the two catalytic transesterification steps and the association between the NTC and the spliceosomal complex during pre-mRNA splicing. Gray boxes and solid black lines represent

648 exons and introns, respectively. U1-U6 small nuclear ribonucleoproteins (snRNPs) are indicated
649 by yellow circles. The NTC is indicated by a green oval.
650
651 **Table 1. XCT physically associates with splicing-related proteins, especially the NTC**
652 **components, in Arabidopsis.** Mass Spectrometry (MS) data showing that spliceosome-
653 associated proteins are significantly more enriched in XCT-YFP-HA-IP than control IP (fold
654 enrichment > 3; *P*-value < 0.05, Welch two sample t.test). The full list of proteins co-purified
655 with XCT are described in Supplemental Dataset 1. Only proteins detected in each biological
656 replicate and with 20 or more total peptides counts are shown.
657
658 **Figure 2. Transcriptome-wide analysis reveals *XCT* as a global pre-mRNA splicing**
659 **regulator.** A, Experimental design and sampling method for the PacBio Iso-Seq experiment.
660 White and grey boxes represent subjective day and subjective night, respectively. Arabidopsis
661 seedlings were grown at constant light and temperature for 10 days before being harvested and
662 pooled. B and C, Numbers of differentially enriched splicing events represented by significantly
663 differentially down- (B) or up-regulated (C) splice junctions in *xct-1*, *xct-2* and *xct-2*
664 complemented with p*XCT*::g*XCT-YFP-HA* compared with Col-0 (false discovery rate < 0.05). D
665 and E, Frequency of different classes of 5' and 3' splice sites among down- (D) or up-regulated
666 (E) splice junctions in *xct-1* and *xct-2*. Known or novel splice sites were classified by comparing
667 to TAIR10 genome annotation.
668
669 **Figure 3. *XCT* is required for the fidelity of 3' splice site selection during pre-mRNA**
670 **splicing.** A, A schematic diagram showing the structure of a typical U2-type splice junction. Y:
671 pyrimidine. B, Distribution of the distance between each pair of novel 3' splice site and its
672 corresponding canonical 3' splice sites in *xct-2*. C, Pictograms showing the frequency of
673 nucleotides in the 23-mers sequences flanking the 3' splice sites in all expressed, down-regulated
674 and up-regulated splice junctions in *xct-2*. D, Maximum Entropy Model scores showing the
675 strength of 3' splice sites of all expressed, down-regulated and up-regulated junctions in *xct-2*. E,
676 Percentage of pyrimidines at the -3 position (i.e. the nucleotide preceding the AG at 3' splice
677 site) in *xct-2*. F, Counts of pyrimidines in the Y10 polypyrimidine tract upstream of the 3' splice
678 sites in *xct-2*. PPT, polypyrimidine tract. The lines in the boxplot represent the 75% quartile,

679 median and 25% quartile of the data, respectively. Statistical significance in (D) and (F) was
680 determined using linear regression model with junction class as a fixed effect and is shown in
681 lower case letters (Tukey's multiple comparison test, $P < 0.05$). Statistical significance in (E) was
682 determined by Fisher's exact test: **: $P < 0.01$; ***: $P < 0.001$.
683
684 **Figure 4. Dawn-phased genes are significantly enriched among all the circadian-clock-**
685 **regulated genes that are aberrantly spliced in *xct-2.*** A, Venn diagram showing the overlaps of
686 aberrantly spliced genes in *xct-1, xct-2*, and total circadian-clock-regulated genes (Romanowski
687 et al., 2020). The core circadian clock genes that are mis-spliced in both *xct* mutants are indicated
688 in red (non-clock-regulated) and blue (clock-regulated) fonts. Only genes considered as
689 'detected' in all three RNA-Seq datasets are included. B and C, Phases of estimated peak
690 expression of all circadian clock-regulated genes (Romanowski et al., 2020) that are detected (B)
691 or significantly aberrantly spliced (C) in *xct-2*. The white and gray backgrounds represent the
692 subjective day and subjective night, respectively. D, Distribution of *P*-values from Fisher's extact
693 tests calculating whether the ratio of the number of aberrantly spliced genes in (B) to total clock-
694 regulated genes in (A) is significantly higher than expected by chance in each 2-hour interval.
695
696 **Figure 5. Time-course qRT-PCR experiments validate the role of *XCT* in regulating pre-**
697 **mRNA splicing of clock-regulated and non-clock-regulated genes.** A - E, Sashimi plots
698 showing PacBio Iso-Seq reads mapped to *LHY* (A)*, LNK2* (B)*, PRR7* (C)*, TOC1* (D) and *SPPA*
699 (E) in Col-0 (teal)*, xct-2* (orange) and *xct-2 XCT* (purple). The red rectangles highlight the
700 aberrantly spliced exon-exon junctions that are examined by qRT-PCR in (F) - (J). F - J,
701 Normalized expression of the aberrantly spliced isoforms of *LHY* (F)*, LNK2* (G)*, PRR7* (H)*,*
702 *TOC1* (I) and *SPPA* (J) in Col-0*, xct-2* and *xct-2 XCT*. Samples were collected every four hours
703 over a 72-h window. Expression levels were examined by qRT-PCR using splice-junction-
704 specific primers and normalized to *PP2*A and *IPP2.* Data points represent mean ± se from three
705 independent biological replicates. For each isoform in each biological replicate, the normalized
706 expression levels were relative to the highest expression levels of their total transcripts in Col-0
707 across all time points. Teal lines, wild type Col-0; orange lines, *xct-2* muants; purple lines, *xct-2*
708 *XCT*. Black background, dark period; white and gray background, light period during subjective
709 day and night, respectively.

710

711 **Figure 6. Loss of *PRP19* function suppresses the short circadian clock period phenotype**
712 **but not the splicing defects of core clock genes in *xct-2*.** A, Circadian periods of Col-0, *xct-2*,
713 *prp19a-1*, *prp19b-1*, *prp19a-1 prp19b-1*, *prp19a-1 xct-2* and *prp19b-1 xct-2* plants. B - F,
714 Normalized expression of the aberrantly spliced isoforms of *LHY, LNK2, TIC, PRR7* and *TOC1*
715 in Col-0, *xct-2*, *prp19a-1*, *prp19b-1*, *prp19a-1 prp19b-1*, *prp19a-1 xct-2* and *prp19b-1 xct-2*.
716 Samples were collected at the estimated peak expression time for each gene. Expression levels
717 were examined by qRT-PCR using splice-junction-specific primers and normalized to *PP2*A and
718 *IPP2.* Data points represent mean ± se from two independent biological replicates. Statistical
719 significance was determined using linear regression model with genotype as a fixed effect and is
720 shown in lower case letters (Tukey's multiple comparison test, $P < 0.05$).
721

722 **Figure 7. Expression of *S.pombe XAP5* in Arabidopsis causes a short circadian clock period**
723 **phenotype without inducing pre-mRNA splicing defects.** A, Circadian period of Col-0,
724 *SpXAP5*, *xct-1* and *xct-1 SpXAP5* plants. B, A scatter plot showing differential splicing events in
725 *xct-1 SpXAP5* compared with *xct-1* indicated by normalized PacBio Iso-Seq read counts of exons
726 and intron elements. Differentially spliced elements that are statistically significant (FDR < 0.05)
727 and insignificant (FDR > 0.05) are indicated with red and grey dots, respectively. The dashed
728 line represents y=x. C - F, Relative abundance of aberrantly-spliced to fully-spliced isoforms of
729 *LNK2*, *TIC*, *PRR7* and *TOC1* in Col-0, *SpXAP5, xct-1, xct-1 SpXAP5* and *xct-2*. Differentially
730 spliced isoforms were amplified by semi-qRT-PCR using primers flanking the examined regions,
731 followed by separation and quantification using a Lab Chip GX bioanalyzer. Statistical
732 significance in (A) and (C) - (F) was determined using linear regression models with genotype as
733 a fixed effect and is shown in lower case letters (Tukey's multiple comparison test, $P < 0.05$).
734

735 **Supplemental Materials**

736

737 **Supplemental Figure 1. *XCT* regulates plant development and the pace of the circadian**
738 **clock in *Arabidopsis thaliana.*** A, Morphology of 52-day-old Col-0, *xct-1*, *xct-2* mutant plants
739 and *xct-2* complemented with p*XCT*::g*XCT-YFP-HA.* Bar = 10 mm. B, Circadian periods of
740 genotypes described in (A). The lines in the boxplot represent the 75% quartile, median and 25%

