1   **Maize inbreds show allelic variation for diel transcription patterns**

2

3   Joseph L. Gage[1,2*], M. Cinta Romay[3], and Edward S. Buckler[3,4,5]

4   [1]Department of Crop and Soil Sciences, North Carolina State University, Raleigh, NC 27695

5   [2]NC Plant Sciences Initiative, North Carolina State University, Raleigh, NC, 27606

6   [3]Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853

7   [4]USDA-ARS, Ithaca, NY 14850

8   [5]School of Integrative Plant Science, Plant Breeding and Genetics Section, Cornell University,

9   Ithaca NY 14853

10

11   [*]To whom correspondence should be addressed; jlgage@ncsu.edu

12

13   **Abstract**

14   Circadian entrainment and external cues can cause gene transcript abundance to oscillate

15   throughout the day, and these patterns of diel transcript oscillation vary across genes and plant

16   species. Less is known about within-species allelic variation for diel patterns of transcript

17   oscillation, or about how regulatory sequence variation influences diel transcription patterns. In

18   this study, we evaluated diel transcript abundance for 24 diverse maize inbred lines. We

19   observed extensive natural variation in diel transcription patterns, with two-fold variation in the

20   number of genes that oscillate over the course of the day. A convolutional neural network trained

21   to predict oscillation from promoter sequence identified sequences previously reported as

22   binding motifs for known circadian clock genes in other plant systems. Genes showing diel

23   transcription patterns that cosegregate with promoter sequence haplotypes are enriched for

24   associations with photoperiod sensitivity and may have been indirect targets of selection as

25   maize was adapted to longer day lengths at higher latitudes. These findings support the idea that

26   cis-regulatory sequence variation influences patterns of gene expression, which in turn can have

27   effects on phenotypic plasticity and local adaptation.

28

29   **Introduction**

30   Because most plants rely on sunlight for energy, they need to entrain their biological

31   processes to the rhythms of the sun. Plant growth, development, physiology, metabolism, and

32   immunity are all governed by diel cues and the internal circadian clock(Lu *et al.* 2017), including

33   temporal patterns of transcription for about one third of genes(Ferrari *et al.* 2019). The number

34   and types of genes showing diel patterns of transcription vary across species in Archaeplastida,

35   but core components of the circadian clock are conserved across green plants(Ferrari *et al.* 2019;

36   Michael 2022; Petersen *et al.* 2022).

37   Despite high conservation of core circadian regulators, plant species have adapted to

38   environments that vary wildly in their diurnal conditions, showing differences in aspects such as

39   day length, light availability, temperature fluctuation, and pathogen pressure. These same

40   conserved circadian regulators underpin molecular processes for circadian-regulated metabolic

41   products that differentiate species from each other. Some of the diversity in form and function

42   related to circadian processes can be attributed to genetic diversity in circadian genes, as

43   previously shown in Arabidopsis(Michael *et al.* 2003) soybean(Wang *et al.* 2022) and

44   maize(Hung *et al.* 2012b), but rewiring of gene interactions by evolution of transcription factor

45   binding sites (TFBS) or chromatin accessibility are other, less explored ways that circadian

46   regulation may have diversified among plant species.

2

47   Transcription factors (TFs) tend to evolve slowly and be conserved at both the sequence and

48   functional level(Lambert *et al.* 2018; Tu *et al.* 2020). Their binding sites, on the order of ~10bp

49   in length(Stewart *et al.* 2012), experience much faster rates of evolution(Stone and Wray 2001;

50   Weirauch and Hughes 2010; Lambert *et al.* 2018). The fact that TFBS evolve much faster than

51   their cognate TFs makes TFBS evolution a strong possible explanation for the observed

52   variability in diel gene expression patterns between species, despite a conserved core circadian

53   network. In addition to between-species variability, differential activity of TFBSs may contribute

54   to within-species diversity for diel gene expression. To date, many studies of diel or circadian

55   gene expression patterns in plants have focused on a representative variety, mutant lines, or

56   comparisons across species. The extent of within-species, natural allelic variation for diel gene

57   expression patterns is less well characterized.

58   Maize is morphologically and genetically diverse, and has adapted to a wide range of

59   environments worldwide, including broad latitudinal adaptation. Maize was originally

60   domesticated in the Balsas valley of Mexico under short day lengths of less than 13

61   hours(Matsuoka *et al.* 2002; Piperno *et al.* 2009). Prior to European colonization, people living

62   in the Americas moved maize to latitudes spanning from Chile to Canada(Sauer and Weatherwax

63   1955; Crawford *et al.* 2006), necessitating selection for growth and reproduction under long day

64   conditions with up to 16 hour days. Short-day-adapted material grown in long-day environments

65   will exhibit delayed flowering or potentially not flower at all; temperate-adapted maize varieties

66   must be photoperiod-insensitive in order to produce grain.

67   Early flowering is a highly polygenic trait critical for adaptation to temperate environments

68   at higher latitudes with longer days(Troyer and Hendrickson 2007; Buckler *et al.* 2009), whereas

69   photoperiod sensitivity, while still polygenic, appears to be governed by fewer loci(Hung *et al.*

3

70 2012b). Genetic variation in the maize circadian clock gene *ZmCCT*, homolog of *Ghd7* in rice

71 and partially similar to CONSTANS in Arabidopsis, has been shown to contribute to photoperiod

72 sensitivity, but adaptation to long days is polygenic and *ZmCCT* accounts for 9% of phenotypic

73 variance for photoperiod sensitivity in the maize Nested Association Mapping population(Hung

74 *et al.* 2012b). Adaptation to long days likely required selection on a standing variation at a large

75 number of loci(Swarts *et al.* 2017), but few of those have been rigorously identified or

76 characterized, and the mechanisms by which selection altered flowering time in maize are not

77 well known.

78 Here we describe natural allelic variation for diel patterns of transcript abundance among 24

79 maize inbred lines. We find that transcription factor binding motifs previously identified in

80 Arabidopsis are predictive of whether or not a gene shows cyclical diel patterns of transcript

81 abundance in maize. In addition, promoter region haplotypes co-segregate with significantly

82 different rhythms of diel transcript abundance for hundreds of genes among our 24 inbreds.

83 Those genes are shown to be enriched for GWAS hits for flowering and photoperiod sensitivity

84 traits, and in some populations, enriched for signals of selection to higher latitudes. We posit that

85 during movement to higher latitudes, people selecting for adaptation to longer days modified

86 circadian regulation networks.

87

88 **Results and Discussion**

89 <u>Natural variation for diel transcriptome rhythmicity</u>

90 Twenty-four diverse inbred lines, a subset of the parents of the maize NAM population(Yu *et*

91 *al.* 2008; Gage *et al.* 2020), were evaluated for transcript abundance every two hours for twenty-

92 four hours (Supplemental Data 1). To reduce reference bias, sequencing data from each inbred

93   were aligned to that inbred's cognate genome assembly(Hufford *et al.* 2021). Each annotated

94   gene in each inbred was evaluated for rhythmicity using diffCircadian(Ding *et al.* 2021), which

95   uses a likelihood-based method to test whether diel transcriptomic data fit a sinusoidal function.

96   This method is less likely to identify significant diel patterns of expression that are not

97   sinusoidal, such as impulses, step changes, or sawtooth patterns.

98   We found variability between inbred lines for the number of genes which display significant

99   diel rhythmicity (likelihood ratio test $p < 0.01$), ranging from 3,345 (Ms71) to 6,730 (Il14H)

100   (Figure 1A, Supplemental File 1). The two genotypes with the highest number of rhythmic

101   genes, P39 and Il14H, are both sweet corn lines which are genetically distinct from the

102   remainder of the NAM parent lines, adapted to northern latitudes, and flower quickly. The

103   number of cycling transcripts in a given inbred was weakly but not significantly correlated with

104   photoperiod sensitivity, flowering time, maturity at sample collection, leaf sampled, or

105   sequencing depth (Supplemental Figure 1), supporting the possibility that this pattern may be due

106   to genetic differences in diel gene regulation and not simply varying levels of adaptation to the

107   northern field site (Aurora, NY) where this study was performed.

