Zinc finger homeobox-3 (ZFHX3) orchestrates genome-wide daily gene expression in the suprachiasmatic nucleus

Akanksha Bafna[1, 2*], Gareth Banks[1, 3], Vadim Vasilyev[4], Robert Dallmann[4, 5], Michael H. Hastings[6], Patrick M. Nolan[1*]

[1]Medical Research Council, Harwell Science Campus, Oxfordshire, OX11 0RD, United Kingdom

[2]Nuffield Department of Clinical Neurosciences, University of Oxford, OX1 3QU, United Kingdom

[3]Nottingham Trent University, Nottingham, NG1 4FQ, United Kingdom

[4]Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Coventry, United Kingdom

[5]Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, University of Warwick, Coventry CV4 7AL, United Kingdom

[6]MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, United Kingdom

*Corresponding author: akanksha.bafna@ndcn.ox.ac.uk , pmnolan10@gmail.com

ABSTRACT

The mammalian suprachiasmatic nucleus (SCN), situated in the ventral hypothalamus, directs daily cellular and physiological rhythms across the body. The SCN clockwork is a self-sustaining transcriptional-translational feedback loop (TTFL) that in turn co-ordinates the expression of clock-controlled genes (CCGs) directing circadian programmes of SCN cellular activity. In the mouse, the transcription factor, ZFHX3 (zinc finger homeobox-3), is necessary for the development of the SCN and influences circadian behaviour in the adult. The molecular mechanisms by which ZFHX3 affects the SCN at transcriptomic and genomic levels are, however, poorly defined. Here, we used chromatin immunoprecipitation sequencing (ChIP-seq) to map the genomic localization of ZFHX3 binding sites in SCN chromatin. To test for function, we then conducted comprehensive RNA sequencing at six distinct times-of-day to compare the SCN transcriptional profiles of control and ZFHX3-conditional null mutants. We show that the genome-wide occupancy of ZFHX3 occurs predominantly around gene transcription start sites (TSS), co-localizing with known histone modifications, and preferentially partnering with clock transcription factors (CLOCK, BMAL1) to regulate clock gene(s) transcription. Correspondingly, we show that the conditional loss of ZFHX3 in the adult has a dramatic effect on the SCN transcriptome, including changes in the levels of transcripts encoding elements of numerous neuropeptide neurotransmitter systems while attenuating the daily oscillation of the clock TF *Bmal1*. Furthermore, various TTFL genes and CCGs exhibited altered circadian expression profiles, consistent with an advanced in daily behavioural rhythms under 12h light- 12h dark conditions. Together, these findings reveal the extensive genome-wide regulation mediated by ZFHX3 in the central clock that orchestrates daily timekeeping in mammals.

INTRODUCTION

Circadian (approximately one day) clocks are internal biological time-keepers co-ordinating coherent 24-h molecular, behavioural and physiological rhythms that are synchronized with changing environmental conditions across day and night. In mammals, the circadian clock mechanism consists of cell-autonomous transcription-translation feedback loops (TTFLs) that drive rhythmic, ~24-h expression patterns of canonical clock components (Takahashi et al., 2008). The primary TTFL involves rhythmic transcription of period (*Per1*, *Per2*, and *Per3*) and cryptochrome genes (*Cry1* and *Cry2*). Thereafter, PER and CRY proteins form heterodimers that act on the dimeric complex of CLOCK (circadian locomotor output cycles kaput) and BMAL1 (brain and muscle ARNT-Like 1) to repress their own transcription, which is mediated by E-box (*CACGTG*) cis-regulatory sequences at their gene promoters. A secondary feedback loop is driven by *Rev-erbα* (*Nr1d1*), which is also directly transcribed through binding of CLOCK:BMAL1 at E-boxes. REV-ERBA feeds back to repress *Bmal1* transcription by competing with a retinoic acid-related orphan receptor (ROR) activator to bind ROR response elements (RREs) in the *Bmal1* promoter. Along with this, CLOCK:BMAL1 heterodimers also trans-activate a series of clock-controlled gene (CCGs), many of which encode transcription factors (TFs) that regulate additional downstream programmes of gene expression (Herzog et al., 2017; Hogenesch et al., 2003).

At an organismal level, mammalian circadian rhythms are directed by the suprachiasmatic nucleus (SCN) of the hypothalamus, which synchronizes robust 24-h oscillations across peripheral tissues (Hastings et al., 2018, 2019; Schibler et al., 2015). In the SCN, the cell-autonomous TTFL molecular clocks in individual cells are synchronized through intercellular coupling to form a coherent oscillatory network (Herzog et al., 2017; Maywood et al., 2021).

This intercellular coupling is maintained through several neuropeptide neurotransmitter systems, including vasoactive intestinal peptide (VIP), arginine vasopressin (AVP) and gastrin-releasing peptide (GRP) systems (Brown et al., 2005), while Prokinetin2 (PROK2) serves as a pace-making element in the SCN (Morris et al., 2021). Moreover, zinc finger homeobox-3, ZFHX3, is a key transcription factor in the SCN. It is highly expressed in discrete populations of adult SCN cells and exerts robust control over circadian behavioural rhythms. Intriguingly, while a dominant missense mutation in *Zfhx3* (short circuit, *Zfhx3*$^{Sci/+}$) results in a short circadian period in mice (Parsons et al., 2015), a conditional null *Zfhx3* mutation results in more severe behaviour with short circadian period and arrhythmic mice (Wilcox et al., 2017).

Given the crucial role of the SCN in orchestrating daily timekeeping in mammals, it is imperative to understand the role of ZFHX3 in regulating the molecular circadian clock, but although the effects of mutations in *Zfhx3* are known at the level of circadian behaviour, its molecular functions that underpin daily timekeeping are not well-characterized. Therefore, we set out to assess systematically the function of ZFHX3 by employing a multi-omics approach in the SCN. First, we analysed the genomic binding pattern of ZFHX3 by conducting ZFHX3 targeted chromatin immunoprecipitation (ChIP-seq) of SCN tissue to identify the genomic scope of ZFHX3 and its potential mechanistic roles. Second, we implemented a detailed transcriptional profiling of the SCN in control and *Zfhx3* conditional null mutants, using vibratome based microdissection (Bafna, Lau, et al., 2023). This allowed us to map the effect of ZFHX3 on tissue-specific expression of both rhythmic and non-rhythmic genes.

Overall, we have defined genome-wide occupancy of ZFHX3 in the SCN, showing that it is found predominantly around actively transcribed genes and is accompanied by modified histones and other SCN TFs. We noted that conditional (adult) deletion of ZFHX3 drastically modifies the SCN transcriptome by affecting genes important for circadian timekeeping, neuropeptide signalling and synaptic processes. An in-depth investigation on time-of-day dependent gene expression clearly highlighted loss of *Bmal1* rhythmic transcription along with aberrant 24-h oscillations for TTFL and non-TTFL genes. This change in rhythmic clock gene transcriptional profiles contributed towards a phase-advanced internal central clock mirroring the advance seen in the daily behavioural activity patterns of the null mutant mice. These findings reveal a pivotal and genome-wide physiological role for ZFHX3 in the maintenance of robust circadian timekeeping in the SCN and identify molecular substrates that underpin this role.

RESULTS

ZFHX3 *binds to active promoter sites in the SCN*

Genome-wide binding of ZFHX3 was profiled in the SCN using ChIP-seq. Briefly, genomic DNA was extracted from C57BL/6J mice raised in standard 12-h/12-h light-dark conditions at two distinct time-points; *zeitgeber* time (ZT) 03 (3 hours after lights on) and ZT15, and subjected to ChIP-seq (Methods). The sequenced reads were aligned to the mouse genome and significant sites of ZFHX3 occupancy were assessed in the context of previously identified modified histone binding regions (Bafna, Banks, et al., 2023) in the SCN. We found a total of 43,236 ZFHX3-bound sites, overlapping with approximately 22% of total modified histone regions in the SCN genome. About 60% of the ZFHX3 bound sites were found near promoters (<= 3 Kb of TSS) and so would be likely to identify direct

transcriptional targets of ZFHX3, whilst ca. 21% were in distal intergenic regions potentially involved in other modes of gene regulation.

When assessed in relation to modified histone binding, ZFHX3 sites fell into four categories. First, the majority of ZFHX3 peaks coincided with both H3K4me3 (trimethylation of Histone H3 at Lysine 4) and H3K27ac (Histone 3 lysine 27 acetylation) occupancy (48.9% of total ZFHX3 peaks at ZT3, Supplementary Table 1), and were focussed near transcription start sites (TSS $\pm$ 3kb) (Fig. 1B). Since such sites are likely responsible for active transcription (Beacon et al., 2021), the ZFHX3 occupancy in these sites highlighted probable transcriptional regulation by ZFHX3 in the SCN. Next, a subset of ZFHX3 sites (20.7%) were bound only by H3K4me3 (no H3K27ac) as illustrated at Fig. 1C. These are indicative of poised transcriptional regulation (Calo & Wysocka, 2013) and were also enriched near the TSS. About 8.2 % of ZFHX3 sites were bound by H3K27ac exclusively (no H3K4me3) (Fig. 1D). These sites were commonly found at intergenic and intronic regions and suggestive of enhancer-mediated gene regulation (Won et al., 2008). Finally, 22% of ZFHX3 sites were found to be devoid of modified histone binding and were also concentrated at distal intergenic regions (39.04%) (Fig. 1d). Overall, the broad occurrence of ZFHX3 alongside histone modifiers near the gene TSS and at distal intergenic regions in the SCN genome suggests its crucial role in gene regulation potentially via promoter-enhancer interaction.

Next, we assigned each ZFHX3 site to its closest transcribed gene ($\pm$ 3Kb) and performed functional annotation upon these genes. The functional annotation of genes which were associated with ZFHX3 and both the H3K4me3 and H3K27ac histone modifications were found to be associated with neurodegeneration and axonal guidance, synaptic processes known to be important in the SCN (glutamatergic, GABAeric and dopaminergic systems) and

various signalling pathways. Importantly, they also included pathways involved with circadian entrainment and timekeeping (Supplementary Fig. 1A), highlighting the reciprocal intersection between ZFHX3 and the SCN cell-autonomous clock. In contrast, genomic regions co-operatively bound by ZFHX3 and H3K4me3 but with no H3K27ac, or sites with ZFHX3 alone, were found to be associated with genes involved in several fundamental biological processes such as DNA replication or MAPK signalling, but notably not circadian rhythm (Supplementary Fig. 1B, C). Therefore, it is plausible that ZFHX3 binds at the TSS and regulates actively transcribing genes, particularly those involved in tissue-specific functions such as daily timekeeping, while its occupancy without histone modifiers and/or at poised state regulates genes responsible for other basal functions. Interestingly, it is worth noting that ZFHX3 occupancy did not differ appreciably between the two time-points tested (ZT03 vs. ZT15): ZFHX3was seen to bind consistently at the same genomic loci (Supplementary Table 1). This indicates that other factors/ mechanisms confer temporal control to these sites whilst the presence of ZFHX3 is permissive to such cis- and trans-regulatory cues.

We next examined the signature DNA binding motifs/ cis-regulatory elements occupied by ZFHX3, either alone or with specific histone modifications (H3K4me3, H3K27ac). As anticipated, this revealed an over-representation of AT-rich motifs (Parsons et al., 2015),. Notably, it also suggested co-occurrence of ZFHX3 with SCN developmental and maintenance TFs such as DLX (distal-less homeobox) and LHX (LIM homeobox) (VanDunk et al., 2011) across the tested categories (Supplementary Fig. 1D). Furthermore, ZFHX3-bound sites were also enriched for non-AT-rich motifs such as the RFX family of TFs, which have been implicated in transcriptional responses to light in the SCN (Araki et al., 2004). Interestingly, we noted enrichment of ARNT (bHLH) domains but only where ZFHX3 peaks

were seen adjacent to H3K27ac and H3K4me3 (i.e., active promoters). This clearly points towards the possibility of shared binding of ZFHX3 with the circadian clock TFs BMAL1:CLOCK at E-boxes and highlights a point of interaction between the core TTFL and ZFHX3 in the regulation of circadian gene transcription in the central pacemaker.

***Zfhx3* knock-out affects circadian behaviour and entrainment of mice**

To explore the contribution of *Zfhx3* to SCN function in the adult animal, we compared circadian wheel-running in cohorts of *Zfhx3*$^{Flox/Flox}$;UBC- mice that had been dosed with tamoxifen (gene deletion, +Tamoxifen group) with mice that had not been dosed with tamoxifen (no deletion genetic control, -Tamoxifen group). Comparisons of gene expression between the –Tamoxifen and +Tamoxifen groups by real time PCR confirmed loss of *Zfhx3* expression in the +Tamoxifen group. (Supplementary Fig. 2A). Behavioural analysis was performed in conditions of constant darkness and in a 12 h light/12 h dark cycle. Consistent with previous studies, this analysis showed that the free-running circadian period was significantly shortened by conditional deletion of *Zfhx3* (Fig. 2A). Furthermore, the overall activity rhythm was less well defined with significantly reduced amplitude, increased intra-daily variability and lower inter-daily stability in the deleted mice (Supplementary Fig. 2B, C, D). The full analysis is given in Supplementary Table 2. The loss of ZFHX3 therefore compromised free-running circadian behaviour.

Visual inspection of the actograms whilst the animals were in light/dark cycles suggested that animals with gene deletion showed poor entrainment to the light cycle, with animals showing increased activity in the hours prior to the light to dark transition (Fig. 2A, right). To quantify this, we analysed the distribution of wheel running activity across 24-h, by

expressing wheel running activity within each 30 minute segment of the day as a proportion of the total daily wheel running activity for that animal (Fig.2C). Repeated Measures ANOVA demonstrated a significant interaction between tamoxifen dosing and time of day (tamoxifen dosing x time: $F(47, 1316) = 9.013$; $p = <0.0001$). Pairwise analysis demonstrated that, compared to genetic control animals, gene deleted animals showed a significant increase in wheel running activity in the hours prior to lights off ($p = <0.05$ for ZT 10.5-12.5) and a significant reduction in wheel running activity in the hours prior to lights on ($p = <0.05$ for ZT 20 -23). Consistent with this, analysis of the activity onset demonstrated that gene deleted animals showed an earlier onset of activity than controls (Supplementary Fig. 2E). This phase advance under 12 h L/12 h D is consistent with the shortening of circadian period seen under DD. Thus, ZFHX3 contributes to circadian time-keeping and entrainment in the adult SCN, independently of any developmental roles.

ZFHX3 knock-out (KO) results in extensive effects across the SCN transcriptome

Having mapped the genomic binding sites of ZFHX3 and validated the behavioural effects of deletion of ZFHX3, we then employed total directional RNA sequencing to define the impact of *Zfhx3* deletion on the SCN transcriptome. Adult SCN tissues were collected from tamoxifen-treated *Zfhx3*$^{Flox/Flox}$;UBC-Cre^{-} (Cre-neg) and *Zfhx3*$^{Flox/Flox}$;UBC-Cre^{+} mice that were not treated with tamoxifen (-Tamoxifen, genetic control) , both serving as independent control groups, and tamoxifen treated *Zfhx3*$^{Flox/Flox}$;UBC-Cre^{+} (+Tamoxifen) mice i.e., *Zfhx3*-deleted, at six distinct times of light/dark cycle (Methods). As an initial analysis we compared overall (i.e., time-independent) gene expression levels between the groups (Cre-neg vs. +Tamoxifen) to examine the global effect of *Zfhx3* deletion on the SCN transcriptome, and found that almost 36% (n = 5,637) of total assessed genes were affected by the acute loss of

Zfhx3 in the SCN. Of these, 2,725 were down-regulated and 2,912 up-regulated in +Tamoxifen animals compared to Cre-neg controls. (Supplementary Table 3, Fig. 3A). Furthermore, as expected, both the control groups (Cre-neg and -Tamoxifen) showed consistent differential gene expression levels when compared to the mutant which mitigated any potential downstream effect of tamoxifen dosing *per se* (Fig. 3B, C).

To better understand the relationship between ZFHX3 occupancy on the genome and changes in gene expression, we compared the results of our initial CHIP-Seq with the RNA-Seq described above. Of the total differentially expressed genes, 80% (n = 4,543) revealed ZFHX3 occupancy at their promoter sites along with histone modifications (H3K4me3 and H3K27ac), showing a direct interaction between ZFHX3 bound TSS ($\pm$ 3Kb) and gene expression (p = 0, odds ratio = 6.3, Jaccard index = 0.4, signifying strong association). In contrast, only 7% of differentially expressed genes (n = 381) were solely bound by ZFHX3 at the TSS (p = 0.37, odds ratio = 1, Jaccard Index = 0.1, signifying weak association), indicating the effect of ZFHX3 on gene transcription is predominantly observed when jointly occupied by histone modification at gene TSS.