741 quartile of the data, respectively. Statistical significance was determined using linear regression
742 model with genotype as a fixed effect and is shown in lower case letters (Tukey's multiple
743 comparison test, $P < 0.05$).
744
745 **Supplemental Figure 2. *XCT* and *PRL1* control the pace of the circadian clock in an**
746 **opposite manner.** A, Morphology of 21-day-old Col-0, *xct-1*, *xct-2* and *prl1-2* mutant
747 Arabidopsis seedlings. Bar = 1 mm. B, Circadian clock periods of plant genotypes described in
748 (A). The lines in the boxplot represent the 75% quartile, median and 25% quartile of the data,
749 respectively. Statistical significance was determined using linear regression model with genotype
750 as a fixed effect and is shown in lower case letters (Tukey's multiple comparison test, $P < 0.05$).
751
752 **Supplemental Figure 3. *prl1-2* mutant mainly induces intron retention defects in**
753 **Arabidopsis.** Frequency of different classes of 5' and 3' splice sites of down- (teal) or up-
754 regulated (orange) junctions in *prl1-2* compared to wild type. Known or novel splice sites were
755 classified based on TAIR10 genome annotation.
756
757 **Supplemental Figure 4. *xct-1* mutant displays reduced fidelity of 3' splice site selection**
758 **during pre-mRNA splicing.** A, Distribution of the distances between each mis-selected novel 3'
759 splice site and its corresponding canonical 3' splice site in *xct-1*. B - C, Pictograms showing the
760 frequency of nucleotides in the 23-mers sequences flanking the 3' splice sites in splice junctions
761 that are down-regulated or up-regulated in *xct-1*.
762
763 **Supplemental Figure 5. Dawn-phased genes are significantly enriched among genes mis-**
764 **spliced in *xct-1*.** A - B, Phases of peak expression of all the circadian-clock-regulated genes
765 (Romanowski et al., 2020) that are detected (A) or aberrantly spliced (B) in *xct-1*. The white and
766 gray backgrounds represent the subjective day and subjective night, respectively. C, Distribution
767 of *P*-values from Fisher's exact test calculating whether the ratio of the number of mis-spliced
768 genes in (B) to all clock-regulated genes in (A) is significantly higher than expected by chance
769 for each 2-hour interval.
770

771 **Supplemental Figure 6. Loss of *XCT* function does not affect overall abundance of**
772 **transcripts of core-circadian-clock genes that are aberrantly spliced in *xct* mutants.** A - F,
773 Normalized expression of the total transcript levels of *LHY* (A)*, LNK2* (B), *PRR7* (C) and *TOC1*
774 (D) in Col-0*, xct-2* and *xct-2 XCT*. Samples were collected every four hours over a 72-h window.
775 Expression levels were examined by qRT-PCR using primers that detect abundance of all
776 transcripts and normalized to *PP2*A and *IPP2.* Data points represent mean ± se from three
777 independent biological replicates. For each isoform in each biological replicate, the normalized
778 expression levels were relative to the highest expression level in Col-0 across all time points.
779 Teal lines, wild type Col-0; orange lines, *xct-2* muants; purple lines, *xct-2* complemented with
780 p*XCT*::g*XCT-YFP-HA*. Black background, dark period; white and gray background, light period
781 during subjective day and night, respectively.
782
783 **Supplemental Figure 7. Mutation of Arabidopsis NTC members lengthens circadian**
784 **period.** A - C, Circadian periods of Col-0 and NTC mutant plants, including (A) *prp19a-1*,
785 *prp19b-1*, *prp19a-1 prp19b-1*, (B) *cdc5-1*, and (C) *skip-1*. The lines in the boxplot represent the
786 75% quartile, median and 25% quartile of the data, respectively. Statistical significance was
787 determined using linear regression model with genotype as a fixed effect and is shown in lower
788 case letters (Tukey's multiple comparison test, $P < 0.05$).
789
790 **Supplemental Figure 8. *PRP19A* (*MAC3A*) and *PRP19B* (*MAC3B*) transcript levels are**
791 **highly reduced in *prp19a-1* and *prp19b-1* mutant backgrounds, respectively.** A - B,
792 Normalized expression of (A) *PRP19A* and (B) *PRP19B* in Col-0, *xct-2*, *prp19a-1*, *prp19b-1*,
793 *prp19a-1 prp19b-1*, *prp19a-1 xct-2* and *prp19b-1 xct-2*. Expression levels were examined by
794 qRT-PCR and normalized to *PP2*A and *IPP2.* Data points represent mean ± se from two
795 independent biological replicates. Samples were collected at ZT0. Statistical significance was
796 determined using linear regression model with genotype as a fixed effect and is shown in lower
797 case letters (Tukey's multiple comparison test, $P < 0.05$).
798
799 **Supplemental Figure 9. Loss of *PRP19* function suppresses the *xct-2* short circadian period**
800 **phenotype.** An independent replicate of the experiment in Figure 6A showing the circadian

801 periods of Col-0, *xct-2*, *prp19a-1*, *prp19b-1*, *prp19a-1 prp19b-1*, *prp19a-1 xct-2* and *prp19b-1*
802 *xct-2* plants.
803
804 **Supplemental Figure 10. *PRP19* and *XCT* additively regulate Arabidopsis rosette size.** A,
805 Morphology of 39-day-old Col-0, *xct-2*, *prp19a-1*, *prp19b-1*, *prp19a-1 prp19b-1*, *prp19b-1 xct-2*
806 and *prp19a-1 xct-2* plants. Bar = 10 mm. B, Rosette diameter of 36-day-old plants grown under
807 long-day conditions at 22˚C (n = 5-11). Data are shown as mean ± se. Statistical significance was
808 determined using linear regression model with genotype as a fixed effect and is shown in lower
809 case letters (Tukey's multiple comparison test, $P < 0.05$).
810
811 **Supplemental Figure 11. Levels of total transcripts of *XCT*-targeted core clock genes are**
812 **not altered in *prp19a-1 xct-2* or *prp19b-1 xct-2* compared to *xct-2.*** A - F, Normalized
813 expression of total transcript levels (A, C, E, F) or specific normally spliced isoforms (B, D) of
814 *LHY, LNK2, TIC, PRR7* and *TOC1* in Col-0, *xct-2*, *prp19a-1*, *prp19b-1*, *prp19a-1 prp19b-1*,
815 *prp19a-1 xct-2* and *prp19b-1 xct-2*. Expression levels were examined by qRT-PCR using primers
816 that detect abundance of all transcripts or normally spliced isoforms and normalized to *PP2*A and
817 *IPP2.* Data points represent mean ± se from two independent biological replicates. Samples were
818 collected at the estimated peak expression time for each gene. Statistical significance was
819 determined using linear regression model with genotype as a fixed effect and is shown in lower
820 case letters (Tukey's multiple comparison test, $P < 0.05$).
821
822 **Supplemental Figure 12. Arabidopsis plants expressing *SpXAP5* don't mimic the**
823 **developmental defects caused by *xct-2* mutant.** Morphology of 50–day-old Col-0, *SpXAP5,*
824 *xct-1*, *xct-1 SpXAP5* and *xct-2*. Bar = 10 mm.
825
826 **Supplemental Dataset 1. A list of proteins co-purified with XCT-IP detected by MS.**
827
828 **Supplemental Dataset 2. Summary of full-length reads obtained from each PacBio Iso-Seq**
829 **library.**
830

831 **Supplemental Dataset 3. Lists of significant differential splicing events identified by**
832 **JunctionSeq.**
833
834 **Supplemental Dataset 4. Lists of total and aberrantly spliced circadian-clock-regulated**
835 **genes described in this study.**
836
837 **Supplemental Dataset 5. A list of primers used in this study.**
838
839 **References**
840

841 **Anver S, Roguev A, Zofall M, Krogan NJ, Grewal SIS, Harmer SL** (2014) Yeast X-
842 chromosome-associated protein 5 (Xap5) functions with H2A.Z to suppress aberrant
843 transcripts. EMBO Rep **15**: 894–902

844 **Bessonov S, Anokhina M, Krasauskas A, Golas MM, Sander B, Will CL, Urlaub H, Stark**
845 **H, Lührmann R** (2010) Characterization of purified human Bact spliceosomal
846 complexes reveals compositional and morphological changes during spliceosome
847 activation and first step catalysis. RNA **16**: 2384–2403

