A novel bioinformatics approach to reveal the role of circadian oscillations in AD development

Zhiwei Ji¹, Dazhi Shang³, Pora Kim¹, Mengyuan Yang¹, Sijia Wu¹, Weiling Zhao¹, Eunju Kim², Marvin Wirianto²,
Zheng Chen², Seung-Hee Yoo^{2,*}, Xiaobo Zhou^{1,*}

¹School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

²Department of Biochemistry and Molecular Biology, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

³West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China

Corresponding Author:

Xiaobo Zhou (Xiaobo.Zhou@uth.tmc.edu), Seung-Hee Yoo (Seung-Hee.Yoo@uth.tmc.edu).

Abstract

Altered circadian gene expression may contribute to Alzheimer's disease (AD) progression. Unfortunately, sampling the central nervous system (CNS) at multiple time points is not feasible. Moreover, there are no AD-related time-series transcriptome datasets available for studying these circadian patterns and their impacts on AD development. In this study, we introduce a novel computational platform, Event-driven Sample Ordering for Circadian Variation Detection (ESOCVD), to reveal rhythmic patterns of gene expression of AD using untimed transcriptome datasets. ESOCVD was applied to 20 untimed gene expression profiles of 16 brain regions from approximately 3000 AD patients in public transcriptome databases. Our analysis revealed five types of circadian alteration patterns in ~2,000 circadian genes in different brain regions of AD patients. Further analyses of additional databases confirmed that our analytical platform can be applied to identify the evolutionary dynamics of circadian variation during the process of AD development. Through the gene expression correlation analysis for our 8 circadian genes identified from AMP-AD MSBB cohorts, we identified stage-specifically enriched biological processes with anticipated context. Gene expression analysis of AD mouse brain tissues further substantiated the predictions of the ESOCVD model. In summary, ESOCVD is highly versatile in bridging circadian research and precision medicine.

Introduction

The circadian clock is a complex biological machine that allows organisms, from cyanobacteria, fruit flies to humans, to predict and prepare for the challenges of daily life¹. The body's clock system maintains 24-h rhythms in physiological functions, including the sleep-wake cycle, and synchronizes them to the light-dark cycle². The clock system is composed of cell-autonomous molecular oscillators^{3,4}, and the molecular oscillator, in turn, consists of conserved transcriptional and translational feedback loops of 'core clock genes'^{5,6}. These clock genes drive circadian oscillations of so-called clock-controlled genes (CCGs) at the cellular level to regulate various cellular and systemic functions^{7,8}. Disturbances of the circadian rhythms have long been associated with many neurological and psychiatric diseases^{9,10}, including Alzheimer's disease (AD)^{1,11}, Parkinson's disease (PD)^{12,13}, and Huntington's disease (HD)^{14,15}. Therefore, mechanistic understanding of the interactions between circadian biology and neurodegenerative pathophysiology may provide new therapeutic opportunities against neurodegenerative diseases.

Alzheimer's disease affects 20-40 million people worldwide and is the most common cause of progressive cognitive dysfunction in adults. Circadian dysfunction phenotypes are observed at the early stage of AD¹. *Drosophila* with overexpressed amyloid- β (A β) peptides also show a severely disturbed circadian function¹⁶. AD mouse models show accelerated disintegration of circadian rhythms with age as compared to normal counterparts¹⁷⁻²¹. Several studies in transgenic mice demonstrate that amyloid deposition in the brain leads to the disruption of normal sleep architecture. Abnormalities in mice include a loss of robust day-night oscillation in brain lactate and altered nocturnal activity level^{17,20,21}. Furthermore, ~80% of AD patients over 65 years old suffer from circadian rhythm disorders, such as disturbances in thermoregulation and sleep-wake cycles^{22,23}. AD patients exhibit disturbances in the timing and duration of the sleep cycle, primarily manifested as increased wakefulness at night and increased sleep during the day, which can progress to a loss of day-night variation¹¹.

Although circadian dysfunction has long been considered as a prevailing symptom of AD, it has more recently been hypothesized as a contributor to the pathogenesis of AD²⁴⁻²⁶. New findings in the fruit fly and rodent models indicate that the deletion of the clock gene *Bmal1*, *Clock*, *Per1*, and *Cry1* may cause an accelerated aging phenotype characterized by an earlier decline in cognitive functions^{5,25,27}. Recent experiments found that sleep deprivation caused a striking increase in the A β plaque burden in mice that express AD-associated mutant forms of human amyloid precursor protein (*APP*) and develop A β plaques with age^{28,29}. A subsequent study using mice that express human *APP*, *PS1*, and human *Tau* transgenes also demonstrated an increase in cortical A β plaque burden following chronic sleep deprivation³⁰. Moreover, human data also support a role for circadian disruption in AD pathogenesis. An epidemiologic study of daily activity from over 1,200 initially cognitively-normal older women demonstrated that diminished circadian rhythm amplitude, robustness, or phase delay were associated with increased risk of developing dementia during the 5-year follow-up period³¹. However, it remains unclear whether or how the circadian system influences AD pathogenesis. Therefore, a better understanding of how circadian dysfunction contributes to neurodegenerative disease mechanisms will facilitate the development of novel therapies for AD.

The abnormalities in circadian-ordered gene expression are linked to an array of downstream pathobiology of AD, including oxidative stress, misfolded proteins, and epigenetic burden^{32,33}. Identifying the dysregulation in the oscillatory gene expression patterns from the existing transcriptome datasets will provide novel mechanistic insights into the complex etiology of AD. Unfortunately, there are few public transcriptome datasets with multiple "timed" control human samples to study such patterns, and none available for AD. However, it has been shown that compiling a collection of single "untimed" gene

expression samples can generate a snapshot of circadian gene expression. This can be achieved by analyzing the existing large transcriptome datasets with computational approaches/models.

In recent years, several computational algorithms have been developed to detect oscillatory patterns of gene expression in health and disease, e.g., supervised learning (ZeitZeiger³⁴, JTK_CYCLE³⁵, TimeSignature³⁶, etc.) and unsupervised learning (Oscope³⁷, and CYCLOPS³⁸, etc.). The limitation of the first three approaches is that they only work on timed gene expression data. Oscope is designed to extract the oscillatory dynamics of gene groups from single-cell RNA-seq data. However, it requires much more computation and is sensitive to the ordering of cells. CYCLOPS can estimate the circadian phase of each sample with untimed gene expression profiles. However, the requirement for a large sample population to fill the entire periodic cycle is the key limitation. Therefore, developing efficient, robust, and data-insensitive algorithms for circadian variation detection on untimed transcriptomic datasets remains a big challenge.

In this study, we proposed a computational approach, Event-driven Sample Ordering for Circadian Variation Detection (ESOCVD), to decipher the oscillatory patterns of gene expression in AD using “untimed” transcriptome datasets. ESOCVD extracts the global pacemakers of the clock system to predict the circadian phase of each sample by modeling time-dependent events with the Bayesian approach. Based on the inferred circadian phases, we tracked five types of circadian variation in gene expression during AD progression. To verify the effectiveness, we applied the ESOCVD model to 20 gene expression profiles collected from public databases AMP-AD³⁹, and GEO (Gene Expression Omnibus), which cover 16 brain regions of ~3000 AD patients. We mainly focused on 1691 seed genes and 857 potential brain-related circadian genes (PBCG) (Supplementary Fig. 1). The outcomes of ESOCVD revealed the heterogeneity of the occurrence of circadian variations in different brain regions of AD patients. Functional analysis indicates that there are five regions associated with circadian rhythm regulation, including superior frontal gyrus (SFG), superior temporal gyrus (STG), etc. The estimated circadian rhythms of well-known clock genes and circadian genes under normal conditions are consistent with these reported in Circadian Gene Database (CGDB)⁴⁰. Moreover, we checked 21 AD-related clock-controlled genes^{25,41} and found that loss or gain of rhythmicity in multiple-regions. In addition, we developed a systems biology approach to reconstruct gene regulatory network (GRN) connecting AD development and circadian rhythm processes. We successfully identified three signaling pathways for brain-region SFG, Visual cortex (VC), and STG, respectively. Finally, we tracked the evolution of the cycling rhythm of circadian genes during the process of AD development. Through studying 8 circadian genes in AMP-AD MSBB cohorts⁴², we identified 16 AD marker SNPs and 85 correlated genes that were significantly associated with these genes in disease stages. These genes showed stage-specifically enriched biological processes with the anticipated context of disease stages. More data sources provide evidence that 8 circadian genes act as intermediators in AD pathogenesis.

Results

The datasets for model validation

In this study, we collected untimed transcriptomic data from existing datasets to validate our ESOCVD approach. These gene expression profiles are related to 16 regions of the brain from 2930 subjects with or without dementia (Supplementary Table 1). The samples from GSE125583, GSE5281, GSE15222, GSE44772, and GSE29378 were labeled as control and AD. Both MSBB and GSE131617 include patients from early to late stage of AD. All the subjects involved in GSE131617 were scored using Braak staging⁴³.