We then sub-selected highly differential genes ($\log_2$FC $\geq$1, FDR $\leq$ 0.05) and conducted functional gene enrichment analysis. Consistent with previous studies (Hughes et al., 2021; Parsons et al., 2015; Wilcox et al., 2021; Wilcox et al., 2017), genes that were downregulated were seen to be associated with neuropeptide signalling pathways (*Vip, Grp*), circadian rhythm, neuron differentiation etc., while those that were upregulated showed functional enrichment for cell differentiation and extracellular matrix organization processes (Fig. 3D). Interestingly, our comprehensive assessment also highlighted an effect of loss of ZFHX3 on

signal transduction by affecting cholinergic receptor (*Chrna3, Chrnb4*), melatonin receptor (*Mtnr1a*), prokineticin receptor (*Prokr2*), GRP receptor (*Grpr*), and the SCN peptide *Nmu*. Therefore, it is likely that loss of ZFHX3 leads to physiological impairment by affecting a diversity of neuropeptide neurotransmitter systems in the SCN and more generally, circadian-relevant transcripts.

.

<u>Loss of Zfhx3 affects rhythmic transcription of TTFL and clock-controlled genes</u>

The present and recent studies have demonstrated severe disturbances of daily and circadian behaviour in mice carrying a missense mutation (*Zfhx3*^{Sci}) (Parsons et al., 2015) or knock-out (*Zfhx3*$^{Flox/Flox}$;UBC-Cre^{+}) (Wilcox et al., 2017) of ZFHX3. In addition to assessing general gene up- or down-regulation in mutants (above), we next wanted to assess how deletion of ZFHX3 might alter transcriptional rhythms. For this, we investigated gene expression levels in the SCN at six distinct times-of-day (Hughes et al., 2017), starting at ZT02, at 4-h intervals (Methods). In total, we analysed the daily transcriptional profile of 20,521 genes in the SCN and compared the rhythmicity index between the control (Cre-neg) and mutant (+Tamoxifen) groups using statistical framework *dryR* (Weger et al., 2021). This enabled us to categorise the genes into one of 5 distinct modules (Methods): Module 1 refers to genes with no detected 24-h rhythm in their expression levels in either condition (n = 16832, Supplementary Fig. 3A); Module 2 consists of genes that were rhythmic in control SCN but lost rhythmicity in the ZFHX3-KO (n =1566, Fig. 4A); Module 3 consists of genes that gained daily rhythmicity after ZFHX3-KO (n = 403, Supplementary Fig. 3B); Module 4 was composed of rhythmic genes that exhibited no change between control and mutants (n = 1653, Fig. 4B); Module 5 included genes that retained rhythmic expression but with a change in either

amplitude or phase (peak expression) in the mutant group (n = 67) (Supplementary Table 4, Fig. 5).

Thus, we noted ~82% of compared genes did not show a daily rhythm (module 1). Conversely, ~18% of gene expressed in the SCN exhibited daily rhythmicity under one or other condition. Of these, ~45% were not affected by loss of ZFHX3, whereas ~55% of rhythmic transcripts showed alteration in their daily pattern following loss of ZFHX3. Indeed, ~42% of rhythmic genes required ZFHX3 to maintain that rhythm, whereas ~2% of genes (403/20521) gained rhythmicity in the absence of ZFHX3. Thus, ZFHX3 contributes, directly or indirectly, to a significant part of the daily rhythmic transcriptome of the SCN.

For each module, we compiled the gene list and examined the functional enrichment pathways. The genes that lost rhythmic transcription after the deletion of ZFHX3 (module 2) were seen to be involved in intracellular protein transport, signal transduction such as WNT-pathway linking circadian and cell-cycle processes (Matsu-Ura et al., 2018), oxidative stress response etc. (Supplementary Fig. 4A). On the contrary, genes that gain rhythmicity (module 3) were associated with RNA processing, stress and metabolic pathways. The genes that did not show any change in rhythmic expression (module 4) were found to be enriched for functional processes such as protein phosphorylation and de-phosphorylation, cytoskeleton organization, DNA repair etc. Interestingly, only a tiny proportion (~1.8%) of the rhythmic genes showed substantial changes in their daily cyclic gene expression profiles (phase and amplitude, module 5), but these were primarily involved in circadian rhythm generation, entrainment and clock-controlled gene expression (Supplementary Fig. 4B). This suggests

that in the absence of ZFHX3 the core circadian mechanism of the SCN is disrupted but is not ablated.

Indeed, almost all the genes driving the cell-autonomous TTFL were severely affected by *Zfhx3* perturbation with *Bmal1* demonstrating a complete loss of 24-h rhythm (Fig. 4A), and its counterpart *Clock* mRNA showing overall reduced expression levels (Supplementary Table 3). PER family genes (*Per1-3*) showed a significant advancement in peak expression with low amplitude (Fig. 5C). Specifically, we observed a phase advance of ca. 2-h for *Per2* and *Per3*, with *Per1* (Kuhlman et al., 2003) which is also known to mediate light-responsive pathways, advancing by only ~0.3-h. Strikingly, there was no obvious change in the rhythmic levels or phase for CRY family genes (*Cry1, 2*) (Fig. 4B), indicating that ZFHX3 acts only on select genes driving the cell-autonomous clock in the SCN. Given that *Bmal1* transcription is under the control of ROR (*Rora, b, c*) and REV-ERB (*Rev-erb a, b*) factors (Liu et al., 2008), we also noted loss of the rhythm in*Rorb* expression, whilst *Rev-erba* showed a phase advance by almost one hour (Table1).

Next, we examined the effect of loss of ZFHX3 on rhythmic expression of the clock-controlled genes (CCGs). Clock-controlled genes are basically genes whose expression is regulated by circadian clock factors (such as CLOCK:BMAL1) that are directly associated with the TTFL. Several CCGs like DBP (D site albumin promoter binding protein) show rhythmic transcription (Lopez-Molina et al., 1997), and in turn control expression of a repertoire of genes, generating a cascade of 24-h cycling genes in the SCN (Herzog et al., 2017). In our mutant group, *Dbp* mRNA levels also presented an advanced peak (~1.66 h) with reduced amplitude (Fig. 5C). Another such TF- *Creb3l1* (cAMP responsive element

binding protein 3-like 1), that has been implicated in intercellular coupling (Bedont et al., 2017), showed a distinctive change in phase and amplitude. Therefore, the alteration seen in the rhythmicity index at the level of TTFL genes due to the loss of ZFHX3 was recapitulated in the diurnal expression of certain CCGs in the SCN.

Finally, we studied the effect of ZFHX3 deletion on the rhythmic pattern of intercellular coupling agents (neuropeptides), whose overall mean expression levels were significantly reduced (Fig.3, *Vip*, *Grp*). Along with lowered baseline expression, we noted severe loss in 24-h rhythm for crucial SCN neuropeptides such as *Avp* (Fig. 4a), *Prok2* and *Npy* (Fig. 5c). For *Avp*, there was no apparent increase in daytime mRNA expression (Arima et al., 2002), and the levels remained intermediate throughout the day. Conversely, for both *Prok2* and *Npy*, transcript levels were reduced to minimal levels, possibly resulting in compromised intercellular-coupling in the SCN.

The effects of ZFHX3 on daily transcriptional profiles therefore varied between different gene categories. Whereas rhythmicity of downstream, clock-controlled genes was either independent of ZFHX3 or required ZFHX3 to be sustained, core circadian function was sustained but with altered dynamics. Intriguingly, this systematic approach not only highlighted the genes that lost 24-h rhythmic oscillations, such as *Bmal1* and the neuropeptide *Avp* (module 2, Fig. 4A), but also identified genes that retained rhythmicity with changes in their waveform (module 5). Given the fact that ZFHX3 deficient mice do not lose rhythmic behaviour in a daily LD cycle (Wilcox et al., 2017) (Fig. 2), but do exhibit a short circadian period or arrhythmicity in free-running (complete darkness: DD) conditions,

we believe our targeted SCN-specific assessment unravels the vital role of ZFHX3 in the mechanistic regulation of the daily molecular clock.

Table 1: *Effect of loss of ZFHX3 on core-clock genes in the SCN*

Gene	Rhythm	Phase	Amplitude
Per1	Altered	-0.36	0.72 (low)
Per2	Altered	-2.0	0.69 (low)
Per3	Altered	-2.1	0.30 (low)
Cry1	No change	-	-
Cry2	No change	-	-
Bmal1	Lost	-	-
Clock	No change	-	-
Nr1d1	Altered	-0.99	0.67 (low)
Nr1d2	No change	-	-
Rorb	Lost	-	-

The state of the daily SCN clock in ZFHX3 deficient mice

In order to link the perturbations of the core molecular clock in the SCN of ZFHX3 deficient mice with their altered behavioural profiles, we used TimeTeller, a recently established method to predict circadian clock phase alongside the parameter 'Theta' which provides an index of clock disruption from transcriptomic data (Vlachou et al., 2024). This approach allows one to assess the clock globally as a system instead of gene by gene. For this, an initial training model was built using the available Cre-neg and -Tamoxifen (n = 35, control) mice data (Fig. 6A). Cross-validation for the training model suggested that the model is informative and predicts the control SCN time well with a median prediction error of -0.08 hours (95% CI: -0.83;1.33) and 0.50 hours (95% CI: -1.67;3.00), i.e., the difference between the TimeTeller predicted time and the known ZT of SCN collection (Supplementary Fig. 5A). This was further assessed by predicting SCN timing in an independent dataset of SCN RNA-Seq (Deota et al., 2023). Similar to the TimeTeller predictions for the control group, median prediction error was 0.00 hours (95% CI: -2.00;1.33) for the *ad libitum* fed and -0.66 hours (95% CI: -1.83;1.00) for the temporary night-time restricted feeding groups. These are not

distinguishable from the control SCN in the current study (Fig. 6B). Similarly, the clock disruption parameter theta is low (Medians: 0.03, 0.03 and 0.04) as is expected for a fully functional and well predictable circadian clock in all three control groups. Furthermore, TT clock functionality parameter theta suggests that +Tamoxifen SCN have a similarly functional daily clock compared to the control SCN (Median 0.04 vs. 0.03 in controls (Supplementary Fig. 5B).

While TT's theta metric suggests the clock in the ZFHX3-deleted SCN is not disrupted as it is indistinguishable from controls, the phase of the +Tamoxifen SCN was consistently advanced by a median 1.83 hours (95.0% CI 0.585, 2.67), when compared to the current controls and the published data (Deota et al., 2023) as illustrated in Fig.6B. Of note, detecting this ~2-h phase difference from data sampled at 4-h intervals showcases the advantages in robustness and sensitivity of assessing the clock as a system rather than by a gene-by-gene approach, and TimeTeller as a tool of inquiry. In summary, according to TimeTeller, the SCN daily clock is intact but advanced by approximately 2 h in comparison to the control groups after loss of ZFHX3. This is consistent with the observed daily behavioural activity in mutant mice as shown in Fig. 2.

DISCUSSION

The SCN is distinctive in its ability to sustain autonomous rhythmicity which is driven by circadian cascades of transcription that direct metabolic, electrophysiological and signalling rhythms. Our current findings highlight the instrumental role of the transcription factor, ZFHX3, in driving the daily timekeeping mechanism. By comparing the daily behavioural activity of ZFHX3-KO mice with the genetic matched control, we clearly noted an advancement in LD activity and shortening of circadian period in mutants. This finding is

consistent with previous studies (Parsons et al., 2015; Wilcox et al., 2021; Wilcox et al., 2017) and prompted a detailed molecular investigation in the SCN. Through targeted ChIP-seq we successfully demonstrated the genome-wide occupancy of ZFHX3 in SCN chromatin with 78% of ZFHX3 bound sites found to be co-localized with modified histones (H3K4me3, H3K27ac), and a vast majority of these concentrated near gene TSS . This provides strong evidence in support of the observed pervasive regulation conferred by ZFHX3 on the SCN transcriptome. That said, a considerable proportion of ZFHX3 sites were also seen devoid of joint histone modifications (H3K4me3 + H3K27ac) and enriched for CTCF binding sites. CTCF is a ubiquitous DNA binding protein that has exceptional roles in genomic organization, chromatin looping and gene regulation (Holwerda & de Laat, 2013; Kim et al., 2015; Mach et al., 2022). This co-occurrence hints on the potential accessory role of ZFHX3 in modifying genome topology and chromosomal looping to bring about changes in target gene expression by regulating promoter-enhancer interactions (Pugacheva et al., 2020). Interestingly, along with CTCF binding sites, ZFHX3 sites were also enriched for other SCN enriched regulatory factors; RFX, LHX, ZIC (Araki et al., 2004; Hatori et al., 2014) (Supplementary fig. 1D), suggesting well-coordinated tissue-specific gene regulation.

Given the crucial role of ZFHX3 in daily timekeeping, we speculated a dynamic change in its genomic occupancy and compared the binding intensity of the transcription factor at two anti-phasic (12-h apart, ZT3 vs. ZT15) time-points (Methods). Surprisingly, we did not detect any differential binding site(s) between the tested time-points, but rather noticed a uniform and coherent localization. This suggests that ZFHX3 belongs to a family of poised transcription factors, whose occupancy at regulatory sites does not change during the day, but could potentially act as a co-activator/mediator to modulate oscillating gene transcription. Indeed, in a recent unpublished study (Del Rocío Pérez Baca et al., 2023), ZFHX3 was shown to be

associated with chromatin remodelling and mRNA processing factors, strengthening the viewpoint as a co-regulator. It is, therefore, highly plausible that ZFHX3 occupies the targeted TSS and partners with clock TFs such as BMAL1 *(CACGTG)*, as observed around active sites, to regulate rhythmic gene transcription.

Along with ChIP-seq, we executed detailed RNA sequencing to study the effect of loss of ZFHX3 specifically in the SCN. Crucially, ZFHX3 affected almost a third of the transcriptome in the SCN. An overall (time-independent) assessment of deletion of *Zfhx3* highlighted a significant impact on expression of genes involved in several biological processes ranging from cellular differentiation to behavioural responses. We then performed a comprehensive time-dependent assessment to delineate the functional role of ZFHX3 in regulating the circadian timekeeping mechanism. We observed about 55% of rhythmic genes showed alteration at distinct levels in their daily transcriptional pattern following loss of ZFHX3.

At the TTFL level, ZFHX3-KO resulted in the loss of rhythmic *Bmal1* gene expression. It remains to be determined whether this is associated with changes at the posttranslational level, given the findings by Abe et al where arrhythmic *Bmal1* transcript did possess posttranslational rhythm (Abe et al., 2022), and could explain the subsequent 24-h driven *Per (1-3)* and *Cry (1-2)* mRNA rhythms noted in our study. Likewise, in a separate study constitutive *Bmal1* expression has been shown to sustain rhythmic levels of transcriptional and translational products of *Per, Cry, Clock, RORc,* and *Rev-erbα* by means of mathematic modelling (Mirsky et al., 2009). Arguably, clear circadian rhythmicity in BMAL1, CLOCK, or REV-ERBα protein levels were not sustained in association with constitutive *Bmal1* levels

under *in-vitro* conditions (Padlom et al., 2022). Therefore, in future it will be interesting to study the daily abundance of protein levels of the clock genes in *Zfhx3*-deficient SCN, with its associated loss of *Bmal1* mRNA 24-h rhythm, to gain a complete understanding of the clock molecular circuitry.