848 **Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R** (2008) Isolation of an active step
849 I spliceosome and composition of its RNP core. Nature **452**: 846–850

850 **Chanarat S, Seizl M, Strässer K** (2011) The Prp19 complex is a novel transcription elongation
851 factor required for TREX occupancy at transcribed genes. Genes Dev **25**: 1147–1158

852 **Chanarat S, Sträßer K** (2013) Splicing and beyond: the many faces of the Prp19 complex.
853 Biochim Biophys Acta **1833**: 2126–2134

854 **Clough SJ, Bent AF** (1998) Floral dip: a simplified method for Agrobacterium-mediated
855 transformation of Arabidopsis thaliana. Plant J **16**: 735–743

856 **Covington MF, Harmer SL** (2007) The circadian clock regulates auxin signaling and responses
857 in Arabidopsis. PLoS Biol **5**: e222

858 **Creux N, Harmer S** (2019) Circadian Rhythms in Plants. Cold Spring Harb Perspect Biol **11**:
859 a034611

860 **Deng X, Lu T, Wang L, Gu L, Sun J, Kong X, Liu C, Cao X** (2016) Recruitment of the
861 NineTeen Complex to the activated spliceosome requires AtPRMT5. Proc Natl Acad Sci
862 U S A **113**: 5447–5452

863 **Fang X, Shi Y, Lu X, Chen Z, Qi Y** (2015) CMA33/XCT Regulates Small RNA Production
864 through Modulating the Transcription of Dicer-Like Genes in Arabidopsis. Mol Plant **8**:
865 1227–1236

866 **Feke A, Liu W, Hong J, Li M-W, Lee C-M, Zhou EK, Gendron JM** (2019) Decoys provide a
867 scalable platform for the identification of plant E3 ubiquitin ligases that regulate
868 circadian function. Elife **8**: e44558

869 **Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, Wong W-K, Mockler TC**
870 (2010) Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome
871 Res **20**: 45–58

872 **Harmer SL** (2009) The circadian system in higher plants. Annu Rev Plant Biol **60**: 357–377

873 **Hartley SW, Mullikin JC** (2016) Detection and visualization of differential splicing in RNA-
874 Seq data with JunctionSeq. Nucleic Acids Res **44**: e127

875 **Hartley SW, Mullikin JC** (2015) QoRTs: a comprehensive toolset for quality control and data
876 processing of RNA-Seq experiments. BMC Bioinformatics **16**: 224

877 **Hatakeyama S, Yada M, Matsumoto M, Ishida N, Nakayama KI** (2001) U box proteins as a
878 new family of ubiquitin-protein ligases. J Biol Chem **276**: 33111–33120

879 **Hogg R, McGrail JC, O'Keefe RT** (2010) The function of the NineTeen Complex (NTC) in
880 regulating spliceosome conformations and fidelity during pre-mRNA splicing. Biochem
881 Soc Trans **38**: 1110–1115

882 **Horowitz DS** (2012) The mechanism of the second step of pre-mRNA splicing. Wiley
883 Interdiscip Rev RNA **3**: 331–350

884 **Hsu PY, Harmer SL** (2012) Circadian phase has profound effects on differential expression
885 analysis. PLoS One **7**: e49853

886 **Hsu PY, Harmer SL** (2014) Wheels within wheels: the plant circadian system. Trends Plant Sci
887 **19**: 240–249

888 **Idrissou M, Maréchal A** (2022) The PRP19 Ubiquitin Ligase, Standing at the Cross-Roads of
889 mRNA Processing and Genome Stability. Cancers (Basel) **14**: 878

890 **James AB, Syed NH, Bordage S, Marshall J, Nimmo GA, Jenkins GI, Herzyk P, Brown**
891 **JWS, Nimmo HG** (2012) Alternative splicing mediates responses of the Arabidopsis
892 circadian clock to temperature changes. Plant Cell **24**: 961–981

893 **Jia T, Zhang B, You C, Zhang Y, Zeng L, Li S, Johnson KCM, Yu B, Li X, Chen X** (2017)
894 The Arabidopsis MOS4-Associated Complex Promotes MicroRNA Biogenesis and
895 Precursor Messenger RNA Splicing. Plant Cell **29**: 2626–2643

896 **Jones MA, Williams BA, McNicol J, Simpson CG, Brown JWS, Harmer SL** (2012) Mutation
897 of Arabidopsis spliceosomal timekeeper locus1 causes circadian clock defects. Plant Cell
898 **24**: 4066–4082

899 **Klaric JA, Wüst S, Panier S** (2021) New Faces of old Friends: Emerging new Roles of RNA-
900 Binding Proteins in the DNA Double-Strand Break Response. Front Mol Biosci **8**:
901 668821

902 **Koncz C, Dejong F, Villacorta N, Szakonyi D, Koncz Z** (2012) The spliceosome-activating
903 complex: molecular mechanisms underlying the function of a pleiotropic regulator. Front
904 Plant Sci **3**: 9

905 **Kumimoto RW, Ellison CT, Toruño TY, Bak A, Zhang H, Casteel CL, Coaker G, Harmer**
906 **SL** (2021) XAP5 CIRCADIAN TIMEKEEPER Affects Both DNA Damage Responses
907 and Immune Signaling in Arabidopsis. Front Plant Sci **12**: 707923

908 **Kwon Y-J, Park M-J, Kim S-G, Baldwin IT, Park C-M** (2014) Alternative splicing and
909 nonsense-mediated decay of circadian clock genes under environmental stress conditions
910 in Arabidopsis. BMC Plant Biol **14**: 136

911 **Lee Y-R, Khan K, Armfield-Uhas K, Srikanth S, Thompson NA, Pardo M, Yu L, Norris**
912 **JW, Peng Y, Gripp KW, et al** (2020) Mutations in FAM50A suggest that Armfield
913 XLID syndrome is a spliceosomopathy. Nat Commun **11**: 3698

914 **Li L, Tian G, Peng H, Meng D, Wang L, Hu X, Tian C, He M, Zhou J, Chen L, et al** (2018)
915 New class of transcription factors controls flagellar assembly by recruiting RNA
916 polymerase II in Chlamydomonas. Proc Natl Acad Sci U S A **115**: 4435–4440

917 **Li Y, Yang J, Shang X, Lv W, Xia C, Wang C, Feng J, Cao Y, He H, Li L, et al** (2019) SKIP
918 regulates environmental fitness and floral transition by forming two distinct complexes in
919 Arabidopsis. New Phytol **224**: 321–335

920 **Liu L, Li X, Yuan L, Zhang G, Gao H, Xu X, Zhao H** (2022) XAP5 CIRCADIAN
921 TIMEKEEPER specifically modulates 3' splice site recognition and is important for
922 circadian clock regulation partly by alternative splicing of LHY and TIC. Plant Physiol
923 Biochem **172**: 151–157

924 **Maréchal A, Li J-M, Ji XY, Wu C-S, Yazinski SA, Nguyen HD, Liu S, Jiménez AE, Jin J,**
925 **Zou L** (2014) PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and
926 drives ATR activation via a ubiquitin-mediated circuitry. Mol Cell **53**: 235–246

927 **Marshall CM, Tartaglio V, Duarte M, Harmon FG** (2016) The Arabidopsis sickle Mutant
928 Exhibits Altered Circadian Clock Responses to Cool Temperatures and Temperature-
929 Dependent Alternative Splicing. Plant Cell **28**: 2560–2575

930 **Martin-Tryon EL, Harmer SL** (2008) XAP5 CIRCADIAN TIMEKEEPER coordinates light
931 signals for proper timing of photomorphogenesis and the circadian clock in Arabidopsis.
932 Plant Cell **20**: 1244–1259

933 **Martin-Tryon EL, Kreps JA, Harmer SL** (2007) GIGANTEA acts in blue light signaling and
934 has biochemically separable roles in circadian clock and flowering time regulation. Plant
935 Physiol **143**: 473–486

936 **Mayas RM, Maita H, Staley JP** (2006) Exon ligation is proofread by the DExD/H-box ATPase
937 Prp22p. Nat Struct Mol Biol **13**: 482–490

938 **de Melo JRF, Gutsch A, Caluwé TD, Leloup J-C, Gonze D, Hermans C, Webb AAR,**
939 **Verbruggen N** (2021) Magnesium maintains the length of the circadian period in
940 Arabidopsis. Plant Physiol **185**: 519–532