108   Using a pangenome(Hufford *et al.* 2021) to establish homologous pan-genes, we found that

109   pan-genes which display diel rhythmicity in at least one inbred are frequently core genes (ie,

110   present in all inbreds), but are not necessarily rhythmic in all inbreds (Figure 1B). This finding

111   indicates that standing variation for rhythmicity exists in otherwise conserved genes and could

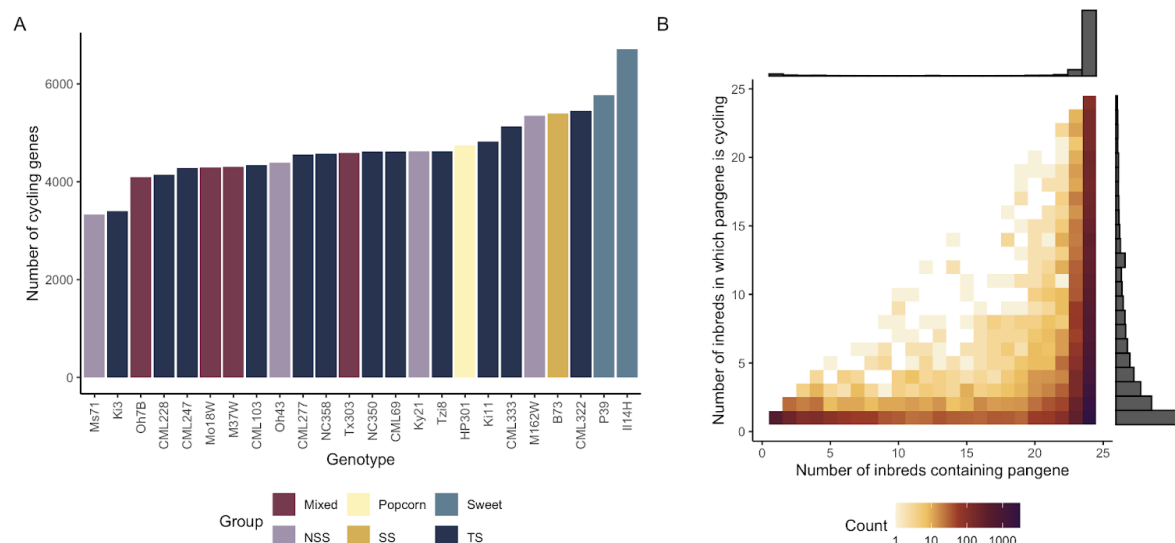112   have been selected upon during breeding and adaptation to new environments.

Figure 1: (A) Number of genes showing significant (p<0.01) sinusoidal diel patterns of transcript abundance in each inbred maize line. (B) Pangenes that show diel cycling are most likely to be present in all 24 inbred lines, but are often cycling only in some inbreds.

Predicting rhythmicity from promoter sequences

Given that there is natural variation in diel transcription patterns, both among pan-genes and between alleles of pan-genes, we trained a convolutional neural network (CNN) to predict whether a given gene is likely to show a cyclical pattern of diel transcription (Supplemental Data 2). Based on previous characterization of the plant circadian clock, we know that diel transcription is subject to temporal feedback loops involving a number of transcription factors (TFs), including CCA1, LHY, PRRs, and ELFs(Sanchez and Kay 2016). We predicted that the presence, absence, and interactions between the binding motifs for these TFs might be identified from the convolutional filters of our model.

Among the motifs identified by our model (Supplemental Figure 2, Supplemental Data 3) was a strong representation of the Evening Element AAATATCT (Nagel *et al.* 2015). We also identified motifs related to other circadian TFs previously characterized in Arabidopsis, including the CONSTANS-responsive element CCACA (CORE2) (Gnesutta *et al.* 2017); the

6

131 TOC1 TCP Binding Site (TBS) binding site motif GGCCC(Gendron *et al.* 2012); the TOC1 GA

132 element(Gendron *et al.* 2012); the E2F/DP binding motif GGCGG(Vandepoele *et al.* 2005)

133 which controls elements of the cell cycle, a process coupled to the circadian clock(Torii *et al.*

134 2022); and the Telo-box motif AAACCCT(Spensley *et al.* 2009; Gardiner *et al.* 2021). Though

135 our model was able to identify known TF binding motifs, most of its predictive power is from

136 predicting across pan-genes; it is limited in its ability to predict differences between pan-gene

137 alleles.

138 These results reinforce conserved aspects of the plant circadian clock(Khan *et al.* 2010;

139 Ferrari *et al.* 2019) as well as results from previous studies which used machine learning to

140 identify sequence motifs related to circadian expression in Arabidopsis(Gardiner *et al.* 2021).

141 This provides evidence for the utility of transferring knowledge from fundamental biology

142 findings in model species to inform understanding about molecular biology in crop plants,

143 despite >100M years of evolutionary divergence.

144

145 <u>Co-segregation of promoter haplotypes and transcript rhythmicity</u>

146 Following the identification of the predictive motifs from our CNN, we hypothesized that

147 sequence variation could be causing differing patterns of diel gene expression among alleles of a

148 gene. For each pangene, we tested for an association between the haplotype sequence of each

149 inbred's promoter region and patterns of diel transcript abundance, identifying 489 pangenes

150 with significant association between promoter haplotype and diel expression pattern (Figure 2).

151 We refer to these genes from here onward as Differential Diel Regulation (DDR) candidates
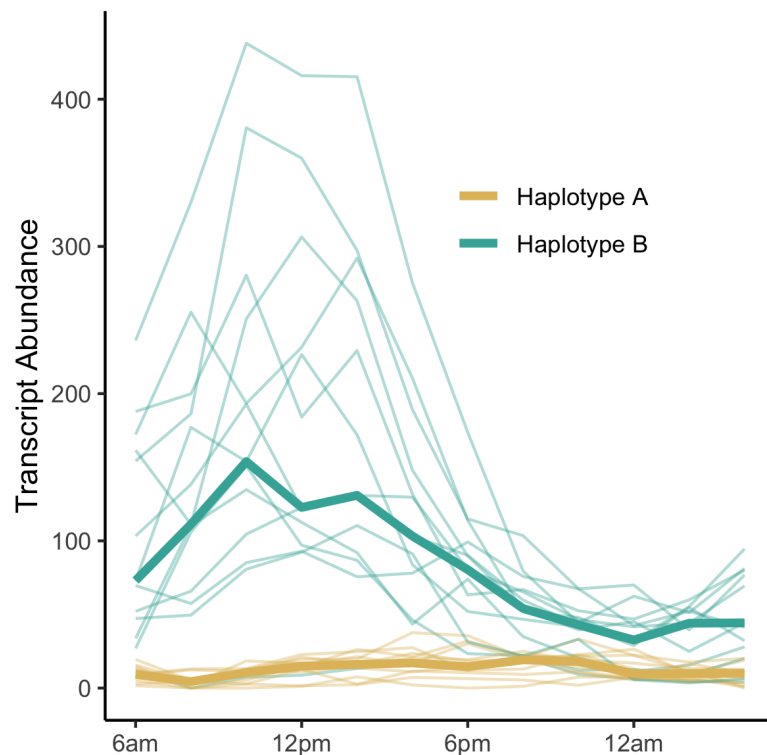
152 (Supplemental Files 2 and 3).