We defined the samples with stage “0” as control, “1-2” as early, “3-4” as mid, and “5-6” as late. The samples in MSBB are grouped based on cognitive status (Clinical Dementia Rating scores, CDR⁴⁴). Thus, we assigned the samples with CDR score “0” as control, “0.5” and “1” as early, “2” as mid, and “3-5” as late. Here, ESOCVD was used to detect the circadian alteration in any stage of AD relative to control.

Eatimation of circadian phases using ESOCVD

For the time series transcriptome, circadian patterns can be easily detected according to the series of sampling time^{35,45-47}. However, the order of samples in an untimed dataset is unknown. ESOCVD optimally combines eigengene patterns to reveal an underlying circadian phase of each sample in a single cycle, and then uses the inferred phases to order the data (Fig. 1). Each eigengene is defined as an oscillating biomarker and evolves following a circadian pattern (see the details in “**Methods**”). Furthermore, the expected expression distribution of an eigengene can be determined based on the relative positions of ‘crest’ and ‘tough’ in a single cycle. Therefore, we defined an event sequence **S**, denoting the occurrence of the ‘crest’ and ‘tough’ (z_{i1} or z_{i2}) for all the eigengenes. ESOCVD quickly searches the optimal sequence **S** by maximizing the likelihood of the expression data. In this study, we chose the number of eigengenes (indexed as $i = 1, \dots, I$) is 12 and the distance between any pair of z_{i1} and z_{i2} is half cycle (12 hours). In other word, each event in **S** denotes 1 hour. The 24 events separate the time axis to 25 progression stages. After maximizing the likelihood in **Eq. (1)**, we obtained the optimal sequence **S**. In the meantime, we got a $N * 25$ matrix pk , which represents the probability of each subject j at each stage ($1 \leq j \leq N$). The subject j is assigned to stage k ($0 \leq k \leq 24$) with maximal probability (**Eq. (2)**). Once all the samples were assigned to corresponding phases in a single cycle, the circadian rhythmicity can be easily detected by using Cosinor model⁴⁸.

We examined GSE125583 as an example. Supplementary Fig. 2 presents the inferred distribution of all the samples across the stages. We found that control samples distributed evenly in a single cycle. However, AD samples mainly distributed in the 1st and 4th quarters. Similarly, the circadian phases of all the samples associated with each data profile were estimated (Supplementary File1).

Detection of the circadian variations of gene expression

Based on the inferred order of samples, we detected the circadian variations of gene expression between control and AD groups with cosinor regression⁴⁸. ESOCVD assigns each sample a phase (stage) along a single cycle ($0 \leq \theta \leq 24h$). Given the ESOCVD-ordered expression data of control or AD samples, we used cosinor regression to fit the expressions of each gene in each group and identified the genes that have significantly different rhythmic patterns between AD and control samples ($P\text{-value} < 0.05$).

Then, we examined five types of circadian variations across the genome: 1) amplitude change; 2) loss of rhythmicity (LR); 3) gain of rhythmicity (GR); 4) Phase shift; 5) Base shift (see the details in “**Methods**”). For each data profile, we mainly focused on two groups of genes (Supplementary Fig. 1) and then examined the gene expression with circadian variations. The first group has 1691 seed genes, including 1670 circadian genes collected from mouse, and 21 AD-related clock-controlled genes. In addition, we searched GO database⁴⁹ and selected 15 core clock genes, which are also included in the seed gene list. The second group includes 631 brain-related circadian genes reported in CGDB⁴⁰, 288 aging-related circadian genes^{45-47,50,51}, and 34 AD risk factors. Supplementary Fig. 1 shows that 700 PBCG genes do not belong to the seed gene list. The details of seed genes and PBCG genes were described in Supplementary File2.

Fig. 2 shows the numbers of seed genes with circadian variation predicted by ESOCVD in 16 brain regions. We found that loss or gain of rhythmicity occurred in multiple regions, while amplitude change and phase shift only appear in the 1 or 2 areas. There are more genes lose the circadian rhythm in SFG, STG, and DPC (dorsolateral prefrontal cortex), etc. In Fig. 2a, the first 16 columns are associated with the AD patients from United States, while the last 3 columns (marked with “*”) are relevant to Japanese patients. It seems that there are more LR genes and fewer GR genes in FC (Frontal cortex) and EC (Entorhinal cortex) regions of Japanese patients, which is opposite to the situations on American patients. Moreover, Fig. 2b shows the periodic rhythms of 10 core clock genes in multiple brain regions of control samples. We found that the circadian phases of “peak” and “trough” for all the clock genes are very close to the evidence reported in the database CGDB, which indicates that our approach is reliable (Supplementary Table 2).

We further explored the circadian patterns of 21 AD-related clock-controlled genes across all the brain regions (Fig. 3a). Eighteen genes show circadian alterations in some brain regions, while 3 genes displayed no changes. Most of them occurred LR or GR in different regions (Supplementary Table 3). Fig. 3b presents 6 selected AD-related clock-controlled genes, which are lost the rhythmicity in a particular region. Moreover, we performed qPCR experiments using cortex from APP/PS1 model mice to validate the cycling patterns of key circadian genes identified by ESOCVD. Fig. 3c indicates that the observed trends of diurnal expression of these genes in wide-type (WT) mice are highly consistent with our prediction of the control group, and the circadian rhythm detected in WT is markedly diminished in the AD group.

Similarly, we summarized the circadian changes of 700 PBCG genes, as shown in Fig. 4a. The distributions of LR and GR are very similar to that in seed genes (Fig. 4a, Fig. 2a). The other three types of patterns appear in only 2-3 regions. Fig. 4b shows 15 selected circadian genes detected by ESOCVD, which are representative samples to elucidate five types of patterns. The genes *VEGFD*, *ALKBH3*, and *MRPL41* in the first row showed amplitude change in the AD group. Particularly, the angiogenesis mediator *VEGFD* has a twofold reduction in amplitude in AD samples, which is consistent with Bridel’s findings that *VEGFD* reveals a tendency towards a decrease with AD severity⁵². *PSMD5*, *KBTD11*, and *SDR39U1* lost rhythmicity in AD, hinting at previous findings that deletion/disruption of these genes may cause behavioral abnormalities in human^{53,54}. Moreover, *ENO3*, *SMAD5*, and *SUGT1* gained rhythmicity in AD samples relative to controls, demonstrating the ability of *ENO3* to rescue neuronal loss⁵⁵. In addition, *DNAJB5*, *SMPD4*, and *ATP13A5* exhibited a clear phase shift in the AD group compared with the control group. Finally, *OMG*, *FRMD4A*, and *CYTB* displayed baseline shift, indicating that the mean values of gene expression are significantly reduced in AD samples. The details of all the genes with circadian variation are described in Supplementary File3.

Pathway analysis of circadian changes in AD samples

After obtaining the signatures with the circadian changes, we conducted pathway-level analysis using bioinformatics platform DAVID⁵⁶ to assess whether these inferred genes are associated with important pathways or biological processes. The top 5 significant brain regions include superior frontal gyrus (SFG), superior temporal gyrus (STG), frontal cortex (FC), temporal cortex (TC), and posterior cingulate (PC). KEGG and GO biological process sets that were overrepresented among those genes that change rhythmicity in AD patients are collected, e.g. microglial cell activation^{57,58}, circadian rhythm, circadian regulation of gene expression, neuron apoptotic process, and insulin signaling pathways^{59,60}, etc. The details of DAVID analysis are shown in Supplementary Tables 4-8.

Exploration of the associations between AD development and circadian rhythm at the network level

Since the molecular mechanism of how circadian dysfunction links AD development remains poorly defined, we developed a systems biology approach to reveal the association between AD development and circadian variation at the network level (Supplementary Fig. 3). Our approach fully utilizes the upstream (cofactor) and downstream (target gene, TG) regulation information around transcription factor (TF) and integrates the functional linkage data and gene expression. By systematically mining the consistency of regulatory relationships, our model infers all the modulatory triplets (cofactor-TF-TG) and constructs gene regulatory network (see the details in “**Methods**”).

Our findings show that 50 AD-risk factors with circadian changes were detected by ESOCVD, e.g., *ABCB1*, *CLU*, *BIN1*, and *TREM2*, etc. (Supplementary Table 9). Based on the information for all possible TF-TG pairs in TRRUST⁶¹, we firstly screened out 2 pairwise transcriptional regulations: 1) *FOS* —| *CLU*^{62,63}; 2) *BIN1*→*MYC*⁶⁴. The analysis on GSE5281 indicates that both *CLU* and its transcription factor *FOS* occur loss of rhythm in SFG (Fig. 4c). *MYC* is the unique target gene of *BIN1*⁶⁵. The transcription factor *BIN1* appears heterogeneity of circadian patterns in VC and STG. Secondly, we identified cofactor-TF regulatory relationships based on the PPI information in STRING⁶⁶. Thirdly, we integrated upstream and downstream regulations of TF in a GRN network. Supplementary Fig. 4 presents the generic GRN pathway maps for *CLU* and *BIN1*. Finally, we inferred the region-specific GRN network by integrating our TILP⁶⁷ approach and gene expression data. Fig. 4c presents the inferred *FOS* pathways connecting circadian rhythm to AD development. The *BIN1*-related pathways are presented in Supplementary Fig. 5. Our analysis shows that the average expressions of *MAPK1* and *FOS* are down-regulated and *CLU* is up-regulated in AD (Supplementary Table 10). Loss of rhythmicity occurred on these 3 genes (Fig. 4d). Our experiment revealed that the observed circadian rhythms on *MAPK1*, *FOS*, and *CLU* are consistent with ESOCVD prediction (Fig. 4e).

