Circadian clock mechanism driving mammalian photoperiodism

Wood S.H.[1,2,8], Hindle, M.M.[3,8], Mizoro, Y.[1], Cheng, Y. [4,5], Saer B.R.C.[1], Miedzinska K.[3], Christian H.C.[6], Begley, N. [1], McNeilly J.[7], McNeilly A.S.[7], Meddle, S.L. [3], Burt, D.W.[3,4], Loudon A.S.I.[1*]

[1] Centre for Biological Timing, Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, UK

[2] Arctic Chronobiology and Physiology research group, Department of Arctic and Marine Biology, UiT – The Arctic University of Norway, Tromsø NO-9037, Norway

[3] The Roslin Institute, and Royal (Dick) School of Veterinary Studies University of Edinburgh, Roslin, Midlothian EH25 9PRG, UK

[4] UQ Genomics Initiative, The University of Queensland, Brisbane, QLD 4072, Australia

[5] School of Life and Environmental Sciences, Faculty of Science, The University of Sydney, New South Wales, Australia

[6] University of Oxford, Department of Physiology, Anatomy and Genetics, Le Gros Clark Building, South Parks Road, Oxford, OX1 3QX, UK

[7] MRC Centre for Reproductive Health, Queen's Medical Research Institute, Edinburgh EH16 4TJ, UK

[8] These authors contributed equally

*Corresponding author

Abstract

The annual photoperiod cycle provides the critical environmental cue synchronizing rhythms of life in seasonal habitats. In 1936, Bünning proposed a circadian-basis for photoperiodic synchronization. Here, light-dark cycles entrain a circadian rhythm of photosensitivity, and the expression of summer or winter biology depends on whether light coincides with the phase of high photosensitivity. Formal studies support the universality of this so-called coincidence timer, but we lack understanding of the mechanisms involved. Here we show in mammals that coincidence timing takes place in the *pars tuberalis* of the pituitary, through a melatonin-dependent flip-flop switch between circadian transcriptional activation and repression. Long photoperiods produce short night-time melatonin signals, leading to induction of the circadian transcription factor *BMAL2*, in turn triggering summer biology through the eyes absent / thyrotrophin (EYA3 / TSH) pathway. Conversely, short photoperiods produce long melatonin signals, inducing circadian repressors including *DEC1*, in turn suppressing BMAL2 and the EYA3/TSH pathway, triggering winter biology. These actions are associated with progressive genome-wide changes in chromatin state, elaborating the effect of the circadian coincidence timer. Hence, circadian clock interactions with pituitary epigenetic pathways form the basis of the mammalian coincidence timer mechanism. Our results constitute a blueprint for circadian-based seasonal timekeeping in vertebrates.

INTRODUCTION

The annual photoperiod cycle provides a critical predictive environmental cue driving annual cycles of fertility, physiology and behaviour in most animal species. In 1936, Erwin Bünning proposed the "external coincidence" hypothesis for a circadian-based timing mechanism driving photoperiodic responses in plants [1]. The proposition was that photoperiod entrains a circadian rhythm of photosensitivity, and the expression of summer or winter biology depends on whether or not light coincides with the phase of high photosensitivity. An "internal coincidence" model has also been proposed where the role of light is to entrain two circadian oscillators and the phase relationship between the two oscillators determines the response to photoperiod [2,3]. In either case the role of the circadian clock is central and formal studies support the universality of a "coincidence timer" in animals [2–7], but we lack understanding of the mechanisms involved.

In mammals, the duration of the night-time pineal melatonin signal is sculpted by the photoperiod and is a critical regulator of annual cycles of reproduction, growth and metabolism [8–14]. The melatonin signal is decoded within the *pars tuberalis* (PT) of the pituitary gland [9,13,15–18]. Day-length dependent changes in hypothalamic thyroid hormone (TH) metabolism control the seasonal changes in physiology and are regulated via the PT, through altered secretion of thryrotrophin (TSH) [19–21]. This TSH circuit is specific to the PT, and distinct from that in the anterior pituitary, which controls normal thyroid gland function. LP-activation of PT-TSH depends on regulation of TSHβ subunit expression by the transcriptional co-activator, EYA3, which operates as a TSH on-switch. EYA3 co-activates the PAR bZIP Transcription Factor TEF (thyrotroph embryonic factor) via a D-box element on the *TSHβ* promoter [22–24] (Fig. 1a). This prior work places EYA3 at the centre of photoperiodic time measurement within the melatonin target tissue, but does not explain how expression is regulated by the seasonal clockwork.

In addition to acute circadian-based mechanisms an epigenetic basis to longer-term seasonal control in flowering in plants is well established. Specifically, the duration of cold temperatures during the winter season alters chromatin

77 accessibility at key genes, and is a requirement for full photoperiod induced
78 flowering (vernalisation) [25–27]. In animals, no evidence for dynamic genome-
79 wide seasonal epigenetic regulation of transcription has been described.

80 Here, we show that coincidence timing depends on a flip-flop switch between
81 the expression of the circadian genes *BMAL2*, (a paralogue of the circadian
82 regulator *BMAL1*) and *DEC1,* and their respective activating or suppressive
83 effects on *EYA3* in the PT. The duration of the melatonin signal is key in
84 sculpting these molecular components. Additionally, the effect of the
85 coincidence timer is elaborated by progressive genome-wide changes in
86 epigenetic status at key seasonal gene promoters. Therefore, circadian clock
87 interactions with pituitary epigenetic pathways form the basis of the mammalian
88 coincidence timer mechanism.

RESULTS

Epigenetic regulation of the seasonal transcriptome within the Pars tuberalis

Seasonally breeding sheep are a well-established photoperiodic model for the study of neuroendocrine mechanisms underpinning seasonal physiology [14,28,29]. Using a study design that compared the effects of transfer from long photoperiod (LP) to short photoperiod (SP) with transfer from SP to LP (Fig. 1b), we collected pars tuberalis (PT) tissue at 1,7 and 28 days after transfer, with collections timed for 4-h after lights on (ZT4), when *EYA3* expression peaks under LP [22,24,29]. Changes in prolactin concentrations [29–31] confirmed photoperiodic hormone responses, as well as LP activation of *EYA3,* and inverse expression patterns of *TSHβ* (LP marker) and *CHGA* (SP marker) (Fig. 1b, Supplementary Fig. 1a-d, n=4), validating this paradigm.

Comparing LP day 28 to SP day 28 by electron microscopy we found an increased nuclear diameter, equating to an approximate doubling in volume of PT thyrotrophs, and a marked reduction in chromatin density on LP (Fig. 1c, Supplementary Fig. 1e & f). These morphological changes were not seen in *Pars distalis* (PD) somatotrophs (Fig. 1c, Supplementary Fig. 1e & f). We hypothesized that increased chromatin accessibility may account for the photoinductive effects of LP and therefore changes in the seasonal transcriptome.

Comparing all seasonal time-points (Fig. 1b) we performed ChIP-seq (histone marker H3K4me3, Supplementary Table 1 & 2, n=2) and RNA-seq (Supplementary Table 3, n=3) to screen for seasonal transcriptional activation. To determine if epigenetic changes in H3K4me3 marks were associated with transcriptional activation we first improved the sheep genome annotation of transcripts and transcription start sites (TSS) with PT RNA and a combination of “Cap Analysis of Gene Expression” (CAGE-seq, Supplementary Table 4) [32] and ISOSEQ long-read RNA-seq (SRA: PRJNA391103, see methods for details). Using this improved annotation we identified H3K4me3 peak distribution on genomic features to be similar to previous studies [33–35], validating the approach (see methods section).

Next we identified seasonally expressed genes, as defined by RNA-seq analysis of differentially regulated genes (DEGs) in the SP to LP and LP to SP transfers (Supplementary Table 3), and observed a strong correlation between seasonal gene expression and H3K4me3 peaks around the transcription start sites (TSS's)(Fig. 1d). Importantly, this correlation was absent in non-seasonally regulated genes (Fig. 1d). Histone modifications are precisely balanced by methyltransferases ("writers"), demethylases ("erasers") and effector proteins ("readers"), therefore we checked the RNA expression of H3K4me3 readers, writers and erasers but found no seasonal changes (Supplementary Table 5). This suggests that changes in protein activity of H3K4me3 modulators may be key in the observed seasonal alterations in H3K4me3 marks.).

We noted that approximately 70% of seasonally DEGs (Supplementary Table 3) had more than one TSS compared to only ~20% of the PT genomic background (Supplementary Fig. 1g, Supplementary Table 4). Next we took genes that were up-regulated in either SP or LP and plotted the proportion of genes with multiple TSS's, and repeated this for H3K4me3 marked TSS's, this revealed that H3K4me3 marks are more likely to occur on genes with multiple TSS's (Supplementary Fig. 1h) and highly expressed seasonal DEGs have a greater prevalence of multiple TSSs than non-seasonal genes expressed at the same level. Furthermore, this phenomenon was more pronounced in LP than SP. This indicates an enrichment of multiple TSS's in LP up-regulated genes which is associated with the H3K4me3 mark.

To investigate which promoter motifs may be actively transcribed we searched for enriched transcription factor binding site motifs in the TSS's marked with H3K4me3 in the LP day 28 vs SP day 28 comparison (Fig. 1e, Supplementary Fig. 2a & b). This showed enrichment of D-box binding motifs, Basic Leucine Zipper ATF-Like transcription related factors and POU domains for in both the up and down-regulated genes (Fig. 1e).

TSHβ, an exemplar seasonal D-box regulated gene [22,23], revealed progressive activation by LP, correlated with expression, and with no detectable H3K4me3 mark at either SP day 28 or 84 (R=0.961, p-value=0.002, Fig. 1f, g &

Supplementary Fig. 1b). In contrast, CHGA revealed an inverse seasonal pattern (R=0.843, p-value=0.009, Supplementary Fig. 2c, d, Supplementary Fig. 1b & c). Global enrichment of D-box sites on LP-activated genes indicates a potentially extensive role for TEF, SIX1 and EYA3 co-activation in seasonal regulation of physiology, and strongly focuses attention on the regulation of EYA3.

In line with our global analysis showing the presence of multiple TSSs in seasonally expressed genes we identified two transcription start sites in *EYA3*, an up-stream TSS with non-canonical E-boxes and a downstream TSS containing two paired canonical E-Boxes (Fig. 1h). CAGE analysis revealed that only the down-stream TSS was actively transcribed on LP (Fig. 1h, Supplementary Table 4). We correlated RNA expression with H3K4me3 marks for both TSS's, which confirmed seasonal regulation specific to the down-stream TSS (EYA3 downstream TSS: R=0.778, P=0.02, Fig. 1h & i, EYA3 upstream TSS: R=0.308, P=0.46, Supplementary Fig. 2e & f). Next, we cloned each EYA3 TSS into luciferase reporters, and using COS7 cells transfected the reporters along with known E-box regulators [22,23] (see methods for details). This revealed significant activation specific to the downstream (seasonal) TSS (Supplementary Fig. 2g), likely due to the presence of multiple canonical e-box pairs.

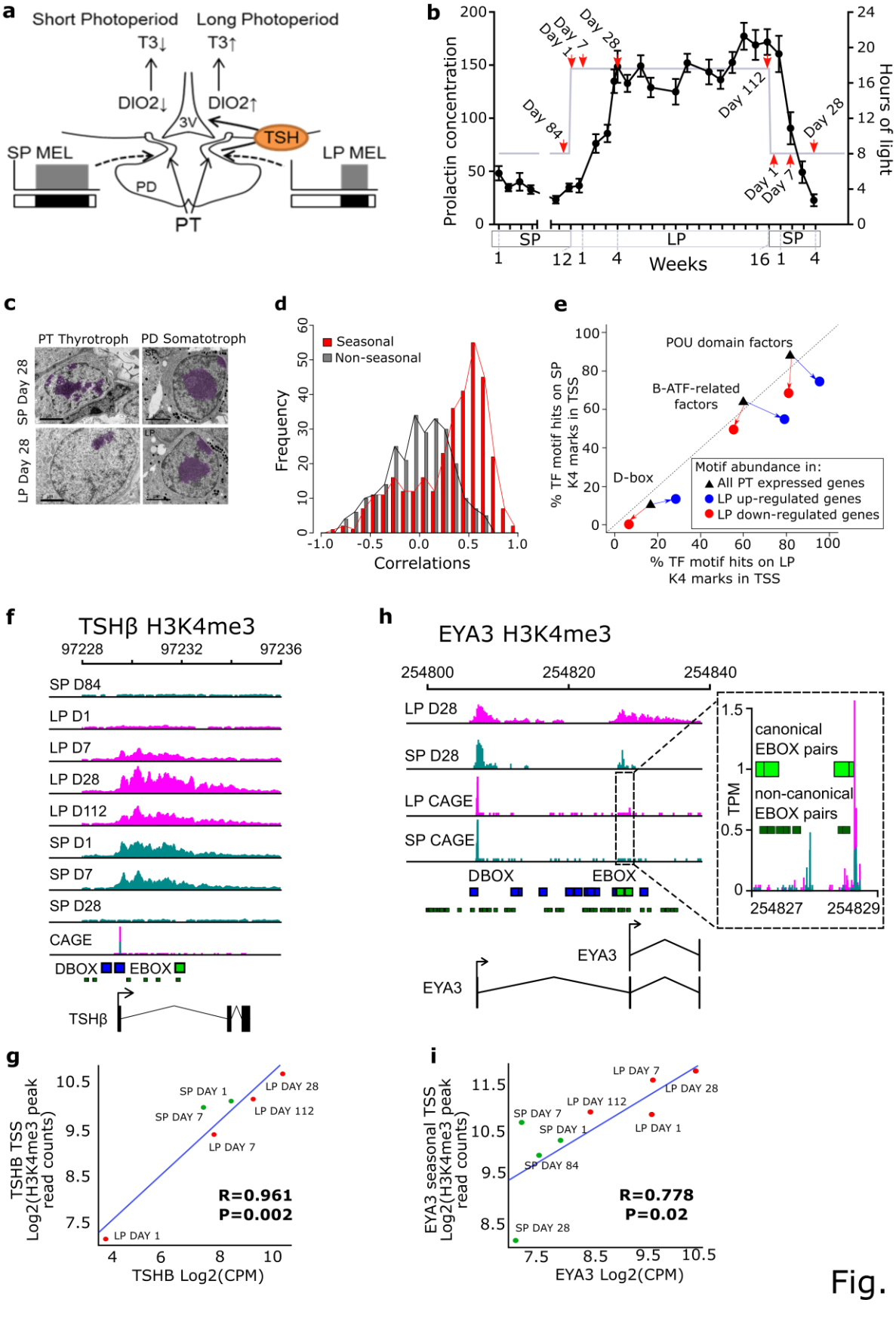


Fig. 1

Figure 1: Photoperiod dependent epigenetic regulation of transcription in the *Pars tuberalis*

a. Current model for the mammalian photoperiodic circuitry. Rhythmic melatonin production from the pineal gland represent short (winter) photoperiods are represented by increased duration of melatonin and long (summer) photoperiods by short duration melatonin. The prime site of melatonin action is the pituitary pars tuberalis driving changes in thyroid hormone availability via day length dependent changes in TSH production. EYA3 is an LP expressed gene that is a strong coactivator of TSHβ expression in the pars tuberalis in synergy with TEF, SIX1 and DBP though D-box binding.

b. Study sampling points indicated by red arrows. Grey line represents number of hours of light the animals received in 24 hour modulo, all animals were sampled at ZT4 except at day 28 where animals in both LP and SP were collected across the day at 4 hour intervals (see data in Figure 3). The black line with error bars corresponds to the prolactin concentrations in plasma for 30 animals. The first double line gap in the graph indicates the gap in weeks between sampling during SPs. Error bars represent the SEM.

c. EM images of PT thyrotrophs and PD somatotrophs at LP day 28 and SP day 28. Dense chromatin in the nucleus is false coloured in purple. Black scale bars = 2um.

d. Histogram revealing frequency distributions of Pearson correlation coefficients between RNA expression (log_2 CPM) and H3K4me3 peak read counts ±200 bp from TSSs (log_2 read counts). Red bars are seasonally expressed genes (log_2 fold change ≥ 1 or ≤ -1 and adjusted p value < 0.05 of SPday84 vs LPday1, 7, 28 and LPday112 vs SPday1, 7, 28 shown in Fig. 2a, differentially expressed genes (DEGs)=480, duplicates were removed) and black bars are non-seasonally expressed genes (log_2 fold change ≤ 0.1 and ≥ -0.1 from the same pairwises above, number of genes=218, duplicates were removed).

e. Plot of the over-represented motifs in the active promoters of both LP day 84 up and down-regulated genes, as compared to SP day 84. The axis plot the percentage of inferred active promoters, containing one or more observed motifs, for a given cohort of genes. Active promoters are defined as contiguous H3K4me3 marked regions within 100bp of a TSS. The black triangles represents the motif abundance on active promoters in all the genes expressed (>0 CPM) in the pars tuberalis (PT) and the percentage of their active promoters containing a binding motif for LP (x-axis) and SP (y-axis) H3K4me3 environments. The blue and red circles are the motif coverage of activate promoters in up-regulated and down-regulated LP genes respectively. See supplementary figure G and H for all motifs that are significantly enriched (FDR < 0.05; fishers two-way exact test) in the active promoters of either up OR down-regulated genes.

f. ChIP-seq tracks for *TSH*β gene H3K4me3 peaks across all experimental time-points. Chromsome 1 region is shown. CAGE-seq track identifying active TSS's in LP day 28 vs SP day 28. Pink represents samples in long photoperiod and marine green represents short photoperiod. Solid green boxes are canonical E-box motifs. Blue boxes are D-box motifs.

g. Correlation plot for *TSH*β log2 H3k4me3 peaks from ChIP-seq versus *TSHb* log2 counts per million (CPM) from RNA-seq. Red symbols are LP sampling points, green are SP sampling points. Correlation coefficient R is shown. R=0.961, p-value=0.002. Note: SP day 28 and 84 are not included because the H3k4me3 peaks are very low.

h. ChIP-seq tracks for *EYA3* gene H3K4me3 peaks. Chromsome 2 region is shown. CAGE-seq track identifying active TSS's in LP day 28 vs SP day 28. Zoom in box provided to identify the downstream promoter.

i. Correlation plot for *EYA3* downstream TSS log2 H3K4me3 peaks from ChIP-seq versus *EYA3* log2 counts per million (CPM) from RNA-seq. Red symbols are LP sampling points, green are SP sampling points. Correlation coefficient R is shown. R=0.778, p-value=0.02.