Broadly speaking, aside from *Bmal1*, the majority of clock genes did not lose 24-h rhythmicity but showed an early peak in expression, ranging between ~0.4 – 2-h that perfectly resonates with the advanced rhythmic behaviour seen in mutants. In this case, it is difficult to delineate a straightforward cause-and-effect axis in terms of the direct targets of the perturbed ZFHX3 activity. Based on the current data, ZFHX3 is possibly partnering with CLOCK:BMAL1 to regulate the daily transcription of clock genes (*Per, Ror, Reverb)* and CCGs. Two possible scenarios could explain the noted effect on the TTFL genes due to lack of ZFHX3 (Fig. 5). First, ZFHX3 preferentially controls *Per* family gene transcription (and not *Cry*) by binding at the promoter (Fig. 1). Loss of ZFHX3 affects the daily rhythm of *Per* gene transcription, and potentially PER protein levels, which in turn disturbs the CLOCK:BMAL1 driven transcription cascades. This leads to an aberrant daily transcriptional profile of the *Ror* (activator)/ *Reverb* (repressor) genes as observed, which are known to control the amplitude and oscillation of *Bmal1* (Liu et al., 2008), and further results in loss of *Bmal1* transcriptional rhythm. Therefore, the observed change in *Ror* and *Reverb* family gene expression levels could be reminiscent of the effect of ZFHX3 on the TTFL. In a second scenario, the constitutive expression of *Bmal1* mRNA could result from the primary effect of ZFHX3 on CLOCK:BMAL1 driven *Ror/Reverb* genes (affecting ROR/REVERB translational products), which does not disrupt the TTFL, but weakens the daily clock. The direct /indirect effect is seen in the cyclic expression of *Per* family genes, with relatively greater phase change in *Per2* and *Per3* than the *Per1* gene (Kuhlman et al., 2003)

(Supplementary fig. 6). It is, however, noteworthy that the CRY family of genes *(Cry1, Cry2)* did not show any obvious change in the daily expression and is supposedly shielding the molecular clock , as recently postulated by Abe et al (Abe et al., 2022).

As previously reported (Parsons et al., 2015) , lack of ZFHX3 also resulted in severe reduction in the SCN neuropeptides such as *Avp, Vip, Grp, Prok2*, that potentially compromises intercellular coupling and synchronization of circadian timekeeping. This universal effect, where loss of ZFHX3 does not deter a particular but all the key SCN neuropeptides, clearly points that 1) ZFHX3 is prevalent in almost all neuronal populations within the SCN, that are largely clustered based on their neuropeptide expression (Morris et al., 2021; Wen et al., 2020) 2) potentially results in loosely coupled intercellular clocks that are still able to drive 24-h behavioural and molecular rhythms but with advanced phase. The prevalence of ZFHX3 in neuronal populations is also supported by a complete independent assessment in which almost 60% of the top 100 categorised neuronal genes in the SCN (as per single-cell atlas) were seen to be affected by ZFHX3 deletion based on the current dataset. In contrast, only 20% of the 100 most expressed genes in the astrocytic cell populations were impacted.

Along with this, lack of *Zfhx3* also leads to aberrant daily transcriptional rhythms of several CCGs, with the majority of them seen to lose daily oscillations. This could potentially result in hindrance of clock-regulated vital biological processes such as regulation of cell-cycle, metabolic pathways etc. Indeed, ZFHX3 is reported to be crucial for the development of a functional SCN (Wilcox et al., 2021), and could act as an ‘integrator’ of afferent signals such as light and feeding to generate timely and robust physiological outputs. Moreover, genetic

ablation of *Zfhx3* has been shown to result in a compromised light sensing mechanism in retina (Hughes et al., 2021), while the missense mutation (*Zfxh3*$^{Sci/+}$) results in reduced growth and altered metabolic gene expression in the arcuate nucleus of the hypothalamus (Nolan et al., 2023) converging on its functional roles in other brain regions. The pleiotropic roles of ZFHX3 further extend in non-brain tissues with its recently described crucial involvement in angiogenesis for tumor development (Fu et al., 2020), cardiac function (Jameson et al., 2023) and spinocerebellar ataxia (Figueroa et al., 2024; Wallenius et al., 2024). It is important to note that although in the present study we used adult-specific *Zfhx3* null mutants resulting in global loss of ZFHX3, the circadian effects observed both at molecular and behavioural levels are independent of its pleiotropic role(s) in other tissues. In future, however, it will be beneficial to carry out the investigation using an inducible, SCN specific genetically ablated model to circumvent any confounding effects. Regardless, the current research is of great importance as it not only maps the genomic binding of ZFHX3, but also demonstrates its critical association with both lineage- and clock-controlling TFs that perfectly resonates with the observed profound effects of ZFHX3 on the SCN transcriptome, regulating a vast majority of genes ranging from circadian timing to synaptic transmission and inflammatory responses. This highlights how ZFHX3 regulates the core-clock (TTFL) in addition to its control over CCGs in order to drive sustained transcriptional rhythms in the SCN. Given the extensive pleiotropy of ZFHX3, it provides a solid substrate to further understand the functional interaction of ZFHX3 in distinct tissues indispensable for overall health and fitness.

METHODS

Mice

All animal studies were performed under the guidance issued by Medical Research Council in Responsibility in the Use of Animals for Medical Research (July1993) and Home Office Project Licence 19/0004. WT C57BL/6J were maintained and provided in-house by MRC Harwell. Animals were group housed (4-5 mice per age) in individually ventilated cages under 12:12h light-dark conditions with food and water available ad libitum. Adult specific knock out *Zfhx3* mice were generated using an inducible Cre line as described by Wilcox et al (Wilcox et al., 2017). Briefly, tamoxifen inducible Cre mice (UBC-Cre) were crossed to *Zfhx3* floxed ($Zfhx3^{Flox}$) mice to produce an initial stock of $Zfhx3^{Flox/+}$;UBC-Cre^{+} mice. These were subsequently crossed to $Zfhx3^{Flox/+}$ mice to generate experimental cohorts (Sun et al., 2012). Out of the 6 possible genotype combinations, 2 genotypes; $Zfhx3^{Flox/Flox}$;UBC-Cre^{-} (mice homozygous for $Zfhx3^{Flox}$ allele but not carrying the Cre allele) and $Zfhx3^{Flox/Flox}$;UBC-Cre^{+} (mice homozygous for $Zfhx3^{Flox}$ allele and hemizygous for Cre allele) were assigned as "control" (Cre-neg) and "experimental" (+Tamoxifen) groups, respectively, and dosed with tamoxifen. In addition, we also included a cohort of $Zfhx3^{Flox/Flox}$;UBC-Cre^{+} mice without tamoxifen treatment as a secondary control (-Tamoxifen). Animals were group housed (4-5 mice per age) in individually ventilated cages under 12:12h light-dark conditions with food and water available ad libitum.

Experimental Design

WT C57BL/6J mice (aged between 8 to 12 weeks) were used for SCN tissue punch collection as described by Jagannath et al. (Jagannath et al., 2013) at ZT3 and ZT15, where lights on at 7 am (ZT0) and lights off at 7 pm (ZT12), for ZFHX3-ChiP-seq. For SCN RNA-Seq, mice from control (Cre-neg and -Tamoxifen) and *Zfhx3* knock out (+Tamoxifen) cohorts were used for SCN tissue collection (Bafna, Lau, et al., 2023) at 6 distinct time-points starting from ZT2 at every 4h. To achieve high resolution transcriptomic dataset, four biological replicates per time-point were collected independently from the control (Cre-neg) and experimental

(+Tamoxifen) groups. For the second control group viz. -Tamoxifen; 2-4 biological replicates were collected due to time constraints. Furthermore, each biological replicate for all the three tested condition constituted 3 individual SCN samples.

Animal behaviour

Circadian wheel running was performed as previously described (Banks & Nolan, 2011). Briefly, mice were singly housed in cages containing running wheels, placed in light controlled chambers and wheel running activity monitored via ClockLab (Actimetrics). Animals were given a five-day acclimatisation period in a 12-h light/dark cycle (100 lux light intensity) before their activity was monitored for 32 days in constant darkness and 24 days under a 12-h light/dark cycle (100 lux light intensity). Circadian activity analysis was performed upon the entire of the period in constant darkness and upon the final seven days in the light/dark cycle. Circadian analysis was performed using ClockLab Version 6 (https://actimetrics.com/products/clocklab/clocklab-analysis-version-6/). *Zfhx3*$^{Flox/Flox}$;UBC-Cre^{+} animals were used for this analysis, with animals being randomly assigned into tamoxifen un-dosed (-Tamoxifen) or tamoxifen dosed (+Tamoxifen) cohorts. The numbers of animals per cohort were: male -Tamoxifen = 8; male +Tamoxifen = 8; female -Tamoxifen = 8; female +Tamoxifen = 8. Statistical analysis and graphing of circadian data was performed using GraphPad Prism 8.

SCN tissue collection for RNA-seq

Animals were sacrificed at the aforementioned time-points by cervical dislocation and brains were removed and flash frozen on dry ice. The frozen brains were placed on dry ice along with the brain matrix (Kent Scientific, Torrington CT, USA) and razor blades to chill. The brain dissection was performed on a vibrating microtome (7000smz-2 Vibrotome, Campden Instruments) under chilled conditions (Bafna, Lau, et al., 2023). First, the frozen brain was

placed on the chilled brain matrix and a ca. 3mm thick region was removed from the caudal end using a razor blade to discard the cerebellum and attain flat surface. Then, the dissected brain was super-glued to a chilled (2 –3°C) metal chuck and placed into the metal specimen tray of the Vibratome. The metal specimen tray was filled with pre-chilled 1X RNase free PBS (phosphate buffered saline, ThermoFisher Scientific, #AM9625). Coronal slices measuring 250 uM in thickness were cut from rostral to caudal at slow speed (0.07mm/s) and pre-tested specifications ; Frequency: 70Hz, Amplitude: 1.00mm. Once cut, the slice with intact optic chiasm, third ventricle and bilateral SCN tissues (Fig. S1) was selected and transferred on to pre-chilled glass slides (ThermoFisher Scientific SuperFrost Plus Adhesion slides) with the help of paint brush. The glass slide was placed on top of cold block and SCN tissues were collected by pre-chilled forceps and dissection needles under dissection microscope (Nikon SMZ645 Stereo Microscope).

ZFHX3 ChIP and sequencing

ChIP was conducted using the Active Motif ChIP□IT enzymatic kit (https://www.activemotif.com/catalog/868/chip-it-high-sensitivity) according to the manufacturer's instructions. Briefly, SCN punches were pooled from 80 mice at each designated times (ZT3, ZT15) corresponding to one biological replicate per timepoint, and fixed for 15 min at room temperature with gentle shaking in 1% formaldehyde. The cross□linking reaction was stopped by addition of glycine to a final concentration of 0.125 M. Cross□linked chromatin was sheared by with Active Motif's EpiShear probe sonicator (cat# 53051). Genomic DNA (Input) was prepared by treating aliquots of chromatin with RNase, proteinase K and heat for de-crosslinking, followed by SPRI beads clean up (Beckman Coulter) and quantitation by Clariostar (BMG Labtech).

24 ug sheared chromatin sample collected from each time point (ZT3, ZT15) was incubated with 12ul of rabbit polyclonal antibody against ZFHX3 (Parsons et al., 2015), together with magnetic beads, on a rotator at 4°C overnight. This was followed by crosslinking reversal and proteinase K digestion as per the manufacturer's protocol. The recovered DNA was purified using PCR purification reagents and subjected to quantitative PCR for quality inspection. ZFHX3 immuno-precipitated genomic DNA along with their corresponding input samples was sequenced on Illumina (75-bp, single end reads: ~30 million reads/sample) NextSeq 500 sequencing platform by Active Motif Epigenomic Services.

ZFHX3 ChIP-seq peak calling and data analysis

Raw sequence data was quality assessed by removing low-quality bases (Phred < 20) and trimmed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic v0.36 (http://www.usadellab.org/cms/index.php?page=trimmomatic), respectively. FASTQ files containing trimmed sequences were then aligned to mm10 genome assembly to generate binary alignment map (BAM) files and used for peak calling by MACS algorithm v2.1.0 (Zhang et al., 2008) at a threshold of $q < 0.01$. Finally, peaks from ZT3 and ZT15 were analysed for differential binding by a Bioconductor package Diffbind v2.10.0 (http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf) at default settings with $p < 0.05$ and fold change > $\pm$ 2. No significant differential ZFHX3 bound peaks were detected between ZT3 and ZT15. In order to investigate the ZFHX3 peaks overlapping with histone modifications (H3K4me3, H3K27ac) in the SCN, data was downloaded from NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) accession number GSE217943. The data was normalized using bamCoverage tools and bigwig files were generated for visual inspection using UCSC Genome Browser and processed using BEDTools suite (Quinlan & Hall, 2010). ChIPSeeker

package (Yu et al., 2015) was used to retrieve the nearest gene around resultant peaks and to annotate the genomic region of the peak.

RNA extraction

Total RNA from each sample replicate was extracted using RNeasy Micro Kit (QIAGEN, #74004) following the protocol for microdissected cryosections as per manufacturer's guidelines and stored at -80°C. Quality and quantity of RNA were measured using an Agilent Bioanalyzer (Pico chip) and a Nanodrop1000 (Thermo Fisher Scientific, Waltham, MA USA), respectively.

RNA Sequencing and data analysis

PolyA RNA-Seq libraries were prepared by Oxford Genomics Centre, University of Oxford using NEBNext Ultra II Directional RNA Library Prep Kit (NEB, #E7760) and sequenced on NovaSeq6000_150PE (150 bp paired-end directional reads: ~40 million reads/sample) platform. Paired-end FASTQ files were quality assessed (Phred<20 removed) with FastQC and Illumina adapters were trimmed with TrimGalore (v0.4.3). Then, the reads were aligned to mm10 genome assembly using STAR (v2.7.8a) with MAPQ value for unique mappers set to 60. Binary alignment map (BAM) files were used to generate read counts per gene by FeatureCounts via Samtools (v1.11). Finally, limma-voom method (Liu et al. 2015) from the Bioconductor package-limma (v3.48.0) was adopted to quantify differential gene expression and normalised logarithmic CPM (counts per million) values were generated for downstream analysis. Genes without more than 0.5 CPM in at least 2 samples are insignificant and filtered out. 8582 of 24421 (35.14%) genes were filtered out for low expression. TMM was the method used to normalise library sizes.

To study the overlap between genes with ZFHX3 occupancy at TSS and those impacted by *Zfhx3* deletion, R-package GeneOverlap

(https://bioconductor.org/packages/release/bioc/html/GeneOverlap.html) was adopted where overlap between two gene lists was tested sing Fisher's exact test with P-value = 0 meaning the overlap is highly significant. The Fisher's exact test also gives an odds ratio which represents the strength of association. If an odds ratio is equal to or less than 1, there is no association between the two lists. If the odds ratio is much larger than 1, then the association is strong. The class also calculates the Jaccard index which measures the similarity between two lists. The Jaccard index varies between 0 and 1, with 0 meaning there is no similarity between the two and 1 meaning the two are identical.

Differential rhythm analysis

The read counts per gene (n = 24,421) for each analysed sample was obtained by FeatureCounts and arranged in sample matrix by using coreutils (v8.25). The matrix file was used as an input for *dryR* (R package for Differential Rhythmicity Analysis) available at https://github.com/neaf-lab/dryR . Normalized gene expression from "control" (Cre-neg) and "experimental" (+Tamoxifen) groups were categorized into 5 modules with a threshold of BICW (Bayes information criterion weight) > 0.4 based on their time-resolved expression pattern.

RT-qPCR

Total RNA was extracted from the SCN tissue as described above from three/ four biological replicates per genotype (-Tamoxifen vs. +Tamoxifen), where each biological replicate constitute one independent SCN tissue. 10 ng RNA per sample replicate was tested by one-step RT_qPCR reagent (PrecisionPLUS OneStep RT-qPCR Master Mix with ROX at a reduced level premixed with SYBRgreen, PrimerDesign) using an Applied Biosystems 7500 realtime PCR instrument (10 min at 55°C, 8 min at 95°C , 40× 10 sec at 95°C, and 60 sec at 60°C, followed by meltcurve analysis 65°C–95°C in 0.5°C increments at 5 sec per step) in

triplicates for *Zfhx3* expression in control and mutant groups . 7500 Applied Biosystems Software v2.3 (https://www.thermofisher.com/uk/en/home/technical-resources/software-downloads/applied-biosystems- 7500-real-time-pcr-system.html) was used to obtain the relative quantification values and examine melt curves. All RT-qPCR data was normalized to the *Rpl32* reference control gene and analysed using the standard curve method (Larionov et al., 2005). The primers used for *Zfhx3* (CTGTGCCAAGACATGCTCAAC [forward] and CAGCAGGGAGGAACATGCTA [reverse]) and *Rpl32* (AGGCACCAGTCAGACCGATA [forward] and TGTTGGGCATCAGGATCTGG [reverse]) were designed and screened for target specificity with the National Center for Biotechnology Information's (NCBI) Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/). The primer sets were validated before use (requirements: no predicted off-targets, primer efficiency 85-120%, no significant signal in nontemplate control, single peak in melt curve, and single band at the predicted size when separated via agarose gel electrophoresis). ZFHX3 gene expression differences from qPCR were plotted and analyzed with GraphPad Prism v9.5.1 for Windows, GraphPad Software (www.graphpad.com). Unpaired-two tailed t-test between control and mutant group across the tested time points was set at alpha= 0.05.