Supplementary Fig. 5a presents the inferred *BIN1* pathways associated with VC and STG. *DNM2* expression level was elevated in the AD group and thus activated the downstream regulation through *BIN1* (Supplementary Table 11). The elevated expression of *DNM2* seems to induce the gain of rhythmicity and cause a base increase in downstream. In addition, Supplementary Fig. 5b shows the *BIN1* pathways associated with STG. *BIN2* and *BIN1* were upregulated in AD patients (Supplementary Table 12) and had lose of rhythmicity.

The evolution of circadian variation during the process of AD development

Since MSBB and GSE131617 include the clinical stage information for each patient, therefore, we explored the dynamic changes in cycling rhythms of seed genes and PBCG genes during the process of AD development. By combining the ESOCVD-inferred cycling rhythm on each condition, we can track the evolution of circadian patterns from early to late-on-set of AD (Fig. 5, Supplementary Fig. 6). Fig. 5a-b show the stage-associated distribution of genes with circadian changes in four brain regions (FP, PG, IFG, and STG) of AD patients in MSBB. Except the superior temporal gyrus (STG) has four types of circadian patterns, other three regions only appear LR and GR. We found that the genes with loss of rhythmicity have a higher overlapping rate in three stages, while the genes with gain of rhythmicity have a lower overlapping rate (Supplementary Fig. 7-8). Our data suggests that LR often occurs at an early stage and continues to the late stage, and GR may happen randomly. Similarly, ESOCVD was also applied to the GSE131617 dataset for detecting the evolution of circadian patterns in other three brain regions (FC, TC, and EC). The results are summarized in Supplementary Fig. 6.

Fig. 5c-d shows the dynamic changes of circadian patterns of 8 representative genes from early to late stage of AD, including *PIGZ*, *PNPLA6*, *RAB11B*, and *CDT1* associated with STG, *SLC5A3*, and *ACAN* from PG, *OMA1* from IFG, and *SLC2A8* from FP. Interestingly, the amplitude of *PIGZ* is altered in the early stage and has a phase shift in the mid stage (Fig. 5c). *ACAN* has obvious rhythm at the early stage and a phase shift at the mid stage (Fig. 5d). *PNPLA6* and *RAB11B* shows attenuated amplitude in the early stage followed by loss of rhythm at the mid-stage. *CDT1* displays an obvious phase shift in the early stage. In addition, *SLC2A8* and *OMA1* gain the rhythmicity from early to mid-stage, while *SLC5A3* lost the rhythm starting from mid stage.

For better understanding the biological meaning of these 8 circadian genes identified above, we examined the genetic backgrounds of individual tissues and disease stages (Fig. 6). Since the single nucleotide polymorphisms (SNPs) are the major genetic information in Alzheimer's disease studies, first, we started from identifying the stage-specific SNPs. We performed correlation study between the variant allele frequency of the known 42 AD-specific SNPs⁶⁸ and expression of 8 genes that have circadian pattern change in their stages (Supplementary Data Table1). Then, we identified 22, 10, and 19 SNPs located near to 28, 10, and 25 genes. Among these, the SNPs located near to *ABCA7*, *CLU*, *CR1*, *EPHA1-AS1*, *MIR6843*, and *SORL1* were common across all AD stages. On the other hand, SNPs near to *ADAM10*, *ALPK2*, *BHMG1*, *CASS4*, and *PICALM* were early-stage specifically correlated with these 8 genes. These genes were enriched in the related pathways with 'amyloid-beta formation'. During mid-stage, *CNTNAP2* was correlated. This gene is known as downregulated in the hippocampus of Alzheimer's disease patients⁶⁹. In the late-stage, *CNN2*, *MS4A6A*, *PILRA*, *SCIMP*, and *ZFP3* showed correlation with 8 genes. These genes were not enriched in any specific pathways. It seems the known AD-specific SNPs are related to our 8 circadian candidate genes in the early-stage to make worsen the condition of the patient. Next, to understand the biological impacts of these 8 genes, we performed correlation studies between the 8 genes and all expressed genes (1.0 ≤ average TPM) per individual tissues and stages of MSBB cohort (Supplementary Data Table2). A total of 672 genes were significantly correlated with these 8 genes (Pearson correlation coefficient >0.9 and p-value < 1.0E-6). Out of 672 genes, 85 genes were matched with the tissue types and stages of the 8 genes, as shown in Fig. 6a. *SLC2A8* has the largest number of correlated genes in FP tissue, including 61, 23, and 4 genes in the early, mid, and late stages, respectively. We performed gene set enrichment analysis focusing on 61, 23, and 4 genes in the early-, mid-, and late-stages using ToppGene⁷⁰ as shown in Fig. 6b. Interestingly, those genes strongly correlated with *SLC2A8* in FP tissue and showed the anticipated enriched biological processes across different AD stages. For example, in the control group, 185 genes are involved in neuron differentiation and development with cell growth. However, in the early stage of AD, 61 correlated genes with *SLC2A8* are involved in the negative regulation of growth, which is opposite to the function in the control group. In the mid stage, 23 correlated genes has main function in the synaptic vesicle uncoating. This observation is consistent with the previous study about Clathrin-mediated endocytosis of synaptic vesicle in Alzheimer's disease. The authors found that the synaptic activity-stimulated endocytosis enhanced the internalization of APP, resulting in the production and release of amyloid-beta (Aβ). In the late stage, the identified 4 genes has function in cardiovascular development. It has been reported that vascular dysfunction is likely to act synergistically with neurodegenerative changes to exacerbate the cognitive impairment in AD⁷¹. Further, a recent study indicated that older people with higher average blood pressure are also more likely to develop tangles and plaques in their brain than their peers⁷². Furthermore, among 85 correlated genes with the matched tissue and stage with the changing circadian rhythms, 11 of them are AD-risk genes. Here, AD genes were downloaded from the International Genomics of Alzheimer's Disease Project (IGAP)⁷³

and Accelerating Medicines Partnership-Alzheimer's Disease (AMP-AD³⁹). Those 11 genes include *AP2A1*, *AP3D1*, *CTIF*, *GAK*, *IL6ST*, *MINK1*, *PTPRS*, *RAB11FIP3*, *SLC2A8*, *SLC5A3*, and *TMEM184B*. Specifically, *MINK1* showed a significant correlation with *SLC2A8* in the FP tissue at the early and mid stages. Broce, *et al.* reported that *MINK1* is a gene with pleiotropic SNPs between AD and cardiovascular disease in a large 'AD-by-proxy' cohort from the UK Biobank and the expression of *MINK1* was differentially altered in the postmortem AD brains⁷⁴. In addition, we also revealed differential expression patterns of 6 of these 8 genes across different stages by one-way ANOVA. Specifically, *SLC5A3* and *RAB11B* showed differential expression patterns in PG and STG, respectively, consistent with their circadian rhythm changes in these tissues (Fig. 6c). Furthermore, we also analyzed the alternative splicing events and RNA A-to-I editing events in these 8 genes. However, there was no specific difference among AD stages. Based on our analyses, we infer that these circadian genes do not change randomly across pseudo time, but serve as intermediate genes in AD pathogenesis.

Discussion

Recent findings have demonstrated that circadian dysfunction might be both a cause and an effect of neurodegenerative disorders during the AD progression^{1,5,16,19,22,27,75}. Much less progress has been made in uncovering the molecular mechanism of how circadian rhythm links to AD development. The lack of time-course transcriptome data for AD has been a key barrier for translational medicine. Our ESOCVD system aims to address this deficiency by using global descriptors of gene expression and Bayesian approach to order unordered data. Based on the predicted circadian phase of each sample, we can order the untimed gene expression data and thus detect the circadian changes and their evolutions in the process of AD progression.

In this study, ESOCVD was applied to 20 untimed gene expression profiles in 16 brain regions of around 3000 AD patients from public transcriptome databases. Together, we examined up to 2391 genes (1691 seed genes and 700 PBCG genes) and observed 1) the distribution of the circadian rhythm changes in all the brain regions of late-onset AD; and 2) the evolution of the circadian rhythm changes in AD progression. Firstly, the preliminary test indicates that the outcomes of ESOCVD are reliable. We found that the inferred circadian patterns of 15 clock genes in the control group are consistent with the evidence reported in CGDB (Supplementary Table 2). Furthermore, we also found that the predicted circadian patterns of 43 PBCG genes were close to the evidences reported in CGDB (Supplementary Table 13). Secondly, our results revealed that loss or gain of rhythmicity occurred with higher frequency rather than the other three types of circadian variation patterns (Fig. 2 and Fig. 4). Thirdly, multiple well-known AD risk factors shows circadian alterations in at least 2 brain regions, e.g. *CR1*, *TREM2*, *PLD3*, and *HLA-DRB5*, *etc.* (Supplementary Table 9). Pathway-level analysis indicates that those circadian changes in AD are mainly associated with five brain regions, including SFG, STG, FC, TC, PC, *etc.* Analysis of circadian rhythm patterns of AD-related clock control genes in all brain regions revealed that 6 of them lost rhythm in certain regions (Fig. 3). Gene expression analysis of AD mouse brain tissues further substantiated the predictions of the ESOCVD model, which provides actionable knowledge for clinical treatment. Thirdly, we developed a systems biology approach to infer GRN to reveal the associations between AD development and circadian dysfunction. Our model fully utilizes the upstream and downstream regulatory information around TF and integrates the functional linkage data and gene expression data. By systematically optimizing the pathway network topology with functional linkage information and gene expression data, our model revealed three TF networks, demonstrating how the upstream of circadian rhythm TF modulate downstream target genes and thus promote AD progression (Fig. 4c-e). We analyzed

the evolution of circadian rhythm changes in AD development. Through the integration of other omics data, eight representative circadian genes were selected to explore their genetic background and dynamic changes across different stages. Our analysis shows that those 8 circadian genes may serve as intermediators in AD pathogenesis.