174

BMAL2 – the missing circadian component in the photoperiodic response

The E-box regulators *CLOCK* and *BMAL1* do not show significant changes in amplitude in the PT under different photoperiods [22,36], therefore we aimed to identify a candidate circadian E-box regulator of the EYA3 downstream TSS. The prediction is that a circadian E-box regulator, would peak in expression only when light falls on the photosensitive phase, as in the early light phase of LP (approx. ZT4), and would be absent on SP where the photosensitive phase is masked by darkness. The RNA-seq dataset revealed progressive up-regulation by LP of multiple transcripts, and a slower inverse pattern on SP (Fig. 2a, Supplementary Table 3).

Importantly these data identified *BMAL2*, an E-box regulator and paralogue of BMAL1, as progressively up-regulated on LP (Fig. 2a & b), closely matching the expression profile for *EYA3* (Fig. 2b). We tested whether BMAL2 was capable of activating the downstream *EYA3* TSS (Supplementary Fig. 2h). In the presence of CLOCK, or alone, BMAL2 had a weak non-significant effect, but in the presence of BMAL1 and CLOCK, we observed significant 4 to 5-fold activation (Supplementary Fig. 2h, Fig. 2c). We then tested whether this depended on direct E-box binding by BMAL2, by mutating the BMAL2 DNA-binding domain (arginine of basic helix-loop-helix domain to alanine R88A [37]), and observed near-identical augmentation of *EYA3* expression (Fig. 2d). The observation that transcription factors can act as co-activators [38] and that bHLH circadian transcription factors require internal PAS domains for functional protein-protein interactions [39], with PAS-B an essential domain for BMAL1 protein-protein interactions [40] led us to mutate the PAS-B domain on ovine BMAL2 (F427R_V439R) [40]. This significantly impaired *EYA3* activation, blocking the co-activation effect with BMAL1 and CLOCK (Fig. 2e). This suggests that BMAL2 operates as a co-activator of *EYA3*, in a CLOCK/BMAL1-dependent manner, requiring PAS-B dependent protein-protein interactions for the mechanism of action.

We assessed the daily profile of *BMAL2* by in-situ hybridization, using archived material [24]. This revealed strong LP-dependent induction of *BMAL2* in the early light phase (ZT3, Fig. 2f), and flat low 24h expression on SP. Collectively this

reveals that BMAL2 operates as a co-activator driving photoperiodic responses, and fulfils the criteria as a candidate activating arm of the circadian-based coincidence timer.

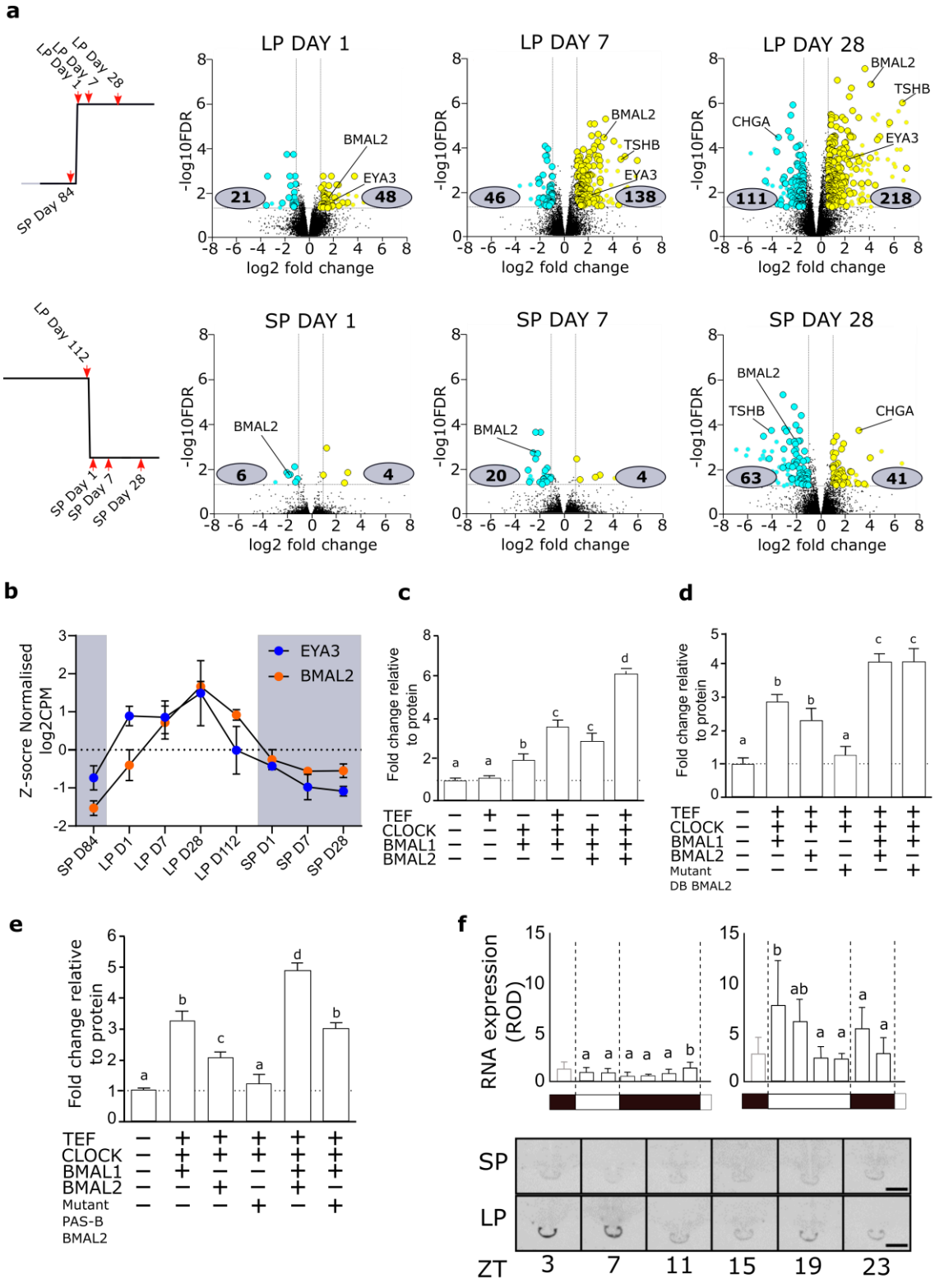


Fig. 2

Figure 2: Identification of BMAL2

a. Volcano plots showing the number of genes up (yellow) and downregulated (blue) in pairwise comparisons. The design for the comparisons is shown on the left. These plots demonstrate the consistency of the BMAL2 seasonal signal. Numbers in grey boxes are the differentially expressed genes in the pairwise comparisons (Supplementary Table 3).

b. Z-score normalized log2CPM RNA-seq plots for EYA3 (blue) and BMAL2 (orange) across the whole experiment. Statistcal significances are found in Supplementary Table 3.

c. Transactivation of the EYA3-downstream-TSS-luc reporter by TEF, CLOCK, BMAL1 and BMAL2. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

d. Transactivation of the EYA3-downstream-TSS-luc reporter by TEF, CLOCK, BMAL1 and BMAL2. The lack of effect of mutating the DNA binding domain of BMAL2 is shown. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

e. Transactivation of the EYA3-downstream-TSS-luc reporter by TEF, CLOCK, BMAL1 and BMAL2. The effect of mutating the PAS-B domain of BMAL2 is shown. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

f. In situ hybridization and quantification for BMAL2 mRNA from archived material from collected every 4 hours from SP and LP. Representative images are shown (n = 4). Error bars represent the SD. Statistical analysis by one-way ANOVA performed, different letters indicate significant differences between groups ($P < 0.01$).

A repressive arm of the circadian clock in SP shuts down the LP transcriptome

A prediction from a coincidence timer model is that repressive mechanisms would emerge on short photoperiods in the late night (i.e. photo-inducible phase masked by darkness). To assess this we conducted a serial RNA-seq experiment on PT tissue collected over at 4-hour intervals from both SP and LP-housed animals (Day 28 Fig. 1b & Fig. 3a). Transcripts at the two ZT0 collections 24h apart (ZT0 and ZT24) in each photoperiod had virtually identical RNA-seq profiles, validating dissection technique and down-stream analyses (Supplementary Fig. 3a & Supplementary Table 6). We identified transcripts

that significantly changed across the day (diurnal genes) within each photoperiod, and defined the phase, amplitude and period [41] (Supplementary Table 6). More transcripts were rhythmic on SP vs LP (880 vs 643 for SP and LP respectively (Fig. 3b), with striking asymmetry in the peak phase of expression on both SP and LP (Fig. 3c). Large numbers of genes were expressed on SP in late night (ZT20, 24) and on LP in the early light phase (ZT4; Fig. 3c), and surprisingly relatively few genes common to both photoperiods (250 transcripts, Fig. 3b & d), but shared genes included the canonical clock genes (Supplementary Fig. 3b).

Amongst diurnal genes shared across photoperiods, dawn-peaking clusters were enriched for the GO term cAMP binding activity, while the dark-onset cluster was marked by the circadian suppressor *CRY1* [42–44] (Fig. 3d & e, Supplementary Fig. 3c). Previous studies have shown *CRY1* is directly induced by onset of melatonin secretion [17,45–47] and that the pattern of induction does not alter between photoperiods. We also observed that regulators of CRY1 protein stability, *FBXL3* and *FBXL21* [48,49], are expressed in the PT, but here there were marked photoperiod-dependent changes in phasing for *FBXL21* (Fig. 3e) .

BMAL2 emerged as the only transcription factor, which showed both phase and amplitude changes in expression on LP (Fig. 3e, Supplementary Table 6). On LP there was a 4 log-fold increase in expression of *BMAL2*, with bi-phasic peaks at ZT4 and ZT20. In contrast, *BMAL1* showed no alteration in amplitude of expression, on either SP or LP (Fig. 3e).

At SP ZT20 we noted a strong functional enrichment for the terms negative regulation of transcription, E-box binding and co-repressor activity (Supplementary Fig. 3c & d). Within this group, we identified 8 genes with known circadian repressor function (*REVERB-α, DEC1, DEC2, CHRONO, FBXL21, KLF10, JUNB, GATA2* and *ERF,* Fig. 3e & Supplementary Fig. 3e). Notably 6 are circadian clock components acting through E-boxes. Cistromic analysis identified over-representation of circadian RORE, D-Box and E-Box sites on the proximal promoter regions in late-night genes on SP (ZT20) and on LP at the dark-light transition (Fig. 3f). Cluster analysis revealed 4 major

clusters of E-box motif-enriched genes associated with late night and dawn transition on SP (ZT20) and LP (ZT4) respectively, with *BMAL2* contained within the LP ZT4 cluster (Fig. 3g), and circadian repressor elements within the SP ZT20 peaking cluster (Fig. 3g). This led us to ask; are SP ZT20 repressors more likely to interact, and therefore potentially repress, LP ZT4 up-regulated genes. We used curated experimental protein-protein interaction (PPI) observations from the STRING database, which contain known protein-protein interactions and functional associations [50]. We found that LP ZT4 up-regulated genes are more enriched within the SP ZT20 repressor network (P-value = 0.001) than down-regulated genes (Supplementary Fig 4a, b). We did not find significant enrichment when considering genes that are differentially up or down regulated across the whole day (P-value = 0.25) (Supplementary Fig. 4b). This suggests that SP is defined by an up-regulation of transcriptional repressors responsible for the suppression of the LP activated transcriptome.

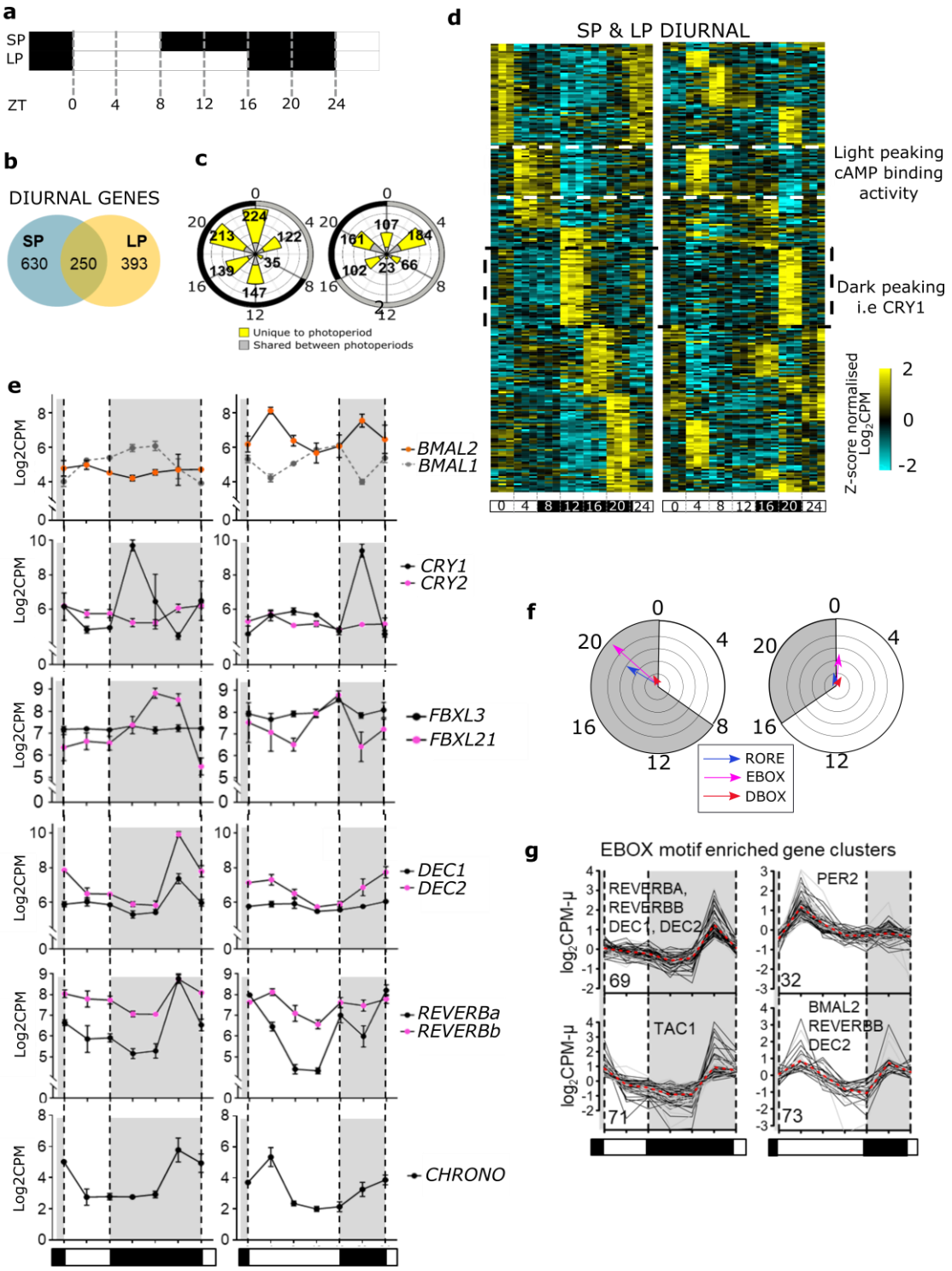


Fig. 3

Figure 3: Short photoperiods are defined by the expression of transcriptional repressors

a. Experimental design shown, at SP day 28 (8:16) and LP day 28 (16:8) the pars tuberalis was collected at 4 hour intervals for 24 hours, zeitgeber times (ZT) are given. N= 3 per time-point.

b. Venn diagram showing the number of diurnal genes (significantly changing throughout the day) in the PT in LP and SP (Supplementary Table 4).

c. Rose plot showing the significant diurnal genes peaking at a particular phase in LP and SP. On the rose plot bars grey indicates genes shared between photoperiods and yellow indicates they are unique to a photoperiod. Around the outside of the plot the durations of light and dark the animals received is plotted in grey and black respectively, the ZT times are also given.

d. Heatmap showing the expression profile of the genes diurnal in both SP and LP (250 genes) for each individual (n=3 per timepoint). Data are ordered by SP peak phase. Yellow is up-regulated and blue down-regulated, the data are scaled. Light dark bars are ZT times are given at the bottom of the heatmap.

e. RNA-seq log2 CPM plots. Light dark bars shown a the bottom and indicated on the graph by dotted line and grey shading. Error bars are SD. Statistical significances are in Supplemental Table 6.

f. Transcription factor binding site analysis by peak phase in SP and LP. Grey and white shading show transition from dark and light and the ZT times are given. Blue arrows represent enrichment of RORE sites in genes peaking at the indicated phase. The pink arrow represents E-boxes and the red D-boxes. The direction of arrows indicate the mean expression peak of genes containing the motif in their promoter and the length of the arrow is the -log10 p-value from a Rayleigh tests of uniformity.

g. Cluster analysis of the RNA-seq data identified ZT20 SP peaking and ZT4 LP peaking gene clusters that were enriched for E-boxes. Plots shown are mean normalised log2CPM of expression profiles to visualise cluster trends. The number of genes in each cluster is indicated in the bottom left of each plot. The light dark bars are shown at the bottom and indicated by dotted lines and grey shading. The broken red line represents the most representive medoid gene in the cluster.