153

154

Figure 2: Example of a pangene (Zm00001eb017170; ortholog of Arabidopsis ROQH1) with differential diel regulation based on promoter haplotype. Thin lines show transcript abundance for a single inbred throughout the course of the day and are colored based on which haplotype the inbred has in the promoter region of the gene; thick lines show the median transcript abundance for individuals with each of two haplotypes.

160

161   Function of DDR genes:

162   Among the identified DDR genes are calmodulin2 (cal2; Zm00001eb128040), a calcium

163   signaling gene previously reported as involved in flowering time and adaptation to higher

164   altitudes(Li *et al.* 2016; Hu *et al.* 2022) and MADS-transcription factor 68 (mads68;

165   Zm00001eb006480), an ortholog of rice mads47 implicated in floral development,

166   brassinosteroid signaling, and in maize, endosperm development(Duan *et al.* 2006; Fornara *et al.*

167   2008; Qiao *et al.* 2016).

168   KEGG pathway analysis of DDR genes reveals that they are enriched for a number of KEGG

169   pathways related to metabolism (Figure 3). The highest enrichment observed was for the sulfur

8

170   metabolism pathway, which, along with cysteine and methionine metabolism (fourth highest

171   enrichment), has been previously reported as downregulated in arabidopsis mutants lacking dawn

172   clock components (LHY and CCA1)(Flis *et al.* 2019). We also found enrichment for carbon

173   metabolism and amino acid biosynthesis, which have been shown to be interlinked in diel

174   regulatory loops between photosynthesis, carbon metabolism, and nitrogen metabolism(Stitt *et*

175   *al.* 2002). Finally, enrichment for ribosomes and protein processing agree with previous work

176   demonstrating that translation of ribosomal proteins is diurnally regulated(Missra *et al.* 2015)

177   and that translation and protein degradation rates are tightly diurnally coupled(Mehta *et al.* 2021;

178   Duncan and Millar 2022). Because plant growth relies on effectively coordinated synthesis and

179   degradation of starch during the day and night(Stitt and Zeeman 2012), respectively, it is quite

180   possible that allelic variation for diel transcription of genes in these various metabolic pathways

181   result from, or contribute to, variable (mal)adaptation among these 24 inbred lines to the diurnal

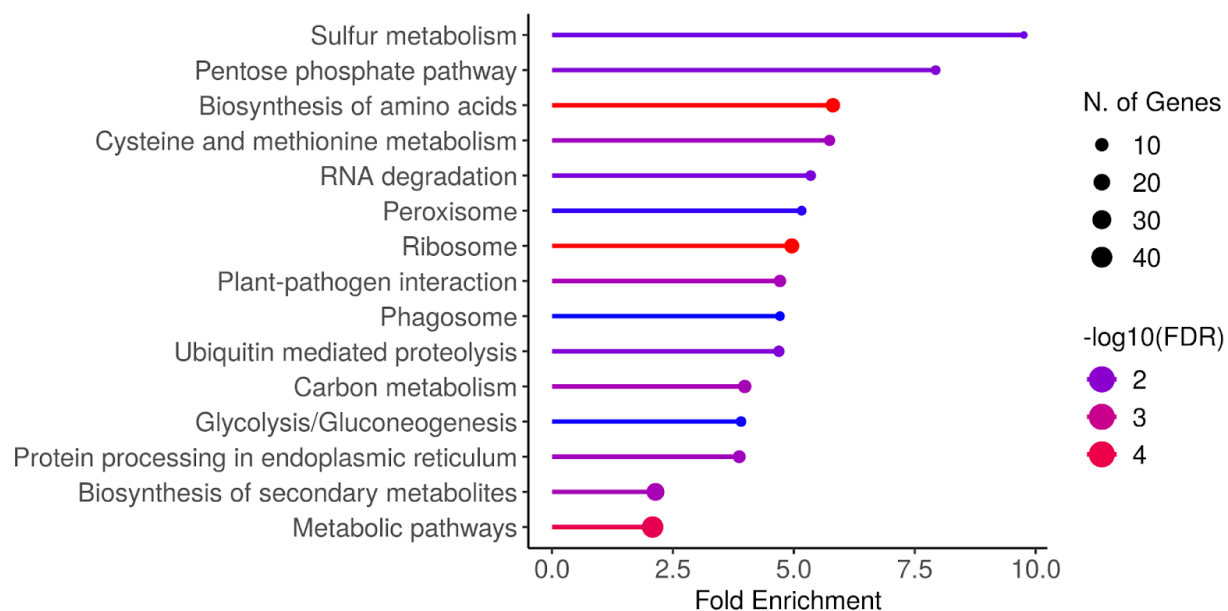182   conditions under which they were grown.

183

Figure 3: KEGG pathway enrichment results from 489 DDR genes, compared to the background set of 33,201 genes transcribed in at least one tissue of at least one NAM parent line.

186

9

187   <u>Comparison of diel transcript abundance in an independent experiment</u>

188   We compared diel patterns of transcript abundance in our field grown plants to three inbreds

189   sampled every three hours in a growth chamber (Supplemental Data 4). We found that among

190   those three inbreds, rank order for the number of significantly cycling genes was preserved. Of

191   all genes that we observed as significantly cycling in the field, 38% (CML103) to 68% (P39)

192   were also significantly cycling in the growth chamber assay (Supplemental Figure 3,

193   Supplemental File 4). This underscores the fact that diel regulation of transcription is consistent

194   across development and environment for some genes, but condition-specific for others. Nearly

195   half of the DDR genes identified in the field experiment were also identified as DDR in the

196   growth chamber, despite the fact that not all haplotypes from the field experiment are

197   represented by the limited genetic diversity of only three inbreds in the growth chamber

198   experiment, as well as the fact that we used a different statistical method for assessing the

199   significance of DDR in growth chamber materials (see Methods). For the following enrichment

200   analyses, we used 205 genes identified as DDR in both field and growth chamber experiments

201   (Supplemental Files 5 and 6).

202

203   <u>DDR candidates are enriched for GWAS hits</u>

204   Because the circadian clock regulates numerous biological processes(Lu *et al.* 2017), we

205   hypothesized that allelic differences in diel transcription patterns contribute to natural variation

206   for traits regulated by the circadian clock. To test this hypothesis, we checked whether DDR

207   candidates are enriched for GWAS hits relative to randomly selected genes. However, choosing

208   an appropriately matched set of genes from which to take random samples is non-trivial, in order

10

209   to ensure that the random samples are truly representative and a fair comparison to the DDR

210   genes under consideration.

211   We compared GWAS hits in DDR genes to random selections from four different subsets of

212   annotated genes: 1) all annotated genes, as a baseline. This null set was likely to overestimate the

213   significance of DDR genes, because the set of all gene annotations contains pseudogenes and

214   incorrectly annotated genes that are likely to have fewer significant GWAS associations. 2) all

215   genes that show evidence of transcription in at least one tissue, in at least one of the 26 NAM

216   parents(Hufford *et al.* 2021). 3) genes that show evidence of diel rhythmicity in at least two

217   inbreds. This gene set was chosen to represent the set of genes that are either segregating for

218   rhythmicity or constitutively rhythmic. 4) genes that show evidence of diel rhythmicity in at least

219   22 inbreds. This gene set was chosen to represent genes that are constitutively, or nearly

220   constitutively, rhythmic.

221   We found that DDR genes have significantly enriched GWAS associations for

222   photoperiod sensitivity traits compared to all genes ($p<0.001$), genes transcribed in at least one

223   tissue/inbred ($p<0.001$), genes segregating for rhythmicity ($p<0.02$), and constitutively rhythmic

224   genes, ($p<0.004$). DDR genes show weaker enrichment for flowering time traits, with significant

225   enrichment only compared to all genes or genes transcribed in at least one tissue/inbred (Figure

226   4A).