ESOCVD modeling is initialized at a random starting point and converges to the global optimal solution. In order to improve the searching efficiency of our model, a novel heuristic search strategy was designed and incorporated into ESOCVD (Supplementary Fig. 9). Supplementary Fig. 10 shows that the likelihood calculation of a new sequence S_{ML} generated from the initial sequence S_{bg} by exchanging the positions of two events only need to consider a few top-ranked cases of S_{bg} (Fig. 1). It indicates that the knowledge gained in the current step can be used to guide the new search. To avoid local minima, the optimal sequence for each dataset was obtained after multiple runs. To ensure good convergence, the ESOCVD model was repeated 50, 100, and 200 times for each profile starting from different initial points. Supplementary Fig. 11 indicates that our model displays good convergence. 200-runs is enough to locate the global optimal solution (Supplementary Table 14).

Individual differences between patients may make it difficult to compare baselines and periodic patterns of gene expression. For each dataset, we normalized gene expression profiles within and between samples before model training. In addition, we used 10-fold cross-validation to evaluate the impact of individual variation on the consistency of the optimal solution (see the details in “**Methods**”). Supplementary Fig. 12 shows that the similarity between the optimal sequence fitted to each fold and the sequence fitted to the whole data set is high.

ESOCVD outperforms other typical approaches. As a unique bioinformatics tool for oscillating rhythm detection on untimed gene expression data, CYCLOPS³⁸ shows better efficiency and robustness than other tools, e.g. OSCOPE³⁷, and ZeitZeiger³⁴, etc. In this study, we mainly compared our ESOCVD with CYCLOPS on timed and untimed transcriptome datasets. Firstly, we applied both methods on three time-course transcriptome datasets (GSE39445, GSE48113, and GSE56931) and compared the accuracy on the prediction of the circadian phase. Our comparison result indicates that the prediction accuracy of ESOCVD is better than CYCLOPS (Supplementary Fig. 13). Secondly, we applied both methods on four untimed AD-related datasets with small to large scale to test the reproducibility of the estimated optimal sequence of events. The 10-fold cross-validation revealed that ESOCVD had the better reproducibility of the estimated optimal sequence of events (Supplementary Fig. 12). Furthermore, Supplementary Fig. 14 presents two typical cases of loss rhythmicity detected by CYCLOPS and ESOCVD, respectively. The performance of ESOCVD was close to CYCLOPS for the large size datasets (Supplementary Fig. 14a). Supplementary Fig. 14b shows that CYCLOPS tends to obtain the circadian phases around 12h and misses at least a half cycle. Moreover, our analysis revealed that the number of samples greatly limited the performance of CYCLOPS. The performance is significantly reduced if the sample size is relatively small (Supplementary Fig. 15), so that the estimated circadian phases of samples cannot fill in underrepresented time of day³⁸, which leads to the circadian patterns in a single cycle cannot be completely detected. The above findings indicate that CYCLOPS is more sensitive to the sample size rather than ESOCVD.

There are several limitations of our ESOCVD model. Light-dark cycle and sleep patterns of patients from which the RNA data may influence gene expression and thus obscure waveform and amplitude of temporal fluctuations. Our model searched for the optimal periodic rhythm by fitting the gene expression across all subjects, therefore, it represents the overall trend of circadian rhythm in a patient population. The selection of seed genes may bring some bias on model outcomes. To deeply understand the association between circadian dysfunction and AD progression, we will further identify the seed genes

suitable for ESOCVD modeling based on reported data and knowledge. Using 12 eigengenes (global descriptors) and 24 events, the resolution of the predicted circadian phase is 1 hour. Increasing both the number of eigengenes and the events can assign the samples more evenly distributed within the entire cycle. Given the relative positions of two events for a certain eigengene in an event sequence, the current model needs to consider two possible periodic patterns (Fig. 1), which will largely increase the computational cost in the likelihood calculation.

Methods

Identification of “seed genes”. Firstly, we identified 8504 circadian genes that are more likely to cycle in multiple tissues by using the data in literature⁴¹ via JTK-cycle³⁵. Secondly, we further screened out 1670 brain-related circadian genes, including 15 well-known core clock genes (CLOCK, PER1, PER2, etc.). Finally, we combined above 1670 genes with 21 AD-related clock-controlled genes^{45-47,50,51}, and thus obtained 1691 seed genes, which are considered as the pacemakers in the clock system of AD. Considering the fact that the expression profiles of the seed genes are still high-dimensional, we use SVD to reduce the dimensionality of the data and extract “eigengenes”, characteristic circadian expression patterns, which span the global expression profiles^{38,76,77}. We normalized each patient sample on SVD matrix by subtracting the mean of control group and dividing by the standard deviation of control group. Finally, our controls have a mean of 0 and a standard deviation of 1, the patients won’t have a mean of 0 or standard deviation of 1, because they are z-scored relative to the controls and they should have some disease signal. Besides the 1691 seed genes, we also explore 857 PBCG genes and their circadian changes (Supplementary Fig. 1).

Five patterns of circadian variation.

In this study, we consider five types of circadian alterations between control and AD groups: (1) amplitude change (AC) in AD: a decrease but not the complete loss in amplitude (2-fold reduction); (2) a significant loss of rhythmicity (LR) in AD; (3) a gain of rhythmicity (GR) in AD but no significant rhythm in control; (4) base shift (BS): a change in levels of expression but not necessarily a change in rhythm; (5) phase shift (PS): no significant changes on the amplitude and mean level but shift the phase at least $\frac{1}{4}$ cycle (6 hours). Based on the above five modes, we can identify the clock genes and the output genes with circadian alteration in AD, respectively.

ESOCVD modeling. We formulate the modeling underlying ESOCVD as groups of samples with distinct circadian patterns of eigengene evolution. The eigengene evolution is described as a periodic z-score model in which each eigengene follows an oscillating trajectory across a common timeframe (24 hours). Our model is based on the concept ‘event’ in the event-based models used for disease progression^{78,79}. Different from a unique z-score value^{78,79} or linear z-score series⁸⁰ described the monotonic transition from a normal to an abnormal level, ESOCVD reformulates the events of each eigengene as continuous periodic dynamics. The model fitting is implemented with a Bayesian approach to search the optimal ordering of the events (sequence **S**) for estimating the possible circadian phase of each sample.

Mathematical basis of ESOCVD. Traditional event-based models describe disease progression as a series of events, where each event corresponds to a biomarker transitioning from normal to abnormal level^{78,81}.

However, the instantaneous normal to abnormal events cannot depict the dynamics of the biomarker evolution trajectory. Young, *et al.* use events that represent the linear accumulation of a biomarker from one z-score to another, to uncover data-driven disease phenotypes with distinct temporal progression patterns⁸⁰. For biomarker i , a set of events was represented as a set of monotonically increasing z-scores $z_{i1}, z_{i2}, \dots, z_{iR}$ indicating that the samples with larger expressions would be expected as severe. The limitation of the above methods is that they cannot address the non-monotonic biomarkers. Our developed ESOCVD uses the periodic z-score strategy (Fig. 1). We assume that the model works on 12 circadian eigengenes, which was extracted from the original expression matrix of seed genes with SVD^{76,77}. For the eigengene i ($i = 1, \dots, I$), we define two events E_{i1} and E_{i2} , with z-score z_{i1} and z_{i2} , which represent the ‘crests’ or ‘troughs’ in a single cycle. There are two cases (evolution patterns) need to be considered only given the relative position of E_{i1} and E_{i2} (Fig. 1). Totally, we have $N = 2 * I = 24$ events locating the x axis 1-24 (Fig. 1). The ‘distance’ between any pair of z_{i1} and z_{i2} is half cycle (12h). Therefore, a candidate solution of ESOCVD, event sequence $\mathbf{S}$, needs to follow the constraints shown as below (Eq. (1-3)):

$$Loc(E_{i1}) \in [1, 12], i = 1, 2, \dots, 12 \quad (1)$$

$$Loc(E_{i2}) \in [13, 24], i = 1, 2, \dots, 12 \quad (2)$$

$$Loc(E_{i2}) - Loc(E_{i1}) = 12 \quad (3)$$

where function $Loc()$ defines the position of a z-score event in the sequence $\mathbf{S}$. Eq. (1-3) assume that the event E_{i2} occurred later than event E_{i1} . The occurrence of an event E_i , is informed by the measurements x_{ij} of eigengene i in subject (sample) j , $j = 1 \dots J$. The dataset $X = \{x_{ij} | i = 1, \dots, I, j = 1, \dots, J\}$ is the set of observations of each eigengene in each subject. The aim of ESOCVD is to find the most likely ordering (sequence $\mathbf{S}$ in Fig. 1) of the events that maximize the data likelihood $P(X|\mathbf{S})$. We consider a single cycle as a continuous time axis, t , which we choose to go from $t = 0$ to $t = 1$. At each stage k , which goes from $t = k/(N + 1)$ to $t = (k + 1)/(N + 1)$, and a z-score event E_{iq} (q equals to 1 or 2) occurs. Given the positions of E_{i1} and E_{i2} in sequence $\mathbf{S}$, the eigengene i at the beginning of stage k can be determined by Eq. (4):

$$ZS_i(k) = A \cos\left(\frac{(k - Loc(E_{i1})) * \pi}{Loc(E_{i1}) - Loc(E_{i2})}\right), k \in [0, N] \quad (4)$$