271

Melatonin-regulated circadian repression

Earlier studies have defined the role for melatonin both as an acute inducer of *CRY1* [17,36,45,47], but also determined clear melatonin duration-dependent effects for photoperiodic responses [8,51]. Our data suggest that prolonged SP-like signals may trigger late-night induction of a separate cohort of circadian repressors genes. To test this, we maintained a cohort of animals on LP and then transferred animals to constant light (LL), a regimen known to suppress the endogenous rise in pineal melatonin secretion [52]. Simultaneously, and at the time of normal lights off (i.e. expected onset of the melatonin signal) we treated animal with an intradermal "Regulin" melatonin implant [47], which mimics the endogenous dark-onset rise of this hormone [53], validated by RIA for melatonin (Supplementary Fig. 5a, n=6). PT tissue was collected for in-situ hybridization analysis at +1.5 hours (ZT17.5), +3.5 hours (ZT19.5), +6.5 hours (ZT22.5) and +9.5h after hormone treatment, the latter time point being equivalent to 1.5 hours into the predicted light-onset phase on LP (i.e. ZT1.5; Fig. 4a, n=4).

CRY1 expression in the PT was completely suppressed by constant light, but rapidly induced by melatonin treatment, as previously observed [17,45–47] (Fig. 4b). By ZT1.5, expression levels had dropped to basal. In LL conditions *DEC1, REVERBα* and *CHRONO* were rapidly elevated in the early "subjective dark" phase, and in an abnormal pattern when compared to endogenous profiles in LP conditions (Fig. 3e vs Fig 4c-e). In marked contrast, exposure to melatonin initially repressed all 3 genes, which peaked at +9.5h, equivalent to ZT1.5 on LP (Fig.s 4c - e). For *DEC1* and *CHRONO*, this pattern closely resembled the phase of the endogenous rise of these genes in animals housed in SP photoperiods (Fig. 3e & 4f). From this, we conclude that multiple repressor arms of the circadian clock are direct melatonin targets and act as a read-out of long duration melatonin signals. This defines a molecular basis for a hormone-regulated circadian-based coincidence timer mechanism, capable of discriminating LP and SP responses within the melatonin target tissue.

DEC1 suppression of EYA3 activation by BMAL2

DEC proteins are known E-box suppressors of the circadian clock [54], so we selected DEC1 for further analysis. Quantitative in-situ hybridization confirmed the temporal pattern of *DEC1* expression (Fig. 4f) in the PT, and using immunohistochemistry of short-photoperiod derived PT tissue, we showed co-localisation of DEC1 protein to αGSU expressing thyrotroph cells in the PT (Fig. 4g). DEC1 appears to be specific to thyrotrophs, and was not detected in a common cell-type in the PT, the folliculo-stellate cells. We cloned ovine *DEC1*, and then tested action on BMAL2-mediated induction of *EYA3.* This showed that DEC1 significantly suppressed the action of both CLOCK /BMAL1 (Supplementary Fig. 5b & c) and CLOCK/BMAL1/BMAL2-mediated expression of *EYA3* (Fig. 4h, Supplementary Fig. 5b & c). We also tested DEC2, which was without effect (Supplementary Fig. 5d). Hence, DEC1 and BMAL2 exert mutually antagonistic effects on EYA3 expression.

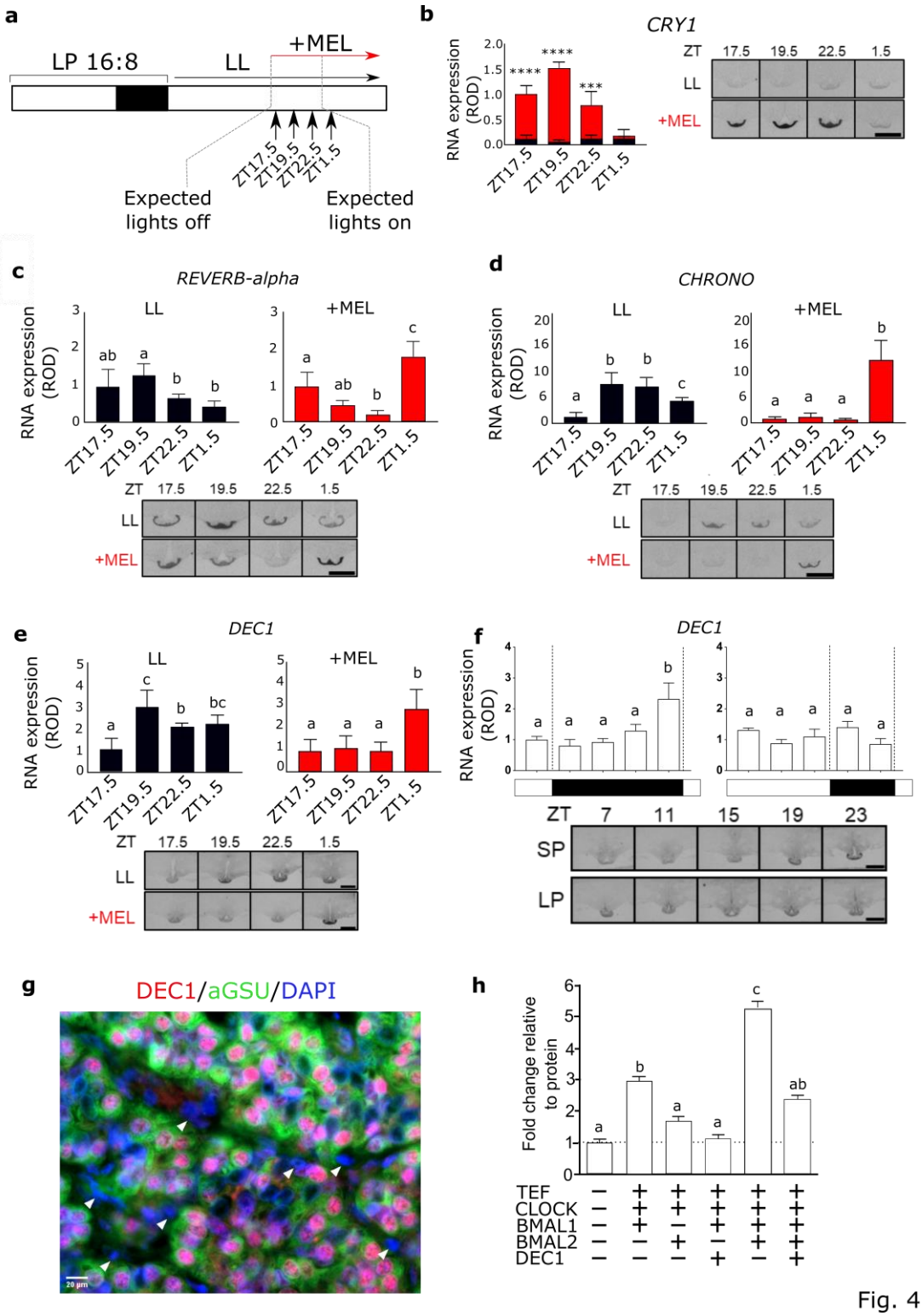


Fig. 4

Figure 4: Melatonin duration defines the expression of repressor genes

a. Study design. Animals were sampled at 1.5 hour intervals with either a sham implant or melatonin implant. LL= constant light.

b. In situ hybridization and quantification for *CRY1* mRNA. Red bars are the value in the presence of melatonin, black bars are in constant light with a sham implant. Representative images are shown (n = 4). Statistical significances from a one-way ANOVA and are between the melatonin and sham implant at each time-point. Error bars represent the SD. P-value; * = <0.01, **=<0.001, ***=<0.0001, ****=<0.00001

c. In situ hybridization and quantification for *REVERb-alpha* mRNA. Red bars are the value in the presence of melatonin, black bars are in constant light with a sham implant. Representative images are shown (n = 4). Statistical analysis by one-way ANOVA performed, different letters indicate significant differences between groups ($P < 0.01$).

d. As in c for *CHRONO* mRNA.

e. As in c for *DEC1* mRNA.

f. In situ hybridization and quantification for *DEC1* mRNA from archived material from collected every 4 hours from SP and LP. Representative images are shown (n = 3). Error bars represent the SD. Statistical analysis by one-way ANOVA performed, different letters indicate significant differences between groups ($P < 0.01$).

g. Double immunohistochemistry for DEC1 (red), aGSU (green) and dapi nuclear stain (blue). White arrow heads indicate FS cells. Scale bar 20um.

h. Transactivation of the EYA3-downstream-TSS-luc reporter by TEF, CLOCK, BMAL1 and BMAL2 is repressed by DEC1. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

319

Discussion

Our study reveals a circadian coincidence timer within the PT, sculpted by the nocturnal melatonin signal that encodes the mammalian photoperiodic response (Fig. 5). Under LP, *BMAL2* exhibits a high-amplitude peak timed approximately 12h after the onset of the preceding dark phase, coincident with expression of *EYA3* in the early light phase [22]. This 12h interval from dark onset to the dawn peak is also remarkably close to the critical photoperiod required to activate a long-day response in sheep [28,55]. We show that BMAL2 acts as a co-activator of *EYA3* with CLOCK and BMAL1, likely in a complex through the PAS-B domain interaction [56] . Co-activators are recognized as important rapid-response functional integrators of multiple transcription factors, driving distinct biological programmes and environmental responses, including adaption to cold, rapid diet change and disease [38]. Within the PT, BMAL2 and EYA3 may therefore operate as a photoperiodic co-activator cascade. It remains unclear how *BMAL2* is regulated, and current models for the regulation of BMAL2 by BMAL1 may not be applicable [57]. Our promoter motif analysis does indicate the presence of E-boxes in BMAL2 but this presents a circular argument and further work on the regulation and evolution of BMAL2 function is required.

While we cannot differentiate between an internal or external coincidence model of photoperiodic time measurement in this study, generally the concept of a photoperiodic coincidence timer predicts that on short photoperiods, repressor activity would dominate in the late night. In line with this we show that long-duration melatonin signals elicit repressor gene transcription in the late night, timed approximately 12h after dark onset/rise of melatonin (ZT20 SP). Impressively, these repressors are directly implicated in the negative regulation of LP ZT4 induced genes (Supplementary Fig. 4). Amongst these, we show the circadian repressor DEC1 [58] blocks induction, by BMAL2, of *EYA3*. Therefore, DEC1 and BMAL2 act as a circadian flip-flop switch for photoperiodic time measurement (Fig. 5). It remains unclear whether DEC-mediated repression is via a direct action on E-box sites occupied by BMAL2 or indirect, leading to modification of a co-activator complex. The discovery of transcriptional repressors in the latter half of the night on SP contrasts with previously described acute and photoperiod-independent induction of *CRY1* by melatonin

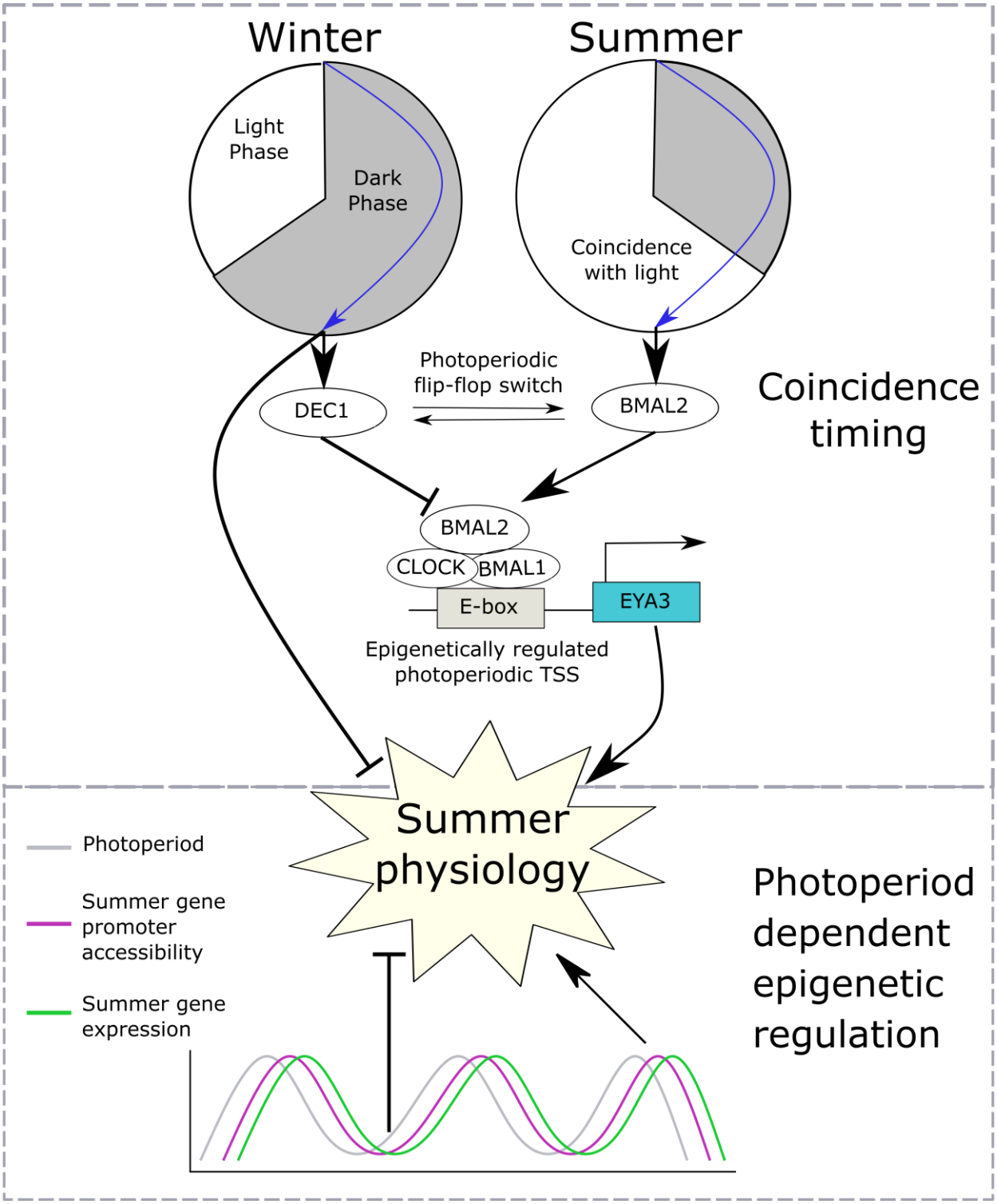
353 onset [17,36,45,47]. However, constant light and therefore CRY1 repression seems
354 to induce the “late-night repressor genes”, suggesting that these are direct
355 targets of CRY1-mediated repression in the early to mid nocturnal phase. This
356 observation along with altered phasing by photoperiod of the CRY1 protein
357 stability regulator FBXL21 [48,49] indicate that the dynamics of protein degradation
358 of CRY1 could play a role in discriminating melatonin signal duration.

359 Our study reveals a progressive increase in H3K4me3 marks at transcription
360 start sites of seasonally expressed genes, especially in long photoperiods (Fig.
361 1d & Supplementary Fig. 1h). Furthermore, these long photoperiod induced
362 genes are more likely to have multiple transcription start sites marked with
363 H3K4me3 (Supplementary Fig. 1g & h). *EYA3* is a particular example of this
364 showing seasonal regulation of only one out of two transcription start sites,
365 presumably seasonally modulating the transcription of EYA3 via BMAL2. On
366 transfer to long photoperiods the *EYA3* and *TSHβ* locus show a progressive
367 increase over a number of weeks in H3K4me3 marks matching the incease in
368 expression (Fig. 1f-i & Supplementary Fig. 2e). Our earlier work shows that at
369 the individual cell level the transition between winter and summer physiology is
370 a binary, all-or-nothing phenomenon [29,59]. Integrating these two findings, we
371 suggest that individual thyrotroph cells of the PT exhibit a distribution of critical
372 day length requirements/sensitivity for circadian triggering of the summer
373 physiology leading to a binary switch in cell phenotype, which in a whole tissue
374 assay would appear as a progressive change in epigenetic status.