227   GWAS hits for control traits (tassel length, tassel branch number, and leaf angle; chosen

228   because we predicted they would not be influenced by DDR genes) were also all significantly

229   enriched in DDR genes when compared to all annotated genes ($p<0.001$) or genes transcribed in

230   at least one tissue/inbred ($p<0.002$). Compared to genes segregating for rhythmicity and genes

231   with constitutive rhythmicity, DDR genes showed no enrichment in GWAS signal for leaf angle

11

232   or tassel branch number (p>0.09) (Figure 4A). Overall, the three control traits showed lower

233   mean GWAS p-values at DDR genes than flowering time and photoperiod sensitivity, indicating

234   DDR genes are more strongly associated with photoperiod and flowering than with control traits.

235   We also checked whether DDR genes overlap with previously identified QTL for photoperiod

236   sensitivity(Hung *et al.* 2012b), but found no evidence for a greater rate of overlap than randomly

237   chosen genes (p=0.51).

238   These findings demonstrate that DDR genes are enriched for GWAS hits related to

239   photoperiod sensitivity, and that GWAS hit enrichment for flowering time, tassel traits, and leaf

240   angle depends on choice of null gene set. Comparing to genes that show evidence of segregating

241   for transcript rhythmicity (null set #2) may be overly conservative: even though many of those

242   genes were not picked up as DDR candidates in our haplotype-based testing, it may that their

243   segregation for rhythmicity contributes to complex trait variation (ie, the null set may be contain

244   false negative DDR genes). On the other hand, comparing to genes that show any evidence of

245   transcription may be an overly liberal test because the null set contains genes that are transcribed

246   but not relevant to variation in the traits that we tested. The fact that DDR candidates show

247   stronger GWAS associations than control genes for flowering time and photoperiod sensitivity

248   traits support the hypothesis that allelic variation for diel transcription patterns contributes to
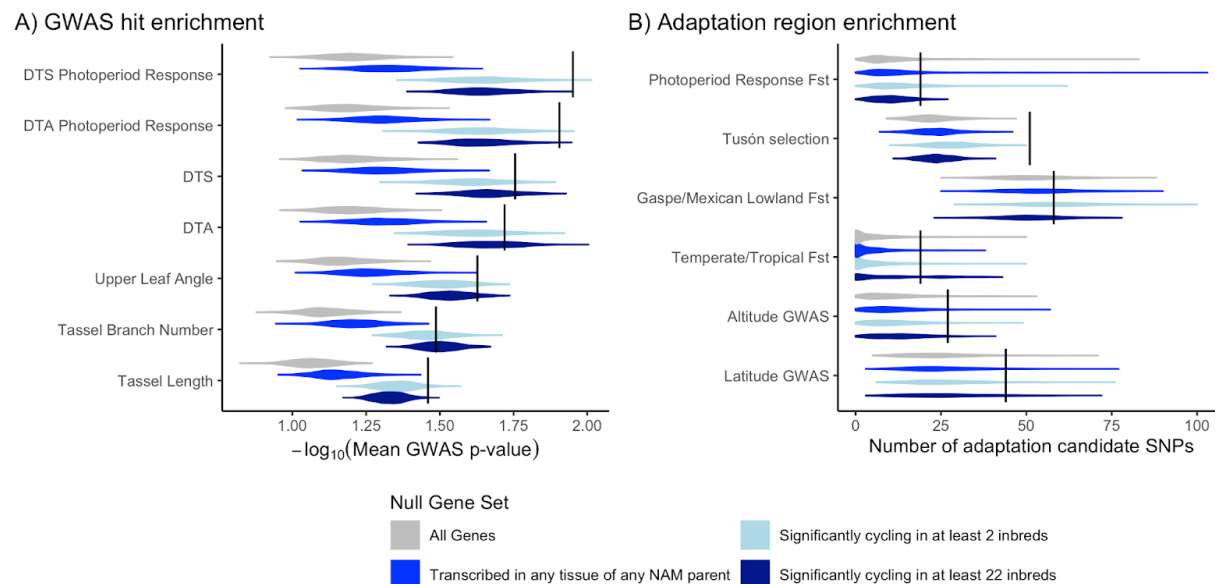
249   longitudinal adaptation to longer days.

250

Figure 4: (A) Enrichment for GWAS signal in 205 DDR genes (black vertical lines) compared to random samples of 205 genes from four different null sets (gray and blue distributions). (B) Overlap between candidate adaptation SNPs from five previous scans for selection and DDR genes (black lines). Gray and blue distributions show overlap between candidate adaptation SNPs and randomly selected genes from different null gene sets.

Night-Day eQTL are enriched for GWAS hits

Next, using an independent dataset we tested our hypothesis that variable cis-regulation of diel transcription patterns contributes to natural variation for traits regulated by the circadian clock. Using RNAseq data from mature leaf tissue sampled in the Goodman Association Panel (GAP)(Kremling *et al.* 2018), we calculated the difference between nighttime and daytime transcript abundances as a rough estimator for how much transcript abundance varies diurnally, and used the difference as the response variable for eQTL mapping.

We found that Night-Day cis-eQTL are highly enriched in GWAS hits for flowering time and photoperiod response for days to anthesis, compared to random SNPs sampled from within the same genomic context (<5,000bp away from the focal gene) and matched for minor allele frequency (Figure 5). There was no enrichment for tassel traits, as expected, or days to silk

13

269   photoperiod response, which was unexpected given the strong enrichment for days to anthesis

270   photoperiod GWAS hits. These results also support the idea that cis-regulation of diel

271   transcription may be contributing to changes in the adaptive traits of flowering time and
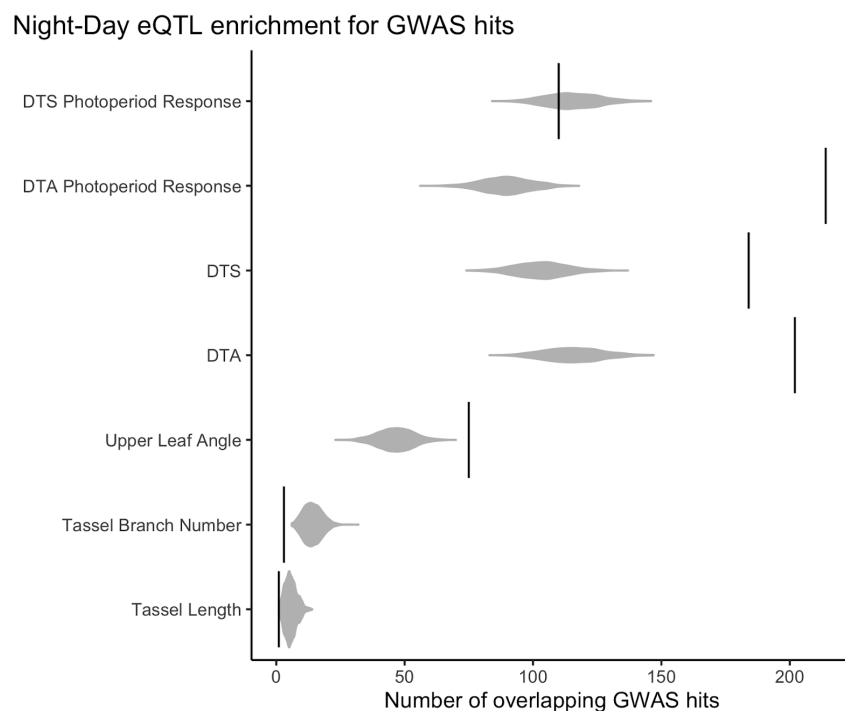
272   photoperiod sensitivity.

273

Figure 5: Overlap between GWAS hits and cis-eQTL for the difference between daytime and nighttime transcript abundance (black vertical lines). Gray distributions represent overlap between GWAS hits and SNPs randomly sampled while controlling for minor allele frequency and distance from the nearest gene.

