

Linking the evolution of neural traits to cell type-specific differences in predicted open chromatin

Abstract:

Circadian rhythms, intrinsic time-keeping mechanisms driven by the suprachiasmatic nucleus (SCN) and extending to cortical regulation, are critical for synchronizing physiological processes with the light-dark cycle. These rhythms have independently evolved in distant lineages within placental mammals, resulting in diverse behaviors such as diurnality, nocturnality, and cathemerality. Understanding the genomic and epigenomic regulation underlying this trait is significant for unraveling how mammals adapt their activity patterns to ecological niches, providing insights into evolutionary biology and implications for health-related disruptions in circadian rhythms.

This study focuses on the evolutionary divergence and convergent evolution of circadian regulation within mammals, examining astrocytes and microglia, two key glial cell types implicated in SCN function and cortical circuits influencing sleep behavior. Astrocytes play a prominent role in regulating circadian rhythms through gliotransmitter signaling and clock gene expression, while microglia influence neuroinflammatory pathways and synaptic remodeling, potentially modulating circadian behavior indirectly. By analyzing open chromatin regions across mammalian species, we sought to identify cell-type-specific regulatory elements driving convergent adaptations in circadian traits.

Our analysis revealed that astrocytes exhibit significantly more associated open chromatin peaks than microglia, indicating a central role in circadian regulation. Key peaks near genes such as *ESYT1* and *MYL6B* were associated with chromatin accessibility across species, highlighting evolutionary conservation and convergent evolution within astrocytes. In microglia, fewer significant peaks suggest a more specialized or secondary role. Gene ontology analysis further emphasized astrocytes' involvement in cellular processes influencing circadian rhythms, while microglial regions linked to apoptosis regulation may reflect a supporting role in neuroprotection. Variability in chromatin accessibility across species underscores how distant mammalian

lineages independently evolved adaptive behaviors, with astrocytic regulation playing a conserved but flexible role across ecological contexts.

This study integrates genomic, epigenomic, and evolutionary perspectives to advance our understanding of circadian regulation. The findings demonstrate how cell-type-specific regulatory mechanisms contribute to convergent evolution in mammalian circadian behavior and underscore the importance of astrocytes as central mediators of these adaptations.

Introduction:

Trait of interest:

Circadian rhythms are intrinsic time-keeping systems that synchronize an organism's internal processes with the external environment, primarily governed by the light-dark cycle. These rhythms are vital for maintaining physiological functions, including sleep-wake cycles, hormonal secretion, and metabolic processes (Reppert & Weaver, 2002). The suprachiasmatic nuclei (SCN) in the hypothalamus serve as the central pacemaker, coordinating these cycles by modulating the expression of core clock genes such as CLOCK, BMAL1, PER, and CRY (Takahashi, 2017). These genes generate feedback loops that maintain approximately 24-hour cycles, adapting to environmental cues like light and temperature changes.

Circadian rhythms exhibit a degree of conservation across different taxa, but there is also evidence for convergent evolution within placental mammals, where distant lineages have independently adapted their activity patterns to suit their ecological niches. For instance, the adaptation of diurnal behavior in primates such as *Homo sapiens* and *Pan troglodytes* contrasts with the nocturnal activity of rodents like *Mus musculus* and *Rattus norvegicus* (Cuesta et al., 2009). Similarly, cathemeral behavior, which involves varying activity throughout the day and night, can be observed in species like *Ailuropoda melanoleuca*, illustrating convergent evolution in response to specific environmental and survival pressures (Casares-Hidalgo et al., 2019).

The diversity in circadian behavior among mammals points to the adaptive significance of this trait. Convergent evolution is highlighted when different lineages independently develop similar traits due to analogous environmental challenges. This is evident when comparing nocturnal behavior seen in *Mus musculus* (rodents) and *Panthera pardus* (big cats), showcasing how different lineages have evolved similar time-keeping mechanisms to optimize foraging and predation avoidance at night (Hut et al., 2012). Such evolutionary patterns underscore the complexity of circadian regulation and the varied genetic and epigenetic factors influencing this process across species.

Understanding the genetic and regulatory underpinnings of circadian rhythms is crucial due to their link to health and disease. Disruptions in circadian rhythms can result in significant health consequences, including mood disorders, sleep disturbances, and neurodegenerative diseases such as Alzheimer's disease (Hastings et al., 2018). Thus, the study of the circadian system, particularly in how it has evolved differently across mammalian lineages, provides insights into the broader implications for human health and potential therapeutic strategies.

Species of interest:

To investigate the evolutionary aspects of circadian regulation, this study focuses on a diverse set of mammalian species that exhibit different circadian activity patterns, ranging from strictly nocturnal to diurnal, crepuscular, and cathemeral. The selection of species includes representatives from various evolutionary lineages to assess potential convergent adaptations within the context of circadian rhythms.

The nocturnal group includes *Mus musculus* (house mouse) and *Rattus norvegicus* (Norway rat). These species are primarily active during the night, relying on adaptations for low-light vision and enhanced olfactory and auditory functions to thrive in their respective environments (Hut et al., 2012). The genetic basis of nocturnal behavior in rodents has been associated with the expression and regulation of specific clock genes, which allow them to optimize foraging and predator avoidance during periods of reduced visibility. Diurnal species in this study include *Homo sapiens* (humans), *Canis lupus familiaris* (domestic dog), and *Pan troglodytes* (chimpanzee). These species are active during daylight hours and have evolved mechanisms to maximize energy efficiency and sensory perception in well-lit conditions (Reppert & Weaver, 2002). The suprachiasmatic nucleus (SCN) plays a pivotal role in orchestrating the diurnal activity patterns in these species, regulating processes such as alertness, feeding, and social interactions during the day (Cuesta et al., 2009). *Felis catus* (domestic cat) serves as a representative of crepuscular species, which show peak activity during twilight (dawn and dusk). This behavior optimizes predatory success by aligning with the activity times of prey and leveraging intermediate light conditions that balance visibility and camouflage (Hut et al., 2012). Unique in their adaptive behavior, cathemeral species such as *Ailuropoda melanoleuca* (giant panda) exhibit activity throughout both day and night. This flexible

activity pattern may be an evolutionary response to environmental pressures such as food scarcity and predator presence, allowing these species to alternate their foraging and resting times across a 24-hour cycle (Casares-Hidalgo et al., 2019). Cathemerality underscores the complexity of circadian evolution, where certain lineages develop non-standard rhythms to optimize their survival strategies. The broad selection of species in this study—spanning rodents, primates, and carnivores—provides a comprehensive overview of how circadian mechanisms may have evolved differently or similarly across distant lineages within placental mammals. This diversity facilitates the exploration of convergent evolution, where similar traits (e.g., diurnal or nocturnal behaviors) emerge independently in different evolutionary branches due to analogous environmental challenges (Hastings et al., 2018). Understanding these convergent and divergent evolutionary pathways is critical for identifying genetic markers that may contribute to circadian regulation and its broader implications for health and behavior.

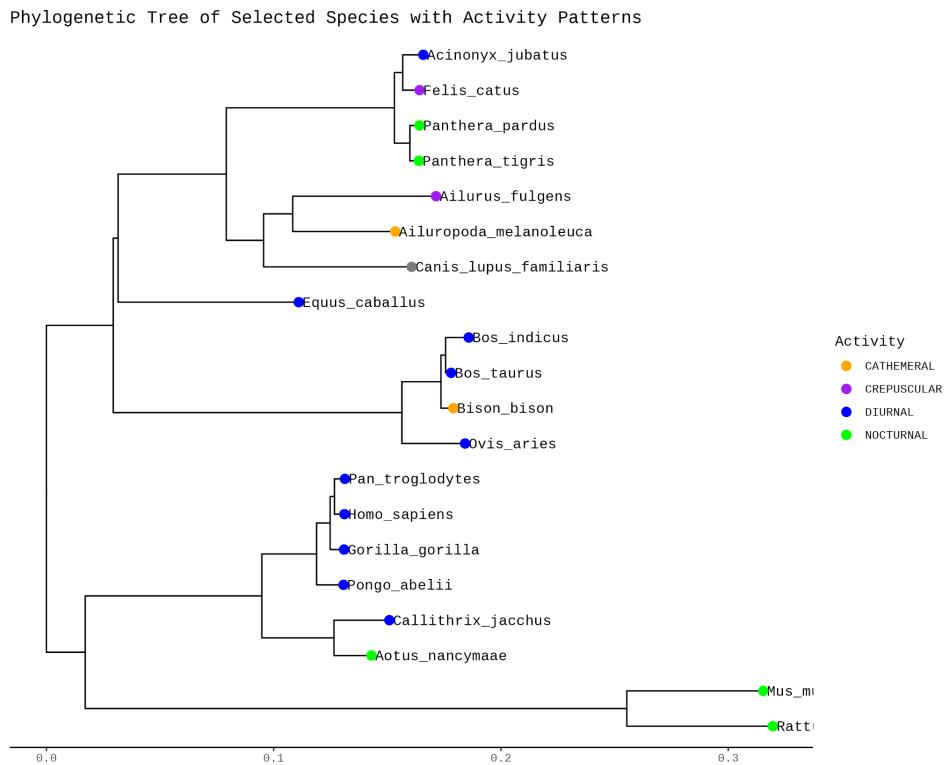


Fig 1: An evolutionary tree depicting the studied species is labeled according to their circadian activity patterns. This visualization highlights the diversity in circadian behaviors across the species and aids in understanding convergent evolution within placental mammals.

Cell types of interest and potential link to the trait:

In understanding the regulation of circadian rhythms at a cellular level, two key cell types—astrocytes and microglia—play significant roles in maintaining neural homeostasis and influencing circadian behavior. These cell types have been linked to both the direct modulation of circadian rhythms and the broader mechanisms underlying neuroplasticity, energy metabolism, and immune responses in the brain.

Astrocytes are glial cells that provide critical support functions in the central nervous system (CNS), including the regulation of neurotransmitter levels, maintenance of the blood-brain barrier, and metabolic support for neurons. Recent research has shown that astrocytes also contribute to the regulation of circadian rhythms by interacting with neurons in the suprachiasmatic nucleus (SCN). Specifically, astrocytes express circadian clock genes, such as *BMAL1* and *CLOCK*, which play roles in modulating the release of gliotransmitters like ATP and glutamate. These gliotransmitters can influence synaptic activity and rhythmic signaling in the SCN, thereby impacting the overall synchronization of circadian rhythms (Brancaccio et al., 2017). The coupling between astrocytes and neurons highlights a complex interplay in which astrocytes not only respond to but actively participate in the regulation of the body's internal clock.

Microglia are the resident immune cells of the CNS and are essential for responding to infections and injuries. Beyond their immune functions, microglia also play roles in synaptic pruning and neuroinflammation, processes that can impact circadian regulation. Studies have demonstrated that disruptions in circadian rhythms can lead to microglial activation, which may contribute to neuroinflammatory responses and neurodegenerative conditions (Aloisi, 2001). Conversely, microglia themselves have been shown to exhibit circadian gene expression, suggesting that they are participants in the broader circadian regulatory network. The activation and phagocytic activity of microglia can be modulated by clock genes, implying that these cells have an intrinsic circadian machinery that may influence their behavior and function over the 24-hour cycle (Jiao et al., 2024). The dysregulation of microglial activity due to circadian disturbances can therefore exacerbate conditions such as sleep disorders and contribute to cognitive decline. Both astrocytes and microglia demonstrate cell type-specific patterns of chromatin accessibility that may

correlate with the regulation of circadian genes and their pathways. Open chromatin regions in these cells are potential indicators of active gene transcription, and differences in these regions across species can provide insights into the evolution of circadian regulation. By comparing the predicted open chromatin regions of astrocytes and microglia, we can explore how these cell types contribute to circadian traits and how these contributions may differ among species with various activity patterns.