375 We therefore envisage a photoperiodic time measurement model in which a
376 melatonin-regulated circadian-based interval timer interacts with the underlying
377 cellular chromatin state, rhythmically recapitulating the final developmental
378 stages of a thyrotroph endocrine cell in response to long days, leading to an
379 adaptive summer-like physiological state. Defining how a “flip-flop” circadian
380 timer (Fig. 5) interacts with the chromatin state at the level of the single PT cell
381 remains a future challenge.

382 In summary, the use of photoperiod to synchronise life history cycles in a
383 variable environment is an ancestral feature observed in a large majority of
384 species. The EYA3-TSH circuitry is conserved amongst vertebrates [9,13,16,21], as

385 is the use of a circadian/coincidence based system for photoperiodic time
386 measurement [2,4–7]. Here, we have defined how photoperiodic changes in
387 EYA3/TSH expression stem from a molecular circadian coincidence timer
388 mechanism. We expect this model to be widely applicable across the vertebrate
389 linage.



390

Figure 5: Model for coincidence timing and epigenetic regulation in mammalian photoperiodism

The photoperiod sculpts the duration of the melatonin signal, long duration on short photoperiods (SP, winter) and short duration on long photoperiods (LP, summer). In the pars tuberalis approximately 12 hours after the onset of darkness there is a photosensitive phase. When this phase coincides with light BMAL2 is expressed. When this phase is coincident with darkness DEC1 is expressed. This forms a photoperiodic flip-flop switch between two stable states. The role of BMAL2, CLOCK and BMAL1 is to co-activate EYA3 and subsequently seasonal physiological changes. DEC1 suppresses the E-box based activation of EYA3. Also occurring in the dark phase of SP a number of repressors are expressed which target appear to target LP induced genes potentially inhibiting the expression of summer physiology. Aside from this acute photoperiodic flip-flop switch there is a progressive photoperiod dependent epigenetic regulation over a number of weeks augmenting the coincidence based timer drive on summer physiology.

391

Methods

Animals & experimental design

All animal experiments were undertaken in accordance with the Home Office Animals (Scientific Procedures) Act (1986), UK, under a Project License held by A.S.I.L. Scottish blackface castrate males were housed in artificial light dark cycles, either 8:16 h light／dark cycle for short photoperiod (SP) or 16:8 h light ／dark cycle for long photoperiod (LP).

Two separate photoperiod controlled studies were undertaken; 1. The experiment presented in Fig. 1b (seasonal comparison) & Fig. 3a (diurnal comparison at day 28), and, 2. The experiment presented in Fig. 4a (melatonin implant study). Animals were blood sampled throughout the study and terminally sampled at the indicated time-points (Figure 1b, 3a, 4a). The seasonal experiment was designed to take into account the effects of a photoperiodic switch from SP (SP day 84) to LP and the progressive seasonal changes (LP day 1, 7, 28 and 112), followed by return to SP (SP day 1, 7, 28). Animals were terminally sampled at ZT4 at all time-points. The diurnal comparison was conducted on day 28 animals from this study, they were sampled across the day at 4 hourly intervals for 24 hours. The melatonin implant study was a separate experiment on pre-conditioned LP animals (8 weeks).

All animals were killed by an overdose of barbiturate (Euthatal; Rhone Merieux, Essex, UK) administered intravenously. Hypothalamic blocks with the *pars tuberalis* (PT) and pituitary attached were collected for immunohistochemistry (n=3 per group), electron microscopy (n=3 per group), transcriptomics (n=3 per group), insitu-hybridization (n=4 per group) and epigenomics (n=2 per group). From the omic analyses the PT was dissected to minimise inclusion of transition zone and median eminence. The PT samples were snap frozen on dry ice and stored at -80C.

Hormone assays

Ovine prolactin (oPRL) was measured as in our previous study [29] for 30 animals during the seasonal experiment (Fig. 1b). In brief, a competitive ELISA using purified oPRL (ovine prolactin NIDDK-oPRL-21; AFP10692C; from Dr. A Parlow,

NHPP, Harbor-UCLA Torrance CA, USA) and a highly specific rabbit anti-ovine prolactin (ASM-R50, produced by ASM) were used. Plates were read at 450nm. The Coefficient of Variation for the assay on control plasma samples was <10%.

Ovine melatonin was measured by radioimmunoassay as previously described [47] for animals in the melatonin implant study (Fig. 4a, n=6 per timepoint). In brief, using a rabbit antimelatonin antiserum (PF1288; P.A.R.I.S., Paris, France) and 2-Iodomelatonin (NEX236050UC; PerkinElmer, Boston, Massachusetts) as tracer RIA was performed. All samples were assayed in a single assay with an intra-assay coefficient of variation of 5% and a sensitivity of 5 pg/mL.

Immunohistochemistry

Tissues (n=3 per group) were immersed in Bouin's fixative for 8 hours, transferred to 70% ethanol, then dehydrated and embedded in paraffin wax. Sagittal sheep brain sections were cut from paraffin embedded tissue at 5 µM, floated onto Superfrost Plus slides (J1800 AMNZ, Thermo scientific), dried at 50°C overnight, then dewaxed and rehydrated. Triple immunoflourscence for TSHb, CHGA and EYA3, and the double immunoflourscence for DEC1 /αGSU was performed as in our previous study [29].

DEC1 primary antibody (CW27, gift from Prof. Adrian L Harris, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford) [60] was used at 1:2000. 1:1000 diluted horse radish peroxidase conjugated chicken anti-rabbit IgG antibody was used as a secondary antibody (PI-1000, Vector Laboratories). TSA Plus Cyanine 5 (NEL763001KT, Perkin Elmer) was used to visualize. aGSU immunofluorescence was used at 1:2000 diluted, ASM-HRSU, R20) and treated with TSA Plus Fluorescein (NEL741001KT, Perkin Elmer). Nuclei were stained by Hoechst 33258 (ab228550, abcam) and cover glasses were mounted by VectaMount AQ (H-5501, Vector Laboratories). Images were collected on a Zeiss Axioimager.D2 upright microscope using a 40x / 0.7 Plan neofluar objective and captured using a Coolsnap HQ2 camera (Photometrics) through Micromanager Software v1.4.23. Images were then processed using Fiji ImageJ (http://imagej.net/Fiji/Downloads).

In Situ Hybridization (ISH) and Quantification of Signal

The OaTSHβ plasmid (XM_004002368.2) was kindly provided by David Hazlerigg. The OaEya3 plasmid (NM_001161733.1) was cloned as previously described [22]. The OaCHGA, OaDEC1, OaREVERB-alpha, CHRONO and OaBMAL2 were cloned as 1,948-1,970 of XM_004017959.4, 407-767 of NM_001129741.1, 1,012-1,411 of NM_001131029.1, 83-532 of XM_027974329.1, and 1,518-1,906 of XM_027965976.1, respectively.

Frozen coronal ovine hypothalamic blocks (n=4 per group) for in-situ hybridization were cut into 16µm sections using a cryostat (CM3050s Leica Microsystems, Ltd., Milton Keynes, UK), and thaw mounted onto poly-l-lysine coated slides (VWR International, Lutterworth, UK). Radiolabelled cRNA riboprobes were prepared by plasmid linearization and transcribed using P33 α-UTP (Perkin-Elmer). Fixed sections were hybridized overnight at 60°C with 5 x 105 cpm of probe per slide. Hybridization signals were visualised on autoradiographic film (Kodak Biomax MR Films, Kodak, USA) after one week exposure at -80°C. Signal intensity was quantified by densitometry analysis of autoradiographs using the image-Pro Plus 6.0 software (Media Cybernetics, Inc., Marlow, UK).

Tissue processing and electron microscopy (EM)

Hypothalamo-pituitary tissue blocks were fixed by immersion in 3% paraformaldehyde/0.05% paraformaldehyde in 0.1M phosphate buffer (pH 7.2) for 24 hours at room temperature and transferred to a 1:10 dilution of the fixative in 0.1M phosphate buffer for storage at 4ºC before processing (n=3 per group). Using a scalpel blade, areas from the medial PT and median eminence were cut into 0.5mm3 pieces which were then stained with osmium (1% in 0.1M phosphate buffer), uranyl acetate (2% w/v in distilled water), dehydrated through increasing concentration of ethanol (70 to 100%), followed by 100% acetone and embedded in Spurr's resin (TAAB laboratory equipment, Aldermarston, UK). Ultrathin sections (50-80 nm) were prepared using a Reichart-Jung Ultracut ultramicrotome and mounted on nickel grids (Agar Scientific Ltd., Stanstead, UK). Sections were then counterstained with lead citrate and uranyl acetate and examined on a JOEL 1010 transmission electron

485 microscope (JOEL USA Inc., Peabody, MA, USA). Sections from 3 animals per
486 group were examined.

487 For analysis of PT cell morphology, twenty micrographs per animal (n=3 sheep
488 per group) of individual PT cells were taken at a magnification of x 5,000.
489 Negatives were scanned into Adobe Photoshop CS2 (Adobe Corp., San Jose,
490 CA, USA) and analysed using Axiovison version 4.5 (Zeiss, Oberkochen,
491 Germany) image analysis software. The analyst was blind to the sample code.
492 For measurement of the cell and nuclear areas, margins were drawn around
493 the cell or nucleus respectively and the area was calculated. All morphometric
494 values represent the mean ± SEM (n=3 sheep per group). Means were
495 compared by one way analysis of variance (ANOVA) with post hoc analysis by
496 the Bonferroni test. $P<0.05$ was considered statistically different.

497 *RNA-seq*

498 RNA was extracted from the pars tuberalis from the seasonal experiment
499 (including the diurnal samples)(Fig1b & Fig. 3a) using Qiagen's TissueLyser II
500 and RNeasy tissue kit (n=3 per group). The quality of the extracted RNA was
501 assessed using the Agilent 2100 Bioanalyser; all RNA integrity numbers (RINs)
502 were above 8, indicating that good quality RNA had been extracted. Poly-A
503 selection was used.

504 RNA was prepared with TruSeq Stranded mRNA Sample Preparation Guide,
505 (15031047 Rev. E, Oct 2013) and sequenced on HiSeq 2500 with 125 base
506 pair paired-end reads by Edinburgh Genomics.

507 The FASTQ files were trimmed with TrimGalore v0.4.0 and mapped (TopHat[61]
508 v2.1.0 and Bowtie[62] v2.3.5) to the 5th release of the sheep genome
509 (Oar_rambouillet_v1.0; assembly GCA_002742125.1). StringTie[63] was used to
510 combine RNASeq and IsoSeq full-length transcripts to generate the genome
511 transcript annotation (accessible in GEO, GSE144677). On average 90%
512 (sdev: 8.46) of paired reads generated were mapped to the genome and 73%
513 (stdev: 0.03) of these were assignable to genes using featureCount [64] (Subread
514 v1.6.3).

All sequence data have been submitted to SRA under the BioProject accession PRJNA391103 and processed data to GEO under GSE144677.

A limma-voom [65] analysis pipeline was used to determine the statistical significance of differential expressed genes. Voom was used to generate normalized precision weighted counts per million (CPM) values which were used in the following regression analyses.

Seasonal comparison

The effect of switching from SP to LP was assessed by comparing SP day 84 to LP day 1, 7 and 28, and the effect of switching from LP to SP was assessed by comparing LP day 112 to SP day 1, 7 and 28 (Fig. 1b, Supplementary Table 3). For each gene, we fit a least squared regression model with limma that calculates a single f-test for significance across all model coefficients (mitigating type I errors). Time (days) was treated as a categorical independent variable model for all ZT4 observations in LP and SP (photoperiod x day) in limma and which allowed us to extract from the model the fold change and significance for each pairwise contrasts of interest in limma (Supplemental Table 3; Figure 2a,b). Significance was determined by an FDR < 0.05, >0 log2CPM and a >1 log2 absolute fold change.

Diurnal comparison

To test for diurnal changes, samples collected at day 28 in LP and SP at a 4hr time resolution were used (n=3 per group, 7 time-points, Fig. 3a). We used a polynomial regression model approach similar to that of maSigPro[66].Least squared regression models were then fitted with orthogonal polynomials up to the 5th order for time in each photoperiod to identify significantly changing genes. To test for rapid single time-point changes in gene expression a categorical regression model was also fitted to this dataset. Diurnal genes that were significantly changing across time were identified as FDR significance <0.05, log2CPM > 0 and absolute log2 fold change > 1 (Supplementary Table 6; Figure 3b,c,d). FDR was calculated throughout using the Benjamini & Hochberg method. Gene expression changes between photoperiods were evaluated by fitting a photoperiod x time (orthogonal polynomials up to 5th order) model and

546 extracting the significance and effect size from photoperiod coefficient of the
547 linear model (Supplementary Table 6). In selecting the polynomial we used
548 Akaike information criterion (AIC) to investigate the optimal model selection for
549 expressed genes, balancing model overfitting and underfitting (using the
550 Oshlack and Gordon selectModel implementation in limma). It is not possible to
551 select a single model that is optimal across all genes, however for genes > 0
552 log2CPM and with an amplitude > 1.5 we found that including orthogonal
553 polynomials up to 5th order was optimal for the most genes in both SP and LP
554 time-series. Again we used the thresholds FDR significance <0.05, log2CPM >
555 0 and absolute log2 fold change > 1 (Supplementary Table 6). MetaCycle[67]
556 v1.1.0 was used to evaluate gene expression in the 24hr time series for
557 periodicity. JTKCyle [41] and Lomb-Scargle statistics were calculated for an
558 assumed period of 24 hours (Supplementary Table 6). Rayleigh tests for
559 uniformity were performed with the CircStats 0.2-6 package in R. We tested the
560 uniformity of distribution of peak expression times for genes containing each of
561 the core clock motifs (canonical EBOXs (CACGTG), DBOXs
562 (TTA[CT][GA]TAA) and RORE sites (AANTAGGTCA)) within the H3K4me3
563 marked region proximal (within 500bp) of the TSS.

564 For all gene cohorts described enrichment analysis of GO terms and pathways
565 was conducted using both consensusPathDB [68] and TopGO [69] 3.1.0. Signicant
566 terms extracted by a FDR < 0.05 from a fishers exact test. The weight01
567 algorithm was used for GO term weighting in TopGO. Where gene annotation
568 did not yet exist for novel and unannotated transcripts OrthoMCL [70] v5 was used
569 to predict orthology from protein sequences and annotation, which were then
570 used to transfer annotation from cow, human, mouse and rat genes. Protein
571 sequences were predicted from novel transcripts using TransDecoder v5.3.0
572 guided by Uniprot [71] and Pfam [72] best hits to rank coding frames.

573 We clustered SP and LP time-series profiles using Partitioning Around Medoids
574 (PAM) (Fig. 3g) with the cluster 2.1.0 package in R. Day 28, 24 hour, time-series
575 were mean normalized and scaled and PAM clustered with Euclidian distance.
576 The Davies Bouldin index was used to evaluate the optimum number of clusters
577 (k=15 SP and k=9 LP). Motif enrichment of genes clusters was evaluated using
578 fishers two-way exact test against all PT expressing genes as the background.

579 Motifs we identified within H3K4me3 marked regions within 500bp of a
580 candidate TSS assigned to a gene.

581 *ChIP-seq*

582 The method used for ChIP-seq was adapted from an ultra low cell number
583 native ChIP method [73]. In brief, nuclei were isolated from whole PT tissue (n=2
584 per group) with a dounce homogenizer and sigma nuclear isolation buffer.
585 Mnase digestion was optimized to to give the best discrimination between mono,
586 di and tri histones. Immunoprecipitation was performed with protein A and G
587 beads (Invitrogen). Importantly the beads were pre-incubated with the
588 H3K4me3 antibody (active motif) to reduce background noise. Stringent
589 washing with salt buffers was also used to reduce the background noise. The
590 bound chromatin was elute and extracted using a phenol chloroform method.
591 Ampure bead purification was used to clean up the samples. The qubit and
592 tapestation were used to quantify, a minimum amount of 5ng was required for
593 the library preparation. The Manchester genomic facility prepared the ChIP-seq
594 libraries and sequenced them according to the standard illumina protocol.

595 Massive parallel sequencing were performed by Illumina HiSeq4000 with 75 bp
596 paired-end and converted fastqs by bcl2fastq (ver 2.17.1.14). Fastqs were
597 trimmed by trimmomatic (ver 0.36) [74] with the following parameters,
598 ILLUMINACLIP: TruSeq3-PE-2.fa:2:30:10 SLIDINGWINDOW:4:20
599 MINLEN:35, aligned to the sheep genome (Oar_rambouillet_v1.0) by BWA
600 MEM (ver 0.7.17) [75] with the default parameter and then converted to BAM and
601 sorted by samtools (ver 0.1.19). Peak calling were carried out by MACS2 (ver
602 2.1.0) [76] with following parameters, --format BAMPE –gsize 2.81e9 --keep-dup
603 1 --broad --broad-cutoff 0.1 --bdg --SPMR --qvalue 0.05.