278
279   <u>DDR candidates are moderately enriched for signatures of selection</u>

280   Given that DDR candidates are enriched for GWAS hits related to flowering time and

281   photoperiod sensitivity, and that cis-regulatory sequences are likely to evolve faster than TFs,

282   changing gene regulation and ultimately phenotype(Stone and Wray 2001; Weirauch and Hughes

283   2010), we hypothesized that diel transcription patterns may have been subjected to selection

284   during adaptation of maize to higher latitudes. Using several independent studies of maize

285   adaptation to changes in latitude and altitude(Romero Navarro *et al.* 2017; Swarts *et al.* 2017;

14

286   Gage *et al.* 2017; Wisser *et al.* 2019), we tested DDR candidates for enrichment of putatively

287   selected regions (Figure 4B).

288   Unlike our GWAS enrichment results, tests for enrichment of adaptive regions among DDR

289   genes are not as sensitive to choice of null set; results remain similar regardless of whether we

290   selected random samples from all genes, genes that show any evidence of transcription, or genes

291   that show evidence of rhythmic transcription.

292   DDR candidates are significantly enriched ($p<0.001$) for SNPs showing evidence of selection

293   in Hallauer's Tusón, a population of tropical traditional varieties that were selected for early

294   flowering in Iowa, USA for 10 generations(Teixeira *et al.* 2015; Hallauer and Carena 2016). We

295   observed low to no significance, however, for enrichment of putative adaptation SNPs identified

296   from Fst between high and low photoperiod sensitivity lines or from three other published

297   studies: eGWAS for latitude or altitude of origin(Romero Navarro *et al.* 2017) (p-value 0.04-

298   0.12), Fst between mexican lowland and northern flint accessions(Swarts *et al.* 2017) (p-value

299   0.24-0.51), and Fst between 30 tropical and 30 temperate inbreds(Gage *et al.* 2017) (p-value

300   0.03-0.22).

301   Adaptation is complex, and involves modulation of processes beyond just those related to

302   flowering and photoperiod sensitivity. It is possible that we were able to observe an enriched

303   overlap of DDR genes and selected loci in Hallauer's Tusón because the population was directly

304   selected for early flowering time in a single location for ten generations, a less complicated

305   selection history than experienced by natural populations. In the two studies of Fst between

306   temperate and tropical accessions, the lack of enrichment for signals of selection near DDR

307   genes may be because Fst is detecting allele frequency changes related to traits other than

308   photoperiod sensitivity or flowering, or false positives in alleles that have had many more than

15

309   ten generations to drift. The Romero Navarro study of latitude and altitude was limited in its

310   representation of traditional varieties from latitudes that would induce a photoperiod

311   response(Romero Navarro *et al.* 2017).

312   Although there is limited statistical significance for these tests of enrichment, in all instances

313   the DDR genes had greater overlap with selection candidates than the median of all

314   permutations, and across all tests the mean enrichment of selection candidates was 275% over

315   the median value from permutations. P-values from Fisher's combined test for p-values across all

316   selection experiments ranged from 1.4e-5 (all genes) to 9.5e-5 (significantly cycling in two or

317   more inbreds). These results support the hypothesis that DDR genes have been the subject of

318   selection during adaptation to new environments, though the strength and consistency selection

319   seems variable depending on the population.

320

321   Conclusions

322   In this study, we show that allelic variation for diel patterns of transcript abundance are

323   common in maize and differences in diel transcript abundance rhythms appear at least partially

324   explainable by sequence variation in cis-regulatory regions. It is also possible that transcriptional

325   differences observed in this study are caused by differential chromatin accessibility or variation

326   for trans-acting mechanisms. Observed variation in diel patterns of transcript abundance appear

327   to influence photoperiod sensitivity (phenotypic plasticity for flowering time) and may have been

328   targets of selection when maize was adapted to longer day lengths. These results are an example

329   of how selection on sequence variation impacting gene expression can contribute to phenotypic

330   plasticity and adaptation to new environments.

331

16

Materials and Methods

Germplasm and field experiment

We chose 24 of the 26 parents of the maize nested association panel (NAM)(Yu *et al.* 2008;

Gage *et al.* 2020) to assay for diel patterns of transcript abundance. Inbreds were grown in the

field at the Musgrave Research Farm in Aurora, NY in the summer of 2019. Due to different

rates of development between inbreds, we sampled tissue from the first leaf without epicuticular

wax to standardize the developmental stage of sampled tissue (Supplemental File 7). The last

leaf with epicuticular wax was scored on July 16, 2019. Tissue was collected every two hours

from 6am on July 19, 2019 until 4am on July 20, 2019. Approximately 1 cm^2 of leaf tissue was

collected from the center of the leaf blade, roughly measured by bending the tip back to the leaf

base, on one side of the midrib. To avoid assaying wounding response, a different plant was

sampled at each timepoint. If not enough plants were present to sample a different plant at each

timepoint, a small number of timepoints were sampled from previously sampled plants, but from

the opposite side of the midrib. Tissue samples were collected onto liquid nitrogen and kept cold

throughout the experiment on dry ice.

Growth chamber experiment

Four genotypes (three of which overlap with genotypes grown in the field), B73, Mo17, P39,

and CML103, were sown in 24-cell trays in 1:1 Lambert GPM:Turface. Two pots of each

genotype were planted for each of 11 timepoints (0h through 30h), except for the 12h timepoint,

for which four pots of each genotype were planted. Each pot was planted with 2 seeds on April 9,

2019. An additional 44 pots were planted in the same manner with B73 to be used as a border in

17

354   the growth chamber. Experimental (non-border) pots were randomized every 2-3 days to reduce

355   spatial effects. The growth chamber was set to 14-hour days at 26°C and 10 hour nights at 22°C.

356   Tissue sampling occurred when plants were at approximately v3 growth stage, younger than

357   plants sampled in the field experiment. Tissue was collected from the middle of the second leaf

358   from the healthiest plant in each pot starting on April 22, 2019 at 09:00 (~1 hour after lights

359   come on) and continuing every 3 hours until April 23, 2019 at 15:00. Tissue samples were

360   collected onto liquid nitrogen and kept cold throughout the experiment on dry ice.

361

362   RNA-seq and bioinformatic analysis

363   RNA extraction and 3' RNA sequencing was performed by the Genomics Facility at the

364   Cornell Institute of Biotechnology following methods previously described(Kremling *et al.*

365   2018). Cutadapt(Martin 2011) version 2.3 was used to trim 12 bases from the 3' end as

366   recommended by Lexogen, remove bases with quality <20 from the 3' end, remove TruSeq

367   adapters and any poly-A or poly-Ts. Because all the NAM parents have de novo genome

368   assemblies(Hufford *et al.* 2021), read alignment and transcript abundance estimation was

369   performed using the cognate genome assembly for each of the 24 inbred lines to reduce reference

370   bias. Read alignment and counting wer done using Salmon(Patro *et al.* 2017) v1.8 with the

371   options --libType A --validateMappings --noLengthCorrection. Transcript abundance was

372   normalized using the estimateSizeFactors() and counts() functions from the R(R Core Team

373   2018) package DESeq2(Love *et al.* 2014).

374

375 <u>Predicting cycling status from promoter sequence</u>

376   For the field experiment data, each gene in each inbred was evaluated for rhythmicity in R(R

377 Core Team 2018) using the MetaCycle(Wu *et al.* 2016) function *meta2d()*, as well as the

378 diffCircadian(Ding *et al.* 2021) function *LR_rhythmicity()*. Because the *meta2d()* function

379 aggregates rhythmicity test statistics from three different methods, using both MetaCycle and

380 diffCircadian resulted in testing each gene for rhythmic patterns of transcript abundance via four

381 different methods.