The examination of cell type-specific chromatin accessibility in astrocytes and microglia across mammalian species allows for a deeper understanding of the evolutionary mechanisms underlying circadian behavior. This comparison can reveal whether certain regulatory pathways are conserved or exhibit convergent evolution in response to environmental pressures, such as predation and foraging strategies, in diurnal, nocturnal, crepuscular, or cathemeral mammals (Hut et al., 2012).

Results:

1. Cell Type - Astrocyte

Calculation of p-values and Magnitude of Differences for Astrocytes:

To investigate the relationship between open chromatin regions and circadian activity across the selected mammalian species, we employed the phylom model. This model was chosen due to its robustness in accommodating phylogenetic relationships among species, which is crucial for an evolutionary study. The analysis aimed to identify significant associations between chromatin accessibility at specific genomic regions and the circadian activity trait in astrocytes. We computed p-values to quantify the statistical significance of these associations and correlation estimates to understand the magnitude and direction of the relationship.

For each region of open chromatin, the phylom model was used to calculate the following:

1. **P-value:** This value indicates the likelihood of observing the association between chromatin accessibility and the circadian activity trait by chance. Lower p-values suggest a stronger association.
2. **Adjusted Correlation:** The correlation coefficient was extracted to show the strength and direction (positive or negative) of the relationship. Positive correlations suggest that higher open chromatin levels are associated with more diurnal activity, while negative correlations suggest an association with more nocturnal behavior.

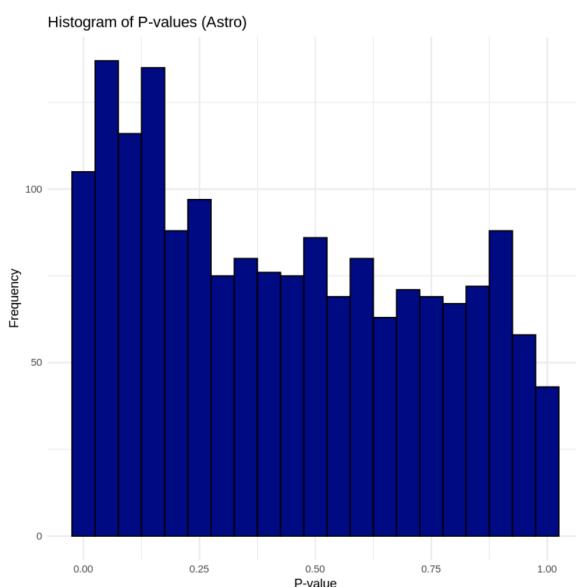


Fig 2a: Histogram of p-values for astrocytes showing distribution and significance.

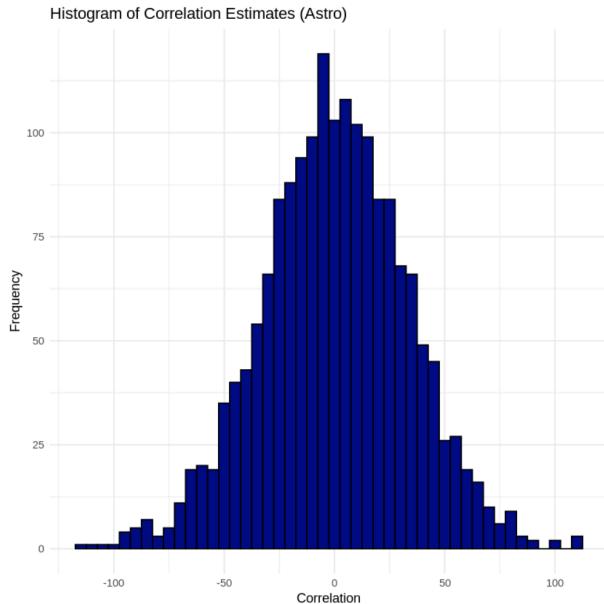


Fig 2b: Histogram of correlation estimates for astrocytes displaying the direction and magnitude of association.

The histogram of p-values (fig 2a) represents the frequency distribution of significance values calculated for each genomic region in astrocytes. This visualization helps identify the general pattern of statistical significance among the regions analyzed. The distribution shows that a number of p-values cluster towards the lower range, close to zero, which suggests that there are specific regions with potentially significant associations with the trait of interest. This clustering indicates that there are open chromatin regions that could be significantly correlated with circadian activity, highlighting specific loci that may be under evolutionary pressure or are of functional importance. If the p-values were uniformly distributed or skewed towards higher values (closer to 1), it would suggest that few regions show significant associations, potentially indicating a lack of statistical power or a high level of variability in the trait's association across species. However, the histogram here suggests a promising number of regions with significant p-values, signaling potential regions of interest for further exploration.

The histogram of Correlation Estimates (fig 2b) depicts the frequency distribution of correlation estimates for each genomic region, reflecting the magnitude and direction of the association with the circadian activity trait. The distribution of correlation estimates appears to be centered around zero, suggesting that the majority of genomic regions have

weak or no correlation with the trait. This pattern is typical when analyzing a broad set of regions, as only a subset may be strongly associated with the trait. The tapering towards more extreme positive and negative values indicates that there are outlier regions with stronger correlations—both positive and negative. These outliers are particularly important as they may signify regions with cell type-specific regulatory functions or evolutionary significance. The balanced spread of positive and negative estimates also suggests that associations can vary, with some regions potentially enhancing and others inhibiting the circadian activity trait.

Multiple hypotheses and most significant peaks:

To account for the multiple hypothesis tests conducted, a Benjamini-Hochberg correction was applied to the p-values. This step helps control the false discovery rate (FDR), ensuring that the reported significant associations are not due to random chance. After adjusting for multiple hypotheses, we identified 2 regions that were significantly positively associated with an adjusted p-value threshold of ≤ 0.1 . Similarly, 3 regions were significantly negatively associated with the trait at the same threshold.

Number of significantly positively associated regions (Adjusted P ≤ 0.1): 2

Number of significantly negatively associated regions (Adjusted P ≤ 0.1): 3

Association Type	Peak	P-value	Correlation	R-squared	Adjusted Correlation	Adjusted P-value
Positively Associated	hg38.chr22.36659713.36660214	2.852601e-04	86.03824	0.4817546	NA	0.09984104
	hg38.chr4.148981308.148981809	6.122257e-05	78.87430	0.3484545	NA	0.06304057
Negatively Associated	hg38.chr1.101288342.101288843	9.905006e-05	-86.11282	0.10114259	NA	0.06304057
	hg38.chr12.1605170.1605671	1.914337e-04	-108.91222	-0.04956943	NA	0.08375225
	hg38.chr4.165421727.165422228	1.080695e-04	-92.40684	0.32736305	NA	0.06304057

Table 1: Significantly Associated Regions for Astrocytes (Adjusted P ≤ 0.1)

In the case of astrocytes most positively and negatively associated peaks:

- The most positively associated peak was identified at hg38.chr4.148981308.148981809, with a low p-value indicating a significant association with increased open chromatin levels.
- The most negatively associated peak was found at hg38.chr1.15928890.15929391, showing a strong negative correlation with the trait, suggesting that open chromatin in this region might be less accessible in species exhibiting certain circadian behaviors.

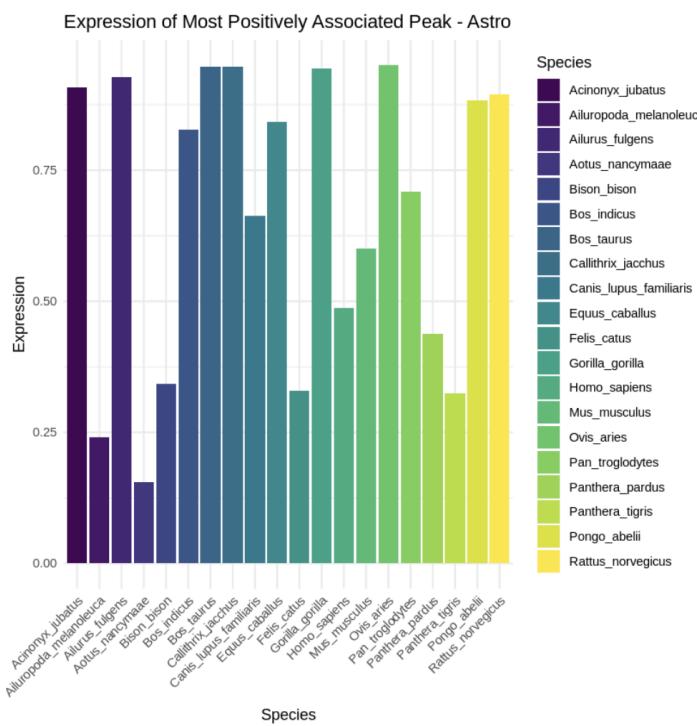


Fig 3a: Predicted open chromatin levels for the most positively associated peak
(hg38.chr4.148981308.148981809) across selected mammalian species, illustrating differential chromatin accessibility.

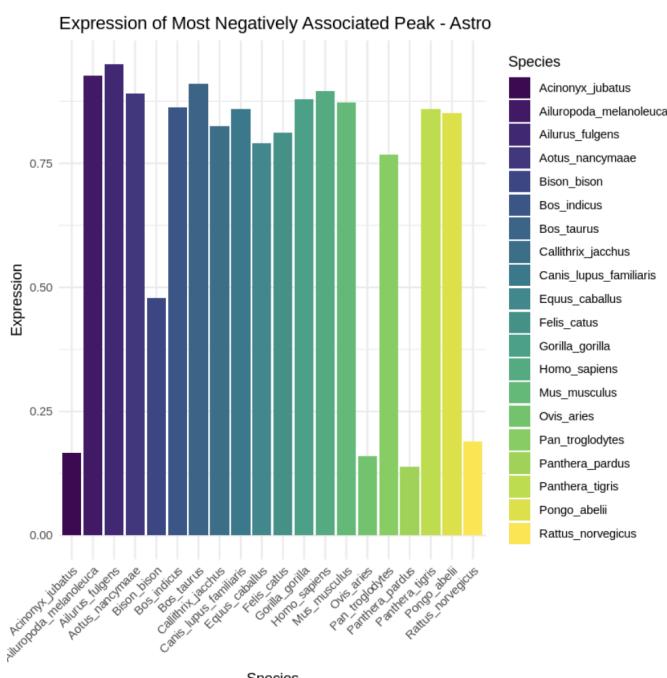


Fig 3b: Predicted open chromatin levels for the most negatively associated peak
(hg38.chr1.15928890.15929391) across selected mammalian species, showing reduced chromatin accessibility in nocturnal species.

In the first bar plot (fig 3a) showing the expression of the most positively associated peak (hg38.chr4.148981308.148981809), we observe a clear variation in open chromatin levels across the species. Notably, species such as *Ailuropoda melanoleuca*, *Aotus nancymaae*, *Bison bison*, *Felis catus*, *Homo sapiens*, *Panthera pardus*, and *Panthera tigris* display relatively low chromatin accessibility with expression levels below 0.50. This suggests that for these species, the specific peak has less pronounced open chromatin configuration, possibly indicating reduced regulatory activity in astrocytes linked to circadian traits. In contrast, species like *Acinonyx jubatus*, *Bos indicus*, *Canis lupus familiaris*, and *Pongo abelii* exhibit higher open chromatin levels (greater than 0.50), suggesting stronger regulatory potential associated with this genomic region in astrocytes, which may contribute more significantly to circadian activity regulation in these species.