604 Read coverages of peaks were calculated by SICER (ver 1.1) [77] with following
605 parameters; window size=200, gap length=200, fdr=0.01. BED files of replicate
606 samples were merged in order to perform SICER analysis which does not allow
607 replicates. Peaks called by SICER were annotated by HOMER (ver 4.10.3) [78]
608 with default parameters. H3K4me3 peaks identified by SICER were validated
609 by monitoring the distributions on the sheep genome. By HOMER annotation,
610 each peak was described as promoter-TSS (1000 bp from TSSs), exon, intron,

TTS, intergenic and the distributions of H3K4me3 peaks were closely resembling to the previous reports [33,34] (Supplementary Fig. 6a). Furthermore, H3K4me3 peaks were well-associated with CpG islands (CGIs) on the sheep genome as described in the previous study (Supplementary Fig. 6b) [35]. We use a standard definition of CpG islands [79]; nucleotides regions with > 50% GC content, extending to > 200 bp and with an observed vs expected CpG ratio >6.5, and detected them using CgiHunterLight 1.0 on Oar_rambouillet_v1.0 (assembly GCA_002742125.1).H3K4me3 peaks of each sampling day were shuffled by bedtools shuffle (ver 2.27.1) [80] with -noOverlapping as negative controls. For correlation analysis with RNA expression, ChIP read counts of peaks overlapped in ± 200 bp from TSSs were used. A workflow diagram can be found in Supplementary Fig. 7.

ISO-seq

For gene annotation, five tissue samples were sequenced over two experimental runs using PacBio Iso-Seq. In the first run PT and PD samples were sequenced from an RNA pool of SP and LP Scottish blackface sheep (N=1) and a pineal from a commercial mule sheep from Manchester, UK. This RNA was sent to GATC Biotech (Konstanz, Germany) for cDNA library preparation using their in-house method with mRNA 5' cap and poly(A) tail selections and sequencing on a PacBio RSII system. GATC made full length normalized RNA libraries.size selected for <2kb, 2kb-4kb, >4kb. sequenced across 75 PacBio RS II SMRT cells (SRX7688275). In a second run, PT from a pool of sheep in LP, and SP (N=3) were sequenced. RNA was extracted using RNeasy Mini Kit (Qiagen) with on-column Dnase digestion. A full-length cDNA library was constructed for each sample using the TeloPrime Full-Length cDNA Amplification Kit V1 (Lexogen) and amplified using PrimeSTAR GXL DNA Polymerase (Takara Bio) with 22 PCR cycles of 98 °C denaturation for 10 seconds, 60 °C annealing for 15 seconds, and 68 °C extension for 10 minutes. PacBio SMRTbell libraries were prepared using SMRTbell Template Prep Kit 1.0 and each library was sequenced on two SMRT Cells v2 LR using 20-hour movies on a Sequel platform at the IMB Sequencing Facility (University of Queensland, SRX7688271). All Iso-Seq data was first processed using software IsoSeq v3.1 to obtain full-length non-concatemer reads with at least 3

644 full sequencing passes, which were then mapped to the sheep reference
645 genome GCA_002742125.1 using GMAP version 2018-05-30. TAMA Collapse
646 from the TAMA tool kit [81] was used to generate unique gene and transcript
647 models, which were further merged with RNAseq-based annotation data using
648 TAMA Merge to incorporate any transcript models that were identified by
649 RNAseq but not Iso-Seq. Functional annotation of transcripts was carried out
650 using Trinotate (v3.1.1).

651 Where multiple transcripts were present for an expressed gene, with more than
652 one transcription start sites (TSS) candidate, the active proximal promoter
653 regions were inferred by selecting contiguous H3K4me3 marked regions within
654 100bp of a TSS.

655 *CAGE-seq*

656 We applied cap analysis gene expression (CAGE) to identify the location and
657 relative expression of TSS regions of the PT across both LP and SP. When
658 combined with IsoSeq and RNASeq derived transcript annotation this provided
659 a comprehensive identification of TSS in the genome which allowed us to more
660 accurately apply DNA binding motif analysis to promoter regions. Libraries were
661 prepared according to Hazuki *et. al* [32] and sequenced on an Illumina HiSeq
662 2500 using V4 chemistry on a 50 cycle Single end sequencing run. We
663 sequenced archived RNA samples from the PT in both SP and LP (ZT4, week
664 12) [29]. We also sequenced RNA from PD (both SP and LP), and Pineal for
665 comparison as outgroups. Reads were trimmed using fastx toolkit 0.0.14 and
666 cutadapt 1.4. Reads were mapped using BWA 0.7.17 to the 5th release of the
667 sheep genome (Oar_rambouillet_v1.0; assembly GCA_002742125.1). CAGEr
668 1.26.0 was used for processing and cluster analysis of TSS (Supplementary
669 Table 4). We filtered reads for a mapping quality > 30 and sequencing quality
670 > 20. Tag counts were normalised using the power law method with an alpha
671 of 1.12 and T of 10^6 (deterimined by plotting the reverse cumulatives of PT
672 samples). We clustering TSS with >1 TPM together using the distclu methods
673 allowing a max distance between TSS of 20 nucleotides.

674

Transcription factor binding site analysis

Transcription Factor binding motifs were identified using the FIMO tool from MEME v4.11.4 with a p-value threshold of $<1x10^{-7}$. for Jasper 2018 core vertebrate database. Fishers one-way exact tests for enrichment were used to identify the significance of motif enrichment within active promoter regions of a gene cohort compared to a background of all PT expressed genes (> 0 CPM). Fishers two-way exact tests were applied to evaluate enrichment and depletion of motifs within active proximal promoter regions of genes with the use of SP and LP H3K4me3 marked regions within 500bp of candidate TSSs.

Protein-protein interaction networks

Using experimentaly evidenced protein-protein interaction (PPI) annotation from the STRING [50] v10 database for cow, sheep, rat, mouse and human we integrated protein interactions in Cytoscape [82] with significantly changing genes in the ZT 4 SP vs ZT20 LP from the 24 hr categorical contrast, which is 12 hrs from dark/melatonin onset (Supplemental Table 6 & Supplemental Figure 4). A threshold of >0.4 for confidence score in experimental evidence and a combined score of > 0.7 was applied, and orphaned proteins removed from the network. The significance of enrichment of PPI repressor connected genes within up-regulated vs down-regulated genes was evaluated using fishers two-way exact tests.

Cloning and constructs

Expression plasmids: PCR fragments of the expected sizes were extracted using a gel extraction kit (Qiagen) and cloned in pGEM-T easy vector (Promega); Four to six positive clones were sequenced (MWG, United Kingdom). To generate expression constructs, a second round of PCR was performed using primers flanked by adequate restriction sites and the pGEM-T clone as template. PCR fragments were extracted as described above, digested by the adequate restriction enzymes, purified with a PCR purification kit (Qiagen) and cloned in the expression vector backbone (pCS2-HIS). In order to generate the mutant expression plasmids for BMAL2 we used the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (210515, Agilent).

706 The bHLH mutant was generated by converting an arginine to alanine
707 (OaBMAL2_R88A) based on a mouse mutagenesis study on BMAL1[37]. The
708 PAS-B mutant was made by converting a phenylalanine to arginine, and a
709 valine to arginine (OaBMAL2_F427R_V439R). Based on a on a mouse
710 mutagenesis study on BMAL1 [40].

711 Sanger sequencing of clones are available in Genbank for BMAL2 cds
712 constructs (Genbank: MT001920), DEC1 cds constructs (Genbank:
713 MT019539), DEC2 cds constructs (Genbank: MT019540), PAS-B-mutated
714 BMAL2 cds constructs (Genbank: MT019541), bHLH-mutated BMAL2 cds
715 constructs (Genbank: MT019542),

716 Promoter reporter constructs: a strategy identical to that described above was
717 applied and fragments were cloned into the pGL4 basic backbone (Promega)
718 digested with the appropriate restriction enzymes. Sequencing was performed
719 to check accuracy of all re-amplified cloned fragments. EYA3 generic and
720 seasonal promoter construct sequences are available on genebank (MT001921
721 and MT001924 respectively).

722

723 *Cell culture, transfection and luciferase reporter assays*

724 The procedure was as previously reported [22]. In brief, COS-7 cells were grown
725 in Dulbecco's modified eagle's medium supplemented with 10% fetal bovin
726 serum, 1% penicillin/streptomycin at 5% CO_2 and 37℃. Cells were plated in 24-
727 well plates at a density of 4x104 cells per ml and incubated for 24 hours prior
728 to transfection. Transfection was performed using Genejuice (Novagen) and
729 the concentration was optmised to transfect the greatest number of cells without
730 compromising cell survival, this was assessed using a luciferase positive
731 control pGL3 containing SV40 (Promega) and trypan blue staining. We
732 recorded a 90% cell survival and a high transfection efficiency. The EYA3
733 promoter constructs were used at 50ng per well, as in a previous study (ref).
734 The expression plasmids were used at different doses based on a previous
735 study and optimization of the assay: TEF = 12.5ng, DEC1 = 25ng, CLOCK,
736 BMAL1, BMAL2 and mutant BMAL2 were all used at 50ng, unless otherwise
737 stated. The total transfected DNA amount was set to an equal amount between

all conditions by addition of the corresponding empty vector. The luciferase assays were performed 48 hours after transfection using the luciferase assay kit (Promega) and the Glomax luminometer (Promega). Thetotal protein per well, assessed by Bradford assay was used to normalize the values to total protein content (a proxy for cell number). All data (in Relative Luminescence Units, RLU) represent fold induction once normalized to total protein content and relative to an inert control transfection. Each experiment contained 4 replicate wells and was repeated 4 times giving similar results. An one-way ANOVA using Tukey's multiple comparisons test was performed for each separate experiment conducted in Graphpad prism 7.05.Representitive plots (n=4) are shown.

Data availability statement

All sequence data have been submitted to SRA under the BioProject accession PRJNA391103 and processed data to GEO under GSE144677.

Code availability statement

Code used is available from the authors on request.

References

1. Bünning, E. Die endogene Tagesperiodik als Grundlage der photoperiodischen Reaktion. *Ber Dtsch Bot Ges* **54**, 590–608 (1936).

2. Pittendrigh, C. S. & Minis, D. H. The entrainment of circadian oscillations by light and their role as photoperiodic clocks. *Am. Nat.* **98**, 261–299 (1964).

3. Nanda, K. K. & Hamner, K. . Studies on the nature of the endogenous rhythm affecting photoperiodic response of Biloxi soybean. *Bot. Gaz.* **120**, 14–28 (1958).

4. Follett, B. K., Mattocks, P. W. & Farner, D. S. Circadian function in the photoperiodic induction of gonadotropin secretion in the white-crowned sparrow, Zonotrichia leucophrys gambelii. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 1666–9 (1974).

5. Goldman, B. D. Mammalian photoperiodic system: formal properties and neuroendocrine mechanisms of photoperiodic time measurement. *J. Biol. Rhythms* **16**, 283–301 (2001).

6. Pittendrigh, C. S., Elliott, J. & Takamura, T. The Circadian Component in Photoperiodic Induction. in 26–47 (Wiley-Blackwell, 2008). doi:10.1002/9780470720851.ch4

7. Ikegami, K. & Yoshimura, T. Circadian clocks and the measurement of daylength in seasonal reproduction. *Mol. Cell. Endocrinol.* **349**, 76–81 (2012).

8. Bartness, T. J., Powers, J. B., Hastings, M. H., Bittman, E. L. & Goldman, B. D. The timed infusion paradigm for melatonin delivery: what has it taught us about the melatonin signal, its reception, and the photoperiodic control of seasonal responses? *J. Pineal Res.* **15**, 161–90 (1993).

9. Wood, S. & Loudon, A. Clocks for all seasons: unwinding the roles and mechanisms of circadian and interval timers in the hypothalamus and

783 pituitary. *J. Endocrinol.* **222**, R39-59 (2014).

784 10. Woodfill, C. J. I., Wayne, N. L., Moenter, S. M. & Karsch, F. J.
785 Photoperiodic Synchronization of a Circannual Reproductive Rhythm in
786 Sheep : Identification of Season-Specific Time Cues ' Reproductive
787 Sciences Program and Department of Physiology , University of
788 Michigan. *Biol. Reprod.* **50**, 965–976 (1994).

789 11. Lincoln, G. A. & Ebling, F. J. P. Effect of constant-release implants of
790 melatonin on seasonal cycles in reproduction, prolactin secretion and
791 moulting in rams. *Reproduction* **73**, 241–253 (1985).

792 12. Bittman, E. L., Dempsey, R. J. & Karsch, F. J. Pineal melatonin
793 secretion drives the reproductive response to daylength in the ewe.
794 *Endocrinology* **113**, 2276–83 (1983).

795 13. Dardente, H., Wood, S., Ebling, F. & Sáenz de Miera, C. An integrative
796 view of mammalian seasonal neuroendocrinology. *J. Neuroendocrinol.*
797 **31**, e12729 (2019).

798 14. Lincoln, G. A. Melatonin Entrainment of Circannual Rhythms.
799 *Chronobiol. Int.* **23**, 301–306 (2006).

800 15. Wood, S. & Loudon, A. The pars tuberalis: The site of the circannual
801 clock in mammals? *Gen. Comp. Endocrinol.* **258**, (2018).

802 16. West, A. C. & Wood, S. H. Seasonal physiology: making the future a
803 thing of the past. *Curr. Opin. Physiol.* **5**, 1–8 (2018).

804 17. Johnston, J. D. *et al.* Multiple Effects of Melatonin on Rhythmic Clock
805 Gene Expression in the Mammalian Pars Tuberalis. *Endocrinology* **147**,
806 959–965 (2006).

807 18. West, A. *et al.* Npas4 is activated by melatonin, and drives the clock
808 gene Cry1 in the ovine pars tuberalis. *Mol. Endocrinol.* **27**, 979–989
809 (2013).

810 19. Hanon, E. A. *et al.* Ancestral TSH mechanism signals summer in a
811 photoperiodic mammal. *Curr. Biol.* **18**, 1147–52 (2008).

812 20. Nakao, N. *et al.* Thyrotrophin in the pars tuberalis triggers photoperiodic
813 response. *Nature* **452**, 317–22 (2008).

814 21. Nakane, Y. & Yoshimura, T. Universality and diversity in the signal
815 transduction pathway that regulates seasonal reproduction in
816 vertebrates. *Front. Neurosci.* **8**, 115 (2014).

817 22. Dardente, H. *et al.* A molecular switch for photoperiod responsiveness
818 in mammals. *Curr. Biol.* **20**, 2193–8 (2010).

819 23. Masumoto, K. *et al.* Acute Induction of Eya3 by Late-Night Light
820 Stimulation Triggers TSHβ Expression in Photoperiodism. *Curr. Biol.* **20**,
821 2199–2206 (2010).

822 24. Dupré, S. M. *et al.* Identification of Eya3 and TAC1 as long-day signals
823 in the sheep pituitary. *Curr. Biol.* **20**, 829–35 (2010).

824 25. Angel, A., Song, J., Dean, C. & Howard, M. A Polycomb-based switch
825 underlying quantitative epigenetic memory. *Nature* **476**, 105–8 (2011).

826 26. Song, J., Irwin, J. & Dean, C. Remembering the prolonged cold of
827 winter. *Curr. Biol.* **23**, R807-11 (2013).

828 27. Satake, A. & Iwasa, Y. A stochastic model of chromatin modification:
829 Cell population coding of winter memory in plants. *J. Theor. Biol.* **302**,
830 6–17 (2012).

831 28. Dardente, H., Hazlerigg, D. G. & Ebling, F. J. P. Thyroid hormone and
832 seasonal rhythmicity. *Front. Endocrinol. (Lausanne).* **5**, 19 (2014).

833 29. Wood, S. H. *et al.* Binary switching of calender cells in the pituitary
834 defines the phase of the circannual cycle in mammals. *Curr. Biol.* **25**,
835 (2015).

836 30. Karsch, F. J., Robinson, J. E., Woodfill, C. J. & Brown, M. B. Circannual
837 cycles of luteinizing hormone and prolactin secretion in ewes during
838 prolonged exposure to a fixed photoperiod: evidence for an endogenous
839 reproductive rhythm. *Biol. Reprod.* **41**, 1034–46 (1989).

840 31. Lincoln, G. A., Clarke, I. J., Hut, R. A. & Hazlerigg, D. G. Characterizing
841 a mammalian circannual pacemaker. *Science* **314**, 1941–4 (2006).

842 32. Takahashi, H., Kato, S., Murata, M. & Carninci, P. CAGE (Cap analysis
843 of gene expression): A protocol for the detection of promoter and
844 transcriptional networks. *Methods Mol. Biol.* **786**, 181–200 (2012).

845 33. Hussey, S. G., Loots, M. T., van der Merwe, K., Mizrachi, E. & Myburg,
846 A. A. Integrated analysis and transcript abundance modelling of
847 H3K4me3 and H3K27me3 in developing secondary xylem. *Sci. Rep.* **7**,
848 3370 (2017).

849 34. Ershov, N. I. *et al.* Consequences of early life stress on genomic
850 landscape of H3K4me3 in prefrontal cortex of adult mice. *BMC*
851 *Genomics* **19**, 93 (2018).

852 35. Rose, N. R. & Klose, R. J. Understanding the relationship between DNA
853 methylation and histone lysine methylation. *Biochim. Biophys. Acta -*
854 *Gene Regul. Mech.* **1839**, 1362–1372 (2014).