Gene annotation

The gene list derived from Bioconductor-based package ChIPseeker v1.28.3 (Yu et al., 2015) was fed into the Database for Annotation, Visualization, and Integrated Discovery (DAVID) tool (Dennis et al., 2003). The functional annotation chart based on KEGG pathway, Gene Ontology (GO): biological processes (BP) was plotted with the help of the ggplot2 package in Rv4.0.5 (https://ggplot2.tidyverse.org/).

HOMER motif analysis

The findMotifsGenome.pl function within the HOMER v4.11 package (Heinz et al., 2010) was used to identify enriched motifs and their corresponding TFs with options size 200 –len 8, 10, 12 –mask –preparse –dumpfasta, with default background regions.

TimeTeller modelling of the SCN clock

TimeTeller (TT) analysis was conducted using all 35 available SCN time-course RNA-Seq samples from control (Cre-neg and -Tamoxifen) mice to train the model (Vlachou et al., 2024). Data from +Tamoxifen SCN were then projected into the model as test samples. Based on the available training data, we selected the 10 most rhythmic and consistent genes to build the TimeTeller model, in this case: *Bmal1, Per2, Cry1, Tef, Hlf, Dbp, Nr1d2, Npas2, Ciart*, and *Bhlhe41*. As a first step to cross-validate the model, we used a "leave-one-out" approach. Iteratively, 5-6 randomly chosen samples from all Cre-neg and -Tamoxifen SCN were left out from the training data to build the TT model, and the left out samples were treated as test samples to determine the TT prediction for all Cre-neg and -Tamoxifen SCN. In addition, we used published SCN RNA-Seq data to independently validate the full model with all 35 control SCN from this study (Deota et al., 2023). TT analysis predicts phase and clock functionality (theta) based on the averaged likelihood for the projected test sample. For the SCN model, we used time-course normalization with log threshold selected at -6. Further, to compare TT phase predictions between genotypes, we used estimation statistics (Ho et al., 2019) and p-values were generated using a permutation testing with 5000 bootstrap samples.

FUNDING

PMN and GB were supported by Medical Research Council (MC_U142684173). AB was supported by Medical Research Council (MC_U142684173) and Biotechnology and Biological Sciences Research Council (BB/Z514792/1). MHH was supported by Medical

Research Council (MC_U105170643). RD was supported by Cancer Research UK and EPSRC (C53561/A19933), and VV was funded by the UK Medical Research Council Doctoral Training Partnership (MR/N014294/1).

AUTHOR CONTRIBUTION

AB, GB and PMN conceived the study. AB and GB were involved in sample collection and processing. AB and PMN carried out the investigation with AB conducting the formal data analysis. Data visualization was performed by AB, GB, MHH and PMN. RD and VV carried the data modelling and TT prediction. Supervision of the research was carried out by MHH and PMN. The original draft was written by AB. Manuscript was reviewed, edited and finalised by AB, GB, RD, VV, MHH and PMN.

COMPETING INTERESTS

Authors declare that they have no competing interests.

ACKNOWLEDGEMENTS

We thank the staff of the Mary Lyon Centre and core services at MRC Harwell Institute for assistance with mouse studies. We thank Dr Benjamin Weger, University of Queensland for his valuable suggestions and timely inputs during data analysis using dryR. We also thank Michelle Simon and Richard Reeves for ChIP-seq data processing and accessibility on UCSC genome browser.

DATA AVAILIBILITY

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE261795) under accession number GSE261795 .

REFERENCES

Abe, Y. O., Yoshitane, H., Kim, D. W., Kawakami, S., Koebis, M., Nakao, K., . . . Fukada, Y. (2022). Rhythmic transcription of Bmal1 stabilizes the circadian timekeeping system in mammals. *Nat Commun*, *13*(1), 4652. https://doi.org/10.1038/s41467-022-32326-9

Araki, R., Takahashi, H., Fukumura, R., Sun, F., Umeda, N., Sujino, M., . . . Abe, M. (2004). Restricted expression and photic induction of a novel mouse regulatory factor X4 transcript in the suprachiasmatic nucleus. *J Biol Chem*, *279*(11), 10237-10242. https://doi.org/10.1074/jbc.M312761200

Arima, H., House, S. B., Gainer, H., & Aguilera, G. (2002). Neuronal activity is required for the circadian rhythm of vasopressin gene transcription in the suprachiasmatic nucleus in vitro. *Endocrinology*, *143*(11), 4165-4171. https://doi.org/10.1210/en.2002-220393

Bafna, A., Banks, G., Hastings, M. H., & Nolan, P. M. (2023). Dynamic modulation of genomic enhancer elements in the suprachiasmatic nucleus, the site of the mammalian circadian clock. *Genome Res*, *33*(5), 673-688. https://doi.org/10.1101/gr.277581.122

Bafna, A., Lau, P., Banks, G., & Nolan, P. M. (2023). Harvesting mouse suprachiasmatic nucleus by vibrating microtome for diurnal transcriptome analysis. *STAR Protoc*, *4*(4), 102618. https://doi.org/10.1016/j.xpro.2023.102618

Banks, G. T., & Nolan, P. M. (2011). Assessment of Circadian and Light-Entrainable Parameters in Mice Using Wheel-Running Activity. *Curr Protoc Mouse Biol*, *1*(3), 369-381. https://doi.org/10.1002/9780470942390.mo110123

Beacon, T. H., Delcuve, G. P., López, C., Nardocci, G., Kovalchuk, I., van Wijnen, A. J., & Davie, J. R. (2021). The dynamic broad epigenetic (H3K4me3, H3K27ac) domain as a mark of essential genes. *Clin Epigenetics*, *13*(1), 138. https://doi.org/10.1186/s13148-021-01126-1

Bedont, J. L., LeGates, T. A., Buhr, E., Bathini, A., Ling, J. P., Bell, B., . . . Blackshaw, S. (2017). An LHX1-Regulated Transcriptional Network Controls Sleep/Wake Coupling and Thermal Resistance of the Central Circadian Clockworks. *Curr Biol*, *27*(1), 128-136. https://doi.org/10.1016/j.cub.2016.11.008

Brown, T. M., Hughes, A. T., & Piggins, H. D. (2005). Gastrin-releasing peptide promotes suprachiasmatic nuclei cellular rhythmicity in the absence of vasoactive intestinal polypeptide-VPAC2 receptor signaling. *J Neurosci*, *25*(48), 11155-11164. https://doi.org/10.1523/JNEUROSCI.3821-05.2005

Calo, E., & Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? *Mol Cell*, *49*(5), 825-837. https://doi.org/10.1016/j.molcel.2013.01.038

Del Rocío Pérez Baca, M., Jacobs, E. Z., Vantomme, L., Leblanc, P., Bogaert, E., Dheedene, A., . . . consortium, Z. (2023). A novel neurodevelopmental syndrome caused by loss-of-function of the Zinc Finger Homeobox 3 (ZFHX3) gene. *medRxiv*. https://doi.org/10.1101/2023.05.22.23289895

Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*, *4*(5), P3.

Deota, S., Lin, T., Chaix, A., Williams, A., Le, H., Calligaro, H., . . . Panda, S. (2023). Diurnal transcriptome landscape of a multi-tissue response to time-restricted feeding in mammals. *Cell Metab*, *35*(1), 150-165.e154. https://doi.org/10.1016/j.cmet.2022.12.006

Figueroa, K. P., Gross, C., Buena-Atienza, E., Paul, S., Gandelman, M., Kakar, N., . . . Pulst, S. M. (2024). A GGC-repeat expansion in ZFHX3 encoding polyglycine causes spinocerebellar ataxia type 4 and impairs autophagy. *Nat Genet*. https://doi.org/10.1038/s41588-024-01719-5

765 Fu, C., An, N., Liu, J., A, J., Zhang, B., Liu, M., . . . Dong, J. T. (2020). The transcription factor ZFHX3 is
766 crucial for the angiogenic function of hypoxia-inducible factor 1α in liver cancer cells. *J Biol*
767 *Chem*, *295*(20), 7060-7074. https://doi.org/10.1074/jbc.RA119.012131
768 Hastings, M. H., Maywood, E. S., & Brancaccio, M. (2018). Generation of circadian rhythms in the
769 suprachiasmatic nucleus. *Nat Rev Neurosci*, *19*(8), 453-469. https://doi.org/10.1038/s41583-
770 018-0026-z
771 Hastings, M. H., Maywood, E. S., & Brancaccio, M. (2019). The Mammalian Circadian Timing System
772 and the Suprachiasmatic Nucleus as Its Pacemaker. *Biology (Basel)*, *8*(1).
773 https://doi.org/10.3390/biology8010013
774 Hatori, M., Gill, S., Mure, L. S., Goulding, M., O'Leary, D. D., & Panda, S. (2014). Lhx1 maintains
775 synchrony among circadian oscillator neurons of the SCN. *Elife*, *3*, e03357.
776 https://doi.org/10.7554/eLife.03357
777 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . Glass, C. K. (2010). Simple
778 combinations of lineage-determining transcription factors prime cis-regulatory elements
779 required for macrophage and B cell identities. *Mol Cell*, *38*(4), 576-589.
780 https://doi.org/10.1016/j.molcel.2010.05.004
781 Herzog, E. D., Hermanstyne, T., Smyllie, N. J., & Hastings, M. H. (2017). Regulating the
782 Suprachiasmatic Nucleus (SCN) Circadian Clockwork: Interplay between Cell-Autonomous
783 and Circuit-Level Mechanisms. *Cold Spring Harb Perspect Biol*, *9*(1).
784 https://doi.org/10.1101/cshperspect.a027706
785 Ho, J., Tumkaya, T., Aryal, S., Choi, H., & Claridge-Chang, A. (2019). Moving beyond P values: data
786 analysis with estimation graphics. *Nat Methods*, *16*(7), 565-566.
787 https://doi.org/10.1038/s41592-019-0470-3
788 Hogenesch, J. B., Panda, S., Kay, S., & Takahashi, J. S. (2003). Circadian transcriptional output in the
789 SCN and liver of the mouse. *Novartis Found Symp*, *253*, 171-180; discussion 152-175, 102-
790 179, 180-173 passim.
791 Holwerda, S. J., & de Laat, W. (2013). CTCF: the protein, the binding partners, the binding sites and
792 their chromatin loops. *Philos Trans R Soc Lond B Biol Sci*, *368*(1620), 20120369.
793 https://doi.org/10.1098/rstb.2012.0369
794 Hughes, M. E., Abruzzi, K. C., Allada, R., Anafi, R., Arpat, A. B., Asher, G., . . . Hogenesch, J. B. (2017).
795 Guidelines for Genome-Scale Analysis of Biological Rhythms. *J Biol Rhythms*, *32*(5), 380-393.
796 https://doi.org/10.1177/0748730417728663
797 Hughes, S., Edwards, J. K., Wilcox, A. G., Pothecary, C. A., Barnard, A. R., Joynson, R., . . . Nolan, P. M.
798 (2021). Zfhx3 modulates retinal sensitivity and circadian responses to light. *FASEB J*, *35*(9),
799 e21802. https://doi.org/10.1096/fj.202100563R
800 Jagannath, A., Butler, R., Godinho, S. I. H., Couch, Y., Brown, L. A., Vasudevan, S. R., . . . Peirson, S. N.
801 (2013). The CRTC1-SIK1 pathway regulates entrainment of the circadian clock. *Cell*, *154*(5),
802 1100-1111. https://doi.org/10.1016/j.cell.2013.08.004
803 Jameson, H. S., Hanley, A., Hill, M. C., Xiao, L., Ye, J., Bapat, A., . . . Ellinor, P. T. (2023). Loss of the
804 Atrial Fibrillation-Related Gene, . *Circ Res*, *133*(4), 313-329.
805 https://doi.org/10.1161/CIRCRESAHA.123.323029
806 Kim, S., Yu, N. K., & Kaang, B. K. (2015). CTCF as a multifunctional protein in genome regulation and
807 gene expression. *Exp Mol Med*, *47*(6), e166. https://doi.org/10.1038/emm.2015.33
808 Kuhlman, S. J., Silver, R., Le Sauter, J., Bult-Ito, A., & McMahon, D. G. (2003). Phase resetting light
809 pulses induce Per1 and persistent spike activity in a subpopulation of biological clock
810 neurons. *J Neurosci*, *23*(4), 1441-1450. https://doi.org/10.1523/JNEUROSCI.23-04-
811 01441.2003
812 Larionov, A., Krause, A., & Miller, W. (2005). A standard curve based method for relative real time
813 PCR data processing. *BMC Bioinformatics*, *6*, 62. https://doi.org/10.1186/1471-2105-6-62
814 Liu, A. C., Tran, H. G., Zhang, E. E., Priest, A. A., Welsh, D. K., & Kay, S. A. (2008). Redundant function
815 of REV-ERBalpha and beta and non-essential role for Bmal1 cycling in transcriptional

816 regulation of intracellular circadian rhythms. *PLoS Genet*, *4*(2), e1000023.
817 https://doi.org/10.1371/journal.pgen.1000023
818 Lopez-Molina, L., Conquet, F., Dubois-Dauphin, M., & Schibler, U. (1997). The DBP gene is expressed
819 according to a circadian rhythm in the suprachiasmatic nucleus and influences circadian
820 behavior. *EMBO J*, *16*(22), 6762-6771. https://doi.org/10.1093/emboj/16.22.6762
821 Mach, P., Kos, P. I., Zhan, Y., Cramard, J., Gaudin, S., Tünnermann, J., . . . Giorgetti, L. (2022). Cohesin
822 and CTCF control the dynamics of chromosome folding. *Nat Genet*, *54*(12), 1907-1918.
823 https://doi.org/10.1038/s41588-022-01232-7
824 Matsu-Ura, T., Moore, S. R., & Hong, C. I. (2018). WNT Takes Two to Tango: Molecular Links between
825 the Circadian Clock and the Cell Cycle in Adult Stem Cells. *J Biol Rhythms*, *33*(1), 5-14.
826 https://doi.org/10.1177/0748730417745913
827 Maywood, E. S., Chesham, J. E., Winsky-Sommerer, R., & Hastings, M. H. (2021). Restoring the
828 Molecular Clockwork within the Suprachiasmatic Hypothalamus of an Otherwise Clockless
829 Mouse Enables Circadian Phasing and Stabilization of Sleep-Wake Cycles and Reverses
830 Memory Deficits. *J Neurosci*, *41*(41), 8562-8576. https://doi.org/10.1523/JNEUROSCI.3141-
831 20.2021
832 Mirsky, H. P., Liu, A. C., Welsh, D. K., Kay, S. A., & Doyle, F. J. (2009). A model of the cell-autonomous
833 mammalian circadian clock. *Proc Natl Acad Sci U S A*, *106*(27), 11107-11112.
834 https://doi.org/10.1073/pnas.0904837106
835 Morris, E. L., Patton, A. P., Chesham, J. E., Crisp, A., Adamson, A., & Hastings, M. H. (2021). Single-cell
836 transcriptomics of suprachiasmatic nuclei reveal a Prokineticin-driven circadian network.
837 *EMBO J*, *40*(20), e108614. https://doi.org/10.15252/embj.2021108614
838 Nolan, P. M., Banks, G., Bourbia, N., Wilcox, A. G., Bentley, L., Moir, L., . . . Dumbell, R. (2023). A
839 missense mutation in zinc finger homeobox-3 (ZFHX3) impedes growth and alters
840 metabolism and hypothalamic gene expression in mice. *FASEB J*, *37*(10), e23189.
841 https://doi.org/10.1096/fj.202201829R
842 Padlom, A., Ono, D., Hamashima, R., Furukawa, Y., Yoshimura, T., & Nishiwaki-Ohkawa, T. (2022).
843 Level of constitutively expressed BMAL1 affects the robustness of circadian oscillations. *Sci*
844 *Rep*, *12*(1), 19519. https://doi.org/10.1038/s41598-022-24188-4
845 Parsons, M. J., Brancaccio, M., Sethi, S., Maywood, E. S., Satija, R., Edwards, J. K., . . . Nolan, P. M.
846 (2015). The Regulatory Factor ZFHX3 Modifies Circadian Function in SCN via an AT Motif-
847 Driven Axis. *Cell*, *162*(3), 607-621. https://doi.org/10.1016/j.cell.2015.06.060
848 Pugacheva, E. M., Kubo, N., Loukinov, D., Tajmul, M., Kang, S., Kovalchuk, A. L., . . . Lobanenkov, V. V.
849 (2020). CTCF mediates chromatin looping via N-terminal domain-dependent cohesin
850 retention. *Proc Natl Acad Sci U S A*, *117*(4), 2020-2031.
851 https://doi.org/10.1073/pnas.1911708117
852 Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic
853 features. *Bioinformatics*, *26*(6), 841-842. https://doi.org/10.1093/bioinformatics/btq033
854 Schibler, U., Gotic, I., Saini, C., Gos, P., Curie, T., Emmenegger, Y., . . . Franken, P. (2015). Clock-Talk:
855 Interactions between Central and Peripheral Circadian Oscillators in Mammals. *Cold Spring*
856 *Harb Symp Quant Biol*, *80*, 223-232. https://doi.org/10.1101/sqb.2015.80.027490
857 Sun, X., Fu, X., Li, J., Xing, C., Martin, D. W., Zhang, H. H., . . . Dong, J. T. (2012). Heterozygous
858 deletion of Atbf1 by the Cre-loxP system in mice causes preweaning mortality. *Genesis*,
859 *50*(11), 819-827. https://doi.org/10.1002/dvg.22041
860 Takahashi, J. S., Hong, H. K., Ko, C. H., & McDearmon, E. L. (2008). The genetics of mammalian
861 circadian order and disorder: implications for physiology and disease. *Nat Rev Genet*, *9*(10),
862 764-775. https://doi.org/10.1038/nrg2430
863 VanDunk, C., Hunter, L. A., & Gray, P. A. (2011). Development, maturation, and necessity of
864 transcription factors in the mouse suprachiasmatic nucleus. *J Neurosci*, *31*(17), 6457-6467.
865 https://doi.org/10.1523/JNEUROSCI.5385-10.2011