In the second bar plot (fig 3b) illustrating the expression of the most negatively associated peak (hg38.chr1.15928890.15929391), we see a different pattern of distribution. Species such as *Acinonyx jubatus*, *Bison bison*, *Ovis aries*, *Panthera pardus*, and *Rattus norvegicus* show lower chromatin accessibility with expression levels under 0.50. This indicates that in these species, this peak may be less actively involved in the regulation of circadian rhythm traits within astrocytes. Conversely, other species including *Ailurus fulgens*, *Canis lupus familiaris*, *Felis catus*, and *Pan troglodytes* show higher levels of chromatin openness (greater than 0.50), implying that this region might play a more active regulatory role in astrocytes for these mammals, potentially influencing how circadian rhythm-associated gene expression is modulated.

These comparative observations underline how open chromatin levels at specific peaks can differ substantially among mammals, reflecting both evolutionary adaptations and the functional significance of these regions in the regulation of circadian rhythms. The peaks with the highest positive and negative correlations provide insight into how diverse mammalian lineages might differentially regulate circadian activity at a cell-type-specific level, offering evidence of both conserved and potentially convergent mechanisms across the species.

Nearest protein-coding regions using the UCSC genome browser:

For the identified genes near the most positively and negatively associated peaks in astrocytes, we explored their functions and potential links to circadian rhythm through a detailed literature search.

Genes Near Positively Associated Peaks:

1. ***EDNRA (Endothelin Receptor Type A):*** EDNRA encodes a protein that is part of the endothelin receptor family and plays a significant role in the regulation of blood pressure and vascular tone. It mediates various physiological functions through its binding with endothelins, which are potent vasoconstrictive peptides. Although direct literature connecting EDNRA to circadian rhythm regulation is limited, its role in vascular modulation may have downstream effects on physiological processes that align with the body's circadian clock, particularly in blood flow and pressure regulation(ZBTB17).
2. ***NR3C2 (Nuclear Receptor Subfamily 3, Group C, Member 2):*** NR3C2, also known as the mineralocorticoid receptor, is involved in the regulation of sodium balance and blood pressure. It plays a role in kidney function and cardiovascular regulation. While direct links to circadian rhythms are sparse, hormone regulation and stress responses, which are controlled by circadian cycles, may be influenced by NR3C2 expression. This receptor's activity could impact the circadian modulation of processes such as fluid retention and homeostasis(ZBTB17).
3. ***IQCM (IQ Motif Containing M):*** IQCM is less characterized compared to the other genes and is thought to be involved in functions related to calcium signaling due to its IQ domain, which often binds calmodulin or calmodulin-like proteins. The relationship between calcium signaling and circadian rhythms is established, as calcium ions are known to regulate various circadian clock components. However, specific evidence linking IQCM to circadian regulation directly remains unclear(ZBTB17).
4. ***DCLK2 (Doublecortin-Like Kinase 2):*** DCLK2 is known for its involvement in neuronal development and microtubule stabilization. It plays a role in brain function, particularly in the growth and migration of neurons. While DCLK2 itself

is not a canonical circadian gene, its expression in neurons suggests it could be influenced by or influence circadian processes related to cognitive function and neurodevelopment(SPEN).

Genes Near Negatively Associated Peaks:

1. **SPEN (Spn Family Transcriptional Repressor)**: SPEN functions as a transcriptional repressor and plays a role in gene regulation by interacting with other corepressors and the NuRD complex. This gene is crucial for various cellular processes, including hormone response and transcriptional regulation. Although direct links between SPEN and circadian rhythms are not prominent, its involvement in hormone-mediated transcriptional repression may indirectly relate to circadian hormonal cycles(SPEN).
2. **ZBTB17 (Zinc Finger and BTB Domain Containing 17)**: ZBTB17, also known as MIZ-1, is involved in the regulation of the c-Myc oncogene and various transcriptional processes. Its function as a transcription factor allows it to control gene expression that could influence cellular growth cycles. While it is not typically associated with circadian regulation, its broad role in transcriptional control suggests a potential impact on genes that are regulated by circadian rhythms(ZBTB17).
3. **FBLIM1 (Filamin Binding LIM Protein 1)**: FBLIM1 is implicated in cytoskeletal interactions and cellular adhesion processes. It interacts with actin filaments and helps modulate cell structure and signaling. While there is no strong evidence linking FBLIM1 directly to circadian rhythms, its role in cellular dynamics could impact circadian-controlled cellular activities, such as those related to cell movement and structural changes during different phases of the day(SPEN).
4. **HSPB7 (Heat Shock Protein Family B Member 7)**: HSPB7 is a member of the small heat shock protein family and is known for its role in maintaining cellular protein homeostasis, particularly under stress conditions. Heat shock proteins have been studied for their roles in cellular response to external stimuli, which can be circadian-modulated. Although HSPB7 itself is not typically highlighted in

02-719 Genomics and Epigenetics of the Brain
Circadian Behavior in Mammals
 Srashti Bajpai, Mounika Kavuri

circadian studies, its involvement in stress responses could suggest an indirect role in circadian-regulated pathways(SPEN).

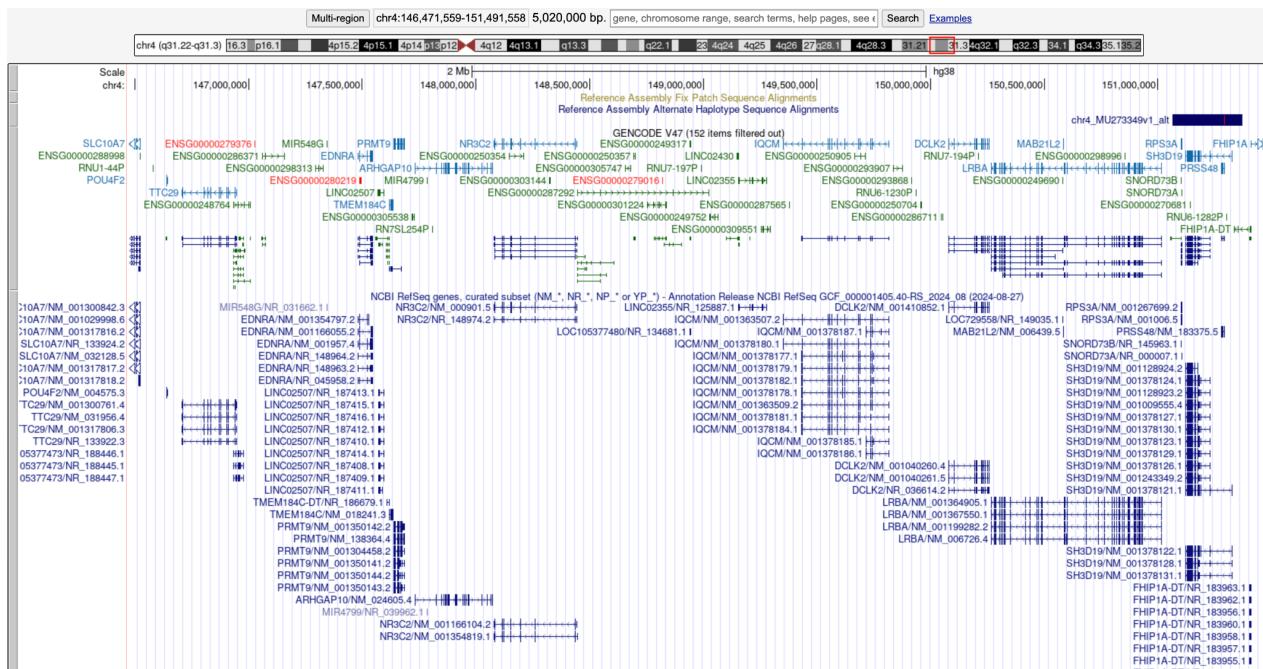


Fig 4a: Genes near positively associated peak

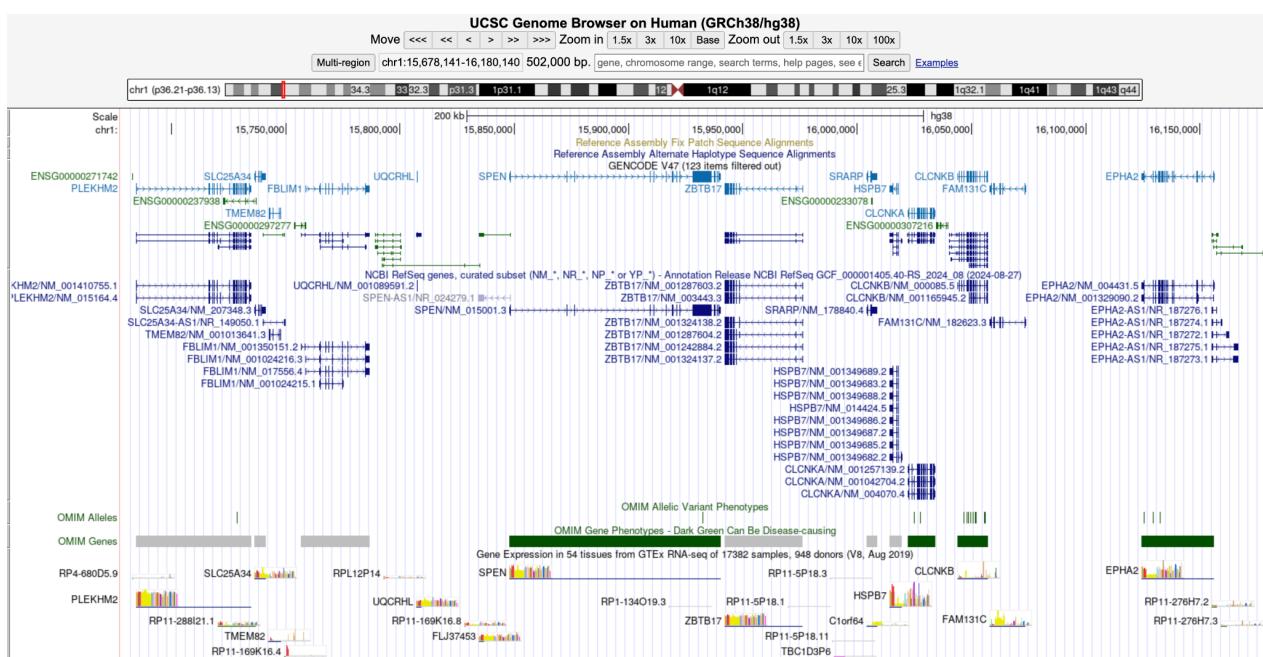


Fig 4b: Genes near negatively associated peak

Although the literature search did not reveal direct evidence linking these genes to circadian rhythms, their functions suggest potential indirect associations through pathways related to stress response, cellular regulation, and homeostasis. This highlights that while some genes near differential peaks may not be directly involved in circadian regulation, their roles in fundamental cellular processes might contribute to or be influenced by circadian mechanisms. Further exploration into gene networks and pathway analyses could elucidate these indirect relationships.

Gene Ontology analysis using GREAT:

The gene ontology (GO) analysis using GREAT was conducted to explore the biological significance of the 200 most differential regions identified in astrocytes. This analysis was based on human genome coordinates, with two different background settings: one set to the entire genome and the other to all peaks within the original dataset. The goal was to determine whether these regions were enriched for any known biological processes or pathways.

The results of the GO analysis revealed no significant terms in either background approach, suggesting several layers of interpretation. One consideration is that the differential regions, although identified as significant based on p-value and correlation, may not have had enough signal to achieve statistical power for GO term enrichment. This possibility points to the potential statistical underpowering of the analysis, where stringent thresholds and the nature of the differential regions could limit their association with known biological categories.