941 **Meng X, Wang Q, Hao R, Li X, Li M, Hu R, Du H, Hu Z, Yu B, Li S** (2022) RNA-binding
942 protein MAC5A interacts with the 26S proteasome to regulate DNA damage response in
943 Arabidopsis. Plant Physiol kiac510

944 **Monaghan J, Xu F, Gao M, Zhao Q, Palma K, Long C, Chen S, Zhang Y, Li X** (2009) Two
945 Prp19-like U-box proteins in the MOS4-associated complex play redundant roles in plant
946 innate immunity. PLoS Pathog **5**: e1000526

947 **Nohales MA, Kay SA** (2016) Molecular mechanisms at the core of the plant circadian oscillator.
948 Nat Struct Mol Biol **23**: 1061–1069

949 **Nolte C, Staiger D** (2015) RNA around the clock - regulation at the RNA level in biological
950 timing. Front Plant Sci **6**: 311

951 **Perez-Santángelo S, Mancini E, Francey LJ, Schlaen RG, Chernomoretz A, Hogenesch JB,**
952 **Yanovsky MJ** (2014) Role for LSM genes in the regulation of circadian rhythms. Proc
953 Natl Acad Sci U S A **111**: 15166–15171

954 **Plautz JD, Straume M, Stanewsky R, Jamison CF, Brandes C, Dowse HB, Hall JC, Kay SA**
955 (1997) Quantitative analysis of Drosophila period gene transcription in living animals. J
956 Biol Rhythms **12**: 204–217

957 **Romanowski A, Schlaen RG, Perez-Santangelo S, Mancini E, Yanovsky MJ** (2020) Global
958 transcriptome analysis reveals circadian control of splicing events in Arabidopsis
959 thaliana. Plant J **103**: 889–902

960 **Sanchez SE, Petrillo E, Beckwith EJ, Zhang X, Rugnone ML, Hernando CE, Cuevas JC,**
961 **Godoy Herz MA, Depetris-Chauvin A, Simpson CG, et al** (2010) A methyl transferase
962 links the circadian clock to the regulation of alternative splicing. Nature **468**: 112–116

963 **Schlaen RG, Mancini E, Sanchez SE, Perez-Santángelo S, Rugnone ML, Simpson CG,**
964 **Brown JWS, Zhang X, Chernomoretz A, Yanovsky MJ** (2015) The spliceosome
965 assembly factor GEMIN2 attenuates the effects of temperature on alternative splicing and
966 circadian rhythms. Proc Natl Acad Sci U S A **112**: 9382–9387

967 **Semlow DR, Blanco MR, Walter NG, Staley JP** (2016) Spliceosomal DEAH-Box ATPases
968 Remodel Pre-mRNA to Activate Alternative Splice Sites. Cell **164**: 985–998

969 **Seo PJ, Park M-J, Lim M-H, Kim S-G, Lee M, Baldwin IT, Park C-M** (2012) A self-
970 regulatory circuit of CIRCADIAN CLOCK-ASSOCIATED1 underlies the circadian
971 clock regulation of temperature responses in Arabidopsis. Plant Cell **24**: 2427–2442

972 **Shakhmantsir I, Sehgal A** (2019) Splicing the Clock to Maintain and Entrain Circadian
973 Rhythms. J Biol Rhythms **34**: 584–595

974 **Shi Y** (2017) Mechanistic insights into precursor messenger RNA splicing by the spliceosome.
975 Nat Rev Mol Cell Biol **18**: 655–670

976 **Uehara TN, Nonoyama T, Taki K, Kuwata K, Sato A, Fujimoto KJ, Hirota T, Matsuo H,**
977 **Maeda AE, Ono A, et al** (2022) Phosphorylation of RNA Polymerase II by CDKC;2
978 Maintains the Arabidopsis Circadian Clock Period. Plant Cell Physiol **63**: 450–462

979 **Wang B-B, Brendel V** (2006) Genomewide comparative analysis of alternative splicing in
980 plants. Proc Natl Acad Sci U S A **103**: 7175–7180

981 **Wang X, Wu F, Xie Q, Wang H, Wang Y, Yue Y, Gahura O, Ma S, Liu L, Cao Y, et al**
982 (2012) SKIP is a component of the spliceosome linking alternative splicing and the
983 circadian clock in Arabidopsis. Plant Cell **24**: 3278–3295

984 **Wei Y, Yee PP, Liu Z, Zhang L, Guo H, Zheng H, Anderson B, Gulley M, Li W** (2020)
985 NEDD4L-mediated Merlin ubiquitination facilitates Hippo pathway activation. EMBO
986 Rep **21**: e50642

987 **Wilkinson ME, Charenton C, Nagai K** (2020) RNA Splicing by the Spliceosome. Annu Rev
988 Biochem **89**: 359–388

989 **Xu Y-J, Lei Y, Li R, Zhang L-L, Zhao Z-X, Zhao J-H, Fan J, Li Y, Yang H, Shang J, et al**
990 (2017) XAP5 CIRCADIAN TIMEKEEPER Positively Regulates RESISTANCE TO
991 POWDERY MILDEW8.1-Mediated Immunity in Arabidopsis. Front Plant Sci **8**: 2044

992 **Yang Y, Li Y, Sancar A, Oztas O** (2020) The circadian clock shapes the Arabidopsis
993 transcriptome by regulating alternative splicing and alternative polyadenylation. J Biol
994 Chem **295**: 7608–7619

995 **Zhan X, Lu Y, Zhang X, Yan C, Shi Y** (2022) Mechanism of exon ligation by human
996 spliceosome. Mol Cell **82**: 2769-2778.e4

997 **Zhang S, Liu Y, Yu B** (2014) PRL1, an RNA-binding protein, positively regulates the
998 accumulation of miRNAs and siRNAs in Arabidopsis. PLoS Genet **10**: e1004841

999 **Zhang S, Xie M, Ren G, Yu B** (2013) CDC5, a DNA binding protein, positively regulates
1000 posttranscriptional processing and/or transcription of primary microRNA transcripts.
1001 Proc Natl Acad Sci U S A **110**: 17588–17593