The model likelihood $P(X|\mathbf{S})$ can be represented as Eq. (5):

$$P(X|\mathbf{S}) = \prod_{j=1}^J \left[\sum_{k=0}^N \left(\int_{t=\frac{k}{N+1}}^{t=\frac{k+1}{N+1}} (p(t) \prod_{i=1}^I p(x_{ij}|t)) dt \right) \right] \quad (5)$$

where $p(x_{ij}|t) = \text{NormFDF}(x_{ij}; g_i(t); \sigma_i)$. $\text{NormFDF}(x, \mu, \sigma)$ is the normal probability density function, with mean μ and standard deviation σ , evaluated at x . $p(t)$ follows a uniform distribution, indicating that a prior individual is equally likely to belong to any stage along the oscillating pattern. $g_i(t)$ is a line segment of the form $g_i(t) = a_i t + b_i$. It means the change of $g_i(t)$ is linear if $\frac{k}{N+1} \leq t \leq \frac{k+1}{N+1}$. Thus, we have $g_i(t) = ZS_i(k)$ if $k = t * (N + 1)$.

Model fitting. Supplementary Fig. 9 presents the flowchart of model fitting in our ESOCVD. Model fitting requires simultaneously the stage membership, and the posterior distributions. The sequence $\mathbf{S}$ is initialized randomly. And then, this sequence is optimized by using our hybrid search algorithm. The hybrid search strategy combines global search and local search to find the optimal position of each event in the

sequence. In our simulation, the sequence $\mathbf{S}$ is optimized from 12 different random starting sequences to find the maximum likelihood solution.

Hybrid Search strategy. The hybrid search algorithm involved in ESOCVD model starts from an initial sequence S_q , then performs global search and local search, and finally return an optimal sequence S_{ML} (Supplementary Fig. 10). For simplicity, we used the integer 1-24 to represent 24 original events: $E_{1,1}, E_{2,1}, \dots, E_{12,1}, E_{1,2}, E_{2,2}, \dots, E_{12,2}$, which are associated with 12 eigengenes (Fig. 1). The inferred sequence S_{ML} (the same as the index sequence in Fig. 1) obtained from our model represents the index of each event in the optimal sequence $\mathbf{S}$ relative to the original event profiles. Constrained by **Eq. (1-3)**, each valid sequence S_p satisfies two conditions: (1) the sub-sequence $S_q(1), \dots, S_q(12)$ is a permutation of the integers from 1 to 12; (2) the sub-sequence $S_q(13), \dots, S_q(24)$ was determined $S_q(1), \dots, S_q(12)$ via **Eq. (3)**. The calculation procedure includes the following steps:

Step 1: Input an initial sequence S_q . To guarantee the variety of start points, we set $1 \leq q \leq 12$. It indicates the first event in sequence S_q is event E_{q1} .

Step 2: Global search is performed based on sequence S_q . A new sequence S_g can be generated through three step3: 1) $S_g(1) = S_p(1)$; 2) permute the elements $S_q(2), \dots, S_q(12)$ of S_q and assign them to $S_g(2), \dots, S_g(12)$; (3) calculate $S_g(13), \dots, S_g(24)$ with **Eq. (3)**. Then, calculate the likelihood of each S_g on data X , and screen the sequence S_{bg} with maximal likelihood. In our simulation, we totally generate 100 new sequence S_g from an initial sequence S_q .

Step 3: By using a heuristic search strategy, local search is further performed based on the results of the global search (Supplementary Fig. 9-10). In this step, we exchange the positions of any two events ($E_{i1} \leftrightarrow E_{j1}$ and $E_{i2} \leftrightarrow E_{j2}$) in sequence S_{bg} each iteration and generate a new sequence S_L , and then evaluate if the likelihood of S_L is larger than S_{bg} . Given the relative positions of two events of each eigengene i (E_{i1} , and E_{i2}) in a sequence $\mathbf{S}$, there are two possible circadian patterns that need to be tested. Totally, we need to test the likelihood values in all the $2^{12} = 4096$ possible cases to determine the maximal likelihood for a sequence $\mathbf{S}$, which may lead to high computational cost. Therefore, we proposed an approximate search strategy to speed up the calculation of likelihood. The *rationale* is: if the distance of two events (E_i and E_j) in sequence S_{bg} is very close, the new sequence S_L , generated from S_{bg} by exchanging E_i and E_j , will overlap the top cases of S_{bg} with a big ratio. Top cases are a subset of all the 4096 cases, which will return good likelihood values. To prove the effectiveness of this rational, we selected GSE125583 and MSBB datasets to test. Firstly, we ranked all the 4096 cases of sequence S_{bg} according to the descending order of 4096 likelihood values. Secondly, we ranked all the 4096 cases of a new sequence S_L in the same way. Thirdly, we calculated the overlap of cases with top 100 if the distance between E_{i1} and E_{j1} is 1. Similarly, calculate the overlap of cases with top 200 if the distance between E_{i1} and E_{j1} is 2. Our result shows that around 90% of the top 100 cases of S_{bg} also rank within top 100 for S_L , which indicates that we can only use the top 100 cases of S_{bg} to calculate the likelihood of S_L if the distance between E_{i1} and E_{j1} is 1 (Supplementary Fig. 16). In the meantime, the overlap ratio will be decreased if the distance between E_{i1} and E_{j1} is large. Therefore, our approximate strategy sharply reduces the computational cost of likelihood calculation. In our model, we use a 12-bit binary number to represent a possible case for the circadian patterns of 12 eigengenes. For example, 101000000000 means only the event E_1 and E_3 in a sequence $\mathbf{S}$ follow the first case of their circadian patterns, and the other events follow the second case.

Step 4: screen the optimal sequence S_{ML} . After performing the local search, we compare all the sequences S_L with S_{bg} , and determine the optimal sequence S_{ML} with the maximal likelihood.

Convergence. As described in the above section, the model fitting performs a search optimization from different starting points and chooses the sequence with maximal likelihood as the final solution. In this study, we repeat the ESOCVD model 50, 100, and 200 times for each dataset in parallel on Texas Advance Computing Center (TACC) to against the local minimum of model fitting. We find that ESOCVD shows good convergence: runs from different start points typically converge to a solution.

Cross-validation. To evaluate the consistency of the inferred optimal sequence, we performed 10-fold cross-validation by dividing the data into ten folds and re-fitting the model to each subset of the data, with one of the folds retained for testing each time. And then, we compute the similarity of any two optimal solutions inferred from the model fitted to each subset and the model fitted to the whole dataset, respectively.

Reconstruction of the GRN network

To reveal the association between circadian dysfunction and AD development, we developed a systems biology approach to reconstruct the circadian rhythm TF regulatory network (Supplementary Fig. 3). The calculation includes the following steps:

Step 1: Inference of pairwise transcriptional regulations. Based on the AD risk factors were detected as circadian alterations, we searched all validated TF-TG pairs from the public platform TRRUST⁶¹. There are two situations: 1) If the risk factor with circadian variation is a TG, we need to firstly screen out all the possible TFs for this TG and only select the TFs with the same type of circadian changes as TG; 2) If the risk factor is a TF, we selected all the possible TGs from TRRUST.

Step 2: Inference of modulatory triplet. To obtain the modulator (cofactor)-TF-TG triple regulatory relationships, we firstly identify all the possible modulators of the TF-TG pairs. In this study, we restrict the candidate modulators in the cofactor gene set that interact with TF in functional linkage network from STRING database⁶⁶. We mainly focus on the top-ranked candidate cofactors, which were reported in literature and were validated by other researchers in the previous experiments.

Step 3: Integrating upstream and downstream regulations. The upstream and downstream regulations for all the TFs in the same tissue need to be integrated as a generic pathway map^{67,82}. Please see two examples shown in Supplementary Fig. 4.