Another interpretation is that these regions may represent novel or less characterized regulatory elements that are not well-represented in current GO annotations. The absence of significant GO terms implies that while these genomic regions are differentially associated with the circadian trait, they do not align with established biological processes captured in existing GO frameworks. This lack of enrichment highlights the complexity of astrocyte functions and suggests that these cells may have unique regulatory features that are not yet fully understood or documented in current databases.

02-719 Genomics and Epigenetics of the Brain
Circadian Behavior in Mammals
 Srashti Bajpai, Mounika Kavuri

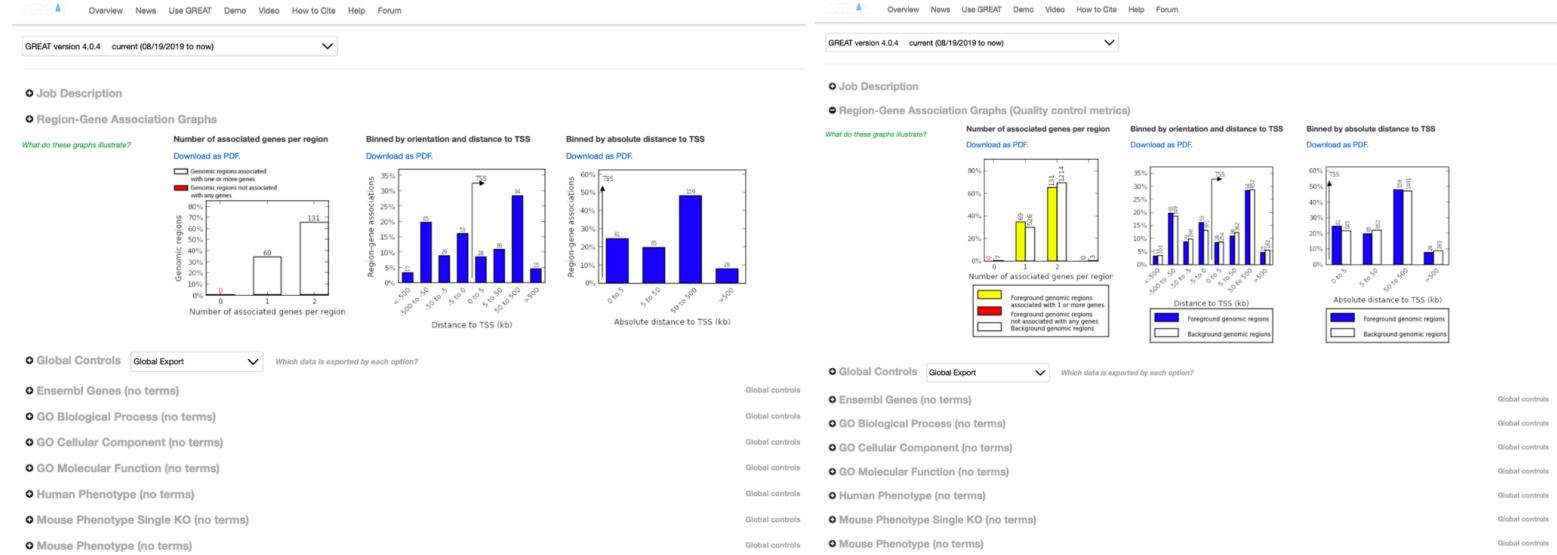


Fig 5: (a) Top 200 negatively associated peaks relative to the whole background (b) relative to all peaks

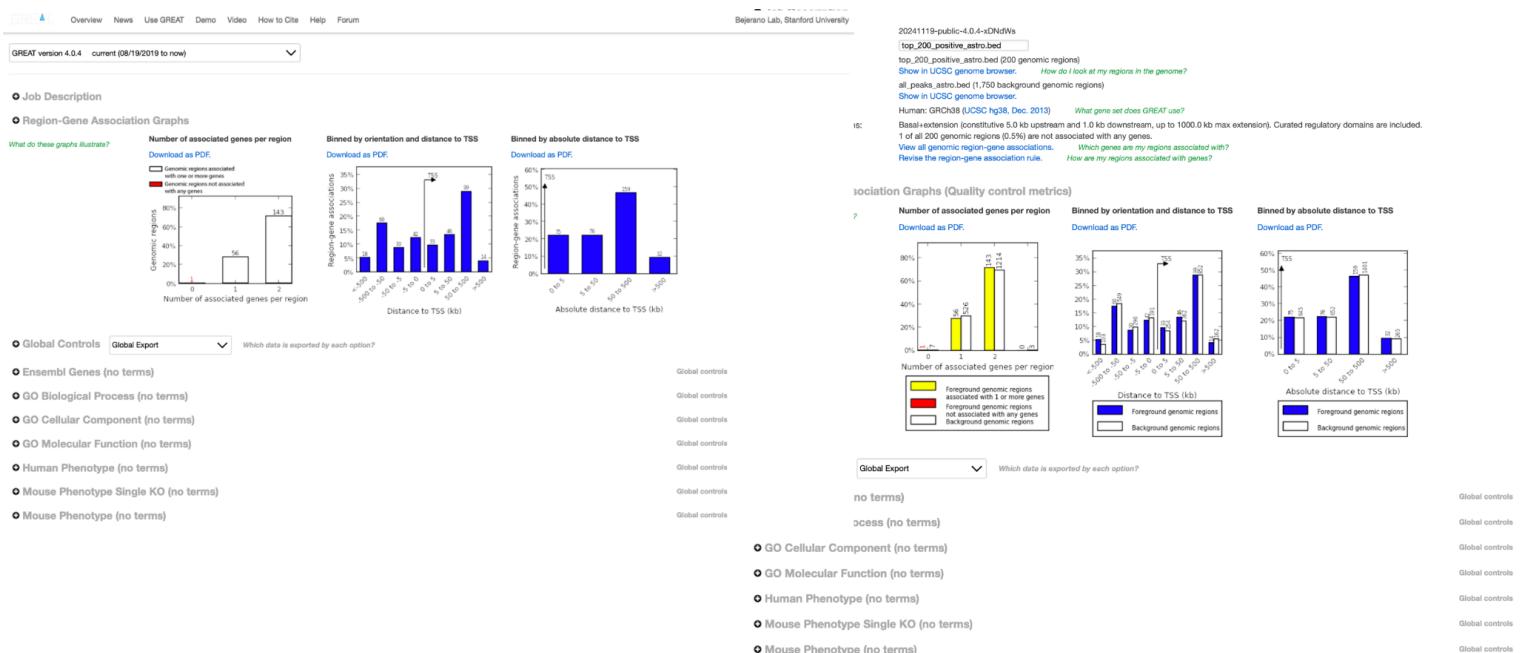


Fig 5: (c) Top 200 positively associated peaks relative to the whole background (d) relative to all peaks

The result of the GO analysis underscores the importance of expanding genomic annotation databases and conducting more in-depth studies into the non-coding genome and its regulatory functions within astrocytes. It brings to light the potential that these

regions could be novel regulators or part of uncharacterized pathways that play roles in circadian rhythm regulation. This aligns with emerging research indicating that astrocytes are not merely supportive cells but have complex, direct roles in neural and circadian processes.

The limitations of this analysis are evident in the reliance on existing GO databases that may not capture the full range of functions, especially for non-coding or lesser-known genomic elements. This emphasizes the need for more targeted studies and potentially integrating multiple data sources to reveal associations that single-platform analyses might miss. Moving forward, a more refined exploration using cell type-specific databases or experimental validation may be necessary to understand the regulatory landscape of astrocytes fully.

In conclusion, while the GO analysis did not reveal significant categories for the astrocyte differential regions, this finding itself is valuable. It suggests that the identified regions may be part of unexplored regulatory networks, warranting further investigation into their roles and contributions to circadian rhythms and astrocyte-specific functions. This outcome highlights the complexity of astrocyte biology and underscores the potential for new discoveries in understanding how these cells contribute to neural regulation and circadian rhythms.

Cell-type specificity:

To address the analysis on cell type-specificity of the chosen genes EDNRA and NR3C2 in astrocytes, I created scatter plots using the Allen Brain Map data. The aim was to visualize the distribution and expression levels of these genes across cortical cell types, focusing on their relevance to astrocytic functions and circadian regulation.

For EDNRA (fig 6a), the scatter plot demonstrates limited and sparse expression across the cortical regions, which may indicate a more selective or region-specific role. This finding aligns with EDNRA's known involvement in vascular function and modulation, suggesting that while it plays a key role, its activity in the brain may be tied to particular zones or under specific conditions.

02-719 Genomics and Epigenetics of the Brain
Circadian Behavior in Mammals
Srashti Bajpai, Mounika Kavuri

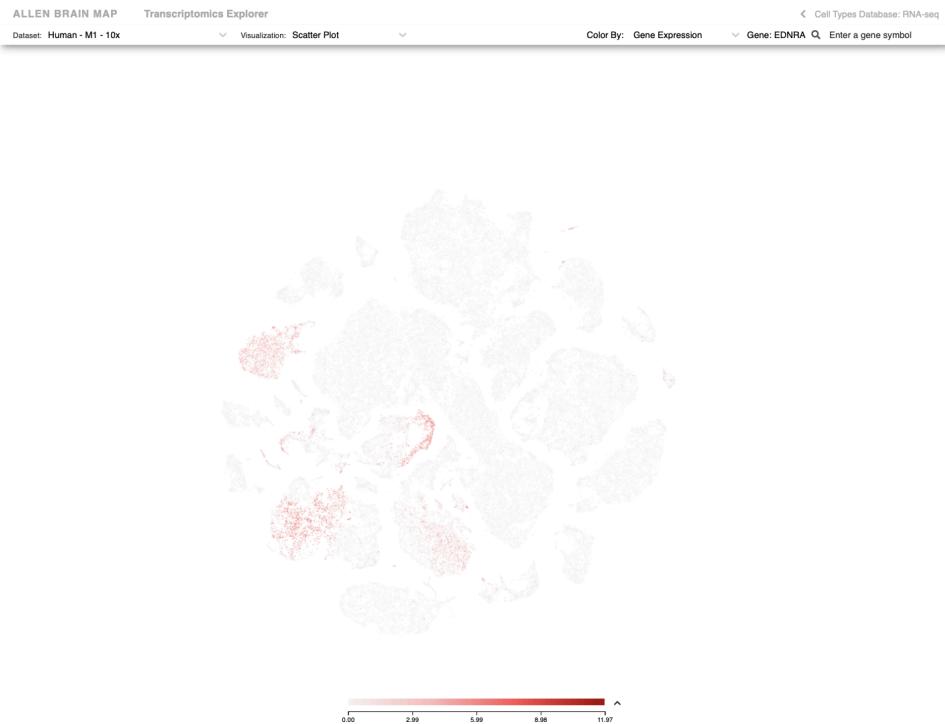


Fig 6a: Cell clusters with EDNRA expression



Fig 6b: Cell clusters with NR3C2 expression

On the other hand, the NR3C2 (fig 6b) scatter plot reveals more widespread and consistent expression across cortical cell types. This pattern is noteworthy given NR3C2's established role in stress response, regulation of the hypothalamic-pituitary-adrenal (HPA) axis, and links to neurogenesis and cognitive functions. The distribution suggests that NR3C2 may contribute broadly to cellular mechanisms involved in circadian regulation and homeostatic balance within astrocytes and related cell types.