382   To model the relationship between promoter sequence and rhythmic diel patterns of

383 transcript abundance, we adapted a modification of the DanQ model(Quang and Xie 2016),

384 originally developed to predict chromatin accessibility. The model architecture consists of

385 convolutional and max-pooling layers followed by a bidirectional LSTM to allow the model to

386 learn DNA sequence motifs and their interactions that are predictive of cyclical transcript

387 abundance patterns.

388   Our model was trained with the sequence from 1000bp upstream to 500bp downstream of the

389 annotated transcription start site of the first transcript ("T001") of each gene as features. Each

390 gene had three continuous labels: relative amplitude computed by MetaCycle, and p-value and

391 R^2 from diffCircadian. We chose these labels to attempt to quantitatively capture the amount of

392 evidence for rhythmic diel transcript abundance patterns.

393   We trained our model for 500 epochs, using all genes from 12 inbreds as training data and all

394 genes from 3 inbreds as validation data. We converted the trained convolutional filters into

395 sequence motifs using methods adapted from(Koo and Eddy 2019). Filter sequence logos were

396 inspected visually and compared to known, previously published Arabidopsis TF binding motifs.

397

19

398   Identification of DDR candidates in field experiment

399   Using the maize PHG v1 (available at https://phg.maizegdb.org, accessed September

400   2024)(Bradbury *et al.* 2022) method name

401   "CONSENSUS_84plusRef_mxDiv_10toNeg3_maxClusers30", we assigned haplotypes for

402   reference ranges that overlapped with the 1000bp upstream of each gene in each inbred. For each

403   pangene, we tested for an association between haplotype ID and the difference in magnitude

404   between maximum and minimum transcript abundance during the 24 hour time course using a

405   Kruskal-Wallis (KW) test. To estimate p-values, we permuted haplotype labels and re-calculated

406   the KW test statistic 10,000 times per reference range.

407   Pangenes were considered to be Differential Diel Regulation (DDR) candidates if the p-value

408   of the KW permutation test was < 0.05, at least six inbreds had significant single-gene cycling

409   tests (p < 0.01), and at least six inbreds had baseline transcript abundance values greater than 50.

410   The additional filtering on single-gene cycling tests ensured that the KW test of change in

411   transcript abundance was including genes with some evidence of diel rhythmicity, and the

412   baseline transcript abundance filtering was to remove noisy, low-expressed pangenes.

413

414   KEGG pathway enrichment

415   Gene ontology enrichment was performed using ShinyGO v0.8.1(Ge *et al.* 2020)

416   (https://bioinformatics.sdstate.edu/go/, accessed December 6, 2024). The web application was

417   used with the Ensembl/STRING-db ID set to "zmays_eg_gene". The 489 DDR genes were used

418   as input, and the background was set to the list of all genes transcribed in at least one NAM

419   parent, in at least one tissue (33,201 genes). All other settings were left at default, and we

420   evaluated results for KEGG pathways.

20

421

422   <u>Identification of DDR genes in growth chamber experiment</u>

423   To robustly filter the DDR genes identified in the field experiment, we tested the 489 field

424   experiment DDR genes for allelic differences in diel transcription in the growth chamber

425   experiment. We grouped P39, B73, and CML103 samples (each genotype replicated twice per

426   timepoint) based on haplotype imputed from the PHG, as described for the field experiment data

427   above. We then tested for differences between haplotypes using the *LR_diff()* function from the

428   diffCircadian package(Ding *et al.* 2021). If all three genotypes had the same promoter haplotype,

429   the gene was not tested. If all three genotypes belonged to two haplotype groups, we simply

430   tested the difference between the two haplotypes. If all three genotypes had three different

431   haplotypes, we performed pairwise tests between all three haplotypes and kept the lowest p-

432   value. 205 genes with a p-value for differences in amplitude less than 0.01 were considered

433   significant and used for GWAS and adaptation enrichment experiments below.

434

435   <u>Enrichment for GWAS results and evidence of adaptation</u>

436   We performed GWAS in the maize Goodman Association Panel (GAP) population, which

437   consists of 282 diverse inbred maize lines(Flint-Garcia *et al.* 2005). Phenotypic data were

438   previously collected (Tian *et al.* 2011; Hung *et al.* 2012a; b) and compiled for ease of

439   use(Khaipho-Burch *et al.* 2023). We followed the methods for fast association described

440   in(Khaipho-Burch *et al.* 2023), using high density genotypes for the GAP on the B73v5

441   reference genome (Grzybowski *et al.* 2023). We chose to test for enrichment of GWAS signal in

442   traits related to flowering time and latitudinal adaptation: growing degree days to flowering for

443   silk (DTS) and anthesis (DTA), and photoperiod response, or the difference between DTS (or

444   DTA) under long days and short days. We also chose to include leaf angle, tassel length, and

445   tassel branch number as control traits that we predicted would not be enriched for GWAS signal

446   in DDR candidates. GWAS was performed using rTASSEL(Monier *et al.* 2022), an R-based

447   front end for TASSEL(Bradbury *et al.* 2007).

448   To compare GWAS p-values from 205 DDR candidates to random samples, we first took the

449   mean of the minimum p-value within each DDR candidate gene +/- 1000bp. This was compared

450   to 1,000 random samples of 205 of genes drawn from a null set. Because results of permutation

451   tests like this can be highly dependent on the choice of null set, we tested four different null sets:

452   1) all annotated gene models; 2) annotated genes with >5 transcripts per million (TPM) in at least

453   one tissue of at least one NAM parent(Hufford *et al.* 2021); 3) annotated genes that show

454   evidence of cycling (p<0.01) in at least two inbreds in this study; 4) annotated genes that show

455   evidence of cycling (p<0.01) in 22 or more inbreds in this study. The second null set was chosen

456   to reduce our sampling of pseudogenes or other non-functional genes relative to the first null set.

457   Null sets #3 and #4 were chosen to capture genes showing segregation for rhythmicity and near-

458   constitutive rhythmicity, respectively.

459   Next, we tested whether DDR candidates showed evidence of enrichment for signatures of

460   selection or adaptive function. SNP variants showing evidence of selection or adaptation were

461   collected from previous studies of: environmental GWAS of maize landrace adaptation to

462   latitude and altitude(Romero Navarro *et al.* 2017); F_{ST} between temperate and tropical maize

463   inbreds(Gage *et al.* 2017); F_{ST} between maize from central Mexican lowlands and Gaspé Flint

464   from Quebec(Swarts *et al.* 2017); and directed, short term evolution in the maize Tuson

465   population for day length adaptation(Wisser *et al.* 2019). In addition, we calculated Fst between

466   30 inbred lines with the lowest photoperiod response values and 30 inbred lines with the highest

467   photoperiod response values in the GAP(Hung *et al.* 2012b) using SNP data from (Grzybowski

468   *et al.* 2023) in VCFtools(Danecek *et al.* 2011) and kept the top 10,000 SNPs as putative

469   adaptation variants.

470   For each set of adaptation candidate SNPs, we first counted the number of adaptation SNPs

471   that overlapped with DDR genes +/-1000bp. We then sampled an equal number of genes from

472   each of our four null sets and counted the number of adaptation candidate SNPs that overlapped

473   with this sample. This sampling procedure was repeated 1,000 times to generate a null

474   distribution, and the p-value for each set of DDR candidates was estimated as the proportion of

475   random samples that had more overlapping adaptation SNPs than the DDR candidate genes.