Step 4: Inference of circadian rhythm TF-specific pathways by TILP. To infer TF-specific pathway networks, the well-established TILP⁶⁷ approach is used to modeling the generic pathway map with gene expression profile. The constraint system involved in TILP is designed to represent the relationship between the states of two linked nodes and their edge. Nodes in the network represent proteins/genes with their discrete values: -1, 1, or 0, where -1, 1, and 0 stands for “down-regulation”, “up-regulation”, and “no-change”, respectively. The signaling events (connections) are encoded by the operations on the nodes, which also have discrete values (‘promotion’ (1) or ‘inhibition’ (-1)). By fitting the TILP model of a generic pathway to the states of observed nodes via the constraints, we can remove the inconsistent connections and obtain TF-specific pathways. In this study, we used the fold change (FC) of the mean value of a gene in AD to control to get the discrete state of each node: if $FC > 1.2$, “up-regulation”; if $FC < 1/1.2$, “down-regulation”; otherwise, “no-change”.

Real-time qPCR

APPswe/PSEN1dE9 (APP/PS1) mice (JAX: 034829) mice were maintained and treated according to IACUC guidelines, and the procedures were conducted as described in animal protocols approved by the IACUC of the University of Texas Health Science Center at Houston (UTHSC-H). APP/PS1 and wild-type littermate mice were collected at the age of 19-20 months at two circadian time points (ZT6 and ZT18). RT-qPCR analysis was conducted as previously described^{83,84}. Briefly, cortex tissues were isolated from the whole brain and pulverized. Total RNA was extracted from frozen cortex powder by using TRizol (Invitrogen). One µg of extracted RNA was used for cDNA synthesis. Gene expression was analyzed by using QuantStudio 7 Flex Real-Time PCR. Primers for qPCRs are: Atxn1, forward: CACGGTCATTCAGACCACACA, and reverse: GGTAGCCGATGACAGGAGGTT; Atxn3, forward: TGTCTTGTTACAGAAAGATCAG, and reverse: GTTACAAGAACAGAGCTGACT; Scna, forward: TGACAGCAGTCGCTCAGA, and reverse: CATGTCTTCCAGGATTCTTC; Scnb, forward: GGAGGAGCTGTGTTCTCTGG, and reverse: TCCTCTGGCTTCAGGTCTGT; Bace2, forward: TGAGGACCTTGTCACCATCCCAAA, and reverse: TGGCCAAAGCAGCATAAGCAAGTC; ApoE, forward: AACCCTTCTGGGATTACCT, and reverse: TGTGTGACTTGGGAGCTCTG; Mapk1, forward: GGAGCAGTATTATGACCCAAGTGA, and reverse: TCGTCCACTCCATGTCAAAC; Fos, forward: CAGAGCGGAATGGTGAAGA, and reverse: CTGTCTCCGCTTGGAGTGTA; Clu, forward: CAGTTCCCAGACGTGGATT, reverse: TCATCTTCAGGCATCCTGTG.

Code availability.

All the raw data used in this study are available from GEO and AMP-AD. Source code is available at github: https://github.com/JakeJiUTHealth/ESOCVD_1.0.

References

- 1 Chauhan, R., Chen, K. F., Kent, B. A. & Crowther, D. C. Central and peripheral circadian clocks and their role in Alzheimer's disease. *Dis Model Mech* **10**, 1187-1199, doi:10.1242/dmm.030627 (2017).
- 2 Bell-Pedersen, D. *et al.* Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat Rev Genet* **6**, 544-556, doi:10.1038/nrg1633 (2005).
- 3 Bass, J. & Takahashi, J. S. Circadian integration of metabolism and energetics. *Science* **330**, 1349-1354, doi:10.1126/science.1195027 (2010).
- 4 Gery, S. & Koeffler, H. P. Circadian rhythms and cancer. *Cell Cycle* **9**, 1097-1103, doi:10.4161/cc.9.6.11046 (2010).
- 5 Hardin, P. E. & Panda, S. Circadian timekeeping and output mechanisms in animals. *Curr Opin Neurobiol* **23**, 724-731, doi:10.1016/j.conb.2013.02.018 (2013).
- 6 Kewley, R. J., Whitelaw, M. L. & Chapman-Smith, A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *Int J Biochem Cell Biol* **36**, 189-204, doi:10.1016/s1357-2725(03)00211-5 (2004).
- 7 O'Neill, J. S. & Feeney, K. A. Circadian redox and metabolic oscillations in mammalian systems. *Antioxid Redox Signal* **20**, 2966-2981, doi:10.1089/ars.2013.5582 (2014).

- 8 O'Neill, J. S. & Reddy, A. B. The essential role of cAMP/Ca²⁺ signalling in mammalian circadian timekeeping. *Biochem Soc Trans* **40**, 44-50, doi:10.1042/BST20110691 (2012).
- 9 Musiek, E. S. Circadian clock disruption in neurodegenerative diseases: cause and effect? *Front Pharmacol* **6**, 29, doi:10.3389/fphar.2015.00029 (2015).
- 10 Mattis, J. & Sehgal, A. Circadian Rhythms, Sleep, and Disorders of Aging. *Trends Endocrinol Metab* **27**, 192-203, doi:10.1016/j.tem.2016.02.003 (2016).
- 11 Hatfield, C. F., Herbert, J., van Someren, E. J., Hodges, J. R. & Hastings, M. H. Disrupted daily activity/rest cycles in relation to daily cortisol rhythms of home-dwelling patients with early Alzheimer's dementia. *Brain* **127**, 1061-1074, doi:10.1093/brain/awh129 (2004).
- 12 Videnovic, A. Disturbances of Sleep and Alertness in Parkinson's Disease. *Curr Neurol Neurosci Rep* **18**, 29, doi:10.1007/s11910-018-0838-2 (2018).
- 13 Choudhury, G. R. & Daadi, M. M. Charting the onset of Parkinson-like motor and non-motor symptoms in nonhuman primate model of Parkinson's disease. *PLoS One* **13**, e0202770, doi:10.1371/journal.pone.0202770 (2018).
- 14 van Wamelen, D. J., Roos, R. A. & Aziz, N. A. Therapeutic strategies for circadian rhythm and sleep disturbances in Huntington disease. *Neurodegener Dis Manag* **5**, 549-559, doi:10.2217/nmt.15.45 (2015).
- 15 Diago, E. B. *et al.* Circadian Rhythm, Cognition, and Mood Disorders in Huntington's Disease. *J Huntingtons Dis* **7**, 193-198, doi:10.3233/JHD-180291 (2018).
- 16 Buhl, E., Higham, J. P. & Hodge, J. J. L. Alzheimer's disease-associated tau alters *Drosophila* circadian activity, sleep and clock neuron electrophysiology. *Neurobiol Dis* **130**, 104507, doi:10.1016/j.nbd.2019.104507 (2019).
- 17 Sterniczuk, R., Dyck, R. H., Laferla, F. M. & Antle, M. C. Characterization of the 3xTg-AD mouse model of Alzheimer's disease: part 1. Circadian changes. *Brain Res* **1348**, 139-148, doi:10.1016/j.brainres.2010.05.013 (2010).
- 18 Stevanovic, K. *et al.* Disruption of normal circadian clock function in a mouse model of tauopathy. *Exp Neurol* **294**, 58-67, doi:10.1016/j.expneurol.2017.04.015 (2017).
- 19 Khabirova, E., Chen, K. F., O'Neill, J. S. & Crowther, D. C. Flyglow: Single-fly observations of simultaneous molecular and behavioural circadian oscillations in controls and an Alzheimer's model. *Sci Rep* **6**, 33759, doi:10.1038/srep33759 (2016).
- 20 Roh, J. H. *et al.* Disruption of the sleep-wake cycle and diurnal fluctuation of beta-amyloid in mice with Alzheimer's disease pathology. *Sci Transl Med* **4**, 150ra122, doi:10.1126/scitranslmed.3004291 (2012).
- 21 Duncan, M. J. *et al.* Effects of aging and genotype on circadian rhythms, sleep, and clock gene expression in APPxPS1 knock-in mice, a model for Alzheimer's disease. *Exp Neurol* **236**, 249-258, doi:10.1016/j.expneurol.2012.05.011 (2012).
- 22 Blazer, D. G., Hays, J. C. & Foley, D. J. Sleep complaints in older adults: a racial comparison. *J Gerontol A Biol Sci Med Sci* **50**, M280-284, doi:10.1093/gerona/50a.5.m280 (1995).
- 23 Van Someren, E. J. Circadian and sleep disturbances in the elderly. *Exp Gerontol* **35**, 1229-1237, doi:10.1016/s0531-5565(00)00191-1 (2000).
- 24 Musiek, E. S., Xiong, D. D. & Holtzman, D. M. Sleep, circadian rhythms, and the pathogenesis of Alzheimer disease. *Exp Mol Med* **47**, e148, doi:10.1038/emm.2014.121 (2015).
- 25 Kress, G. J. *et al.* Regulation of amyloid-beta dynamics and pathology by the circadian clock. *J Exp Med* **215**, 1059-1068, doi:10.1084/jem.20172347 (2018).
- 26 Musiek, E. S. *et al.* Circadian clock proteins regulate neuronal redox homeostasis and neurodegeneration. *J Clin Invest* **123**, 5389-5400, doi:10.1172/JCI70317 (2013).
- 27 Koh, K., Zheng, X. & Sehgal, A. JETLAG resets the *Drosophila* circadian clock by promoting light-induced degradation of TIMELESS. *Science* **312**, 1809-1812, doi:10.1126/science.1124951 (2006).