These visualizations underscore that while both genes are connected to important regulatory functions in the brain, EDNRA exhibits a more specialized expression potentially tied to specific astrocytic subfunctions, whereas NR3C2 appears more integrated into general cortical processes, potentially influencing circadian rhythm modulation through its interaction with stress and metabolic pathways.

Associated open chromatin regions:

For the final analysis of astrocytes, two associated open chromatin regions corresponding to the genes EDNRA and NR3C2 (fig 7a, 7b) were examined to assess their open chromatin distribution across cortical cell types using CATlas UMAP visualizations. This step aimed to determine whether the evidence for open chromatin was widespread across cell types or specific to a subset, thus potentially indicating a broader regulatory mechanism or a cell type-specific function.

The EDNRA gene UMAP (fig 7a) from CATlas illustrated its chromatin accessibility across various cell populations. The visual representation indicated that EDNRA expression was distributed with a notable frequency in non-neuronal cell populations, highlighting its relevance in potentially broader cortical functions. This suggests that EDNRA may not be confined to a single cell type but may play a role in different regulatory mechanisms involving multiple cell types in the brain cortex.

Similarly, the NR3C2 UMAP (fig 7b) visualization also demonstrated open chromatin accessibility across a wide range of non-neuronal cells. The distribution pattern suggested that NR3C2 is involved in diverse cellular processes, potentially contributing to different aspects of brain function beyond specific cortical subpopulations. The open

chromatin regions related to NR3C2 revealed accessibility patterns that support its involvement in various regulatory pathways within non-neuronal cells.

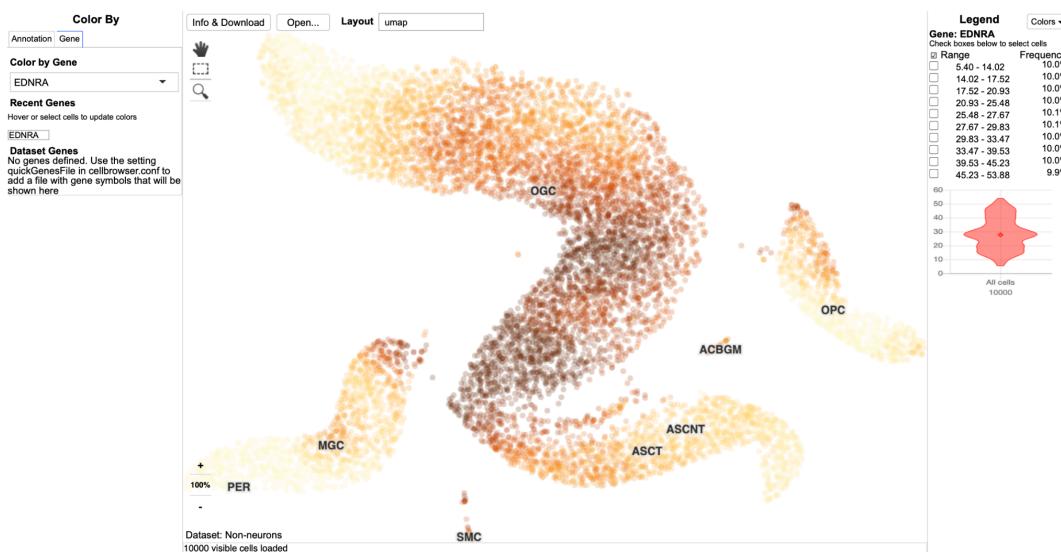


Fig 7a: EDNRA gene UMAP

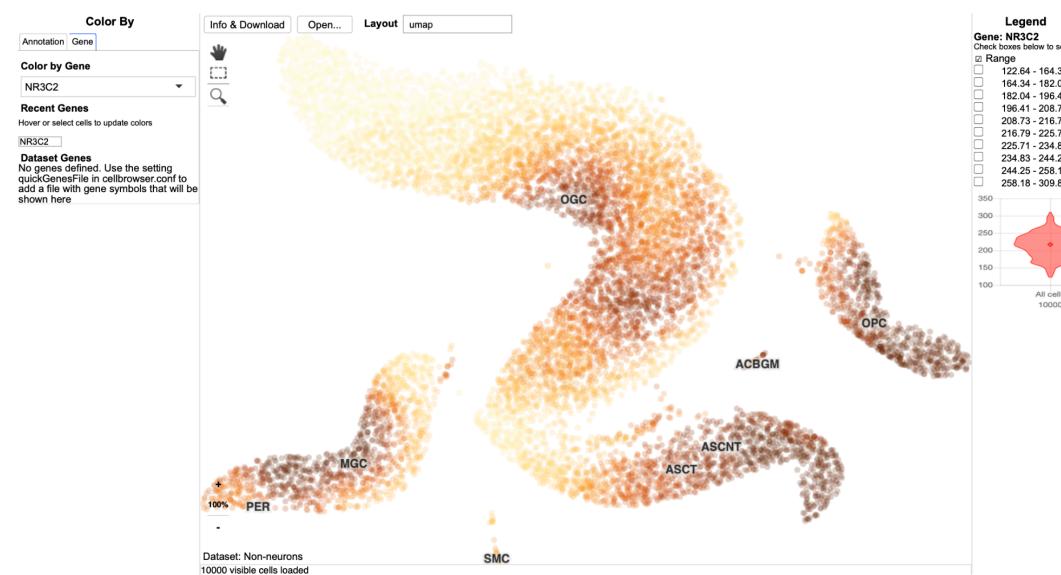


Fig 7b: NR3C2 gene UMAP

Both visualizations indicate that the regulatory regions near EDNRA and NR3C2 are not highly restricted to individual subtypes of cortical cells. Instead, these chromatin accessibility patterns imply that the regulation involving these regions may be part of a broader mechanism impacting multiple cell types. This could suggest their involvement

in more generalized cortical processes or pathways that intersect with multiple functions within the brain's cellular network.

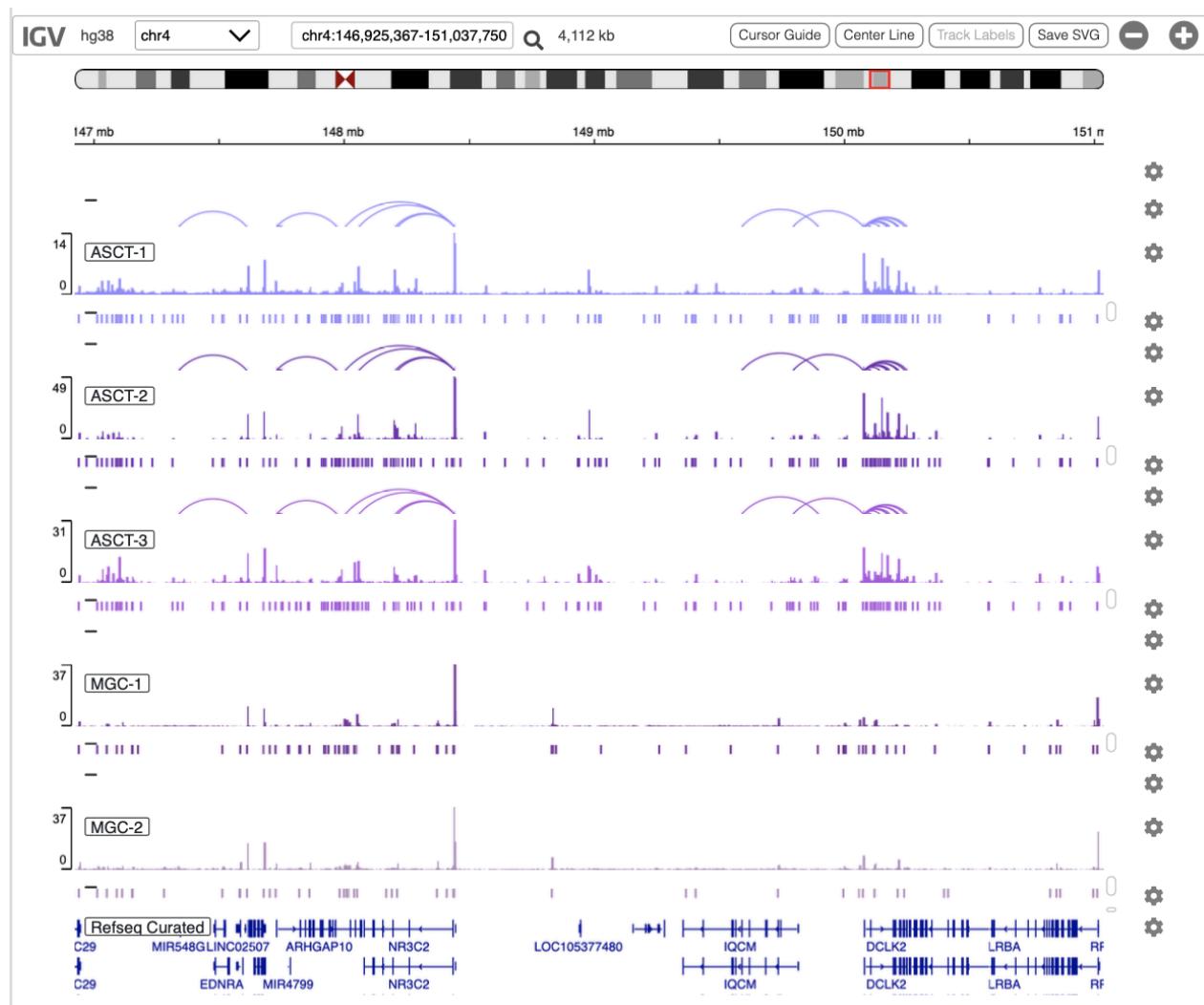


Fig 7c: IGV visualization on chromosome 4

The IGV visualization (Fig 7c) centered on chromosome 4 reveals the open chromatin landscape around the EDNRA and NR3C2 gene promoters across different cell types, specifically astrocyte subtypes (ASCT-1, ASCT-2, ASCT-3) and microglia subtypes (MGC-1, MGC-2). Notably, the astrocyte subtypes (ASCT-1, ASCT-2, and ASCT-3) exhibit prominent peaks at the promoters of EDNRA and NR3C2, indicating higher levels of open chromatin and suggesting that these promoters are more accessible and potentially active in astrocytes compared to microglia. This pattern implies that the transcriptional regulation of EDNRA and NR3C2 may be more prominent in astrocytes, aligning with

their specific functions in these glial cells. On the other hand, the microglia subtypes (MGC-1 and MGC-2) show comparatively lower chromatin accessibility at these sites, suggesting less promoter activity in microglia under the conditions visualized. The differential open chromatin landscape hints at a cell type-specific regulatory mechanism, with astrocytes potentially having a more active transcriptional profile for EDNRA and NR3C2, which could be associated with roles in neural signaling or other astrocyte-specific processes.

2. Cell Type - Microglia

Calculation of p-values and Magnitude of Differences for Microglia:

For the microglia cell type, the distribution of p-values and the correlation estimates provides an insight into the statistical significance and the strength of association for each open chromatin region analyzed across species in relation to the circadian activity trait. The histogram of p-values for microglia (fig 8a) shows that while there is a spread across all significance levels, there are concentrations at the lower end, indicating some regions with potential associations. However, the even distribution suggests that many regions did not meet the threshold for significance. This implies either a limited number of regions with strong associations or statistical limitations due to sample size or data variability.