1002

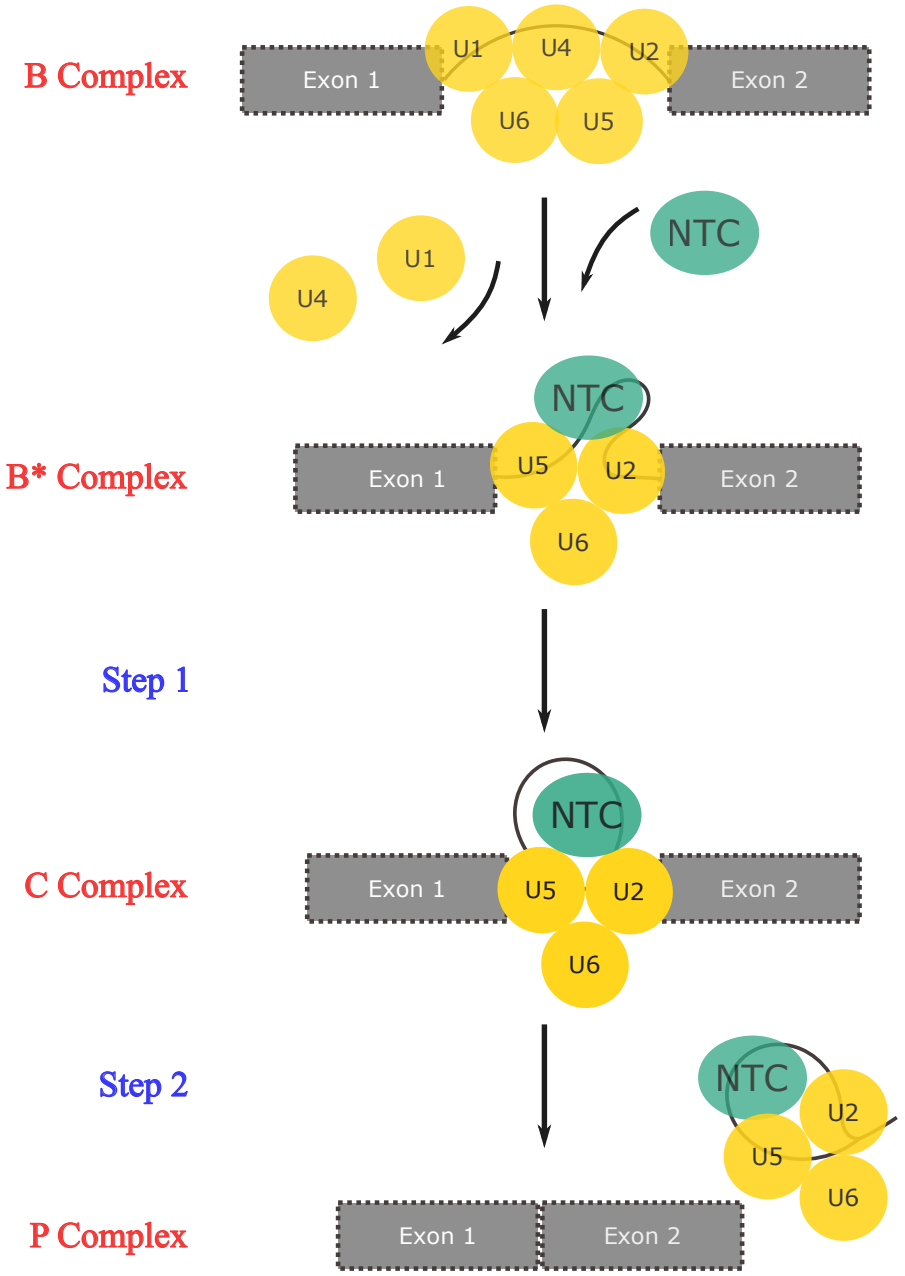

Figure 1. Simplified overview of pre-mRNA splicing reactions. A schematic diagram highlighting the two catalytic transesterification steps and the association between the NTC and the spliceosomal complex during pre-mRNA splicing. Gray boxes and solid black lines represent exons and introns, respectively. U1-U6 small nuclear ribonucleoproteins (snRNPs) are indicated by yellow circles. The NTC is indicated by a green oval.

Table 1. XCT physically associates with splicing-related proteins, especially the NTC components, in Arabidopsis. Mass Spectrometry (MS) data showing that spliceosome-associated proteins are significantly more enriched in XCT-YFP-HA-IP than control IP (fold enrichment > 3; *P*-value < 0.05, Welch two sample t.test). The full list of proteins co-purified with XCT are described in Supplemental Dataset 1. Only proteins detected in each biological replicate and with 20 or more total peptides counts are shown.

AGI	*A. thaliana* gene name	*P*- value	fold enrichment	category	*S. cerevisiae* homolog	*H. sapiens* homolog
AT2G21150	XCT	6.21E-04	Inf			FAM50A
AT1G04510	PRP19A/MAC3A/PUB59	1.52E-02	4.2	NTC	Prp19p	PRPF19; PSO4; SNEV
AT2G33340	PRP19B/MAC3B/PUB60	2.17E-03	4.8	NTC		
AT4G15900	PRL1/MAC2	1.24E-02	7.7	NTC	Prp46p	PLRG1
AT1G77180	SKIP/MAC6	1.36E-03	4.9	NTC	Prp45p	SNW1
AT1G09770	CDC5/MAC1	4.06E-03	6.7	NTC	Cef1p	CDC5L
AT1G07360	MAC5A	2.15E-03	8.8	NTC	Ecm2p	RBM22
AT5G41770	CRNK1/MAC10	5.78E-03	7.7	NTC	Clf1p/Syf3	CRNKL1/SYF3
AT3G18790	ISY1/MAC8	1.26E-03	6.1	NTC	Isy1p	ISY1
AT2G38770	EMB2765/MAC7	1.36E-02	4.4	NTC	Cwf11p	AQR
AT5G28740		8.86E-03	4.6	NTC	Syf1p	XAB2
AT3G18165	MOS4	4.09E-03	4.5	NTC	Snt309p	BCAS2/SPF27
AT5G51280		1.65E-02	4.3	other splicing	Dbp2p	DDX41
AT2G47640		4.93E-03	4.3	other splicing	Smd2p	SNRPD2
AT3G26560		8.15E-03	24	other splicing	Prp22p	DHX8
AT1G09760	U2A'	3.51E-02	3.3	U2 complex	Lea1p	SNRPA1

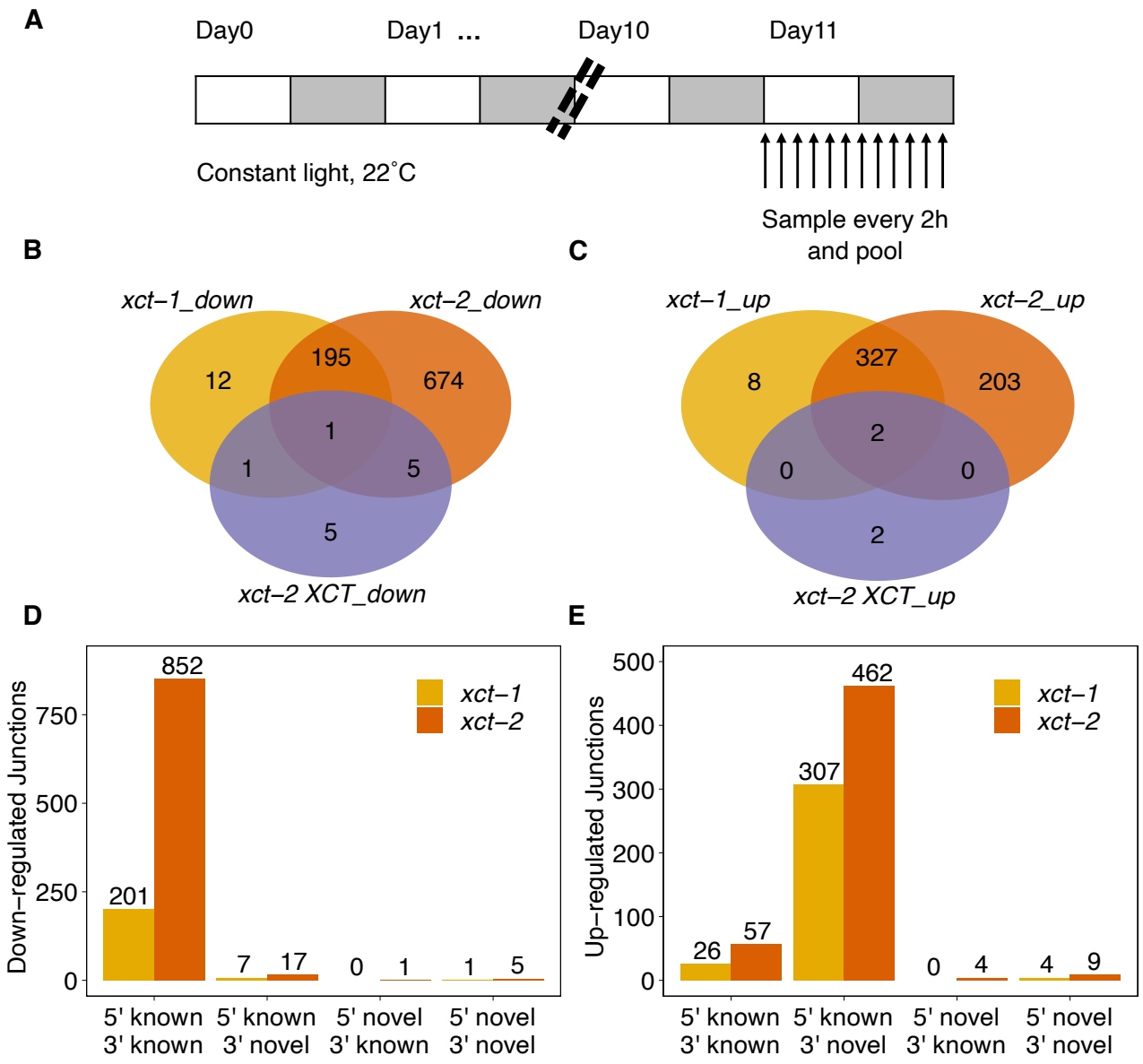

Figure 2. Transcriptome-wide analysis reveals *XCT* as a global pre-mRNA splicing regulator. A, Experimental design and sampling method for the PacBio Iso-Seq experiment. White and grey boxes represent subjective day and subjective night, respectively. Arabidopsis seedlings were grown at constant light and temperature for 10 days before being harvested and pooled. B and C, Numbers of differentially enriched splicing events represented by significantly differentially down- (B) or up-regulated (C) splice junctions in *xct-1*, *xct-2* and *xct-2* complemented with p*XCT*::g*XCT-YFP-HA* compared with Col-0 (false discovery rate < 0.05). D and E, Frequency of different classes of 5' and 3' splice sites among down- (D) or up-regulated (E) splice junctions in *xct-1* and *xct-2*. Known or novel splice sites were classified by comparing to TAIR10 genome annotation.