476

477   <u>eQTL overlap with GWAS hits</u>

478   We performed an additional test for whether changes in cis-regulatory sequences associated

479   with changes in transcript abundance co-localize with GWAS hits for photoperiod sensitivity,

480   flowering time, and control traits. Using daytime and nighttime leaf transcript abundance data

481   from(Kremling *et al.* 2018), we estimated the degree of diel fluctuation in transcripts by

482   subtracting estimates of daytime expression from estimates of nighttime expression. We then

483   performed eQTL mapping in 203 individuals using the same techniques described by(Kremling

484   *et al.* 2018). We then counted the overlaps between cis-eQTL, within 5,000bp of their focal gene,

485   and GWAS hits (variants with p-value < 1e-5) for photoperiod sensitivity, flowering time, leaf

486   angle, tassel length, and tassel branch number. This overlap was compared to 10,000 random

487   draws of the same number of SNPs as cis-eQTL, matched for distance to the nearest gene and

488   distribution of minor allele frequency.

489

Code availability

490 Code will be available upon final publication at https://github.com/joegage/diverse_diel

492

Data availability

494 All sequencing data will be available on NCBI SRI upon final publication. All supplementary

495 data will be made publicly available upon final publication.

496

Acknowledgements

507

References

509 Bradbury P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, *et al.*, 2007 TASSEL:
510 software for association mapping of complex traits in diverse samples. Bioinformatics 23:
511 2633–2635.

512 Bradbury P. J., T. Casstevens, S. E. Jensen, L. C. Johnson, Z. R. Miller, *et al.*, 2022 The Practical
513 Haplotype Graph, a platform for storing and using pangenomes for imputation.
514 Bioinformatics 38: 3698–3702.

24

515 Buckler E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, *et al.*, 2009 The genetic
516   architecture of maize flowering time. Science 325: 714–718.

517 Crawford G. W., D. Saunders, and D. G. Smith, 2006 Pre-contact maize from Ontario, Canada:
518   Context, chronology, variation, and plant association. Histories of maize:
519   Multidisciplinary approaches to the prehistory, linguistics, biogeography, domestication,
520   and evolution of maize 549–559.

521 Danecek P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, *et al.*, 2011 The variant call format
522   and VCFtools. Bioinformatics 27: 2156–2158.

523 Ding H., L. Meng, A. C. Liu, M. L. Gumz, A. J. Bryant, *et al.*, 2021 Likelihood-based tests for
524   detecting circadian rhythmicity and differential circadian patterns in transcriptomic
525   applications. Brief. Bioinform. 22. https://doi.org/10.1093/bib/bbab224

526 Duan K., L. Li, P. Hu, S.-P. Xu, Z.-H. Xu, *et al.*, 2006 A brassinolide-suppressed rice MADS-
527   box transcription factor, OsMDP1, has a negative regulatory role in BR signaling. Plant J.
528   47: 519–531.

529 Duncan O., and A. H. Millar, 2022 Day and night isotope labelling reveal metabolic pathway
530   specific regulation of protein synthesis rates in Arabidopsis. Plant J. 109: 745–763.

531 Ferrari C., S. Proost, M. Janowski, J. Becker, Z. Nikoloski, *et al.*, 2019 Kingdom-wide
532   comparison reveals the evolution of diurnal gene expression in Archaeplastida. Nat.
533   Commun. 10: 737.

534 Flint-Garcia S. A., A.-C. Thuillet, J. Yu, G. Pressoir, S. M. Romero, *et al.*, 2005 Maize
535   association population: a high-resolution platform for quantitative trait locus dissection:
536   High-resolution maize association population. Plant J. 44: 1054–1064.

537 Flis A., V. Mengin, A. A. Ivakov, S. T. Mugford, H.-M. Hubberten, *et al.*, 2019 Multiple
538   circadian clock outputs regulate diel turnover of carbon and nitrogen reserves:
539   Metabolism in five circadian clock mutants. Plant Cell Environ. 42: 549–573.

540 Fornara F., V. Gregis, N. Pelucchi, L. Colombo, and M. Kater, 2008 The rice StMADS11-like
541   genes OsMADS22 and OsMADS47 cause floral reversions in Arabidopsis without
542   complementing the svp and agl24 mutants. Journal of Experimental Botany 59: 2181–
543   2190.

544 Gage J. L., D. Jarquin, C. Romay, A. Lorenz, E. S. Buckler, *et al.*, 2017 The effect of artificial
545   selection on phenotypic plasticity in maize. Nat. Commun. 8: 1348.

546 Gage J. L., B. Monier, A. Giri, and E. S. Buckler, 2020 Ten Years of the maize Nested

25

547   Association Mapping Population: Impact, Limitations, and Future Directions. Plant Cell.
548   https://doi.org/10.1105/tpc.19.00951

Gardiner L.-J., R. Rusholme-Pilcher, J. Colmer, H. Rees, J. M. Crescente, *et al.*, 2021
     Interpreting machine learning models to investigate circadian regulation and facilitate
     exploration of clock function. Proc. Natl. Acad. Sci. U. S. A. 118: e2103070118.

Ge S. X., D. Jung, and R. Yao, 2020 ShinyGO: a graphical gene-set enrichment tool for animals
     and plants. Bioinformatics 36: 2628–2629.

Gendron J. M., J. L. Pruneda-Paz, C. J. Doherty, A. M. Gross, S. E. Kang, *et al.*, 2012
     Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. Proc.
     Natl. Acad. Sci. U. S. A. 109: 3167–3172.

Gnesutta N., R. W. Kumimoto, S. Swain, M. Chiara, C. Siriwardana, *et al.*, 2017 CONSTANS
     Imparts DNA Sequence Specificity to the Histone Fold NF-YB/NF-YC Dimer. Plant Cell
     29: 1516–1532.

Grzybowski M. W., R. V. Mural, G. Xu, J. Turkus, J. Yang, *et al.*, 2023 A common
     resequencing-based genetic marker data set for global maize diversity. Plant J. 113:
     1109–1121.

Hallauer A. R., and M. J. Carena, 2016 Registration of BS39 maize germplasm. J. Plant Regist.
     10: 296–300.

Hu H., T. Crow, S. Nojoomi, A. J. Schulz, J. M. Estévez-Palmas, *et al.*, 2022 Allele-specific
     expression reveals multiple paths to highland adaptation in maize. Mol. Biol. Evol. 39:
     msac239.

Hufford M. B., A. S. Seetharam, M. R. Woodhouse, K. M. Chougule, S. Ou, *et al.*, 2021 De
     novo assembly, annotation, and comparative analysis of 26 diverse maize genomes.
     Science 373: 655–662.

Hung H.-Y., C. Browne, K. Guill, N. Coles, M. Eller, *et al.*, 2012a The relationship between
     parental genetic or phenotypic divergence and progeny variation in the maize nested
     association mapping population. Heredity 108: 490–499.

Hung H.-Y., L. M. Shannon, F. Tian, P. J. Bradbury, C. Chen, *et al.*, 2012b ZmCCT and the
     genetic basis of day-length adaptation underlying the postdomestication spread of maize.
     Proc. Natl. Acad. Sci. U. S. A. 109: E1913-21.

Khaipho-Burch M., T. Ferebee, A. Giri, G. Ramstein, B. Monier, *et al.*, 2023 Elucidating the
     patterns of pleiotropy and its biological relevance in maize. PLoS Genet. 19: e1010664.

26

Khan S., S. C. Rowe, and F. G. Harmon, 2010 Coordination of the maize transcriptome by a conserved circadian clock. BMC Plant Biol. 10: 126.

Koo P. K., and S. R. Eddy, 2019 Representation learning of genomic sequence motifs with convolutional neural networks. PLoS Comput. Biol. 15: e1007560.

Kremling K. A. G., S.-Y. Chen, M.-H. Su, N. K. Lepak, M. C. Romay, *et al.*, 2018 Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. Nature 555: 520–523.

Lambert S. A., A. Jolma, L. F. Campitelli, P. K. Das, Y. Yin, *et al.*, 2018 The Human Transcription Factors. Cell 172: 650–665.

Li Y.-X., C. Li, P. J. Bradbury, X. Liu, F. Lu, *et al.*, 2016 Identification of genetic variants associated with maize flowering time using an extremely large multi-genetic background population. Plant J. 86: 391–402.