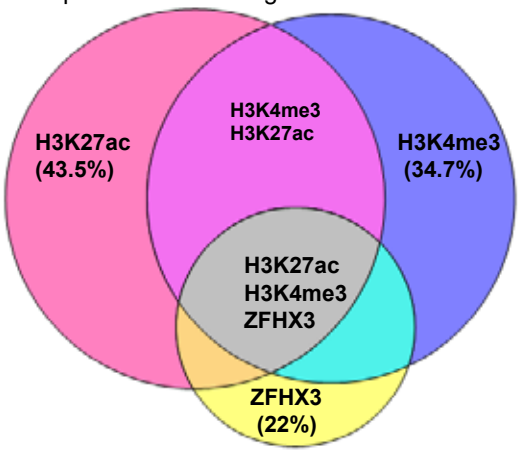
866 Vlachou, D., Veretennikova, M., Usselmann, L., Vasilyev, V., Ott, S., Bjarnason, G. A., . . . Rand, D. A.
867 (2024). TimeTeller: A tool to probe the circadian clock as a multigene dynamical system.
868 *PLoS Comput Biol*, *20*(2), e1011779. https://doi.org/10.1371/journal.pcbi.1011779
869 Wallenius, J., Kafantari, E., Jhaveri, E., Gorcenco, S., Ameur, A., Karremo, C., . . . Puschmann, A.
870 (2024). Exonic trinucleotide repeat expansions in ZFHX3 cause spinocerebellar ataxia type 4:
871 A poly-glycine disease. *Am J Hum Genet*, *111*(1), 82-95.
872 https://doi.org/10.1016/j.ajhg.2023.11.008
873 Weger, B. D., Gobet, C., David, F. P. A., Atger, F., Martin, E., Phillips, N. E., . . . Gachon, F. (2021).
874 Systematic analysis of differential rhythmic liver gene expression mediated by the circadian
875 clock and feeding rhythms. *Proc Natl Acad Sci U S A*, *118*(3).
876 https://doi.org/10.1073/pnas.2015803118
877 Wen, S., Ma, D., Zhao, M., Xie, L., Wu, Q., Gou, L., . . . Yan, J. (2020). Spatiotemporal single-cell
878 analysis of gene expression in the mouse suprachiasmatic nucleus. *Nat Neurosci*, *23*(3), 456-
879 467. https://doi.org/10.1038/s41593-020-0586-x
880 Wilcox, A. G., Bains, R. S., Williams, D., Joynson, E., Vizor, L., Oliver, P. L., . . . Nolan, P. M. (2021).
881 Zfhx3-mediated genetic ablation of the SCN abolishes light entrainable circadian activity
882 while sparing food anticipatory activity. *iScience*, *24*(10), 103142.
883 https://doi.org/10.1016/j.isci.2021.103142
884 Wilcox, A. G., Vizor, L., Parsons, M. J., Banks, G., & Nolan, P. M. (2017). Inducible Knockout of Mouse
885 Zfhx3 Emphasizes Its Key Role in Setting the Pace and Amplitude of the Adult Circadian
886 Clock. *J Biol Rhythms*, *32*(5), 433-443. https://doi.org/10.1177/0748730417722631
887 Won, K. J., Chepelev, I., Ren, B., & Wang, W. (2008). Prediction of regulatory elements in mammalian
888 genomes using chromatin signatures. *BMC Bioinformatics*, *9*, 547.
889 https://doi.org/10.1186/1471-2105-9-547
890 Yu, G., Wang, L. G., & He, Q. Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak
891 annotation, comparison and visualization. *Bioinformatics*, *31*(14), 2382-2383.
892 https://doi.org/10.1093/bioinformatics/btv145
893 Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., . . . Liu, X. S. (2008).
894 Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, *9*(9), R137.
895 https://doi.org/10.1186/gb-2008-9-9-r137

896

Figure 1

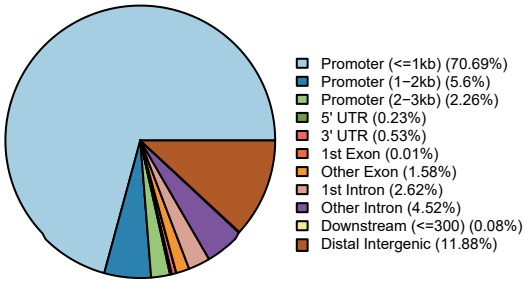
A

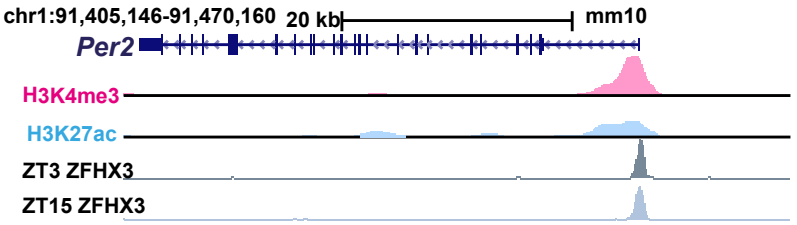
Overlap of ZFHX3 binding with histone modifications



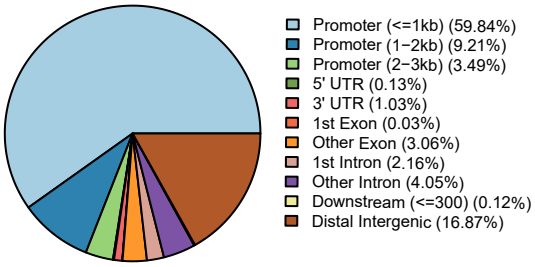
Category	Number of sites
Only H3K27ac	47552
Only H3K4me3	35673
Only ZFHX3	9512
H3K4me3 + ZFHX3	8992
H3K27ac + ZFHX3	3570
H3K27ac + H3K4me3	36819
H3K27ac + H3K4me3 + ZFHX3	21162

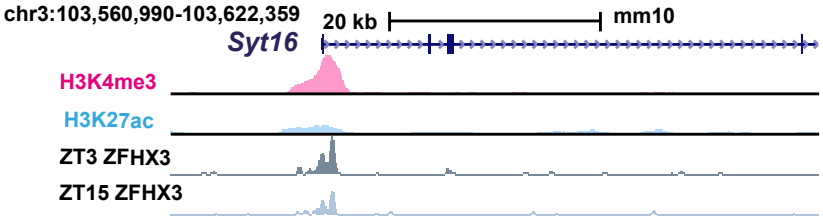
B H3K27ac+H3K4me3+ZFHX3



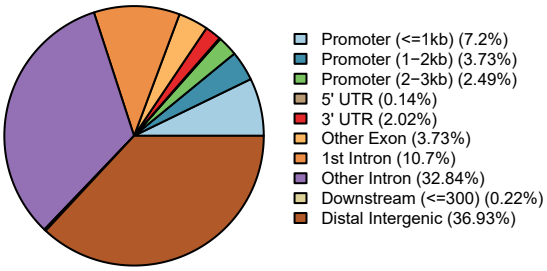


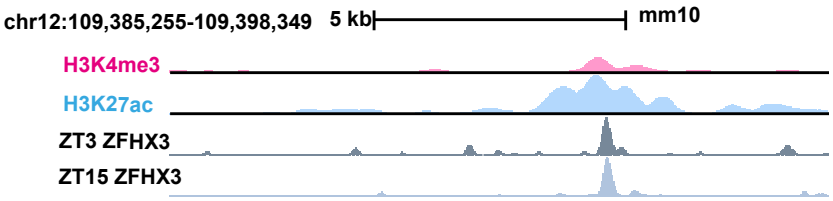
C H3K4me3+ZFHX3



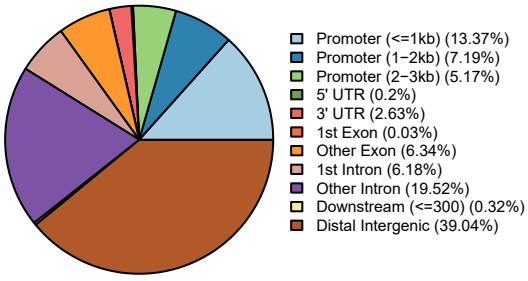


D H3K27ac+ZFHX3





E ZFHX3 ONLY



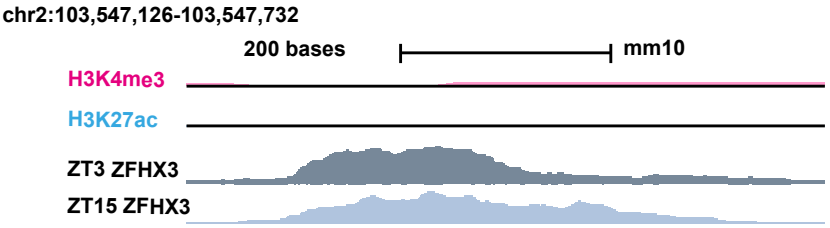


Figure 1. Genomic occupancy of ZFHX3 in the SCN. (A) left- Venn-diagram illustrating overlap of ZFHX3 peaks with histone marks, H3K4me3 and H3K27ac, right- table showing number of sites co-occupied by ZFHX3 and histone marks. (B,C,D and E) left- Genomic feature distribution of ZFHX3 peaks with histone marks (ChIPSeeker), right- UCSC genome browser tracks showing histone modifications H3K4me3(pink), H3K27ac (blue) and ZFHX3 (ZT3 = dark grey and ZT15 = light grey) normalized ChIP-seq read coverage at representative examples for each category. The chromosome location and scale (mm10 genome) indicated at the top.

Figure 2

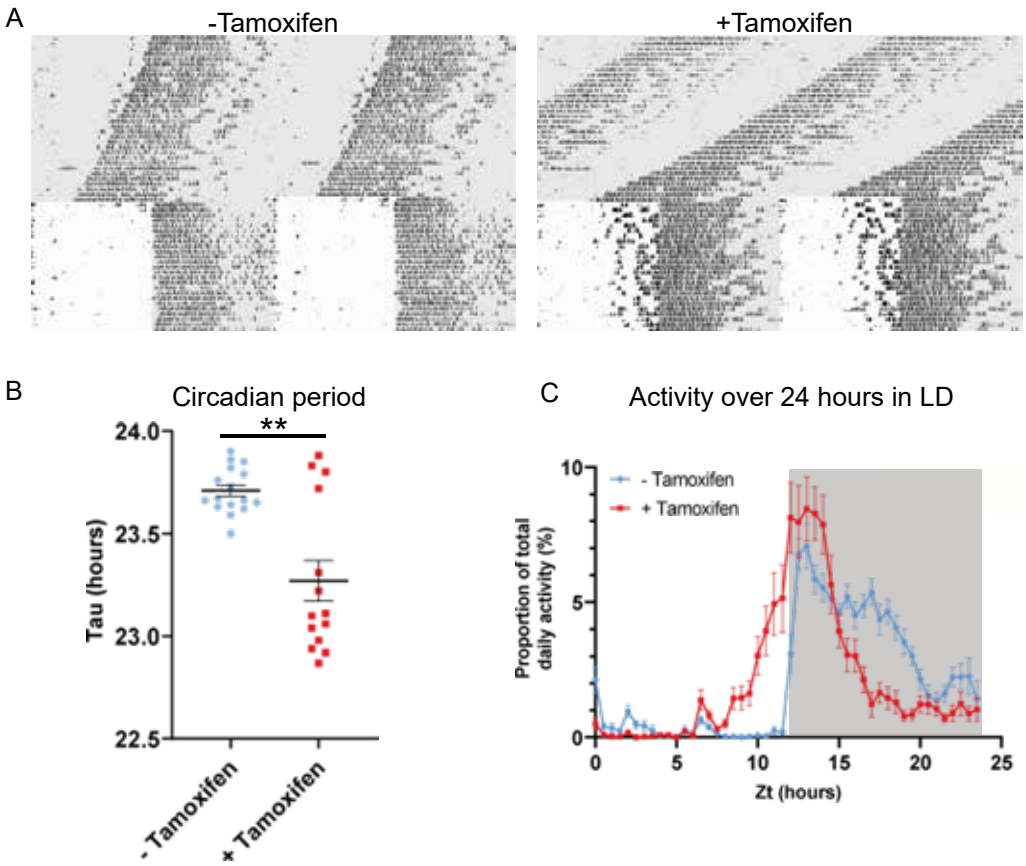


Figure 2. Circadian wheel running analysis of control and mutant animals. A. Double plotted actograms showing wheel running activity of -Tamoxifen and +Tamoxifen animals through 32 days of constant darkness and 24 days under a 12h light/dark cycle. Regions in grey denote periods in darkness. B. +Tamoxifen animals show a reduced circadian period C: Graph showing the distribution of activity in -Tamoxifen and +Tamoxifen animals over 24 hours in a 12h light/dark cycle. Data is taken from the final seven days the animals were in the light/dark cycle shown in A. Grey denotes the dark period of the cycle.