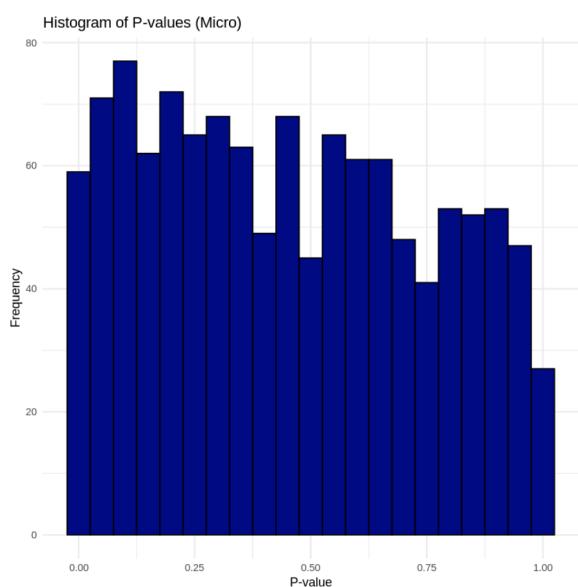


Fig 8a: Histogram of p-values for microglia showing distribution and significance.

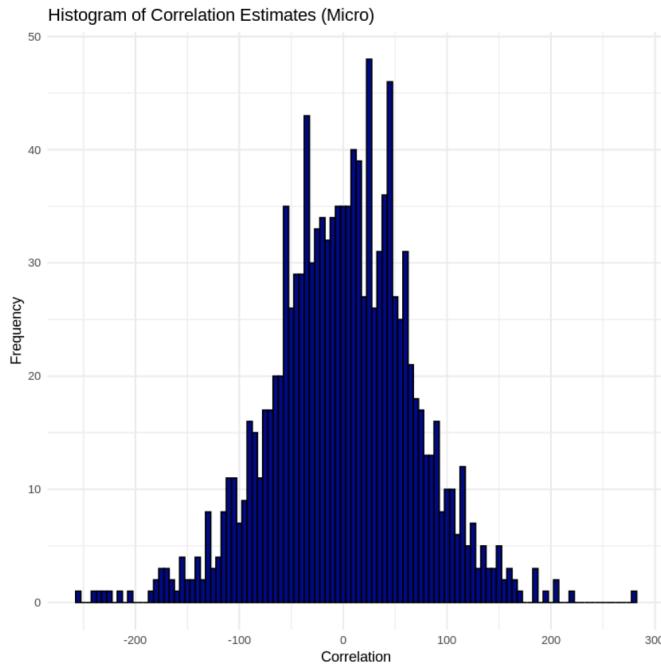


Fig 8b: Histogram of correlation estimates for microglia displaying the direction and magnitude of association.

The correlation histogram (fig 8b) indicates that most correlations cluster around zero, showing that most open chromatin regions have weak or no association with the trait. Notably, the distribution has a peak at or near zero, with fewer instances of strong positive or negative correlations, which is indicative of limited regions showing robust associations.

Multiple hypotheses and most significant peaks:

Upon applying the multiple hypothesis correction, only one region was significantly positively associated with the circadian trait at an adjusted p-value ≤ 0.1 : the peak at "hg38.chr12.56152313.56152814" (p-value: 1.32e-05, adjusted p-value: 0.01595), with a high positive correlation of 219.57 and an R-squared value of 0.38, suggesting a substantial association. This result emphasizes that while a majority of the regions did not pass the threshold for significance after correction, this specific peak stands out with a strong positive relationship to the circadian activity trait.

Number of significantly positively associated regions (Adjusted P ≤ 0.1): 1

Number of significantly negatively associated regions (Adjusted P <= 0.1) : 0

	Peak	P_value	Correlation	R_squared
1	hg38.chr12.56152313.56152814	1.321561e-05	219.5727	0.3755046
	Adjusted_Correlation	Adjusted_P_value		
1	NA	0.01595124		

No significantly negatively associated regions were identified under the same correction threshold, indicating that, in the context of microglia, regions correlating negatively with the trait may be statistically underpowered or absent within the tested dataset.

In the case of microglia most positively and negatively associated peaks:

- The most positively associated peak was identified at hg38.chr12.56152313.56152814, with a low p-value indicating a significant association with increased open chromatin levels.
- The most negatively associated peak was found at hg38.chr1.28248390.28248891, showing a strong negative correlation with the trait, suggesting that open chromatin in this region might be less accessible in species exhibiting certain circadian behaviors.

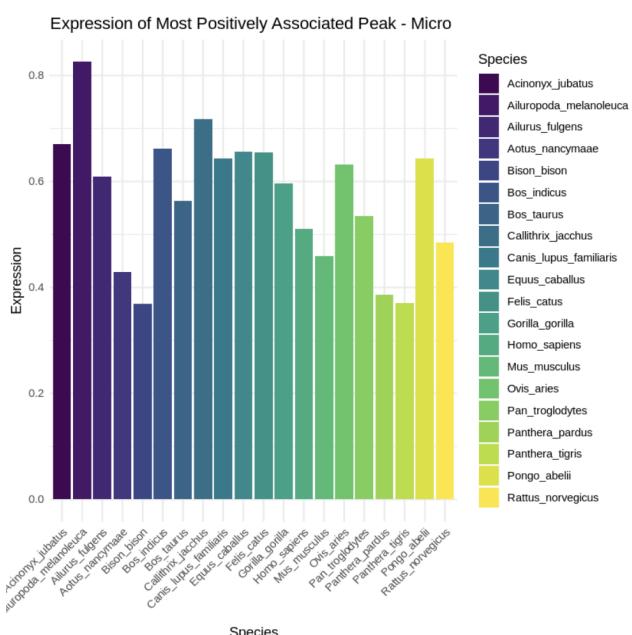


Fig 9a: Predicted open chromatin levels for the most positively associated peak (hg38.chr12.56152313.56152814) across selected mammalian species, illustrating differential chromatin accessibility.

For the most positively associated peak, hg38.chr12.56152313.56152814 (fig 9a), the analysis of chromatin accessibility across species revealed notable differences. Species such as *Ailurus fulgens*, *Bison bison*, *Mus musculus*, *Panthera pardus*, *Panthera tigris*, and *Rattus norvegicus* showed normalized chromatin accessibility values below 0.50, indicating comparatively lower open chromatin levels at this genomic region. These lower levels suggest potential species-specific regulatory mechanisms that might influence gene expression related to circadian rhythms. On the other hand, other species such as *Acinonyx jubatus*, *Bos taurus*, and *Homo sapiens* displayed higher chromatin accessibility levels above 0.50, hinting at more pronounced regulatory activity at this peak that could be associated with circadian regulation.

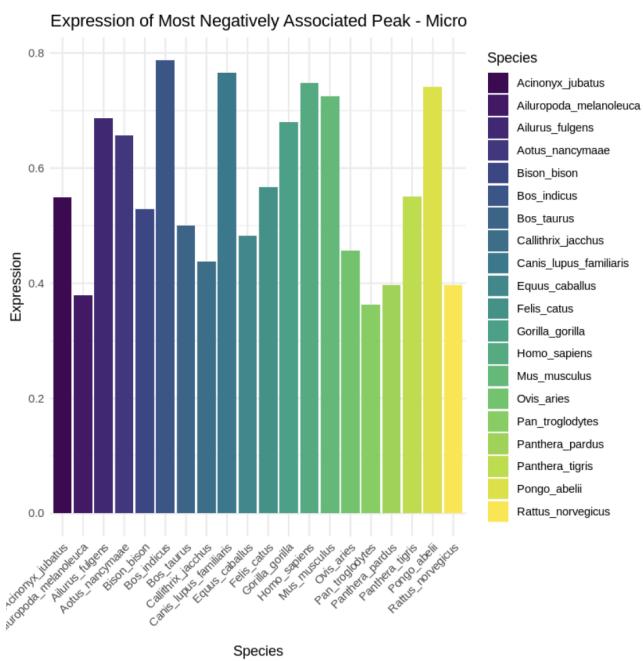


Fig 9b: Predicted open chromatin levels for the most negatively associated peak (hg38.chr1.28248390.28248891) across selected mammalian species, illustrating differential chromatin accessibility.

In contrast, the most negatively associated peak, is hg38.chr1.28248390.28248891 (fig 9b) showed a broader range of chromatin accessibility across the analyzed species. Species including *Ailuropoda melanoleuca*, *Callithrix jacchus*, *Equus caballus*, *Ovis aries*, *Pan troglodytes*, *Panthera pardus*, and *Rattus norvegicus* had accessibility levels below 0.50, indicating less open chromatin and potentially reduced regulatory activity at this site. This suggests a potential repression or lack of association with circadian rhythm regulation in these species. Conversely, species such as *Acinonyx jubatus*, *Bos taurus*,

and *Homo sapiens* exhibited higher chromatin accessibility levels above 0.50, indicating more active regulatory regions at this peak.

These results collectively illustrate that the chromatin accessibility at these peaks varies across species, suggesting that specific regions may play differential roles in the regulation of circadian rhythms within microglial cells. Such variability underscores the complex interplay between genomic regulation and the evolution of circadian mechanisms across mammals.

Nearest protein-coding regions using the UCSC genome browser:

Positively Associated Genes:

MYL6B (Myosin Light Chain 6B): MYL6B encodes a myosin light chain component that is essential for the ATPase motor protein complex, which contributes to muscle contraction and cytoskeletal organization. It is expressed in both slow-twitch skeletal muscle and various non-muscle tissues. Although MYL6B is integral to maintaining structural cellular integrity and movement, there is no direct evidence linking it to circadian rhythms or their regulation (MYL6B).

SMARCC2 (SWI/SNF Related, Matrix Associated, Actin Dependent Regulator of Chromatin Subfamily C Member 2): SMARCC2 is a vital subunit of the SWI/SNF chromatin remodeling complex, which modulates chromatin structure to regulate transcription. Its primary function is to facilitate the assembly and operation of the complex, influencing gene expression at a broad scale. While chromatin remodeling is a fundamental aspect of cellular processes that could potentially affect circadian rhythm regulation, current literature does not show a direct connection between SMARCC2 and circadian functions (SMARCC2).

ESYT1 (Extended Synaptotagmin-1): ESYT1 plays a role in endoplasmic reticulum (ER)-plasma membrane tethering and is involved in lipid transfer and cellular signaling processes. This gene is significant for cellular response mechanisms and signaling pathways that maintain homeostasis. Although ESYT1's role in cellular signaling could

theoretically impact circadian biology by influencing signaling cascades involved in timekeeping, no direct evidence links ESYT1 to circadian rhythm modulation (ESYT1).

IKZF4 (IKAROS Family Zinc Finger 4): Also known as Eos, IKZF4 is a transcription factor primarily associated with immune response regulation and T-cell differentiation. It acts as a DNA-binding protein to modulate gene transcription. While transcription factors like IKZF4 have the potential to influence circadian gene expression by regulating core pathways or genes indirectly, no established research has linked IKZF4 directly to circadian rhythms (IKZF4).



Fig 10a: Genes near positively associated peak

Negatively Associated Genes:

ATP5IF1 (ATP Synthase Inhibitory Factor 1): ATP5IF1 is known for its regulatory role in mitochondrial ATP synthase, where it inhibits the ATP hydrolysis function under conditions of low energy availability. This action is crucial for maintaining cellular energy balance. While mitochondrial activity and energy homeostasis are essential for cellular processes influenced by circadian rhythms, ATP5IF1 itself has not been directly associated with circadian regulation (ATP5IF1).

SESN2 (Sestrin 2): SESN2 is a stress response gene that plays a significant role in autophagy and protection against oxidative stress. It helps in maintaining cellular homeostasis by modulating pathways related to energy and metabolic stress. Although SESN2's role in cellular stress response is vital and could intersect with circadian processes that regulate stress adaptation, literature reviews do not highlight a direct link to circadian rhythm regulation (SESN2).