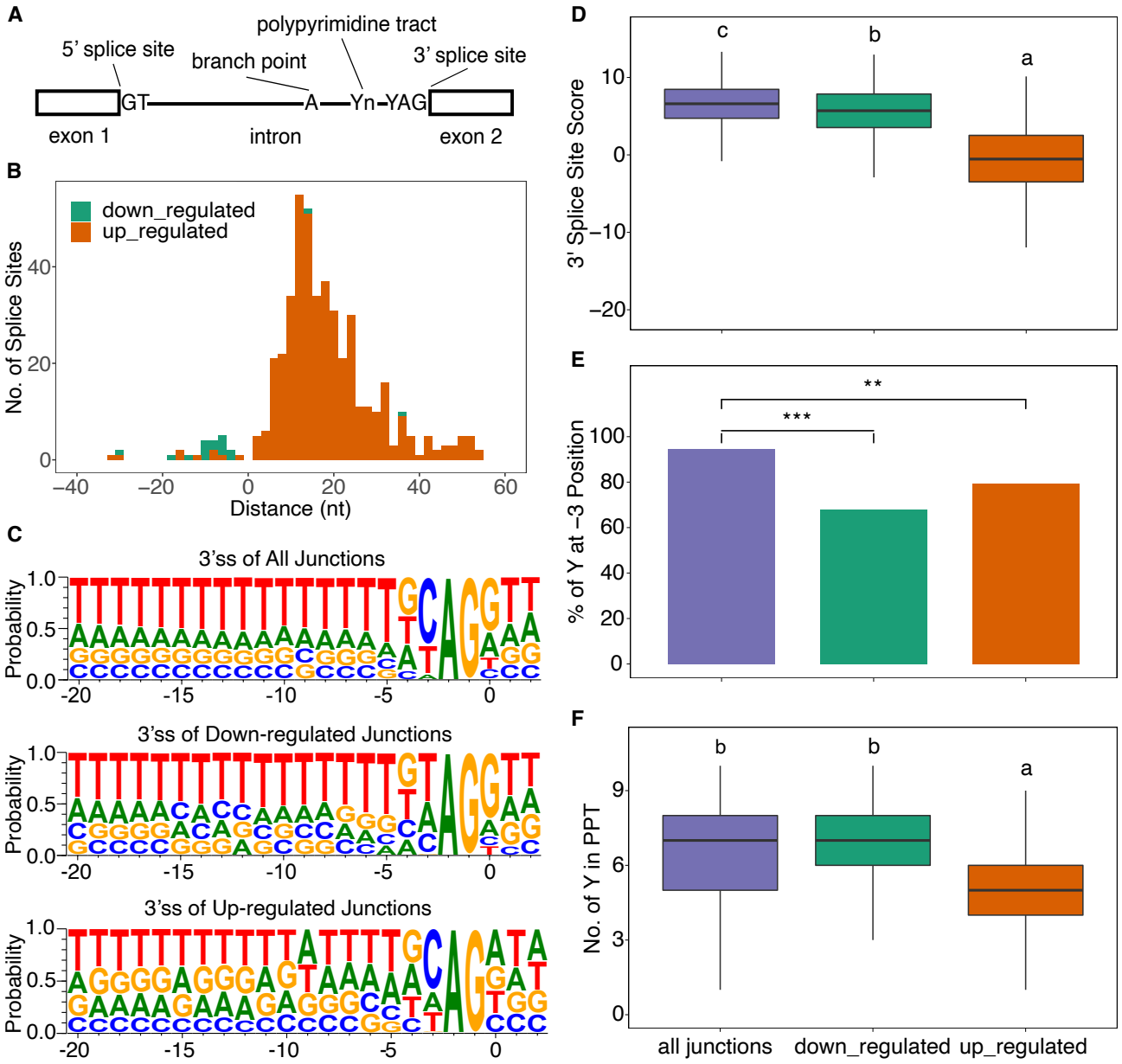

Figure 3. *XCT* is required for the fidelity of 3' splice site selection during pre-mRNA splicing. A, A schematic diagram showing the structure of a typical U2-type splice junction. Y: pyrimidine. B, Distribution of the distance between each pair of novel 3' splice site and its corresponding canonical 3' splice sites in *xct-2*. C, Pictograms showing the frequency of nucleotides in the 23-mers sequences flanking the 3' splice sites in all expressed, down-regulated and up-regulated splice junctions in *xct-2*. D, Maximum Entropy Model scores showing the strength of 3' splice sites of all expressed, down-regulated and up-regulated junctions in *xct-2*. E, Percentage of pyrimidines at the -3 position (i.e. the nucleotide preceding the AG at 3' splice site) in *xct-2*. F, Counts of pyrimidines in the Y10 polypyrimidine tract upstream of the 3' splice sites in *xct-2*. PPT, polypyrimidine tract. The lines in the boxplot represent the 75% quartile, median and 25% quartile of the data, respectively. Statistical significance in (D) and (F) was determined using linear regression model with junction class as a fixed effect and is shown in lower case letters (Tukey's multiple comparison test, $P < 0.05$). Statistical significance in (E) was determined by Fisher's exact test: **: $P < 0.01$; ***: $P < 0.001$.