Figure 3

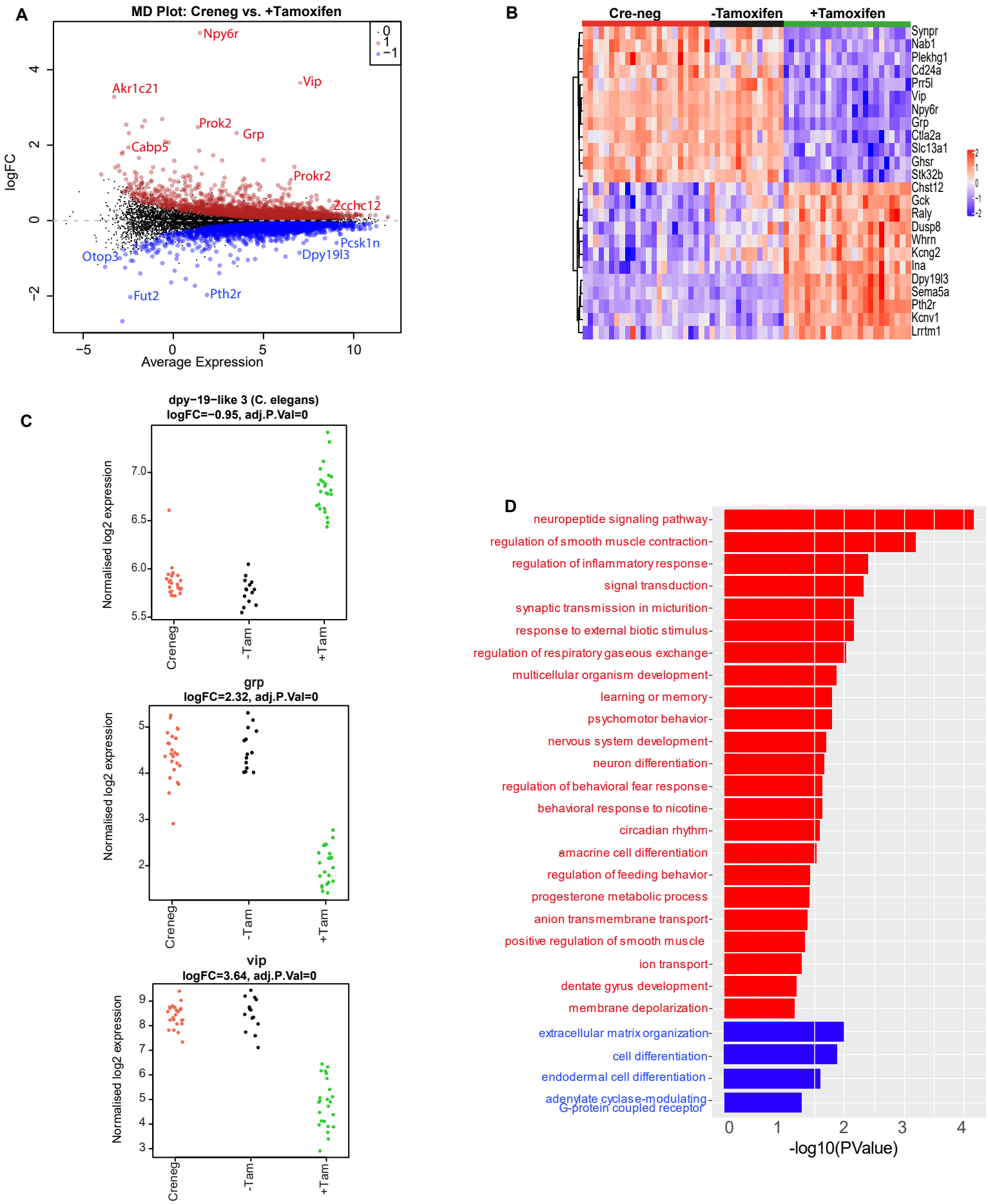


Figure 3. Effect of ZFHX3-KO on SCN transcriptome. (A) Mean-difference plot (MD-plot) showing downregulated (red) and upregulated (blue) gene expression after the loss of ZFHX3. (B) Heatmap showing normalized expression of top 25 (by adj pvalue) differential genes for the compared groups. (C) Stripcharts of differentially expressed genes (*Dpy19l13, Grp, Vip)*. (D) Functional annotation of downregulated (red) and upregulated (blue) genes after ZFHX3-KO using the GO::BP (biological processes) terms by DAVID.

Figure 4

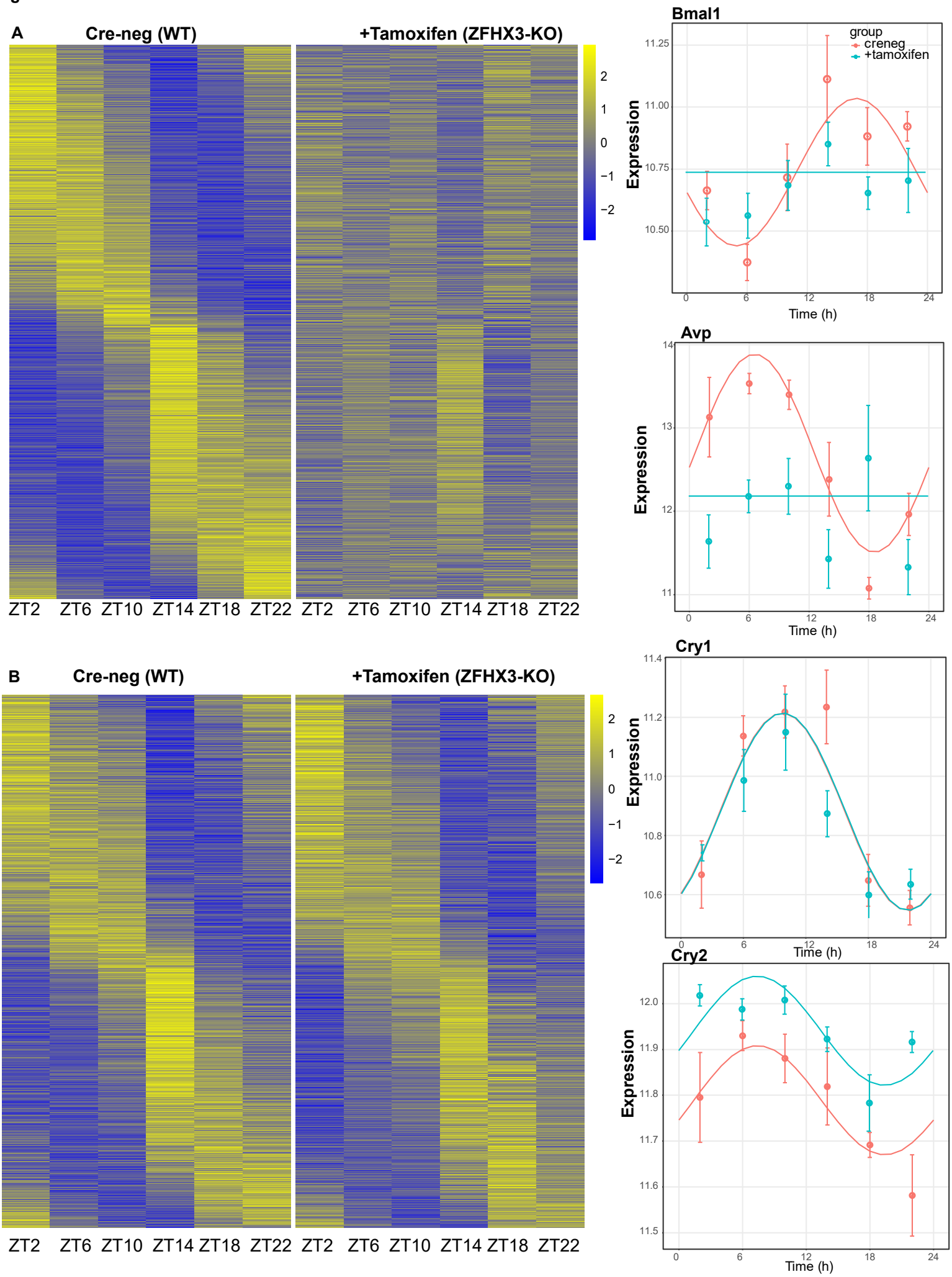


Figure 4. Effect of ZFHX3-KO on rhythmic gene expression . (A) Left- Heatmap showing loss of rhythmic gene expression after *Zfhx3-KO* (module 2) as computed by dryR statistical framework , Right- Illustrative examples of daily abundance of module 2 genes, *Bmal1 and Avp,* in Cre-neg and +Tamoxifen conditions. (B) Left- Heatmap showing no effect on rhythmic gene expression after *Zfhx3-KO* (module 4) , Right- Illustrative examples of module 4 genes, *Cry1 and Cry2,* in Cre-neg and +Tamoxifen conditions.

Figure 5

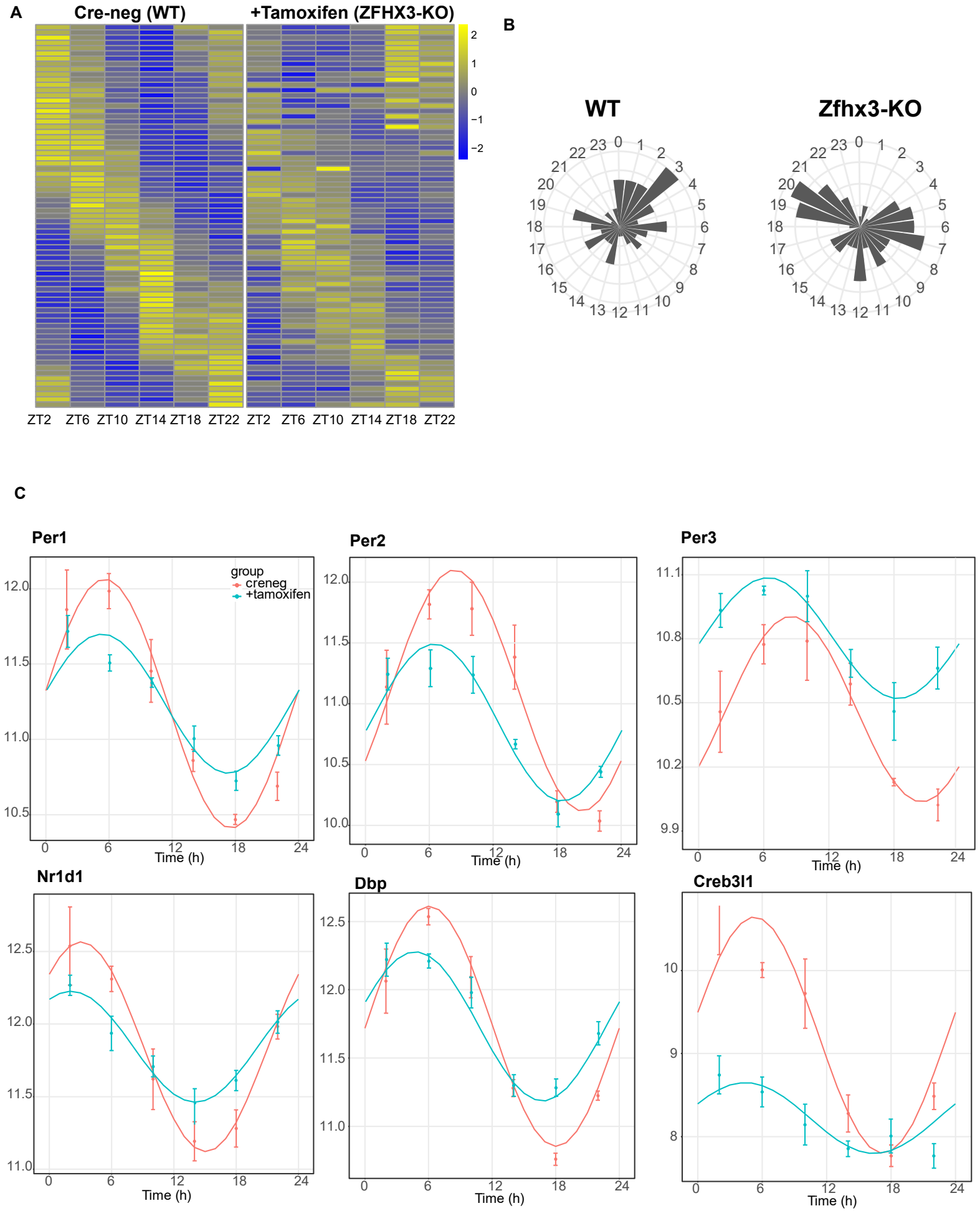


Figure 5. Effect of ZFHX3-KO on clock gene expression clustered in module 5 . (A) Heatmap showing change in rhythmic gene expression after *Zfhx3-KO* (module 5) as computed by dryR statistical framework , (B) 24-h radial plots showing effect on phase (peak expression) of rhythmic genes in Cre-neg (WT) vs. +Tamoxifen (ZFHX3-KO) conditions (C)Illustrative examples of daily abundance of module 5 genes, showing change in phase and amplitude in ZFHX3-KO condition (blue) as compared to control (pink). The y-axis values represents normalised expression in the SCN.

Figure 6

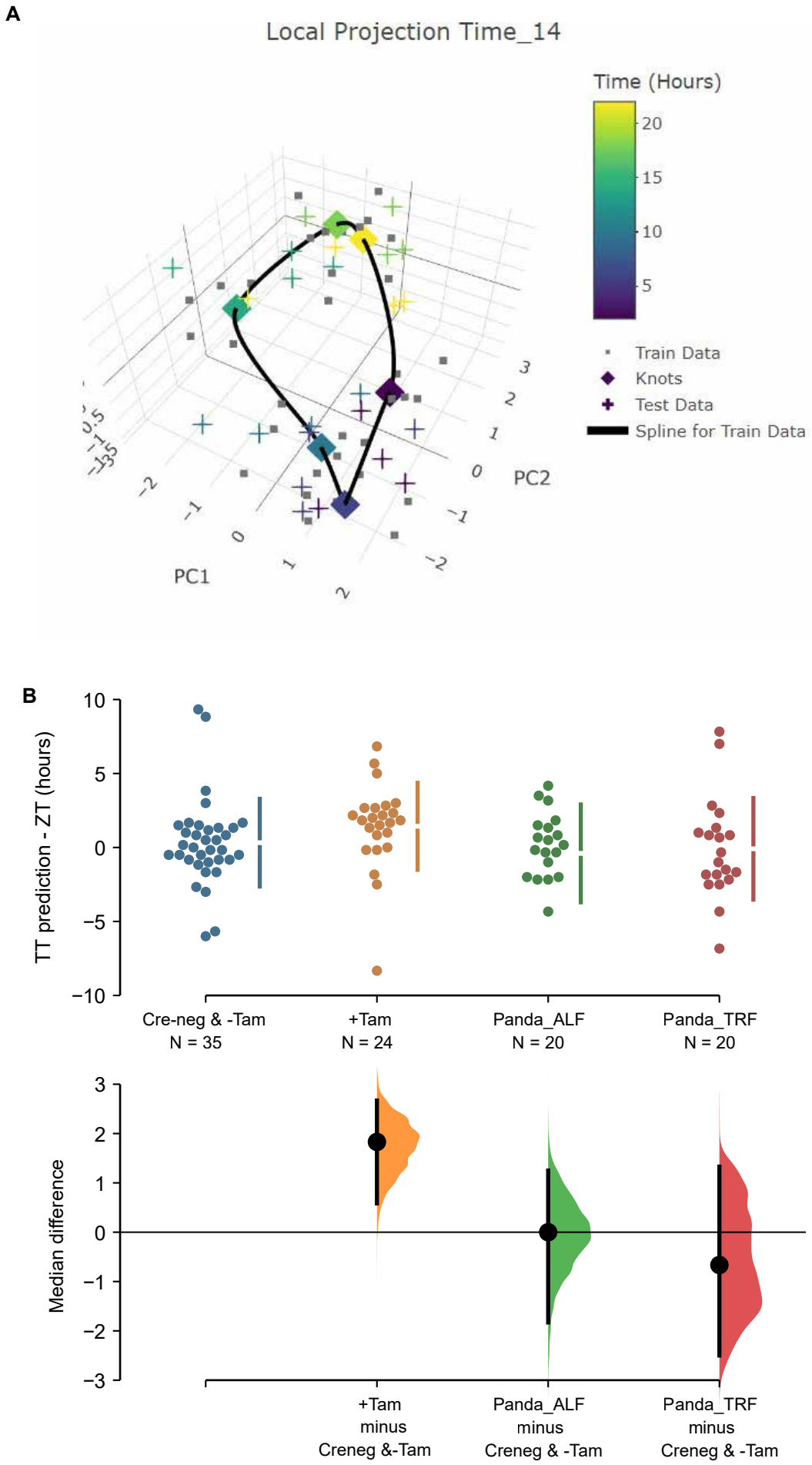


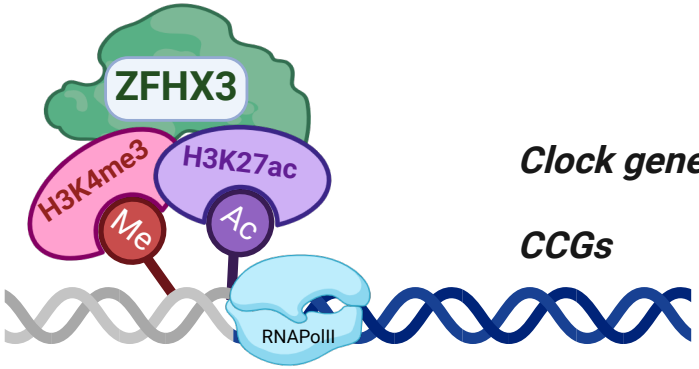
Figure 6. TimeTeller model SCN circadian clock in control and ZFHX3-KO mice. (A) TimeTeller (TT) model build with the data of all 35 available SCN from Creneg and -Tam mice. Only the knots of the model at the collected ZTs, and predictions for test data of +Tam SCN according to when they were collected are shown. (B) Quantitation of SCN phases as difference between ZT to TT prediction for all genotypes of this study and previously published wild-type SCNs from ad libitum or night fed mice (Panda_ALF and Panda_TRF). Only the +Tam group is different from Creng & -Tam control training data (p=0.0028).