DNAJC8 (DnaJ Heat Shock Protein Family (Hsp40) Member C8): DNAJC8 is part of the Hsp40 family, known for its function in molecular chaperoning and stress response. It assists in protein folding and the maintenance of protein homeostasis during stress conditions. Although heat shock proteins and related stress responses may interact with circadian mechanisms under specific conditions, no direct association between DNAJC8 and circadian rhythms has been established (DNAJC8).

MED18 (Mediator Complex Subunit 18): MED18 is a component of the mediator complex, which acts as a critical interface between transcriptional regulators and RNA polymerase II, facilitating transcription. Given its essential role in gene expression regulation, MED18 could theoretically be involved in modulating circadian clock-related gene expression; however, no studies have shown a direct relationship between MED18 and circadian rhythm regulation (MED18).

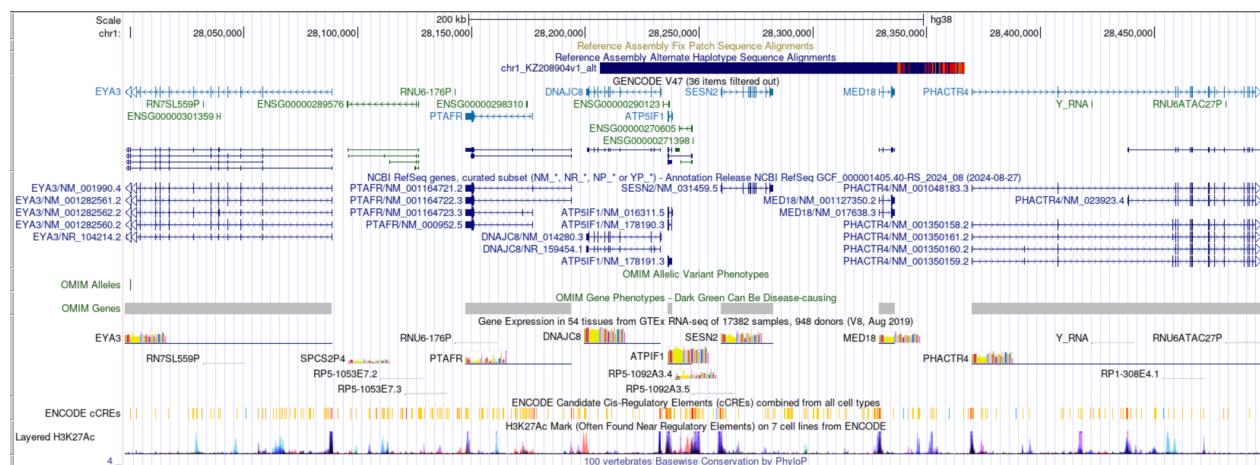


Fig 10b: Genes near negatively associated peak

Gene Ontology analysis using GREAT:

The GREAT gene ontology (GO) analysis for the top 200 most differentially associated regions in microglia, using the whole genome as the background, provided notable insights into potential biological roles linked to chromatin accessibility. When examining positively associated peaks with this broad background, certain GO terms with moderate significance were highlighted. These terms suggested potential biological processes involving gene regulation and cellular responses, although none reached the stringent threshold for highly significant enrichment. This implies that while these positively associated regions are differentially open, their specific functional implications may not align neatly with well-annotated or highly characterized processes in current databases. The presence of these moderately significant terms hints at involvement in complex or multifaceted regulatory networks, possibly related to microglial activation or interaction with other neural cell types in maintaining circadian functions.

For negatively associated peaks, the analysis using the whole genome as a background was more revealing. One significant biological process that emerged was the regulation of apoptotic signaling pathways. This finding suggests that certain chromatin regions with reduced accessibility might play roles in controlling apoptosis, a process critical for maintaining cell homeostasis and protecting neural environments from excessive damage or inflammatory responses. In microglia, which are essential for immune surveillance and neuroprotection, the regulation of apoptosis can be a vital aspect of their function. The association of these chromatin regions with apoptotic pathways could imply that changes in open chromatin states in these areas may affect the microglial capacity to respond to neural damage, potentially influencing the overall circadian-linked processes within the brain.

These findings emphasize the importance of using a comprehensive background like the whole genome, as it allows for the identification of broader biological themes that may be missed when using narrower backgrounds, such as the set of all peaks in the dataset. The positively associated peaks point to general regulatory mechanisms, while the negatively associated peaks, linked to apoptosis regulation, offer more concrete insights into how chromatin states might intersect with microglial function. This suggests that while not all

02-719 Genomics and Epigenetics of the Brain
Circadian Behavior in Mammals
 Srashti Bajpai, Mounika Kavuri

differentially accessible regions are directly linked to well-characterized biological functions, key pathways like apoptosis could be involved in shaping how these cells contribute to neural regulation and possibly circadian rhythm maintenance.



Fig 11: (a) Top 200 negatively associated peaks relative to the whole background (b) relative to all peaks

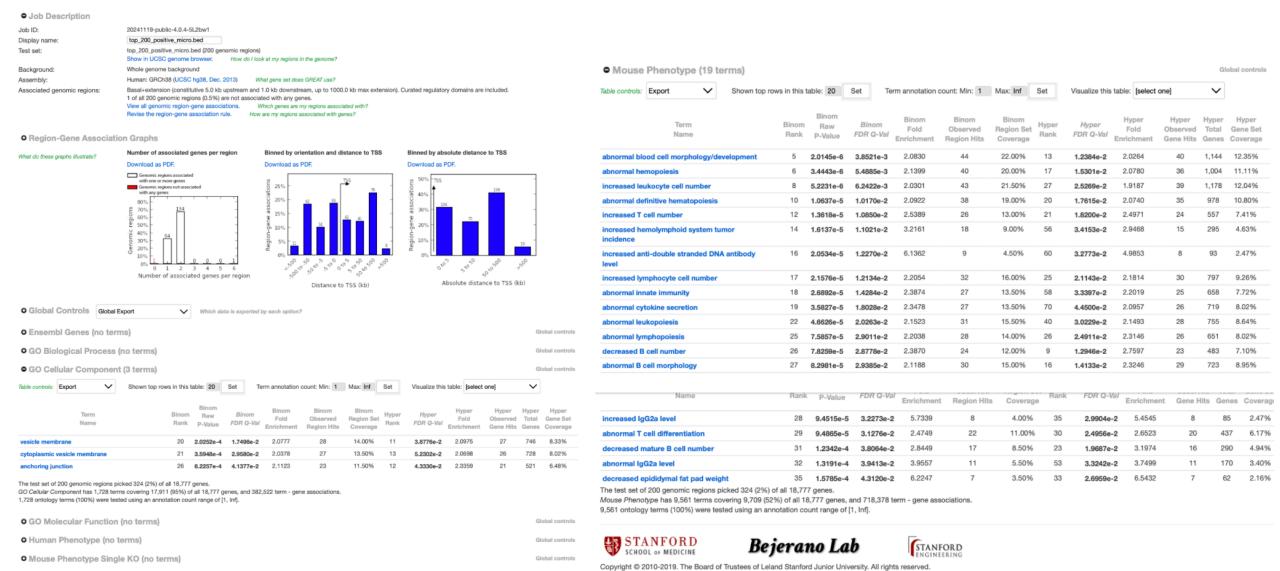


Fig 11: (c) Top 200 positively associated peaks relative to the whole background

Circadian Behavior in Mammals

Srashti Bajpai, Mounika Kavuri

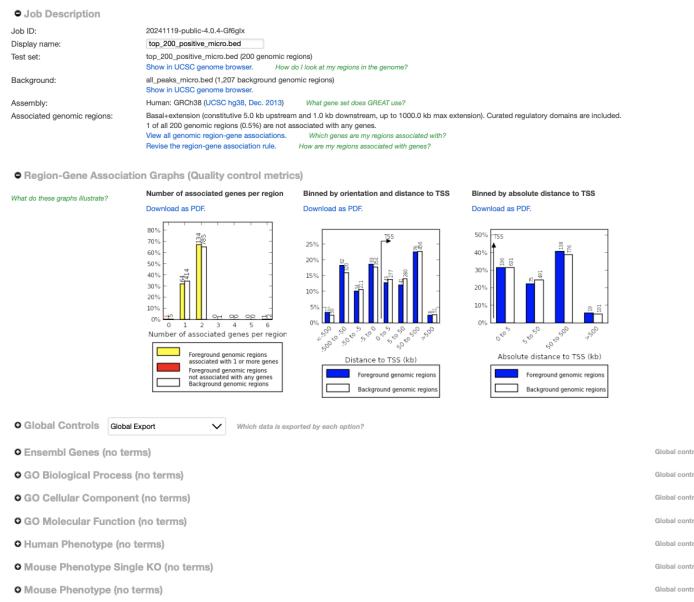


Fig 11: (d) Top 200 positively associated peaks relative to all peaks

Cell-type specificity:

The scatter plots generated using the Allen Brain Map for ESYT1 and MYL6B provide insights into their cell type-specific expression within the human cortex. In these plots, the density and intensity of color indicate the level of gene expression across different cell clusters, suggesting which cell types predominantly express these genes.

For ESYT1 (fig 12a), the scatter plot demonstrates moderate expression levels across the cortical cell types, with certain regions showing more concentrated expression. This could indicate that ESYT1 is involved in functions shared by multiple cell types in the cortex, suggesting a broader role in cortical cellular processes. The overall distribution hints at involvement in cell membrane dynamics or vesicle trafficking, as ESYT1 is known to be involved in membrane contact sites and calcium-dependent lipid binding.

Similarly, the scatter plot for MYL6B (fig 12b) shows a more widespread expression compared to ESYT1. MYL6B, which encodes a myosin light chain involved in muscle contraction and cellular motility, exhibits expression that spans multiple cortical regions, implying a potential involvement in cytoskeletal organization or cell migration. Its

expression pattern aligns with roles in microglial function, where cytoskeletal rearrangements are crucial for their movement and response to environmental cues.

Both genes' expression profiles in the scatter plots suggest they are not restricted to microglia alone but are expressed in various cell types in the cortex. This could imply that while they are associated with peaks of interest found in microglial data, their functions might extend to other cortical cells, supporting the possibility of a more generalized cellular mechanism in cortical function rather than one strictly confined to microglia.

ALLEN BRAIN MAP Transcriptomics Explorer
Dataset: Human - M1 - 10x Visualization: Scatter Plot Color By: Gene Expression Gene: ESYT1 Enter a gene symbol



Fig 12a: Cell clusters with *ESYT1* expression

ALLEN BRAIN MAP Transcriptomics Explorer
Dataset: Human - M1 - 10x Visualization: Scatter Plot Color By: Gene Expression Gene: MYL6B Enter a gene symbol

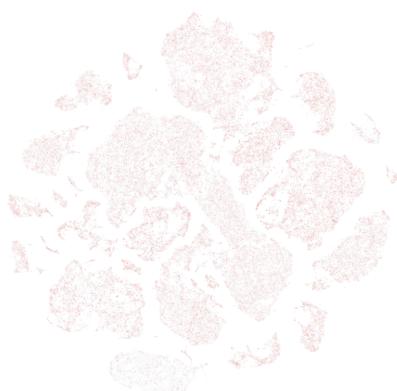


Fig 12b: Cell clusters with *MYL6B* expression

Associated open chromatin regions:

In analyzing the open chromatin regions associated with the genes ESYT1 and MYL6B, their chromatin accessibility was visualized across various cortical cell types using Catlas. The results provide insight into whether the chromatin regions related to these genes support a broader mechanistic role across different cells or are specific to a particular subset.