- 28 Ambree, O. *et al.* Activity changes and marked stereotypic behavior precede Abeta pathology in TgCRND8 Alzheimer mice. *Neurobiol Aging* **27**, 955-964, doi:10.1016/j.neurobiolaging.2005.05.009 (2006).
- 29 Kang, J. E. *et al.* Amyloid-beta dynamics are regulated by orexin and the sleep-wake cycle. *Science* **326**, 1005-1007, doi:10.1126/science.1180962 (2009).
- 30 Rothman, S. M., Herdener, N., Frankola, K. A., Mughal, M. R. & Mattson, M. P. Chronic mild sleep restriction accentuates contextual memory impairments, and accumulations of cortical Abeta and pTau in a mouse model of Alzheimer's disease. *Brain Res* **1529**, 200-208, doi:10.1016/j.brainres.2013.07.010 (2013).
- 31 Tranah, G. J. *et al.* Circadian activity rhythms and risk of incident dementia and mild cognitive impairment in older women. *Ann Neurol* **70**, 722-732, doi:10.1002/ana.22468 (2011).
- 32 Calabrese, V., Cornelius, C., Dinkova-Kostova, A. T., Calabrese, E. J. & Mattson, M. P. Cellular stress responses, the hormesis paradigm, and vitagenes: novel targets for therapeutic intervention in neurodegenerative disorders. *Antioxid Redox Signal* **13**, 1763-1811, doi:10.1089/ars.2009.3074 (2010).
- 33 Perla-Kajan, J. & Jakubowski, H. Dysregulation of Epigenetic Mechanisms of Gene Expression in the Pathologies of Hyperhomocysteinemia. *Int J Mol Sci* **20**, doi:10.3390/ijms20133140 (2019).
- 34 Hughey, J. J., Hastie, T. & Butte, A. J. ZeitZeiger: supervised learning for high-dimensional data from an oscillatory system. *Nucleic Acids Res* **44**, e80, doi:10.1093/nar/gkw030 (2016).
- 35 Hughes, M. E., Hogenesch, J. B. & Kornacker, K. JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms* **25**, 372-380, doi:10.1177/0748730410379711 (2010).
- 36 Braun, R. *et al.* Universal method for robust detection of circadian state from gene expression. *Proc Natl Acad Sci U S A* **115**, E9247-E9256, doi:10.1073/pnas.1800314115 (2018).
- 37 Leng, N. *et al.* Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. *Nat Methods* **12**, 947-950, doi:10.1038/nmeth.3549 (2015).
- 38 Anafi, R. C., Francey, L. J., Hogenesch, J. B. & Kim, J. CYCLOPS reveals human transcriptional rhythms in health and disease. *Proc Natl Acad Sci U S A* **114**, 5312-5317, doi:10.1073/pnas.1619320114 (2017).
- 39 Hodes, R. J. & Buckholtz, N. Accelerating Medicines Partnership: Alzheimer's Disease (AMP-AD) Knowledge Portal Aids Alzheimer's Drug Discovery through Open Data Sharing. *Expert Opin Ther Targets* **20**, 389-391, doi:10.1517/14728222.2016.1135132 (2016).
- 40 Li, S. *et al.* CGDB: a database of circadian genes in eukaryotes. *Nucleic Acids Res* **45**, D397-D403, doi:10.1093/nar/gkw1028 (2017).
- 41 Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111**, 16219-16224, doi:10.1073/pnas.1408886111 (2014).
- 42 Lo, I., Hill, J., Vilhjalmsen, B. J. & Kjems, J. Linking the association between circRNAs and Alzheimer's disease progression by multi-tissue circular RNA characterization. *RNA Biol*, 1-9, doi:10.1080/15476286.2020.1783487 (2020).
- 43 Braak, H., Alafuzoff, I., Arzberger, T., Kretschmar, H. & Del Tredici, K. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol* **112**, 389-404, doi:10.1007/s00401-006-0127-z (2006).
- 44 Katsel, P., Li, C. & Haroutunian, V. Gene expression alterations in the sphingolipid metabolism pathways during progression of dementia and Alzheimer's disease: a shift toward ceramide accumulation at the earliest recognizable stages of Alzheimer's disease? *Neurochem Res* **32**, 845-856, doi:10.1007/s11064-007-9297-x (2007).

- 45 Koike, N. *et al.* Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338**, 349-354, doi:10.1126/science.1226339 (2012).
- 46 Chen, C. Y. *et al.* Effects of aging on circadian patterns of gene expression in the human prefrontal cortex. *Proc Natl Acad Sci U S A* **113**, 206-211, doi:10.1073/pnas.1508249112 (2016).
- 47 Li, J. Z. *et al.* Circadian patterns of gene expression in the human brain and disruption in major depressive disorder. *Proc Natl Acad Sci U S A* **110**, 9950-9955, doi:10.1073/pnas.1305814110 (2013).
- 48 Cornelissen, G. Cosinor-based rhythmometry. *Theor Biol Med Model* **11**, 16, doi:10.1186/1742-4682-11-16 (2014).
- 49 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29, doi:10.1038/75556 (2000).
- 50 Ye, Y. *et al.* The Genomic Landscape and Pharmacogenomic Interactions of Clock Genes in Cancer Chronotherapy. *Cell Syst* **6**, 314-328 e312, doi:10.1016/j.cels.2018.01.013 (2018).
- 51 Mure, L. S. *et al.* Diurnal transcriptome atlas of a primate across major neural and peripheral tissues. *Science* **359**, doi:10.1126/science.aao0318 (2018).
- 52 Bridel, C. *et al.* Glutaminyl cyclase activity correlates with levels of Abeta peptides and mediators of angiogenesis in cerebrospinal fluid of Alzheimer's disease patients. *Alzheimers Res Ther* **9**, 38, doi:10.1186/s13195-017-0266-6 (2017).
- 53 Rahman, M. R. *et al.* Identification of molecular signatures and pathways to identify novel therapeutic targets in Alzheimer's disease: Insights from a systems biomedicine perspective. *Genomics* **112**, 1290-1299, doi:10.1016/j.ygeno.2019.07.018 (2020).
- 54 Rahman, M. R. *et al.* Discovering Biomarkers and Pathways Shared by Alzheimer's Disease and Ischemic Stroke to Identify Novel Therapeutic Targets. *Medicina (Kaunas)* **55**, doi:10.3390/medicina55050191 (2019).
- 55 Feinberg, K. *et al.* A neuroprotective agent that inactivates prodegenerative TrkA and preserves mitochondria. *J Cell Biol* **216**, 3655-3675, doi:10.1083/jcb.201705085 (2017).
- 56 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
- 57 Prinz, M. & Priller, J. Microglia and brain macrophages in the molecular age: from origin to neuropsychiatric disease. *Nat Rev Neurosci* **15**, 300-312, doi:10.1038/nrn3722 (2014).
- 58 Solito, E. & Sastre, M. Microglia function in Alzheimer's disease. *Front Pharmacol* **3**, 14, doi:10.3389/fphar.2012.00014 (2012).
- 59 Koh, S. H., Noh, M. Y. & Kim, S. H. Amyloid-beta-induced neurotoxicity is reduced by inhibition of glycogen synthase kinase-3. *Brain Res* **1188**, 254-262, doi:10.1016/j.brainres.2007.10.064 (2008).
- 60 Reger, M. A. *et al.* Intranasal insulin administration dose-dependently modulates verbal memory and plasma amyloid-beta in memory-impaired older adults. *J Alzheimers Dis* **13**, 323-331, doi:10.3233/jad-2008-13309 (2008).
- 61 Han, H. *et al.* TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res* **46**, D380-D386, doi:10.1093/nar/gkx1013 (2018).
- 62 Jin, G. & Howe, P. H. Transforming growth factor beta regulates clusterin gene expression via modulation of transcription factor c-Fos. *Eur J Biochem* **263**, 534-542, doi:10.1046/j.1432-1327.1999.00533.x (1999).
- 63 Foster, E. M., Dangla-Valls, A., Lovestone, S., Ribe, E. M. & Buckley, N. J. Clusterin in Alzheimer's Disease: Mechanisms, Genetics, and Lessons From Other Pathologies. *Front Neurosci* **13**, 164, doi:10.3389/fnins.2019.00164 (2019).
- 64 Taira, T. *et al.* AMY-1, a novel C-MYC binding protein that stimulates transcription activity of C-MYC. *Genes Cells* **3**, 549-565, doi:10.1046/j.1365-2443.1998.00206.x (1998).