Lost rhythm

Avp Syt17 Wdr1 Bmal1 Prok2

Intracellular transport
Protein import to nucleus
mTOR signalling pathway
Spinocerebellar ataxia
Rhythmic process

49% ZFHX3 sites localized with modified histones



Clock genes

CCGs

Gained rhythm

Hdac3 Sox4 Aldh1 Notch1 Galnt4

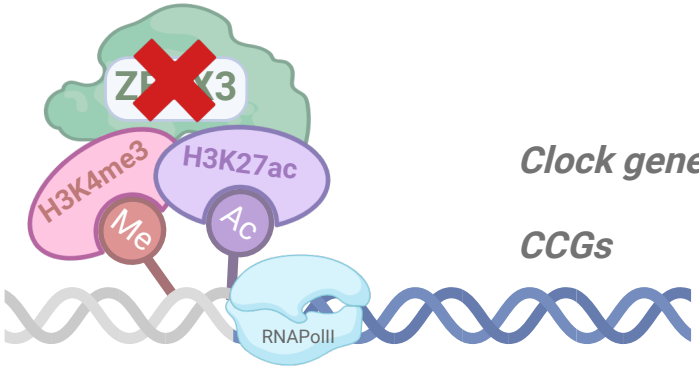
RNA processing
Notch Signalling
Angiogenesis
Metabolic pathways
Cell proliferation

Retained rhythm

Dusp16 Hlf Cry1 Dlk1 Rfx5

Cell cycle
DNA repair
ER stress
MAPK Signalling
Protein phosphorylation

36% SCN genes affected by loss of ZFHX3



Clock genes

CCGs

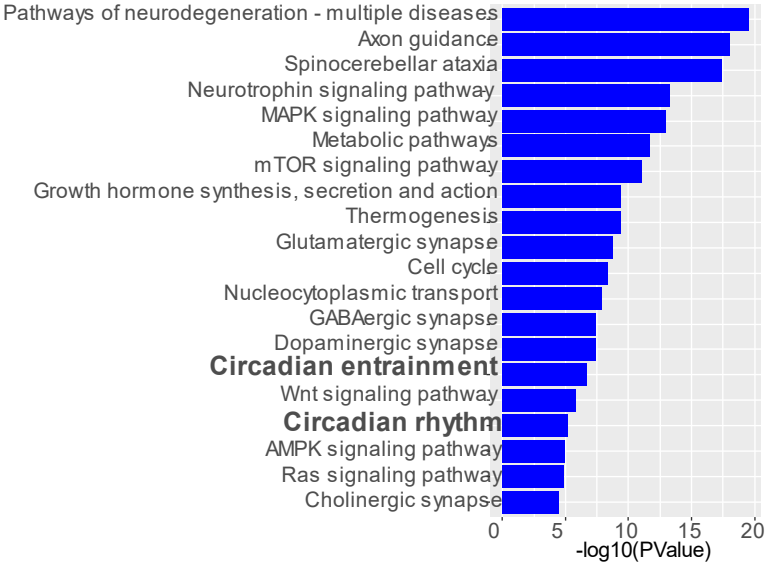
Altered rhythm

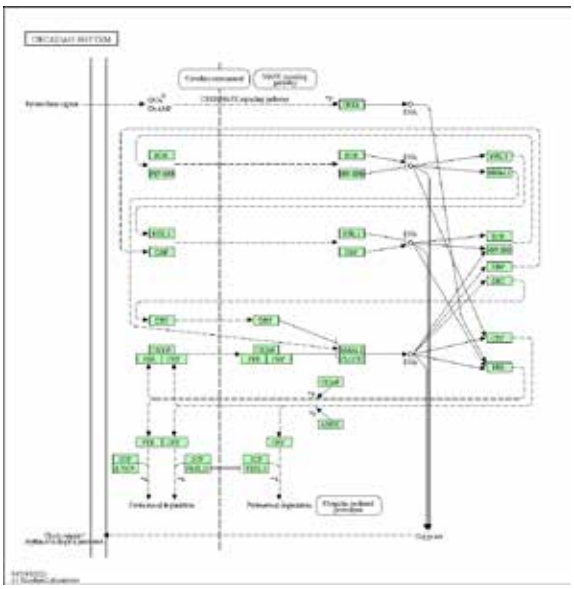
Per2 Dbp Creb3l1 Hif3a Rgs16

Circadian rhythm
Ppar signalling
Response to hypoxia
Butanoate metabolism
Circadian gene regulation

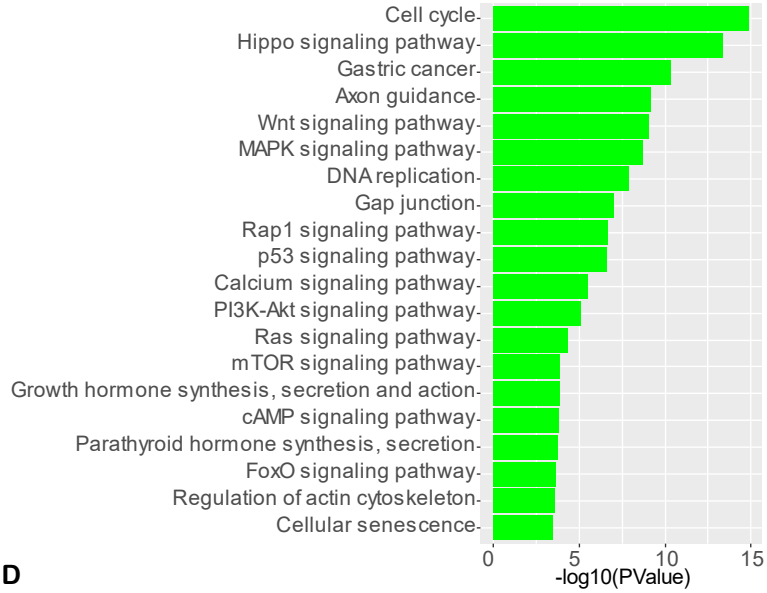
Supplementary Figure 1

A

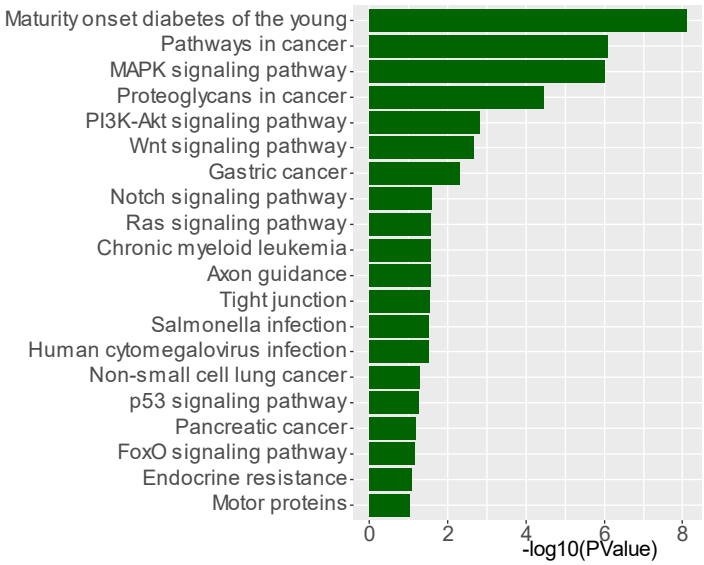




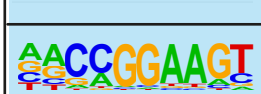
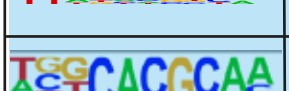
B

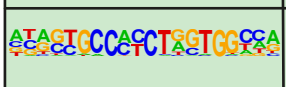


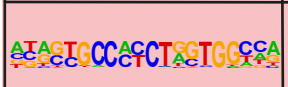
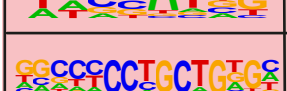
C



D

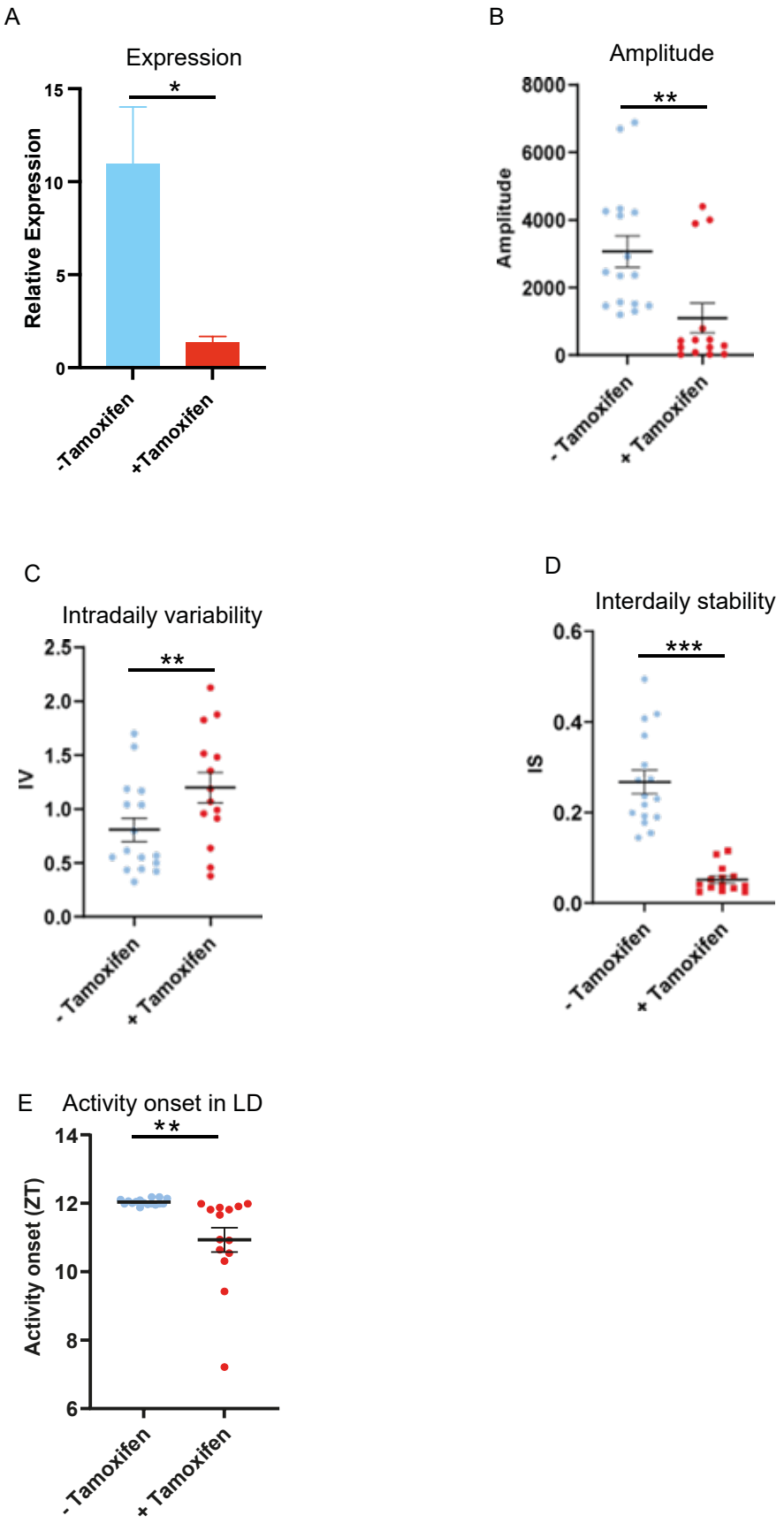
Motif	Name	*p*
	RFX	1e-50
	X-box	1e-34
	DLX	1e-28
	GABPA	1e-27
	ARNT (bHLH)	1e-18

Motif	Name	*p*
	RFX	1e-161
	CTCF	1e-100
	DLX	1e- 51
	LHX	1e- 44
	ATHB33	1e- 27

Motif	Name	*p*
	CTCF	1e-100
	KLF	1e- 48
	LHX	1e- 17
	DLX	1e- 16
	ZIC	1e- 12

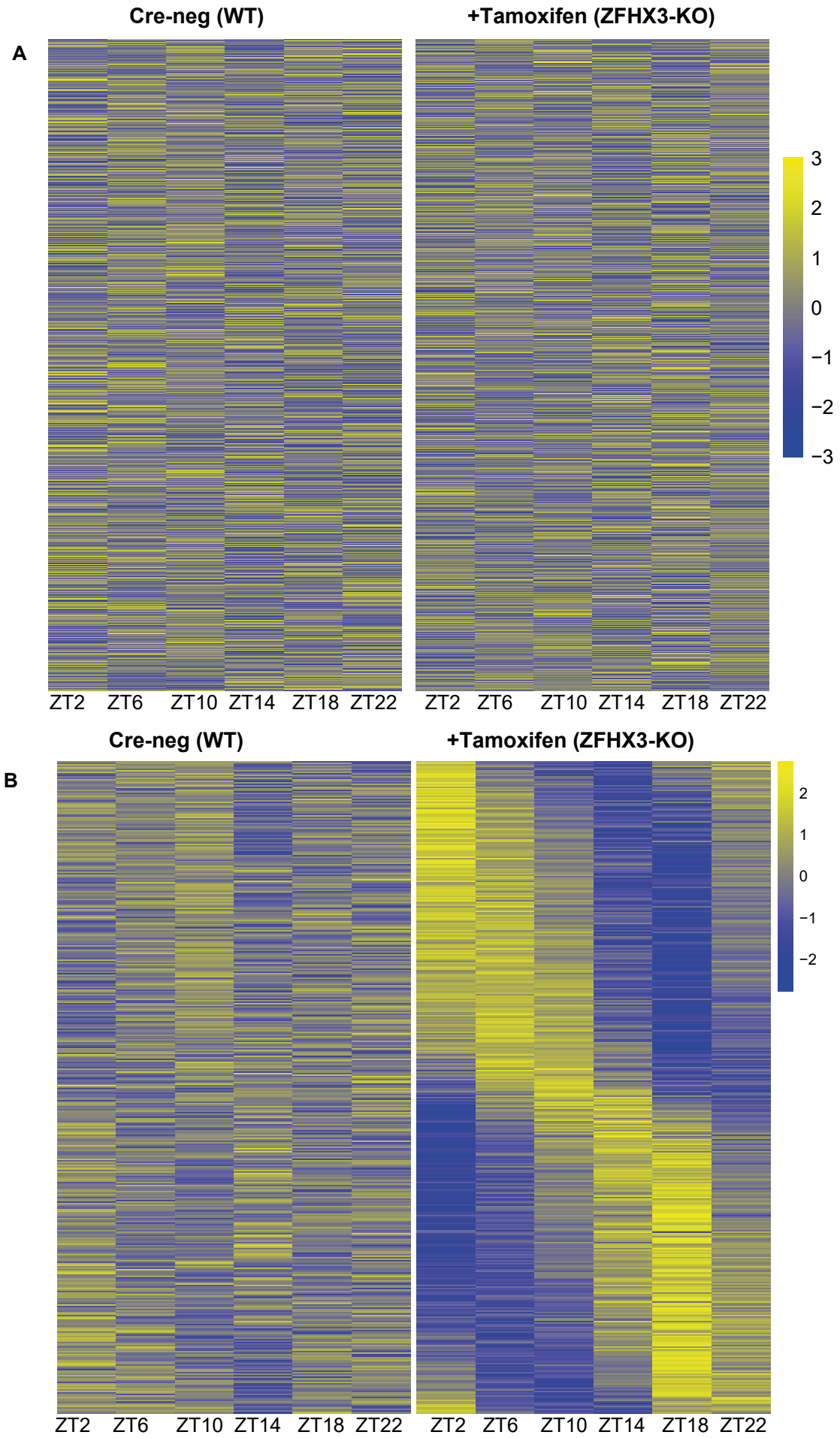
Supplementary figure 1. Gene ontology and motif analysis of ZFHX3 bound sites. (A) left- Functional annotation of nearest neighbouring gene (TSS) to active promoter sites occupied by ZFHX3 and histone marks (H3K4me3 and H3K27ac) using KEGG pathway (DAVID), right- ciradian rhythm pathway genes that showed co-occupancy of ZFHX3 and histone marks at TSS illustrated using KEGG pathway . (B ,C) Functional annotation of nearest neighbouring gene (TSS) to sites occupied by ZFHX3 + H3K4me3 (light green) and ZFHX3 alone (dark green) (D) Over-represented transcription factor binding motifs at ZFHX3+H3K4me3+H3K27ac sites (shaded blue), ZFHX3+H3K27ac sites (shaded green) and ZFHX3+ H3K4me3 (shaded red)by HOMER.