855 36. Lincoln, G., Messager, S., Andersson, H. & Hazlerigg, D. Temporal
856 expression of seven clock genes in the suprachiasmatic nucleus and
857 the pars tuberalis of the sheep: evidence for an internal coincidence
858 timer. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13890–5 (2002).

859 37. Hosoda, H., Motohashi, J., Kato, H., Masushige, S. & Kida, S. A BMAL1
860 mutant with arginine 91 substituted with alanine acts as a dominant
861 negative inhibitor. *Gene* **338**, 235–41 (2004).

862 38. Spiegelman, B. M. & Heinrich, R. Biological Control through Regulated
863 Transcriptional Coactivators. *Cell* **119**, 157–167 (2004).

864 39. Guo, Y., Scheuermann, T. H., Partch, C. L., Tomchick, D. R. & Gardner,
865 K. H. Coiled-coil Coactivators Play a Structural Role Mediating
866 Interactions in Hypoxia-inducible Factor Heterodimerization. *J. Biol.*
867 *Chem.* **290**, 7707–7721 (2015).

868 40. Huang, N. *et al.* Crystal Structure of the Heterodimeric CLOCK:BMAL1

Transcriptional Activator Complex. *Science (80-.).* **337**, 189–194 (2012).

41. Hughes, M. E., Hogenesch, J. B. & Kornacker, K. JTK-CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J. Biol. Rhythms* **25**, 372–380 (2010).

42. Xu, H. *et al.* Cryptochrome 1 regulates the circadian clock through dynamic interactions with the BMAL1 C terminus. *Nat. Struct. Mol. Biol.* **22**, 476–484 (2015).

43. Maywood, E. S. *et al.* Translational switching of Cry1 protein expression confers reversible control of circadian behavior in arrhythmic Cry-deficient mice. *Proc. Natl. Acad. Sci.* **115**, E12388–E12397 (2018).

44. Partch, C. L., Green, C. B. & Takahashi, J. S. Molecular architecture of the mammalian circadian clock. *Trends Cell Biol.* **24**, 90–99 (2014).

45. Dardente, H. *et al.* Melatonin induces Cry1 expression in the pars tuberalis of the rat. *Mol. Brain Res.* **114**, 101–106 (2003).

46. Dupré, S. M. *et al.* Identification of melatonin-regulated genes in the ovine pituitary pars tuberalis, a target site for seasonal hormone control. *Endocrinology* **149**, 5527–39 (2008).

47. West, A. *et al.* Npas4 is activated by melatonin, and drives the clock gene Cry1 in the ovine pars tuberalis. *Mol. Endocrinol.* (2013). doi:10.1210/me.2012-1366

48. Hirano, A. *et al.* FBXL21 Regulates Oscillation of the Circadian Clock through Ubiquitination and Stabilization of Cryptochromes. *Cell* **152**, 1106–1118 (2013).

49. Dardente, H., Mendoza, J., Fustin, J.-M., Challet, E. & Hazlerigg, D. G. Implication of the F-Box Protein FBXL21 in Circadian Pacemaker Function in Mammals. *PLoS One* **3**, e3530 (2008).

50. Szklarczyk, D. *et al.* The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic*

Acids Res. **45**, D362–D368 (2017).

51. Maywood, E. S. *et al.* The effect of signal frequency on the gonadal response of male Syrian hamsters to programmed melatonin infusions. *J. Neuroendocrinol.* **4**, 37–44 (1992).

52. Ebling, F. J., Lincoln, G. A., Wollnik, F. & Anderson, N. Effects of constant darkness and constant light on circadian organization and reproductive responses in the ram. *J. Biol. Rhythms* **3**, 365–84 (1988).

53. Staples, L. D., McPhee, S., Kennaway, D. J. & Williams, A. H. The influence of exogenous melatonin on the seasonal patterns of ovulation and oestrus in sheep. *Anim. Reprod. Sci.* **30**, 185–223 (1992).

54. Honma, S. *et al.* Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature* **419**, 841–844 (2002).

55. Hazlerigg, D., Lomet, D., Lincoln, G. & Dardente, H. Neuroendocrine correlates of the critical day length response in the Soay sheep. *J. Neuroendocrinol.* e12631 (2018). doi:10.1111/jne.12631

56. Gustafson, C. L. & Partch, C. L. Emerging Models for the Molecular Basis of Mammalian Circadian Timing. *Biochemistry* **54**, 134–149 (2015).

57. Bunger, M. K. *et al.* Mop3 Is an Essential Component of the Master Circadian Pacemaker in Mammals. *Cell* **103**, 1009–1017 (2000).

58. Honma, S. *et al.* Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature* **419**, 841–844 (2002).

59. Wood, S. H. How can a binary switch within the pars tuberalis control seasonal timing of reproduction? *J. Endocrinol.* **239**, R13–R25 (2018).

60. Turley, H. *et al.* The hypoxia-regulated transcription factor DEC1(Stra13, SHARP-2) and its expression in human tissues and tumours. *J. Pathol.* **203**, 808–813 (2004).

61. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the

presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).

62. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

63. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).

64. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

65. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

66. Nueda, M. J., Tarazona, S. & Conesa, A. Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. *Bioinformatics* **30**, 2598 (2014).

67. Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K. & Hogenesch, J. B. MetaCycle: An integrated R package to evaluate periodicity in large scale data. *Bioinformatics* **32**, 3351–3353 (2016).

68. Kamburov, A., Stelzl, U., Lehrach, H. & Herwig, R. The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res.* **41**, D793-800 (2013).

69. Alexa, A., Rahnenführer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600–1607 (2006).

70. Li, L., Stoeckert, C. J. & Roos, D. S. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003).

71. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

955 72. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic*
956 *Acids Res.* **47**, D427–D432 (2019).

957 73. Brind'Amour, J. *et al.* An ultra-low-input native ChIP-seq protocol for
958 genome-wide profiling of rare cell populations. *Nat. Commun.* **6**, 6033
959 (2015).

960 74. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer
961 for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

962 75. Li, H. & Durbin, R. Fast and accurate short read alignment with
963 Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

964 76. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome*
965 *Biol.* **9**, R137 (2008).

966 77. Zang, C. *et al.* A clustering approach for identification of enriched
967 domains from histone modification ChIP-Seq data. *Bioinformatics* **25**,
968 1952–1958 (2009).

969 78. Heinz, S. *et al.* Simple Combinations of Lineage-Determining
970 Transcription Factors Prime cis-Regulatory Elements Required for
971 Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).

972 79. Ioshikhes, I. P. & Zhang, M. Q. Large-scale human promoter mapping
973 using CpG islands. *Nat. Genet.* **26**, 61–63 (2000).

974 80. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for
975 comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

976 81. Kuo, R. I., Cheng, Y., Smith, J., Archibald, A. L. & Burt, D. W.
977 Illuminating the dark side of the human transcriptome with TAMA Iso-
978 Seq analysis. *bioRxiv* 780015 (2019). doi:10.1101/780015

979 82. Shannon, P. *et al.* Cytoscape: A software Environment for integrated
980 models of biomolecular interaction networks. *Genome Res.* **13**, 2498–
981 2504 (2003).

982

Acknowledgements

The authors thank Joan Docherty and the staff at the Marshall Building, Roslin, Edinburgh for expert care of their research animals. Past members of the Manchester/Edinburgh team have made contributions to discussions and offered advice, including Sandrine Dupre, and Alex West. We thank David Hazlerigg, Carrie Partch and Jean Michel Fustin for their critical review of earlier drafts of this work. The work was supported by grants from the Biotechnology and Biological Sciences Research Council UK (BB/N015584/1, BB/N015347/1, BBP013759/1) and a Human Sciences Frontier Programme Grant RGP0030/2015 Evolution of Seasonal Timers awarded to ASIL and DWB. ASIL acknowledges the support of the Wellcome Trust, Grant 107851/Z/15/Z

Author contributions

SHW - designed the experiments, collected samples, performed RNA preparation, chromatin immunoprecipitation preparation, luciferase reporter assays, cloning and mutagenesis, analysed/interpreted data, prepared the manuscript and figures. **MH** – Bioinformatic analysis, data analysis and submission, sequencing, prolactin assay, figure preparation and revised the manuscript. **YM** - Bioinformatic analysis, immunohistochemistry, luciferase reporter assays, cloning and mutagenesis, figure preparation and revised manuscript. **YC** – ISO-seq and bioinformatic analysis **KM** – immunohistochemistry and analysis. **HC** – performed EM and analysis, revised manuscript. **BRCS** - collected samples and performed the in situ hybridization, luciferase reporter assays, cloning and mutagenesis. **NB** – collected samples and provided lab support. **JM** - developed the novel prolactin assay and prepared antibodies. **ASM** - collected samples, designed experiments, developed the prolactin assay, and revised the manuscript. **SM** - collected samples, designed experiments, prolactin assay, and revised manuscript. **DWB** - designed experiments, bioinformatics analysis, and revised the manuscript. **ASIL** – conceived the study, designed experiments, collected samples, analysed/interpreted data, and prepared the manuscript.

Supplementary Figure legends

Supplementary Fig. 1: Photoperiod dependent epigenetic regulation of transcription in the *Pars tuberalis*

a. Triple immunofluorescence showing expression of aGSU (red), TSHb (green), and EYA3 (blue) in the PT on SP day 28 and LP day 28. Scale bars, 50 um.

b. RNAseq log2 counts per million (CPM) of TSHb (purple) and CHGA (green) over the experiment. Grey shading represents SP sampling points. Error bars represent the SEM.

c. In situ hybridization and quantification for CHGA mRNA at SP day 84, LP day 1, 7, 28, 112, SP day 1, 7, 28. Representative images are shown (n = 3). Error bars represent the SD. Statistical analysis by one-way ANOVA confirms there is a significant seasonal change. **** = p-value less than 0.00005.

d. As in c for EYA3 mRNA . ** = p-value less than 0.005.

e. PT thyrotroph chromatin density at LP and SP day 28, given as the percentage of dense chromatin contained within the nucleus relative to nucleus size. N=3 individuals from SP and LP. 40 nuclei were measured per group. T-test was used to assess statistical significance. **= FDR less than 0.00001

f. PT thyrotroph, PT follicular stellate (FS) cell, pars distalis (PD) FS cells and PD somatotroph nucleus measurements (um) at LP and SP day 28. N=3 individuals from SP and LP. 40 nuclei were measured per group. T-test was used to assess statistical significance. **= FDR less than 0.00001

g. Percentage of genes with a given number of transcription start sites in the genomic background (all >0 log2CPM expressed genes) of the pars tuberalis (grey bars) as compared to all seasonally differentially expressed genes (white bars).

h. Comparison of the prevalence of multiple (>1) TSS across different gene cohorts. The cohorts are LP 28 days up-regulated DEGs (solid red), SP 28 days up-regulated DEGs (solid blue) and all PT expressed genes as the background (solid black). Prevalence of multiple TSS on genes is shown for all thresholds (%) of the uppermost expressed genes (i.e. increasing thresholds for the upper quantile of gene expression). The equivalent gene expression (log2CPM) values for upper quantiles for a lower threshold cutoff upper x-axis. Dashed lines indicate the proportion of gene in the cohorts with multiple H3K4me3 (>1) marked TSS.

Supplementary Fig. 2: Epigenetic regulation of a seasonal E-box enriched promoter in EYA3

a. Plot of all over-represented motifs in the active promoters of LP day 84 up-regulated genes. The axis plot the percentage of inferred active promoters, containing one or more observed motifs, for a given cohort of genes. Active promoters are defined as contiguous H3K4me3 marked regions within 100bp of a TSS. The black triangles represents the motif abundance on active promoters in all the genes expressed (>0 CPM) in the pars tuberalis (PT) and the percentage of their active promoters containing a binding motif for LP (x-axis) and SP (y-axis) H3K4me3 environments. The circles are the motif coverage of activate promoters in differentially up-regulated LP genes (FDR < 0.05; fishers two-way exact test).

b. As in G for the over-represented motifs in the TSS of LP down-regulated genes when compared to SP day 84.

c. ChIP-seq tracks for *CHGA* gene H3K4me3 peaks for the whole experiment. Chromsome 2 region is shown. Pink represents samples in long photoperiod and marine green represents short photoperiod. Solid green arrows are canonical E-box motifs. Blue arrows are D-box motifs. Gene schematic is also shown.

d. Correlation plot for *CHGA* downstream TSS log2 H3K4me3 peaks from ChIP-seq versus *CHGA* log2 counts per million (CPM) from RNA-seq. Red symbols are LP sampling points, green are SP sampling points. Correlation coefficient R is shown. R=0.843, p-value=0.009.

e. As in A for *EYA3* gene. Chromsome 2 region is shown.

f. As in B for *EYA3* up-stream TSS. R=0.368, p-value=0.458.

g Transactivation of the EYA3-upstream-TSS-luc reporter versus the EYA3-downstream-TSS-luc reporter by TEF, CLOCK and BMAL1. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

h Transactivation of the EYA3-downstream-TSS-luc reporter by TEF, CLOCK, BMAL1 and BMAL2. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

Supplementary Fig. 3: Diurnal gene expression analysis show enrichment for repressors at SP ZT20

a. Volcano plots showing the number of genes up (yellow) and downregulated (blue) across the day in SP and LP. All comparisons were pairwise against ZT0 pairwise comparisons. ZT24 is included to show the consistency between time-points. ZT timepoints are indicated above each plot and the light dark transistion is shown below. Numbers on the plots are the number of significantly up or down regulated genes in that pairwise comparison (Supplementary Table 6).

b. RNA-seq log2 CPM plots. Light dark bars shown a the bottom and indicated on the graph by dotted line and grey shading. Error bars are SD. Statistical significances are in (Supplementary Table 6).

c. Enrichment term analysis using CPDB and radial plots of ZT time against -log10 pvalue (terms further out have lower pvalues) to indicate the functional enrichments at each peak phase.

d. TopGO analysis of genes differentially expressed between SP ZT20 and LP ZT4, i.e. 12 hours after lights off. Enrichment is shown as –Log10 p-value.

e. RNA-seq log2 CPM plots. Light dark bars shown a the bottom and indicated on the graph by dotted line and grey shading. Error bars are SD. Statistical significances are in (Supplementary Table 6).

Supplementary Fig. 4: STRING analysis reveals that the SP ZT20 repressors are high connected to the LP ZT4 transcriptome

a. Repressor genes up-regulated at ZT20 in SP (red boxes) were analysed for their connectivity, and therefore interaction with the genes up-regulated at ZT4 in LP (Yellow boxes and down-regulated (blue boxes). A high degree of protein-protein intereaction (PPI) connectivity was found to the up-regulated genes (see red lines on network). Black boarders on gene boxes indicate presence of E-boxes in the promoter.
b. Number of differentially expressed genes (DEGs), up-regulated (yellow) and down-regulated (blue) for a daily mean (all 24 hour timepoints) between SP vs LP on day 28, compared to LP ZT4 vs SP ZT20 contrast. Connectivity via protein-protein intereactions (PPI), defined by STRING to transcriptional repressors is indicated by the checkered shading (also represented in Fig. 4a). The significance of enrichment (fishers two-way exact test) for PPI connectivity within the up-regulated vs down-regulated genes is shown.

Supplementary Fig. 5: Melatonin implant validation and DEC1 repression

a. Melatonin concentration after implantation compared to sham and pre –implantation. Values are pg/ml. Mean and SD shown.
b. Transactivation of the EYA3-downstream-TSS-luc reporter versus EYA3-upstream-TSS-luc reporter by TEF, CLOCK, BMAL1 and BMAL2, and the effect of DEC1. The up-stream promoter is not significantly repressed by DEC1 but the downstream TSS activation

is significantly suppressed. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

c. Transactivation of the EYA3-downstream-TSS-luc reporter by TEF, CLOCK, BMAL1 and BMAL2, and the suppression by DEC1. The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM. d. Transactivation of the EYA3-downstream-TSS-luc reporter by TEF, CLOCK and BMAL1, and the effect of DEC2 (12.5 to 50ng). The experiment was repeated 4 times (n=4 per experiment), plot displayed is a representative result. A one-way ANOVA was performed on each individual experiment using Tukey's multiple comparisons test. Different letters indicate significant differences between groups ($P < 0.01$). Error bars SEM.

Supplementary Fig. 6: : Quality check of H3K4me3 ChIP-seq.

a. Pie charts revealing distributions of H3K4me3 peaks on each genomic feature. Peaks of promoter-TSS were located on ±1000 bp from TSS.

b. Bar plots revealing percentages of H3K4me3 peaks co-localised with CGIs (left) and CGIs co-localised with H3K4me3 peaks (right) on the sheep genome. Black is observed H3K4me3 peaks in PTs and grey is randomly shuffled peaks with the same fragment sizes as negative contorols.