- 65 De Rossi, P. *et al.* BIN1 localization is distinct from Tau tangles in Alzheimer's disease. *Matters (Zur)* **2017**, doi:10.19185/matters.201611000018 (2017).
- 66 Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613, doi:10.1093/nar/gky1131 (2019).
- 67 Ji, Z. *et al.* A linear programming computational framework integrates phosphor-proteomics and prior knowledge to predict drug efficacy. *BMC Syst Biol* **11**, 127, doi:10.1186/s12918-017-0501-6 (2017).
- 68 Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404-413, doi:10.1038/s41588-018-0311-9 (2019).
- 69 van Abel, D. *et al.* Direct downregulation of CNTNAP2 by STOX1A is associated with Alzheimer's disease. *J Alzheimers Dis* **31**, 793-800, doi:10.3233/JAD-2012-120472 (2012).
- 70 Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* **37**, W305-311, doi:10.1093/nar/gkp427 (2009).
- 71 Rius-Perez, S., Tormos, A. M., Perez, S. & Talens-Visconti, R. Vascular pathology: Cause or effect in Alzheimer disease? *Neurologia* **33**, 112-120, doi:10.1016/j.nrl.2015.07.010 (2018).
- 72 Arvanitakis, Z. *et al.* Late-life blood pressure association with cerebrovascular and Alzheimer disease pathology. *Neurology* **91**, e517-e525, doi:10.1212/WNL.0000000000005951 (2018).
- 73 Andrews, S. J., Fulton-Howard, B. & Goate, A. Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *Lancet Neurol* **19**, 326-335, doi:10.1016/S1474-4422(19)30435-1 (2020).
- 74 Broce, I. J. *et al.* Dissecting the genetic relationship between cardiovascular risk factors and Alzheimer's disease. *Acta Neuropathol* **137**, 209-226, doi:10.1007/s00401-018-1928-6 (2019).
- 75 Loehfelm, A., Boucsein, A., Pretz, D. & Tups, A. Timing Matters: Circadian Effects on Energy Homeostasis and Alzheimer's Disease. *Trends Endocrinol Metab* **30**, 132-143, doi:10.1016/j.tem.2018.12.001 (2019).
- 76 Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* **97**, 10101-10106, doi:10.1073/pnas.97.18.10101 (2000).
- 77 Cho, R. J. *et al.* A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol Cell* **2**, 65-73, doi:10.1016/s1097-2765(00)80114-8 (1998).
- 78 Fonteijn, H. M. *et al.* An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *Neuroimage* **60**, 1880-1889, doi:10.1016/j.neuroimage.2012.01.062 (2012).
- 79 Venkatraghavan, V., Bron, E. E., Niessen, W. J., Klein, S. & Alzheimer's Disease Neuroimaging, I. Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling. *Neuroimage* **186**, 518-532, doi:10.1016/j.neuroimage.2018.11.024 (2019).
- 80 Young, A. L. *et al.* Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat Commun* **9**, 4273, doi:10.1038/s41467-018-05892-0 (2018).
- 81 Young, A. L. *et al.* A data-driven model of biomarker changes in sporadic Alzheimer's disease. *Brain* **137**, 2564-2577, doi:10.1093/brain/awu176 (2014).
- 82 Ji, Z. *et al.* Integrating genomics and proteomics data to predict drug effects using binary linear programming. *PLoS One* **9**, e102798, doi:10.1371/journal.pone.0102798 (2014).
- 83 Yoo, S. H. *et al.* Competing E3 ubiquitin ligases govern circadian periodicity by degradation of CRY in nucleus and cytoplasm. *Cell* **152**, 1091-1105, doi:10.1016/j.cell.2013.01.055 (2013).

- 84 Nohara, K. *et al.* Cardiolipin Synthesis in Skeletal Muscle Is Rhythmic and Modifiable by Age and Diet. *Oxid Med Cell Longev* **2020**, 5304768, doi:10.1155/2020/5304768 (2020).

Acknowledgments

This work was supported by National Institutes of Health U01CA166886, U01AR069395-01A1, R01CA241930 and R01GM123037 (Zhou). Welch Foundation (AU-1971-20180324) and NIH/NIGMS (R01GM114424) to S.-H.Y., The Welch Foundation (AU-1731-20190330) and NIH/NIA (RF1AG061901, R56AG063746 and R01AG065984) to Z.C. Funding for open access charge: Dr Carl V. Vartian Chair Professorship Funds to Dr. Zhou from the University of Texas Health Science Center at Houston. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Z.J. and X.Z. conceived and designed the computational framework. Z.J. implemented programming code and wrote the manuscript. D.S., P.K., S.W, M.Y analyzed the data. S.Y., Z.C. designed the gene expression analysis which was performed by E.K. and M.W. W.Z. advised on the description of some analysis. All authors contributed to reviewing and editing of the report.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Files accompanies this paper at the following links:

Supplementary File1: <https://ccsm.uth.edu/ESOCVD/SupplFile1.rar>.

Supplementary File2: <https://ccsm.uth.edu/ESOCVD/SupplFile2.rar>.

Supplementary File3: <https://ccsm.uth.edu/ESOCVD/SupplFile3.rar>.

Supplementary DataTable1: <https://ccsm.uth.edu/ESOCVD/SupplDTb1.xlsx>.

Supplementary DataTable2: <https://ccsm.uth.edu/ESOCVD/SupplDTb2.xlsx>.

Figures

Fig. 1 The flowchart of ESOCVD model.

a

	Entorhinal cortex	Hippocampus	Medial temporal gyrus	Posterior cingulate	Superior frontal gyrus	Primary visual cortex	Fusiform gyrus	Frontal pole	Parahippocampal gyrus	Inferior frontal gyrus	Superior temporal gyrus	Frontal cortex	Temporal cortex	Dorsolateral Prefrontal cortex	Visual cortex	Cerebellum	Frontal cortex*	Temporal cortex*	Entorhinal cortex*
<i>Amplitude change</i>	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Loss of rhythmicity</i>	24	57	42	63	350	50	19	15	19	10	256	31	119	108	31	67	265	34	126
<i>Gain of rhythmicity</i>	202	60	85	62	21	101	52	39	307	12	3	264	27	96	576	72	18	29	13
<i>Phase shift</i>	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
<i>Base shift</i>	2	1	0	0	0	0	0	0	0	0	0	1	2	1	4	1	0	1	0

b

Fig. 2 ESOCVD analysis of circadian transcriptome in late on-set of AD. **a**, the number of seed genes with circadian changes in multiple brain regions detected by ESOCVD. **b**, Reconstructed expression profiles of selected clock genes are plotted as a function of ESOCVD phase. Each dot represents one sample in control group.

Fig. 3 ESOCVD analysis of AD-related circadian changes in late on set of AD. **a**, the distribution of circadian changes of AD-related clock-controlled genes in multiple brain regions detected by ESOCVD. **b**, Reconstructed expression profiles of six AD-related circadian genes are plotted as a function of ESOCVD phase. ATXN1 and ATX3 are plotted with GSE131617_FC. SNCA and SNCB are related with DPC (GSE44772) and SFG (GSE5281). BACE2 and APOE are plotted with MSBB_STG and GSE15222_TC, respectively. The dots with blue and green color represent control and AD samples, respectively. **c**, Real-time qPCR analysis of six genes presented in **(b)** in the cortex. Error bars represent SEM (n=4-6) # p<0.05, ## p<0.01, unpaired t-test; *p<0.05, **p<0.01, ***p<0.001, two-way ANOVA with Sidak multiple comparisons test.

Fig. 4 ESOCVD analysis of PBCG-related circadian changes in late on set of AD. **a**, the number of PBCG genes with circadian changes in multiple brain regions detected by ESOCVD. **b**, Reconstructed expression profiles of selected PBCG genes are plotted as a function of ESOCVD phase. **c**, Inferred CLU-related signaling pathways in SFG. **d**, Reconstructed expression profiles of genes on CLU pathways are plotted as a function of ESOCVD phase. **e**, qPCR experiments for CLU pathway in the cortex. Error bars represent SEM (n=4-6). *p<0.05, **p<0.01, ***p<0.001, two-way ANOVA with Sidak multiple comparisons test. The dots in (b) and (d) with blue and green color represent control and AD samples, respectively.

Fig. 5 ESOCVD analysis of circadian dynamics in four regions during AD progression. Expression data from MSBB was used. **a**, the number of seed genes with circadian changes in four brain regions detected by ESOCVD. **b**, the number of PBCG genes with circadian changes in four brain regions detected by ESOCVD. **c**, expression of four selected seed genes as a function of ESOCVD phase. PIGZ and PNPLA6 are detected from STG. SLC5A3 and OMA1 are associated with PG and IFG, respectively. **d**, expression of four selected PBCG genes as a function of ESOCVD phase. RAB11B and CDT1 are detected from STG. SLC2A8 and ACAN are associated with FP and PG, respectively. The dots with blue and green color (**c-d**) represent control and AD samples, respectively.

Fig. 6 Genetic backgrounds of the 8 genes with circadian pattern changes in the MSBB cohort. **a**, Overall key genes among stages. The red fonts are the genes that gained the circadian rhythm compared to the

previous stage. The blue fonts are the genes that lost the circadian rhythm. The green fonts are the genes that reduced the amplitude of circadian rhythm or had a phase shift. The bold fonts are the genes that showed the differential gene expression in the tissue matched with the circadian rhythm changes. SLC5A3 and RAB11B had differential gene expressions across stages. There were seven genes that have AD-specific SNPs associated with AD stages. **b**, Enriched biological functions of the correlated genes with SLC2A8 in FP tissue across stages. **c**, Expression difference across stages of SLC5A3 and RAB11B.