For ESYT1 (fig 13a), the chromatin accessibility appears consistently across multiple cell types, showing moderate to high levels in astrocytes (ASCT), microglia (MGC), oligodendrocyte precursor cells (OPC), and other glial populations. This uniform distribution indicates that the regulatory region associated with ESYT1 is accessible across a diverse array of cortical cells, suggesting a broader role in maintaining essential functions shared among non-neuronal cells.

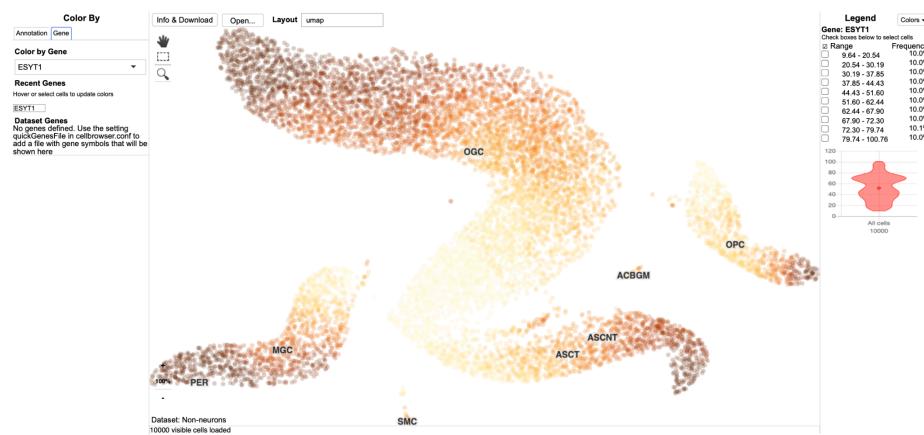
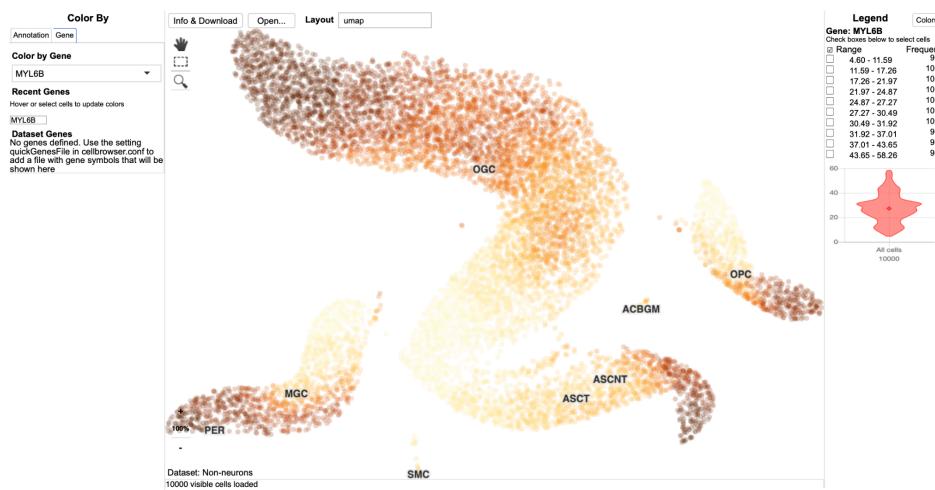


Fig 13a: ESYT1 gene UMAP

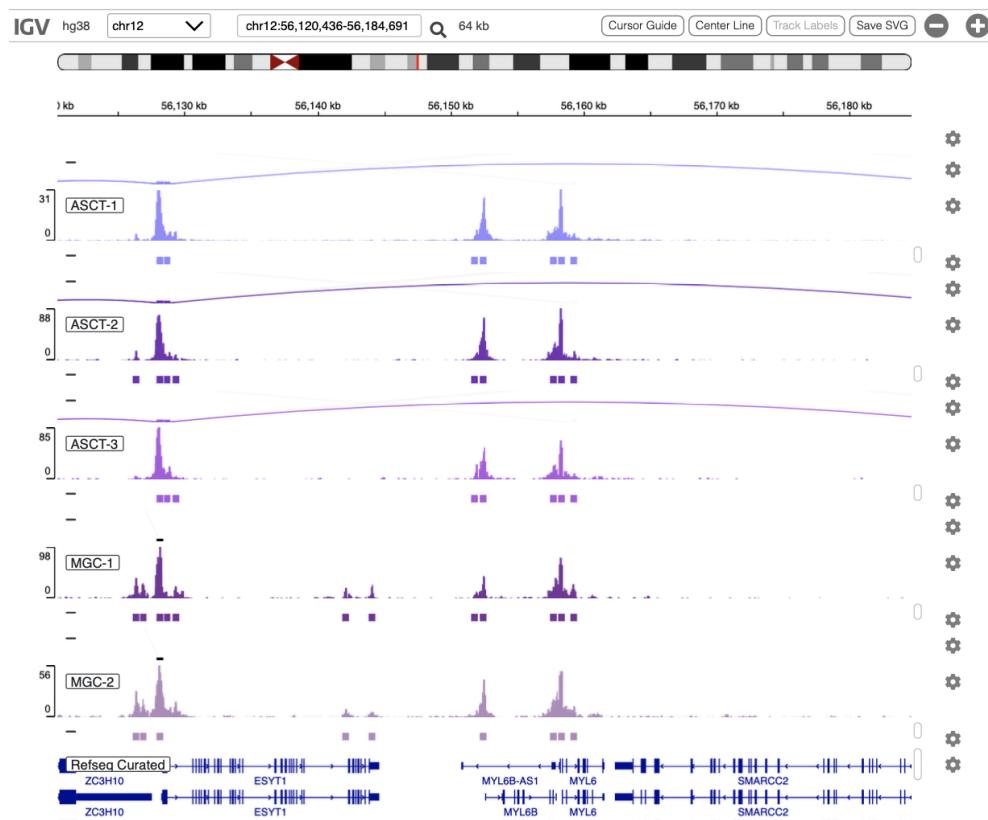
Similarly, for MYL6B (fig 13b), the chromatin landscape was visualized, demonstrating accessibility in various cortical cell types, including microglia and oligodendrocyte cells. The uniform signal across these populations, though with slight variations in intensity, supports the hypothesis that the regulatory functions of this region could extend across multiple cell types. The accessibility being present in microglial cells and glial populations highlights a potential broader mechanism for transcriptional regulation that is not exclusive to a single cell type but rather integrated into shared pathways involved in cellular maintenance and response.

Circadian Behavior in Mammals

Srashti Bajpai, Mounika Kavuri

**Fig 13b: MYL6B gene UMAP**

The Catlas visualizations for both genes, therefore, point toward chromatin regions that are integral to general cell regulatory mechanisms in the cortex, rather than being specific to only microglial functions. This indicates that while these genes may have unique functions, their open chromatin regions are part of broader regulatory elements that facilitate transcription across a range of related cell types.

**Fig 13c: IGV visualization on chromosome 12**

The peaks in the tracks indicate areas of open chromatin, which suggest regions accessible for transcriptional machinery binding. In the ASCT-1, ASCT-2, and ASCT-3 tracks, significant peaks are present, suggesting a high level of chromatin accessibility within the astrocyte subtypes at specific sites near the ESYT1 and MYL6B genes. This indicates that these regions may play an active regulatory role in transcription for astrocytes.

Similarly, in the MGC-1 and MGC-2 tracks, there are notable peaks, though they may be less pronounced compared to the astrocyte subtypes. This reflects that the chromatin around the ESYT1 and MYL6B genes is also accessible in microglial cells but possibly to a lesser degree. The presence of these peaks suggests that while both astrocytes and microglia may share regulatory elements in this region, the expression and regulatory importance could differ between the cell types.

The genomic view also shows that ESYT1 and MYL6B have neighboring non-coding RNAs, such as MYL6B-AS1, which may contribute to regulatory complexity. The shared and distinct chromatin accessibility between the ASCT and MGC tracks implies that these open chromatin regions potentially play a role in regulatory mechanisms pertinent to these cell types, supporting both common and cell type-specific gene expression patterns. This chromatin accessibility indicates an involvement in broader cellular processes shared by glial cells, while still allowing for unique functions between astrocytes and microglia.

Comparing results for astrocytes and microglia:

The comparison of associated open chromatin peaks between astrocytes and microglia reveals distinct regulatory dynamics, with astrocytes exhibiting a more extensive and diverse involvement in circadian regulation compared to microglia. Astrocytes had a greater number of significant peaks after multiple hypothesis correction, with two positively and three negatively associated peaks identified (Adjusted P ≤ 0.1). In contrast, microglia displayed only one significantly positively associated peak, with no significant negatively associated peaks, suggesting a more limited role in circadian-related chromatin accessibility. The distribution of p-values and correlation estimates further

underscores these differences; astrocytes showed a clustering of p-values near zero and a broader range of strong positive and negative correlations, indicating robust associations with circadian traits. Microglia, on the other hand, exhibited a more even distribution of p-values and correlations tightly centered around zero, reflecting weaker and less frequent associations. Visualizations of chromatin accessibility through IGV and scatter plots revealed that astrocyte subtypes demonstrated pronounced peaks and accessibility at key regulatory regions, aligning with their established roles in circadian signaling, such as the expression of clock genes and gliotransmitter regulation. In microglia, chromatin accessibility at significant peaks was less pronounced and more restricted, suggesting a secondary or indirect role in circadian regulation, potentially through neuroinflammation and synaptic remodeling. These findings highlight the evolutionary divergence in the functions of these cell types, with astrocytes playing a conserved and central role in circadian regulation across species, while microglia exhibit a more specialized and context-dependent involvement. This distinction underscores the importance of cell type-specific analyses in uncovering the complexity of circadian regulatory mechanisms.

References:

- Hastings, M. H., Reddy, A. B., & Maywood, E. S. (2018). The mammalian circadian timing system and the suprachiasmatic nucleus as its pacemaker. *Nature Reviews Neuroscience*, 19(8), 453–469.
- Hut, R. A., Kronfeld-Schor, N., van der Vinne, V., & De la Iglesia, H. (2012). In search of a temporal niche: Environmental factors. *Progress in Brain Research*, 199, 281–304.
- Reppert, S. M., & Weaver, D. R. (2002). Coordination of circadian timing in mammals. *Nature*, 418(6901), 935–941.
- Cuesta, M., Clesse, D., Pévet, P., & Challet, E. (2009). From daily behavior to hormonal and neurotransmitters rhythms: comparison between diurnal and nocturnal rat species. *Hormones and behavior*, 55(2), 338–347.
<https://doi.org/10.1016/j.yhbeh.2008.10.015>

- Takahashi, J. S. (2017). Transcriptional architecture of the mammalian circadian clock. *Nature Reviews Genetics*, 18(3), 164–179.
- Casares-Hidalgo, C., Pérez-Ramos, A., Forner-Gumbau, M., Pastor, F. J., & Figueirido, B. (2019). Taking a look into the orbit of mammalian carnivorans. *Journal of Anatomy*, 234(5), 622-636.
- Brancaccio, M., Patton, A. P., Chesham, J. E., Maywood, E. S., & Hastings, M. H. (2017). Astrocytes control circadian timekeeping in the suprachiasmatic nucleus via glutamatergic signaling. *Neuron*, 93(6), 1420-1435.
- Jiao, H., Kalsbeek, A., & Yi, C. X. (2024). Microglia, circadian rhythm and lifestyle factors. *Neuropharmacology*, 110029.
- Aloisi, F. (2001). Immune function of microglia. *Glia*, 36(2), 165-179

Other References:

- UCSC genome browser
- GREAT
- Allen Brain Map
- CATlas
- ChatGPT