Supplementary Fig. 7: Overview the analysis workflow for ChIP-seq

List of Supplementary Tables

Supplementary Table 1: Seasonal epigenetic changes, MACs peak calling ChIP-seq H3k4me3

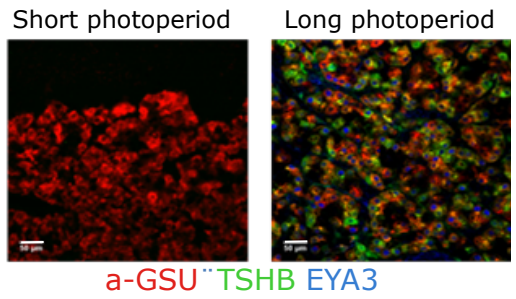
1164 Supplementary Table 2: Seasonal epigenetic changes, SICER differential peak
1165 analysis for H3K4me3

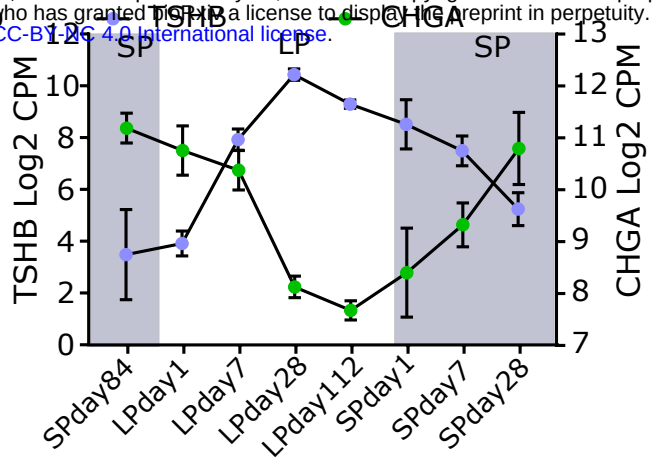
1166 Supplementary Table 3: Seasonal gene expression, pairwise contrasts for
1167 RNA-seq

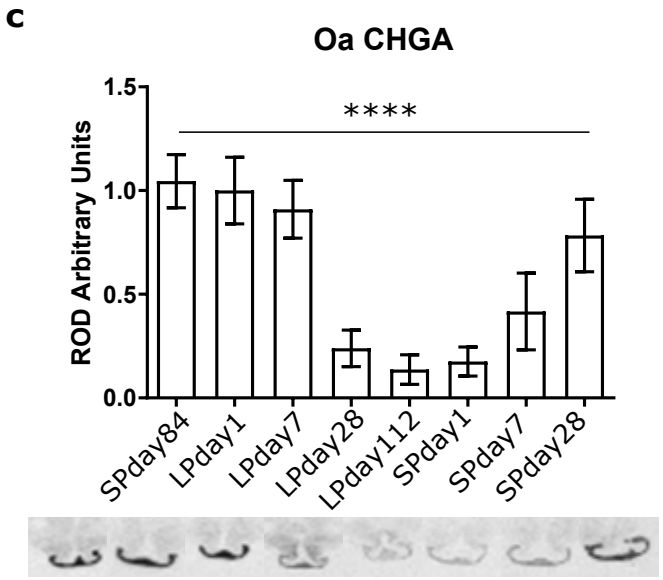
1168 Supplementary Table 4: CAGE transcription start site clusters and their relative
1169 seasonal abundance

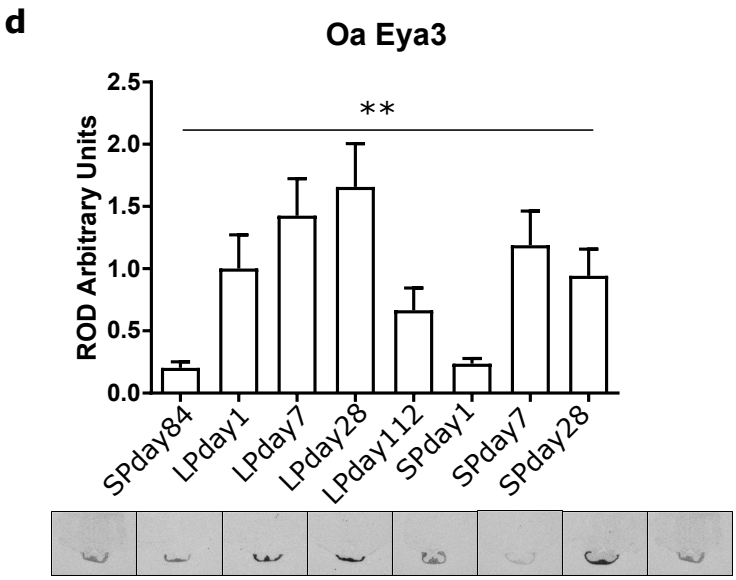
1170 Supplementary Table 5: RNA-seq expression for known H3K4me3 modulators

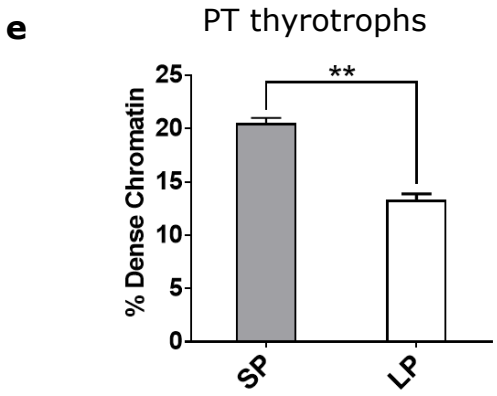
1171 Supplementary Table 6: Diurnal gene expression, statistical analysis of 24 hour
1172 profiles from RNA-seq