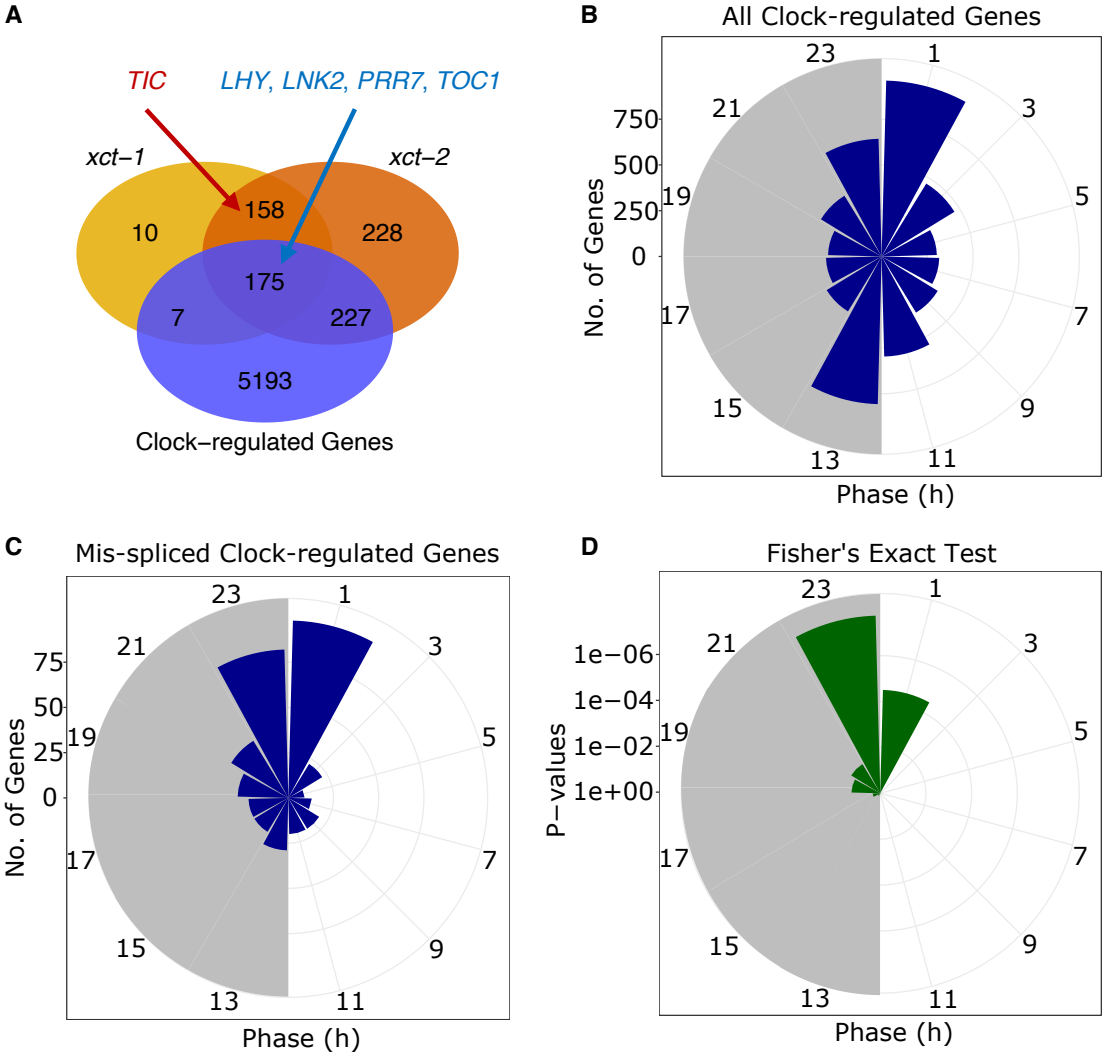

Figure 4. Dawn-phased genes are significantly enriched among all the circadian-clock-regulated genes that are aberrantly spliced in *xct-2*. A, Venn diagram showing the overlaps of aberrantly spliced genes in *xct-1*, *xct-2*, and total circadian-clock-regulated genes (Romanowski et al., 2020). The core circadian clock genes that are mis-spliced in both *xct* mutants are indicated in red (non-clock-regulated) and blue (clock-regulated) fonts. Only genes considered as 'detected' in all three RNA-Seq datasets are included. B and C, Phases of estimated peak expression of all circadian clock-regulated genes (Romanowski et al., 2020) that are detected (B) or significantly aberrantly spliced (C) in *xct-2*. The white and gray backgrounds represent the subjective day and subjective night, respectively. D, Distribution of *P*-values from Fisher's extact tests calculating whether the ratio of the number of aberrantly spliced genes in (B) to total clock-regulated genes in (A) is significantly higher than expected by chance in each 2-hour interval.

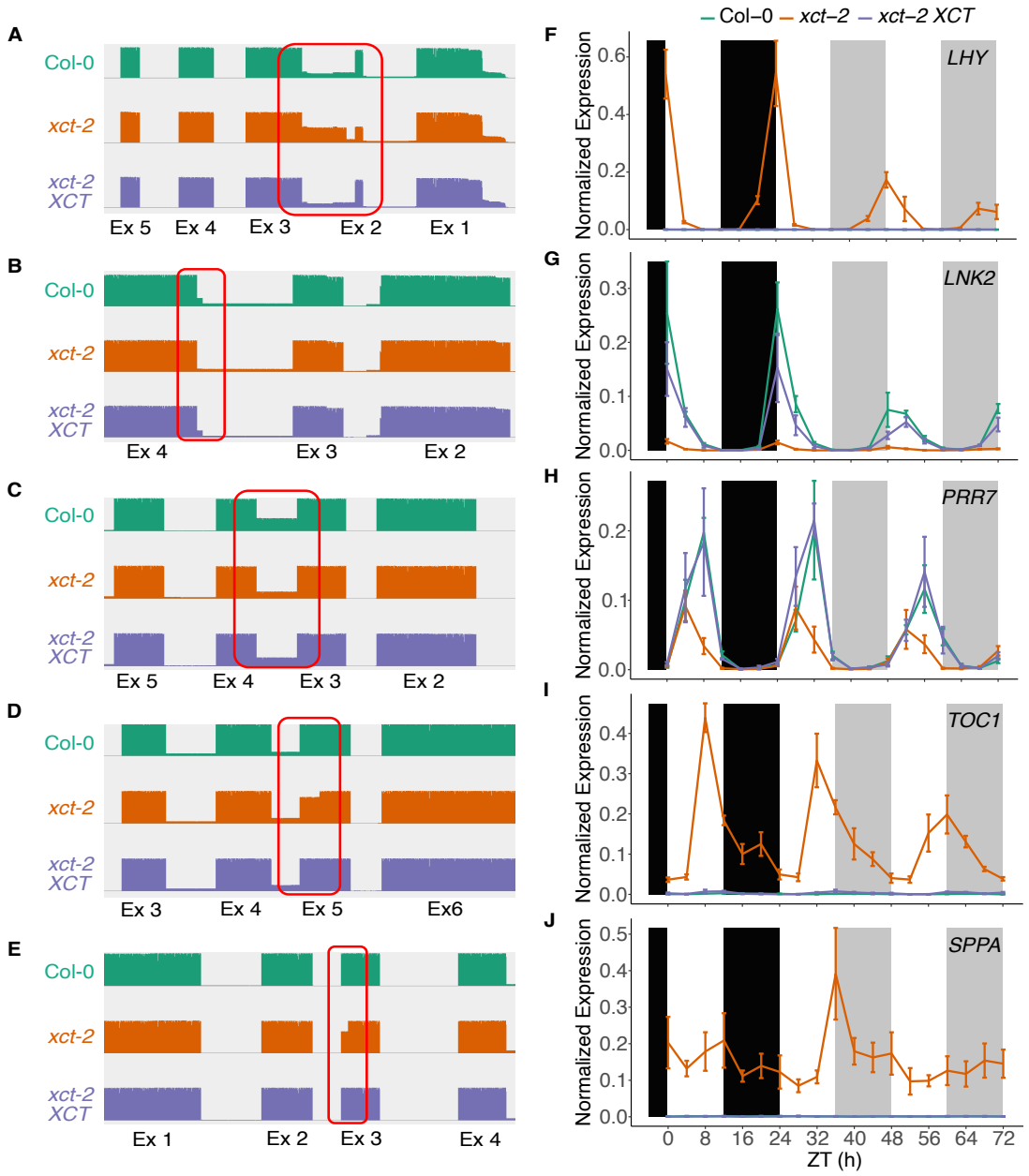

Figure 5. Time-course qRT-PCR experiments validate the role of *XCT* in regulating pre-mRNA splicing of clock-regulated and non-clock-regulated genes. A - E, Sashimi plots showing PacBio Iso-Seq reads mapped to *LHY* (A), *LNK2* (B), *PRR7* (C), *TOC1* (D) and *SPPA* (E) in Col-0 (teal), *xct-2* (orange) and *xct-2 XCT* (purple). The red rectangles highlight the aberrantly spliced exon-exon junctions that are examined by qRT-PCR in (F) - (J). F - J, Normalized expression of the aberrantly spliced isoforms of *LHY* (F), *LNK2* (G), *PRR7* (H), *TOC1* (I) and *SPPA* (J) in Col-0, *xct-2* and *xct-2 XCT*. Samples were collected every four hours over a 72-h window. Expression levels were examined by qRT-PCR using splice-junction-specific primers and normalized to *PP2A* and *IPP2*. Data points represent mean ± se from three independent biological replicates. For each isoform in each biological replicate, the normalized expression levels were relative to the highest expression levels of their total transcripts in Col-0 across all time points. Teal lines, wild type Col-0; orange lines, *xct-2* muants; purple lines, *xct-2 XCT*. Black background, dark period; white and gray background, light period during subjective day and night, respectively.