Supplementary Figure 2



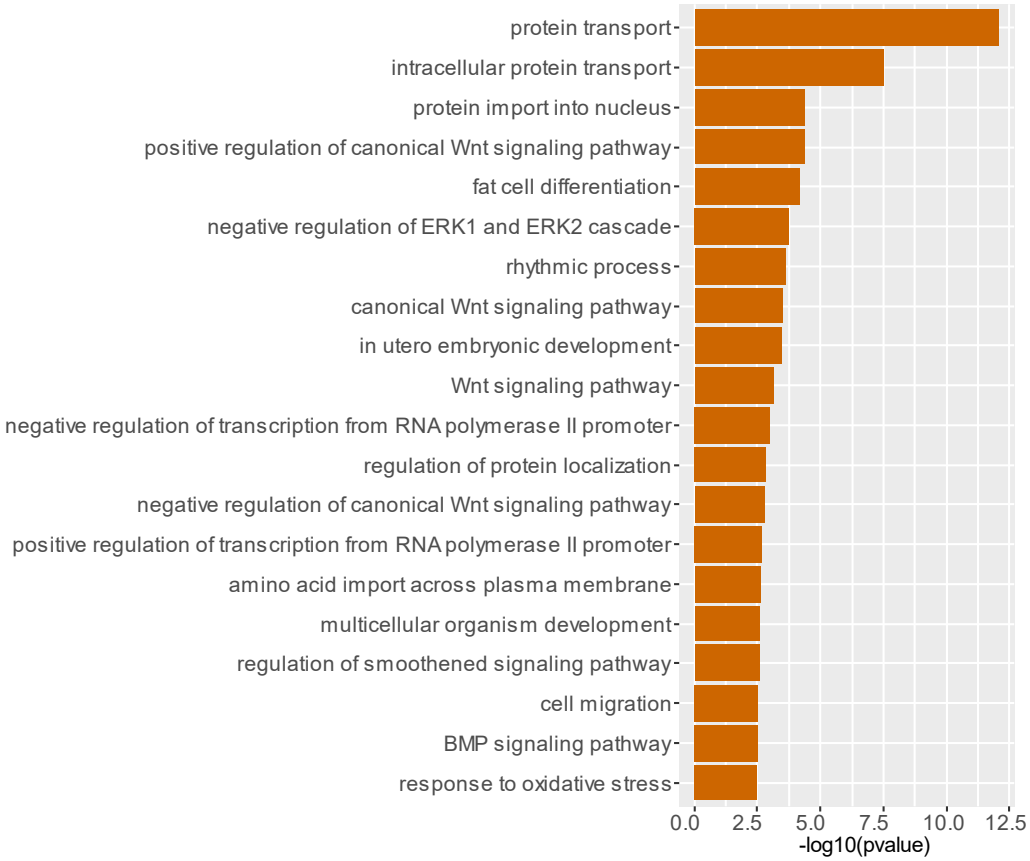
Supplementary Figure 2. +Tamoxifen animals show changes in circadian wheel running parameters in conditions of constant darkness. Compared to -Tamoxifen animals, +Tamoxifen show a significantly (A) reduced *Zfhx3* gene expression ($p<0.02$, unpaired t-test) (B) reduced circadian amplitude, (C) increased Intradaily variability (IV) , (D) reduced Interdaily stability(IS) and (E) advanced activity onset in LD.

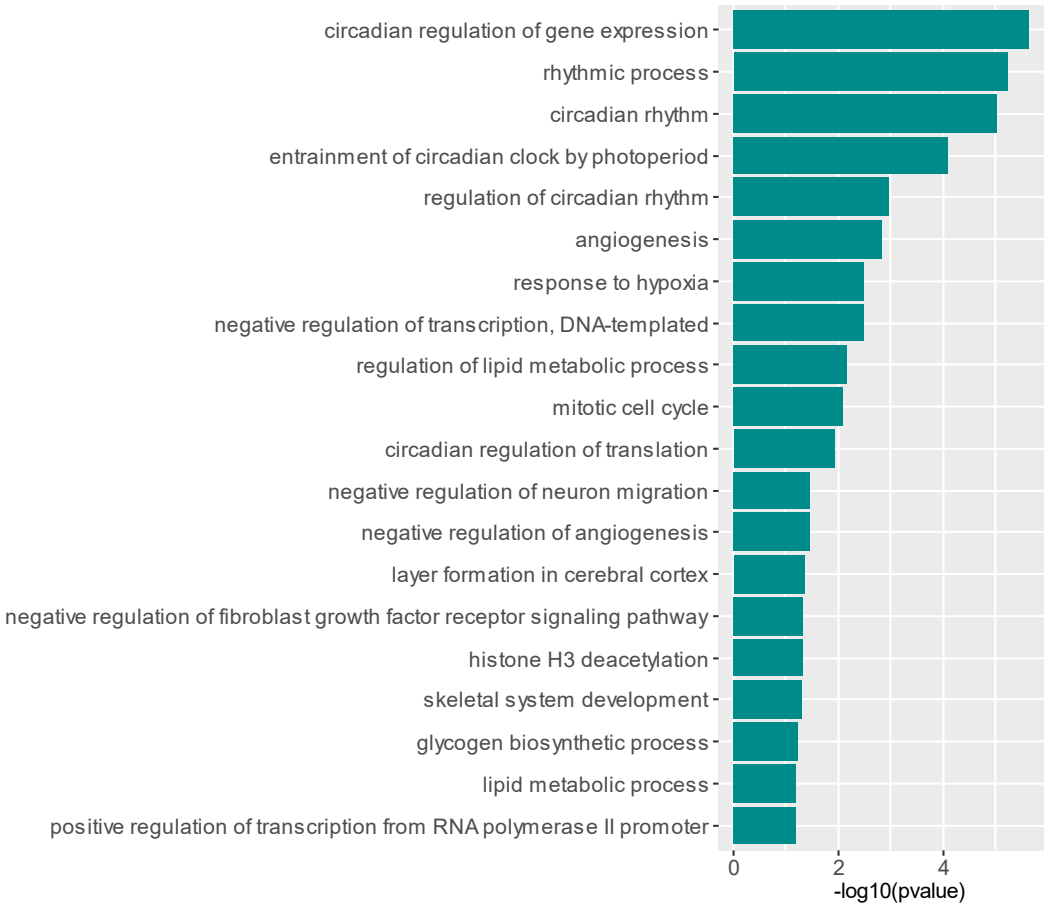
Supplementary Figure 3



Supplementary Figure 3. Effect of *Zfhx3*-KO on SCN transcriptome. (A) Heatmap showing module 1 genes (n = 16832) that show no rhythmic expression before and after *Zfhx3*-KO as computed by dryR statistical framework (B)Heatmap showing module 3 genes (n = 403) that showed rhythmic expression after *Zfhx3*-KO.

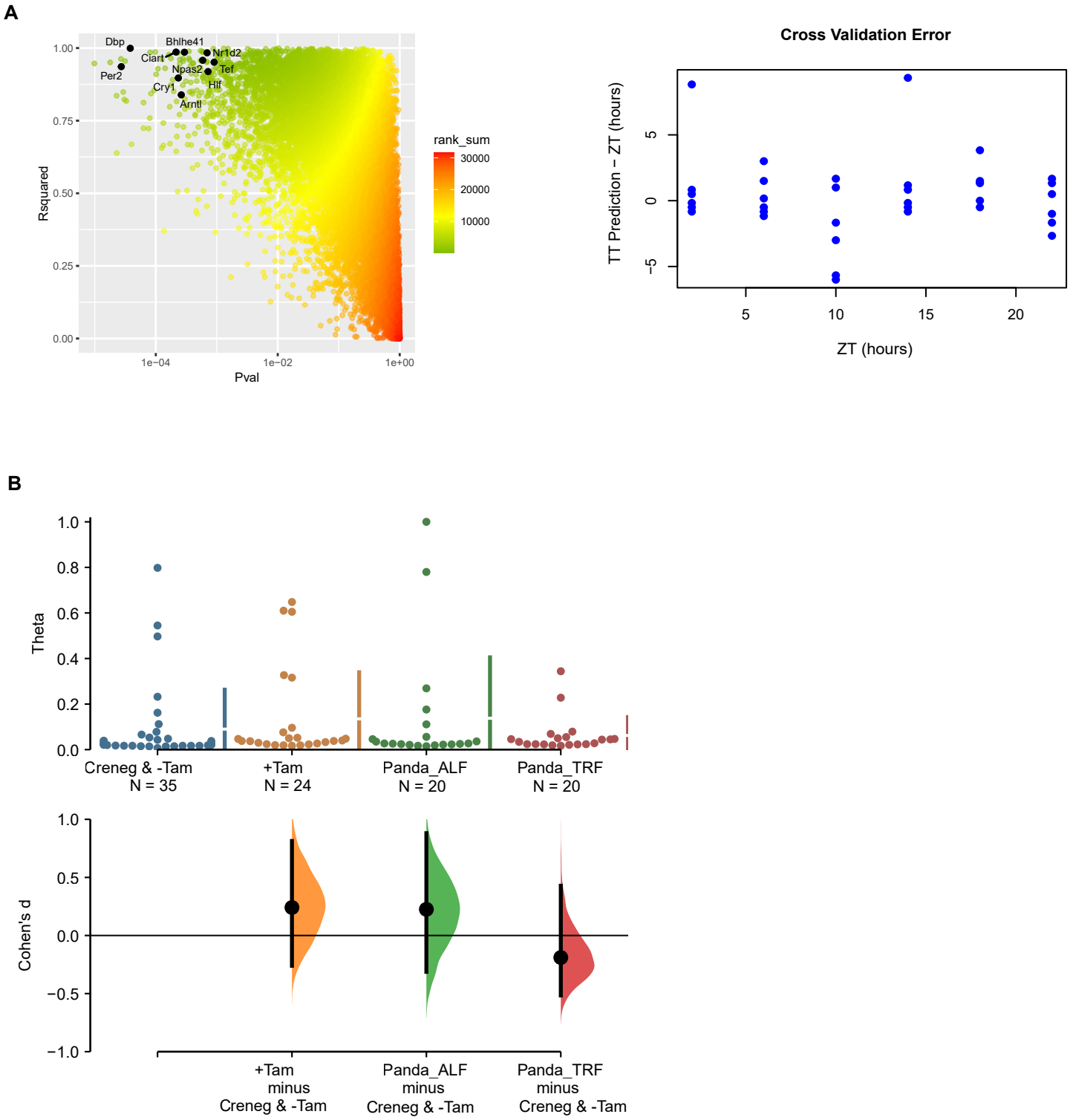
Supplementary Figure 4





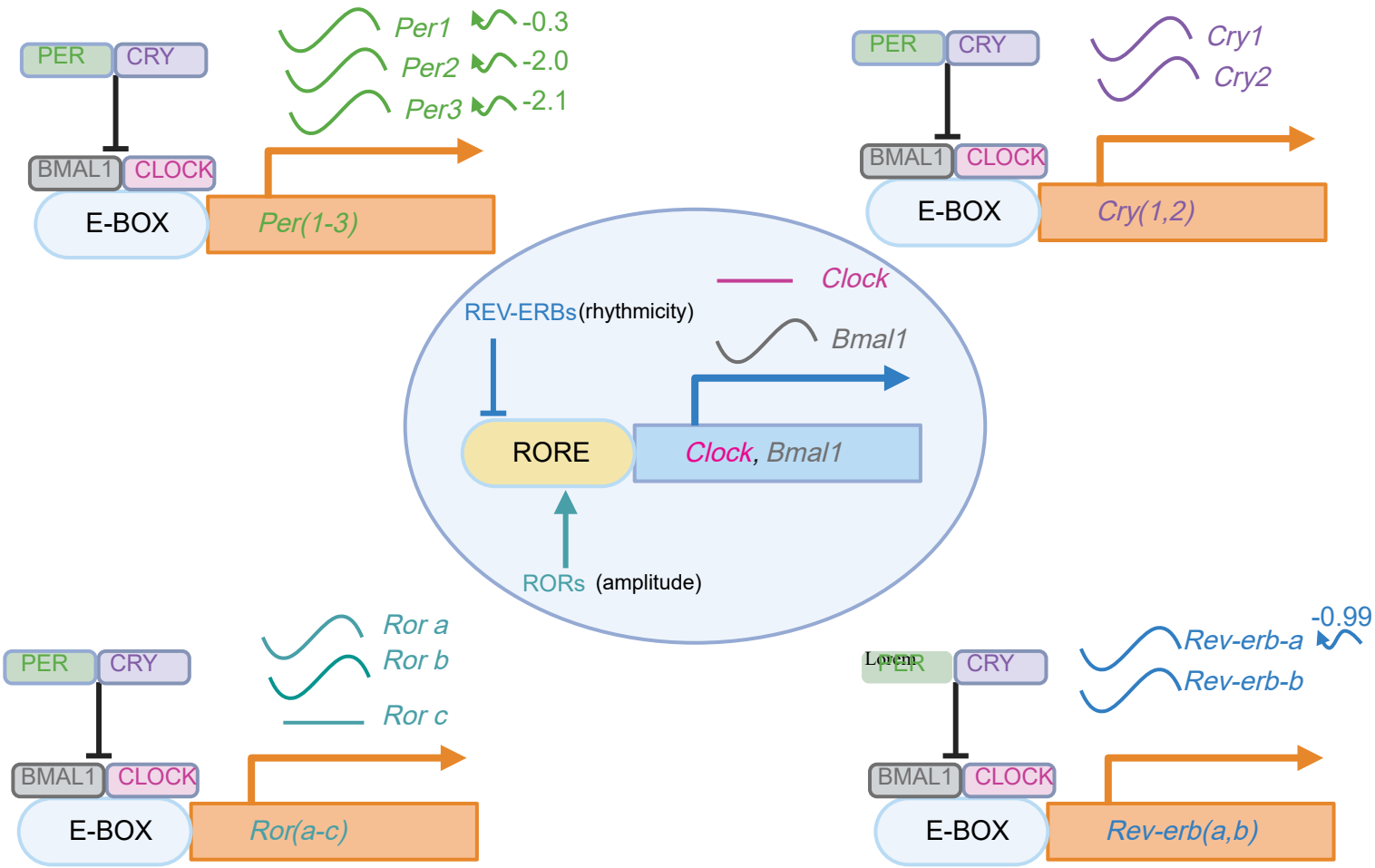
Supplementary Figure 4. Functional annotation of rhythmic genes in the SCN. (A) Top 20 enriched functional pathways from genes that lost rhythmic transcription after gene deletion (module 2) and (B) genes that showed change in phase and/or amplitude (module 5) using the GO::BP (biological processes) terms by DAVID.

Supplementary Figure 5



Supplementary Figure 5. SCN circadian clock using TimeTeller. (A) (Left) Choosing genes for the TimeTeller model. Looking for good rhythmicity p-val and high R squared in population cosinor analysis leads to selection of a limited panel of well-behaved clock genes (named), i.e., clock genes with adjusted pval < 0.10 and R squared > 0.80.(Right) Results of the cross-validation of the TT model using the indicated genes in the left panel. (B) TimeTeller theta parameter is indistinguishable between the genotypes. Theta is a measure of certainty for phase prediction in TimeTeller, i.e., how closely the test samples can be projected into the training model.

Supplementary Figure 6



Supplementary Figure 6. Effect of ZFHX3-KO on circadian clock . The effect of loss of ZFHX3 on the primary and ancillary TTFL in the SCN. Blue and red arrow depicts decreased/increased average gene expression in *Zfhx3* deficient mice as compared to the control, respectively. The waveform symbol next to the gene name depicts detected 24-h rhythmic expression in our study. The curly arrows (right) of the gene illustrates the change in phase of the rhythmic expression.