Love M. I., W. Huber, and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15: 550.

Lu H., C. R. McClung, and C. Zhang, 2017 Tick Tock: Circadian Regulation of Plant Innate Immunity. Annu. Rev. Phytopathol. 55: 287–311.

Martin M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10–12.

Matsuoka Y., Y. Vigouroux, M. M. Goodman, J. Sanchez G, E. Buckler, *et al.*, 2002 A single domestication for maize shown by multilocus microsatellite genotyping. Proc. Natl. Acad. Sci. U. S. A. 99: 6080–6084.

Mehta D., J. Krahmer, and R. G. Uhrig, 2021 Closing the protein gap in plant chronobiology. Plant J. 106: 1509–1522.

Michael T. P., P. A. Salomé, H. J. Yu, T. R. Spencer, E. L. Sharp, *et al.*, 2003 Enhanced fitness conferred by naturally occurring variation in the circadian clock. Science 302: 1049–1053.

Michael T. P., 2022 Core circadian clock and light signaling genes brought into genetic linkage across the green lineage. Plant Physiol. 190: 1037–1056.

Missra A., B. Ernest, T. Lohoff, Q. Jia, J. Satterlee, *et al.*, 2015 The circadian clock modulates global daily cycles of mRNA ribosome loading. Plant Cell 27: 2582–2599.

Monier B., T. M. Casstevens, P. J. Bradbury, and E. S. Buckler, 2022 rTASSEL: An R interface

610   to TASSEL for analyzing genomic diversity. J. Open Source Softw. 7: 4530.

611 Nagel D. H., C. J. Doherty, J. L. Pruneda-Paz, R. J. Schmitz, J. R. Ecker, *et al.*, 2015 Genome-
612   wide identification of CCA1 targets uncovers an expanded clock network in Arabidopsis.
613   Proc. Natl. Acad. Sci. U. S. A. 112: E4802-10.

614 Patro R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford, 2017 Salmon provides fast and
615   bias-aware quantification of transcript expression. Nat. Methods 14: 417–419.

616 Petersen J., A. Rredhi, J. Szyttenholm, and M. Mittag, 2022 Evolution of circadian clocks along
617   the green lineage. Plant Physiol. 190: 924–937.

618 Piperno D. R., A. J. Ranere, I. Holst, J. Iriarte, and R. Dickau, 2009 Starch grain and phytolith
619   evidence for early ninth millennium B.P. maize from the Central Balsas River Valley,
620   Mexico. Proc. Natl. Acad. Sci. U. S. A. 106: 5019–5024.

621 Qiao Z., W. Qi, Q. Wang, Y. Feng, Q. Yang, *et al.*, 2016 ZmMADS47 regulates Zein gene
622   transcription through interaction with Opaque2. PLoS Genet. 12: e1005991.

623 Quang D., and X. Xie, 2016 DanQ: a hybrid convolutional and recurrent deep neural network for
624   quantifying the function of DNA sequences. Nucleic Acids Res. 44: e107.

625 R Core Team, 2018 R: A Language and Environment for Statistical Computing

626 Romero Navarro J. A., M. Willcox, J. Burgueño, C. Romay, K. Swarts, *et al.*, 2017 A study of
627   allelic diversity underlying flowering-time adaptation in maize landraces. Nat. Genet. 49:
628   476–480.

629 Sanchez S. E., and S. A. Kay, 2016 The Plant Circadian Clock: From a Simple Timekeeper to a
630   Complex Developmental Manager. Cold Spring Harb. Perspect. Biol. 8.
631   https://doi.org/10.1101/cshperspect.a027748

632 Sauer J. D., and P. Weatherwax, 1955 Indian corn in old America. Geogr. Rev. 45: 446.

633 Spensley M., J.-Y. Kim, E. Picot, J. Reid, S. Ott, *et al.*, 2009 Evolutionarily conserved regulatory
634   motifs in the promoter of the Arabidopsis clock gene LATE ELONGATED
635   HYPOCOTYL. Plant Cell 21: 2606–2623.

636 Stewart A. J., S. Hannenhalli, and J. B. Plotkin, 2012 Why transcription factor binding sites are
637   ten nucleotides long. Genetics 192: 973–985.

638 Stitt M., C. Müller, P. Matt, Y. Gibon, P. Carillo, *et al.*, 2002 Steps towards an integrated view of
639   nitrogen metabolism. J. Exp. Bot. 53: 959–970.

640 Stitt M., and S. C. Zeeman, 2012 Starch turnover: pathways, regulation and role in growth. Curr.
641   Opin. Plant Biol. 15: 282–292.

642 Stone J. R., and G. A. Wray, 2001 Rapid evolution of cis-regulatory sequences via local point
643   mutations. Mol. Biol. Evol. 18: 1764–1770.

644 Swarts K., R. M. Gutaker, B. Benz, M. Blake, R. Bukowski, *et al.*, 2017 Genomic estimation of
645   complex traits reveals ancient maize adaptation to temperate North America. Science
646   357: 512–515.

647 Teixeira J. E. C., T. Weldekidan, N. de Leon, S. Flint-Garcia, J. B. Holland, *et al.*, 2015
648   Hallauer's Tusón: a decade of selection for tropical-to-temperate phenological adaptation
649   in maize. Heredity (Edinb.) 114: 229–240.

650 Tian F., P. J. Bradbury, P. J. Brown, H. Hung, Q. Sun, *et al.*, 2011 Genome-wide association
651   study of leaf architecture in the maize nested association mapping population. Nat. Genet.
652   43: 159–162.

653 Torii K., K. Inoue, K. Bekki, K. Haraguchi, M. Kubo, *et al.*, 2022 A guiding role of the
654   Arabidopsis circadian clock in cell differentiation revealed by time-series single-cell
655   RNA sequencing. Cell Rep. 40: 111059.

656 Troyer A. F., and L. G. Hendrickson, 2007 Background and importance of 'Minnesota 13' corn.
657   Crop Sci. 47: 905–914.

658 Tu X., M. K. Mejía-Guerra, J. A. Valdes Franco, D. Tzeng, P.-Y. Chu, *et al.*, 2020
659   Reconstructing the maize leaf regulatory network using ChIP-seq data of 104
660   transcription factors. Nat. Commun. 11: 5089.

661 Vandepoele K., K. Vlieghe, K. Florquin, L. Hennig, G. T. S. Beemster, *et al.*, 2005 Genome-
662   wide identification of potential plant E2F target genes. Plant Physiol. 139: 316–328.

663 Wang P., L. Wang, L. Zhang, T. Wu, B. Sun, *et al.*, 2022 Genomic Dissection and Diurnal
664   Expression Analysis Reveal the Essential Roles of the PRR Gene Family in Geographical
665   Adaptation of Soybean. Int. J. Mol. Sci. 23. https://doi.org/10.3390/ijms23179970

666 Weirauch M. T., and T. R. Hughes, 2010 Conserved expression without conserved regulatory
667   sequence: the more things change, the more they stay the same. Trends Genet. 26: 66–74.

668 Wisser R. J., Z. Fang, J. B. Holland, J. E. C. Teixeira, J. Dougherty, *et al.*, 2019 The Genomic
669   Basis for Short-Term Evolution of Environmental Adaptation in Maize. Genetics 213:
670   1479–1494.

29

671   Wu G., R. C. Anafi, M. E. Hughes, K. Kornacker, and J. B. Hogenesch, 2016 MetaCycle: an
672   integrated R package to evaluate periodicity in large scale data. Bioinformatics 32: 3351–
673   3353.

674   Yu J., J. B. Holland, M. D. McMullen, and E. S. Buckler, 2008 Genetic design and statistical
675   power of nested association mapping in maize. Genetics 178: 539–551.