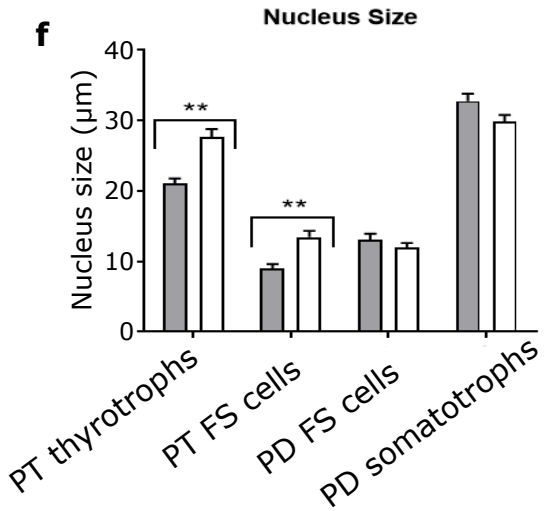


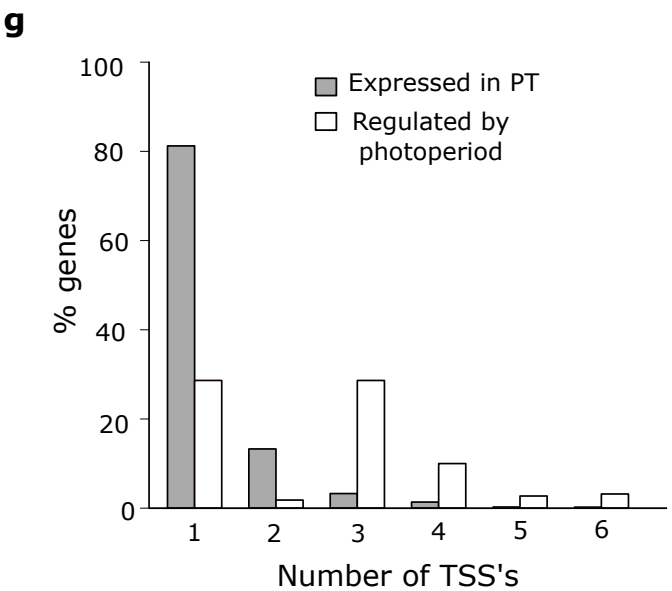


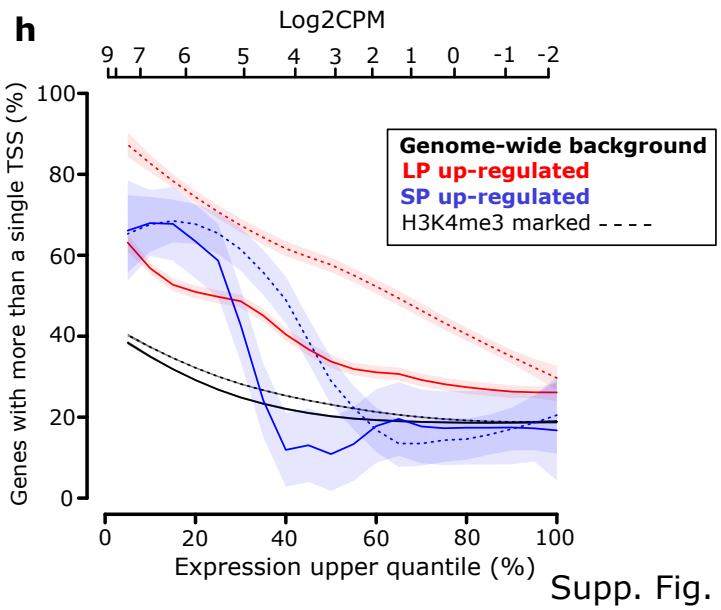




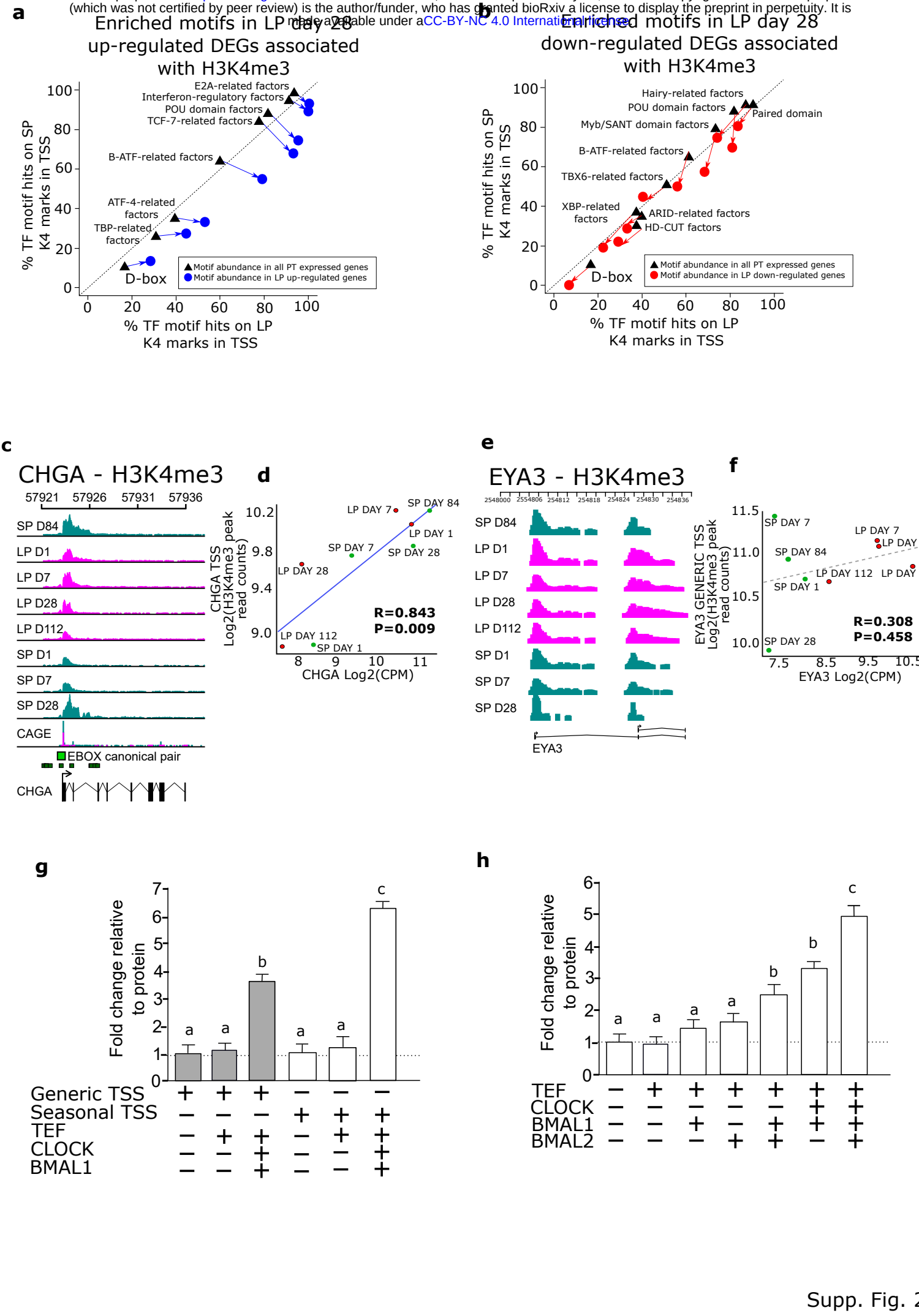




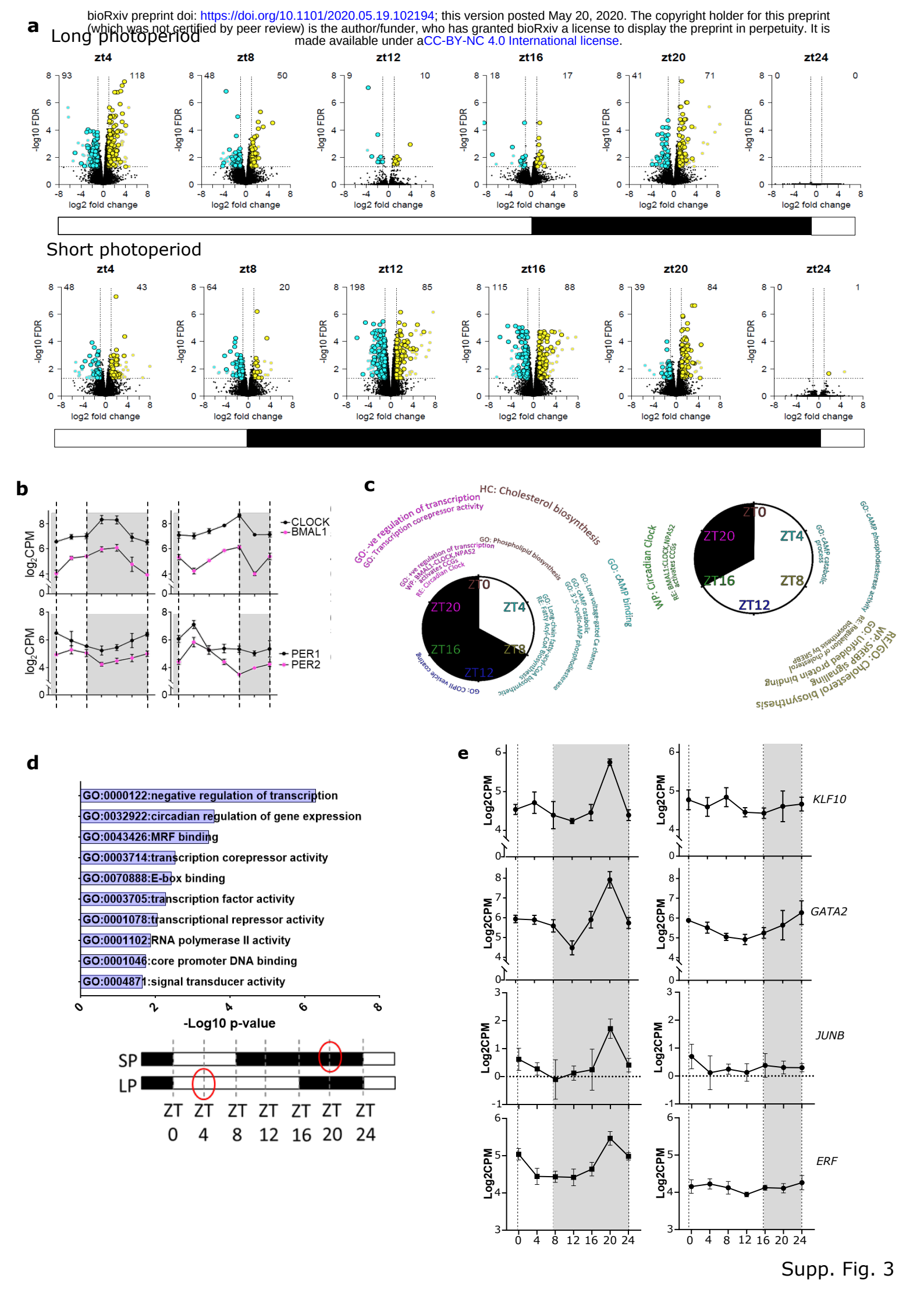




Supp. Fig. 1

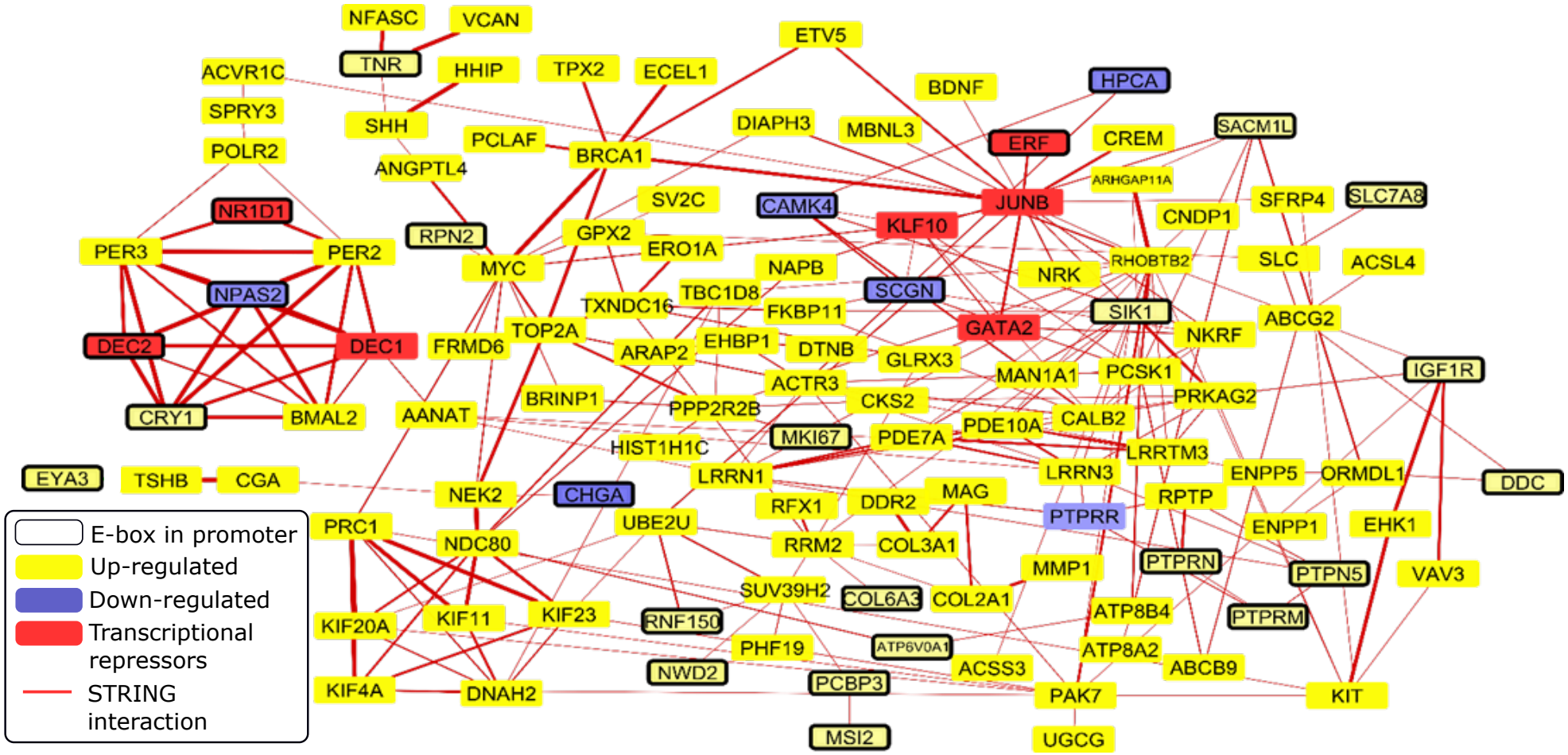


Supp. Fig. 2

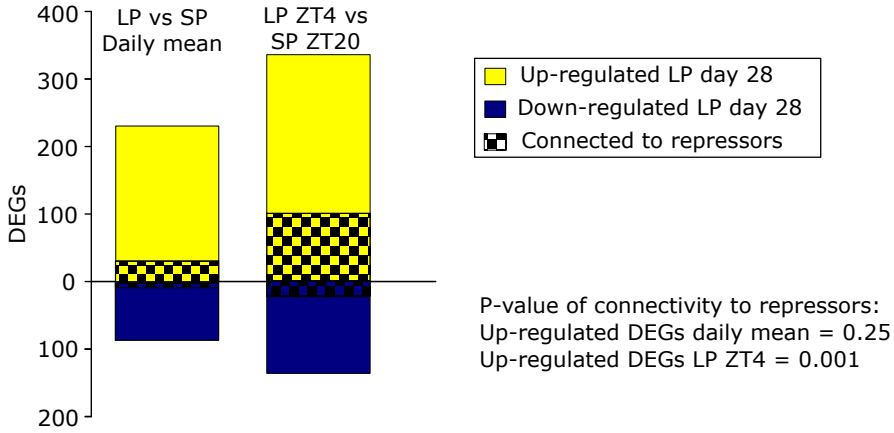


Supp. Fig. 3

a

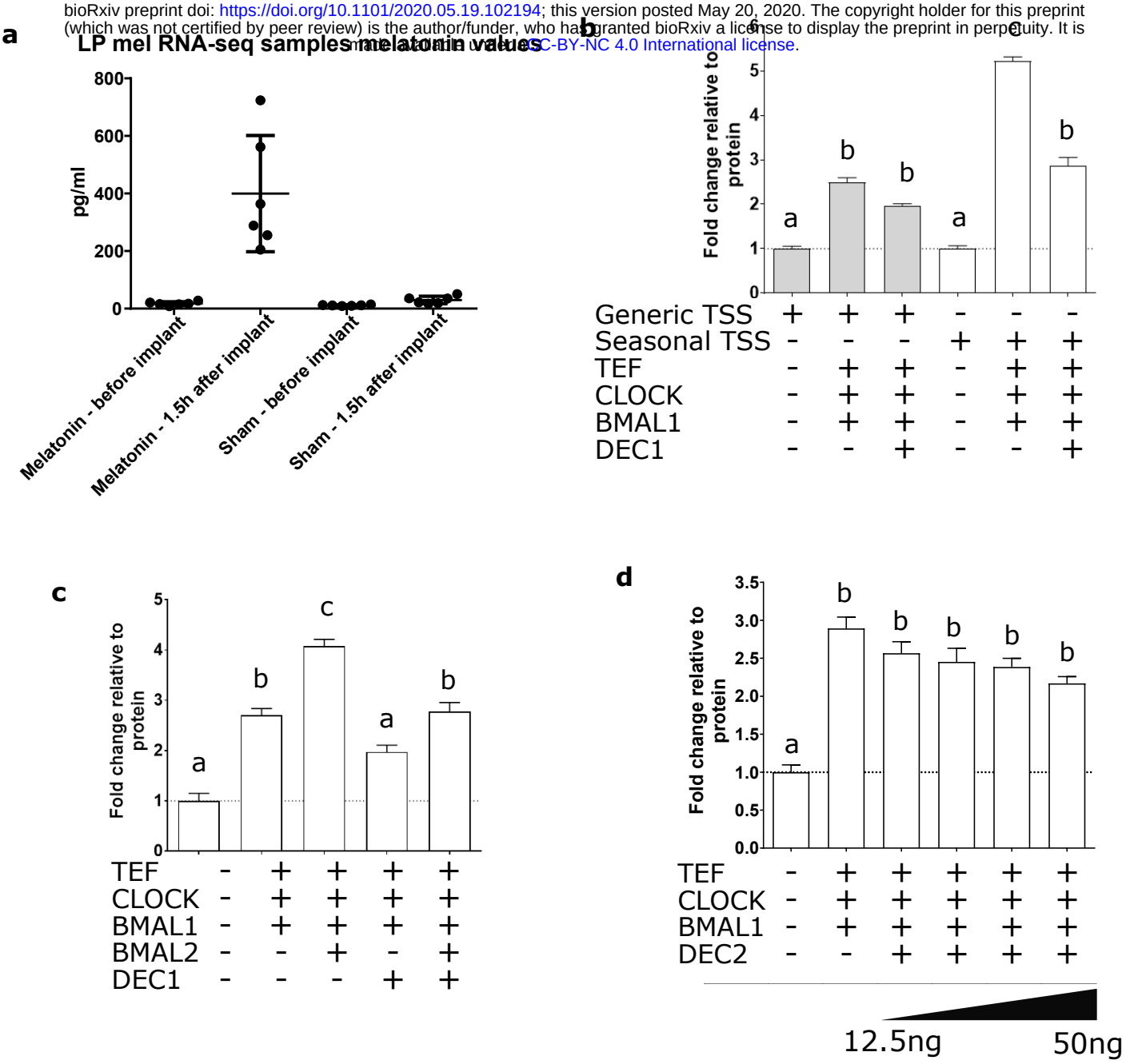


b



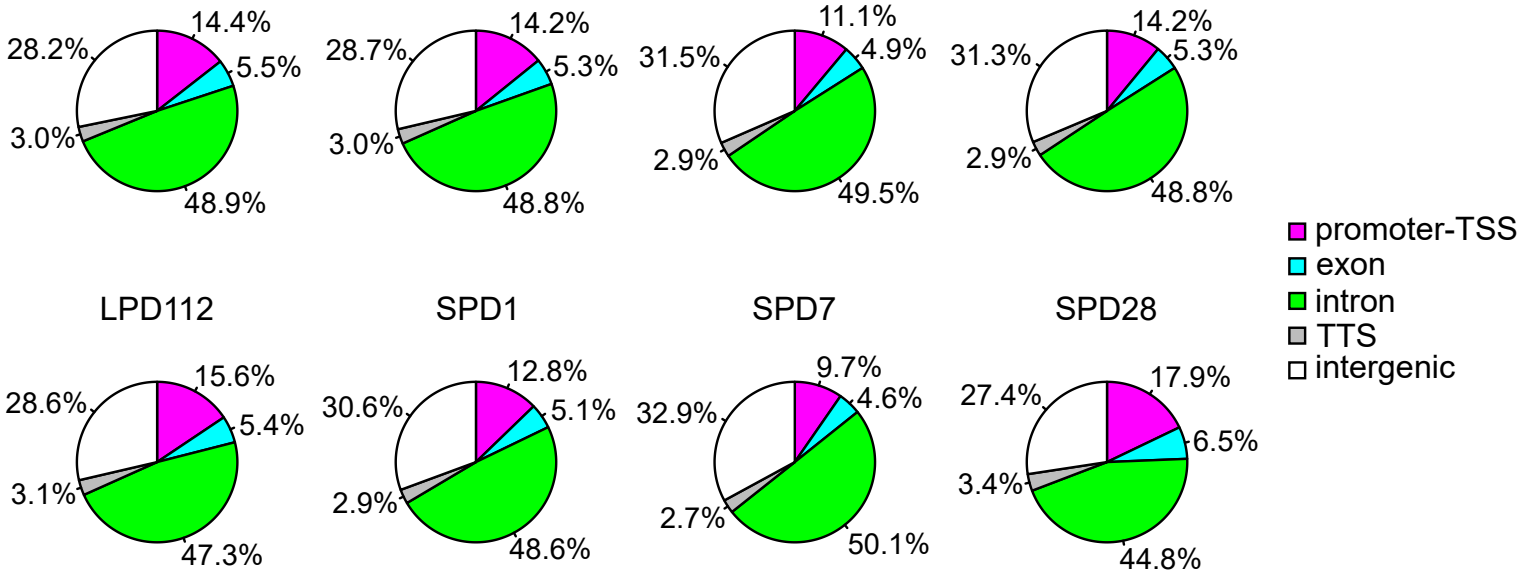
P-value of connectivity to repressors:
Up-regulated DEGs daily mean = 0.25
Up-regulated DEGs LP ZT4 = 0.001

Supp. Fig. 4

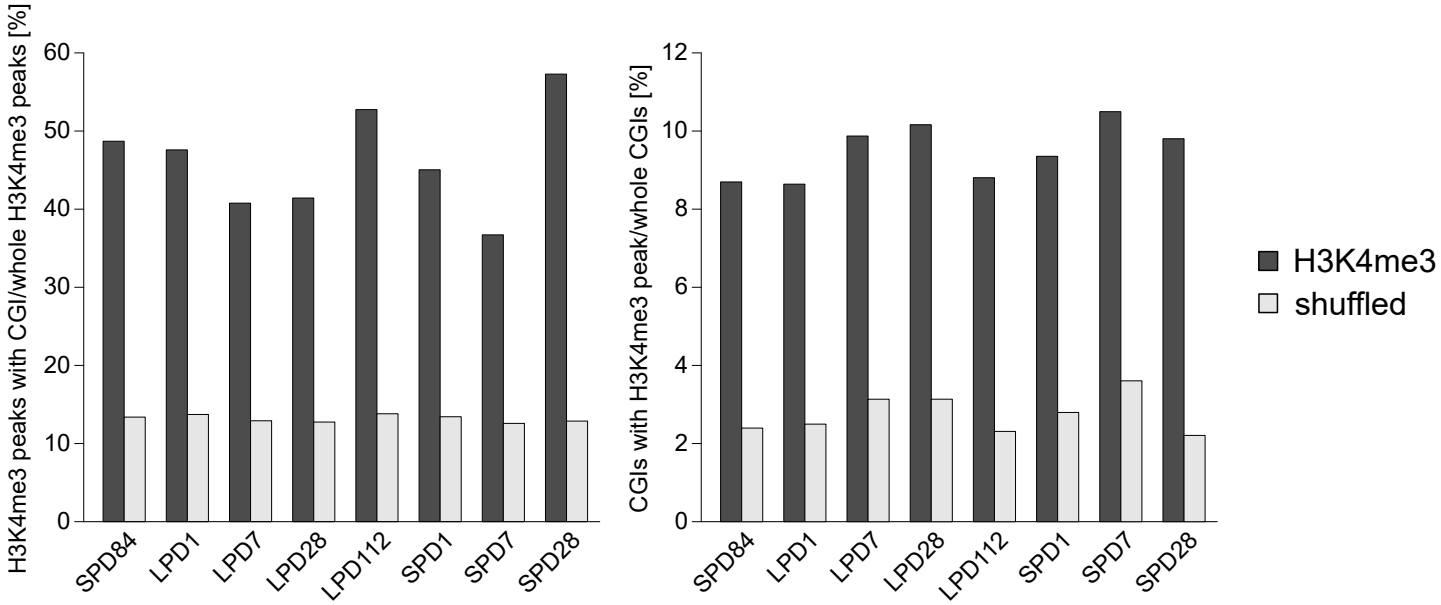


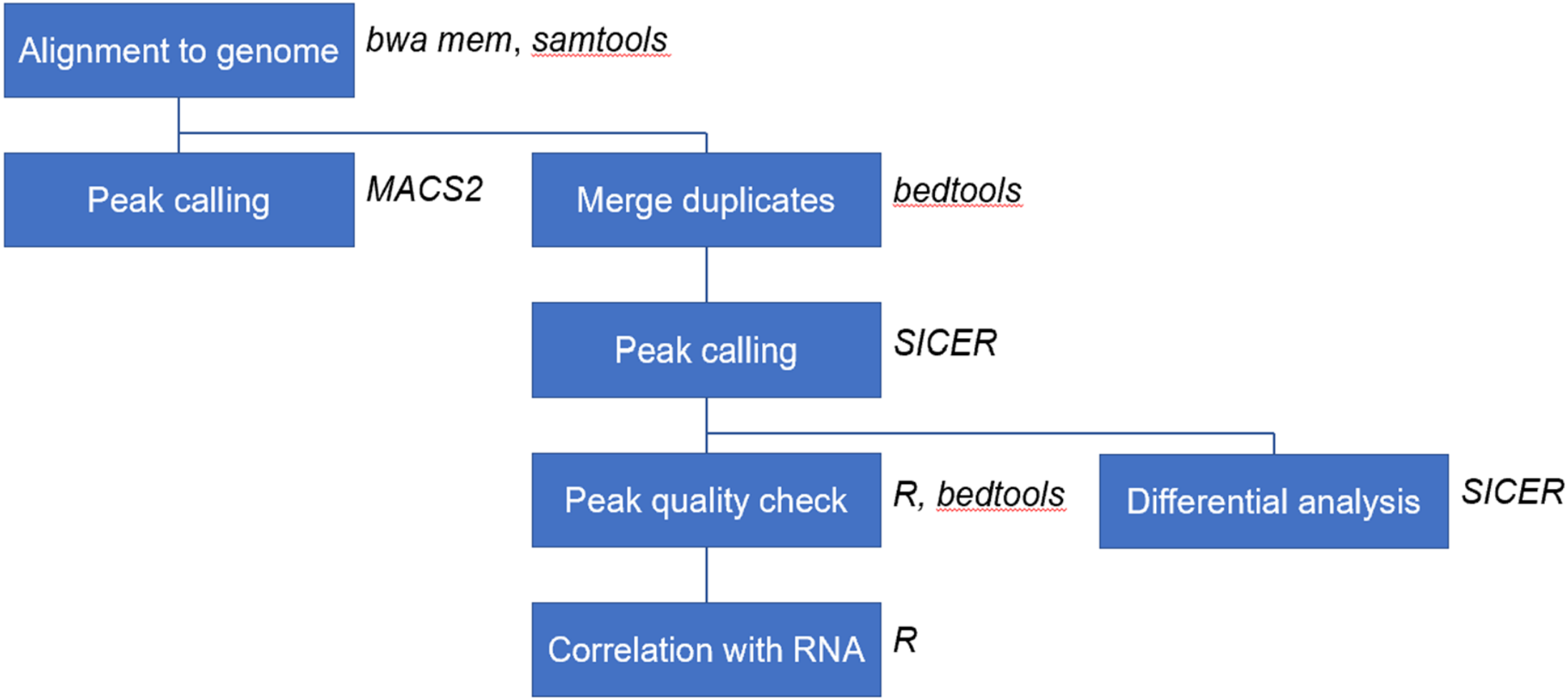
Supp. Fig. 5

a



b




Alignment to genome
bwa mem, samtools
Peak calling
MACS2
Merge duplicates
bedtools
Peak calling
SICER
Peak quality check
R, bedtools
Differential analysis
SICER
Correlation with RNA
R