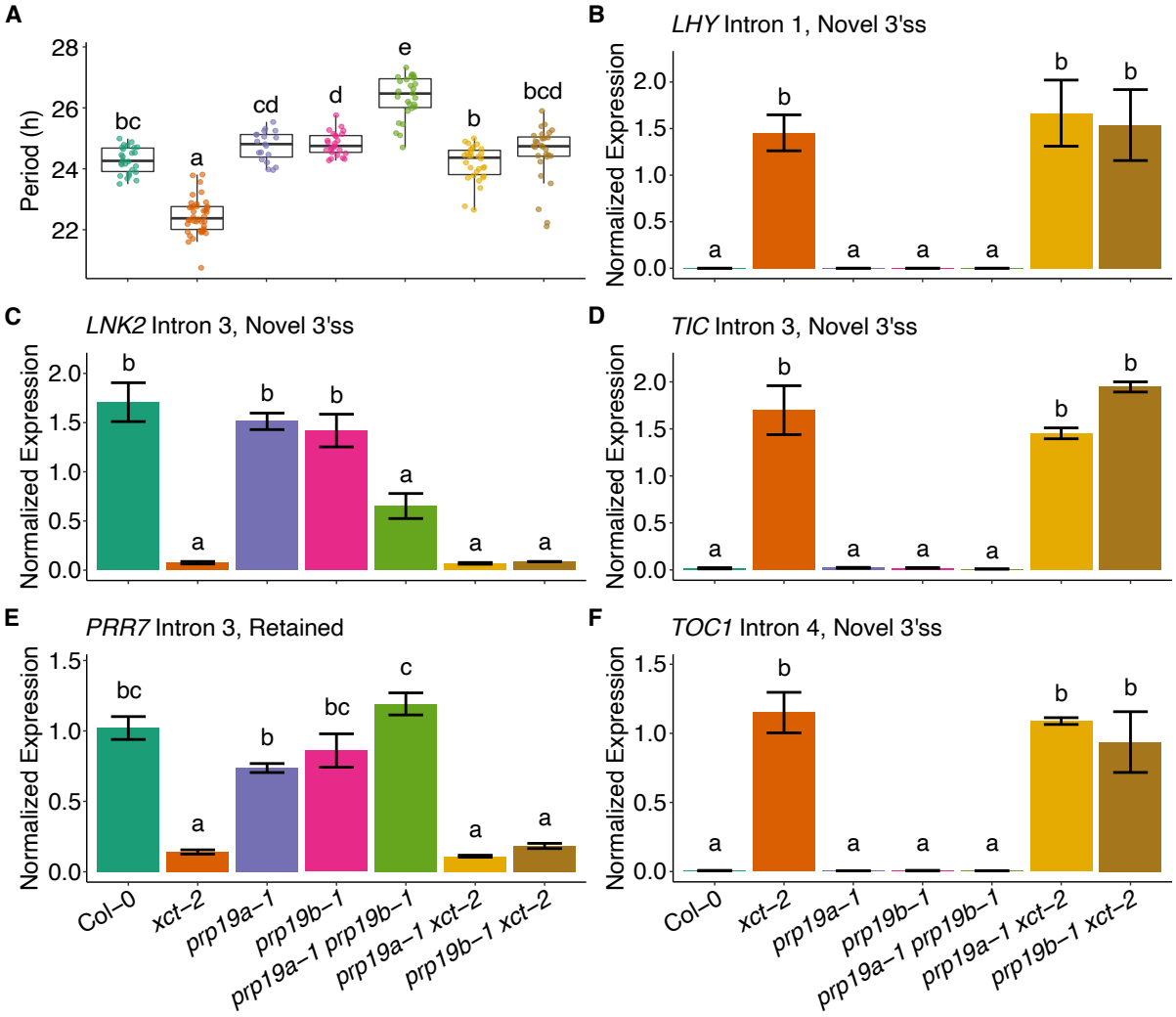

Figure 6. Loss of *PRP19* function suppresses the short circadian clock period phenotype but not the splicing defects of core clock genes in *xct-2*. A, Circadian periods of Col-0, *xct-2*, *prp19a-1*, *prp19b-1*, *prp19a-1 prp19b-1*, *prp19a-1 xct-2* and *prp19b-1 xct-2* plants. B - F, Normalized expression of the aberrantly spliced isoforms of *LHY*, *LNK2*, *TIC*, *PRR7* and *TOC1* in Col-0, *xct-2*, *prp19a-1*, *prp19b-1*, *prp19a-1 prp19b-1*, *prp19a-1 xct-2* and *prp19b-1 xct-2*. Samples were collected at the estimated peak expression time for each gene. Expression levels were examined by qRT-PCR using splice-junction-specific primers and normalized to *PP2A* and *IPP2*. Data points represent mean ± se from two independent biological replicates. Statistical significance was determined using linear regression model with genotype as a fixed effect and is shown in lower case letters (Tukey's multiple comparison test, $P < 0.05$).

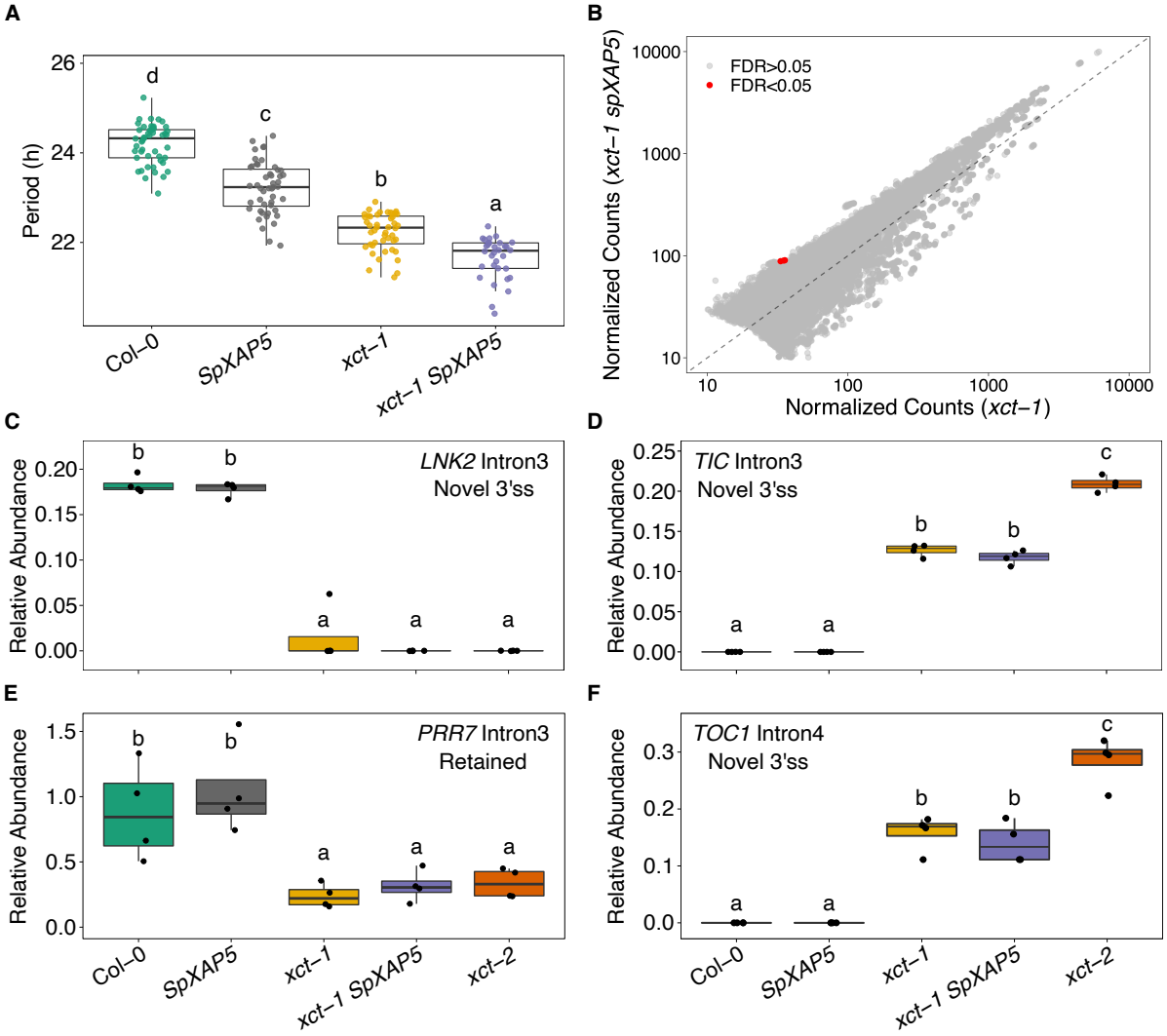

Figure 7. Expression of *S.pombe XAP5* in Arabidopsis causes a short circadian clock period phenotype without inducing pre-mRNA splicing defects. A, Circadian period of Col-0, *SpXAP5*, *xct-1* and *xct-1 SpXAP5* plants. B, A scatter plot showing differential splicing events in *xct-1 SpXAP5* compared with *xct-1* indicated by normalized PacBio Iso-Seq read counts of exons and intron elements. Differentially spliced elements that are statistically significant (FDR < 0.05) and insignificant (FDR > 0.05) are indicated with red and grey dots, respectively. The dashed line represents y=x. C - F, Relative abundance of aberrantly-spliced to fully-spliced isoforms of *LNK2*, *TIC*, *PRR7* and *TOC1* in Col-0, *SpXAP5*, *xct-1*, *xct-1 SpXAP5* and *xct-2*. Differentially spliced isoforms were amplified by semi-qRT-PCR using primers flanking the examined regions, followed by separation and quantification using a Lab Chip GX bioanalyzer. Statistical significance in (A) and (C) - (F) was determined using linear regression models with genotype as a fixed effect and is shown in lower case letters (Tukey's multiple comparison test, $P < 0.05$).