Nature and Science of Sleep

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Machine Learning Analyses Reveal Circadian Features Predictive of Risk for Sleep Disturbance

Rebeccah Overton[1], Aziz Zafar[1,2], Ziad Attia[1,3], Ahmet Ay[1,2], Krista K Ingram[1]

[1]Department of Biology, Colgate University, Hamilton, NY, USA; [2]Department of Mathematics, Colgate University, Hamilton, NY, USA; [3]Department of Computer Science, Colgate University, Hamilton, NY, USA

Correspondence: Krista K Ingram, Email kingram@colgate.edu

Introduction: Sleep disturbances often co-occur with mood disorders, with poor sleep quality affecting over a quarter of the global population. Recent advances in sleep and circadian biology suggest poor sleep quality is linked to disruptions in circadian rhythms, including significant associations between sleep features and circadian clock gene variants.

Methods: Here, we employ machine learning techniques, combined with statistical approaches, in a deeply phenotyped population to explore associations between clock genotypes, circadian phenotypes (diurnal preference and circadian phase), and risk for sleep disturbance symptoms.

Results: As found in previous studies, evening chronotypes report high levels of sleep disturbance symptoms. Using molecular chronotyping by measuring circadian phase, we extend these findings and show that individuals with a mismatch between circadian phase and diurnal preference report higher levels of sleep disturbance. We also report novel synergistic interactions in genotype combinations of *Period 3, Clock* and *Cryptochrome* variants (PER3B (rs17031614)/ CRY1 (rs228716) and CLOCK3111 (rs1801260)/ CRY2 (rs10838524)) that yield strong associations with sleep disturbance, particularly in males.

Conclusion: Our results indicate that both direct and indirect mechanisms may impact sleep quality; sex-specific clock genotype combinations predictive of sleep disturbance may represent direct effects of clock gene function on downstream pathways involved in sleep physiology. In addition, the mediation of clock gene effects on sleep disturbance indicates circadian influences on the quality of sleep. Unraveling the complex molecular mechanisms at the intersection of circadian and sleep physiology is vital for understanding how genetic and behavioral factors influencing circadian phenotypes impact sleep quality. Such studies provide potential targets for further study and inform efforts to improve non-invasive therapeutics for sleep disorders.

Keywords: circadian clock, chronotype, sleep disturbance, sleep quality, machine learning, circadian misalignment

Introduction

Adequate regulation of sleep is important to both mental and physical human health yet sleep disturbances plague nearly one-quarter of the world's population.[1] Disruptions in sleep can negatively impact productivity, lowering motivation and job performance, as well as the quality of life, impacting healthy relationships and general happiness.[2] Individuals who experience sleep disturbance and poor sleep efficiency are also more likely to suffer from mental health disorders, particularly major depressive disorder. A better understanding of the mechanisms underlying sleep disturbances is vital to providing better treatments and outcomes for individuals experiencing sleep disruptions and associated mood disorders. A complex interplay of factors influence why some individuals experience sleep disturbance more frequently than others. Recent advances in circadian and sleep science underscore the influence of the inernal circadian clock on sleep processes and propose both direct and indirect mechanisms for how circadian factors may intersect with sleep physiology.

There are two primary drivers of sleep regulation, the sleep homeostat (Process S) and the circadian rhythm (Process C) (Figure 1).[3] These two processes involve distinct mechanisms that control different aspects of the sleep-wake cycle. The sleep homeostat is the body's drive for sleep. The drive builds during the day as our time awake increases and get stronger the longer we are awake. After a full night of sleep, the drive reaches its lowest point when we have no more need for sleep. The circadian clock is an endogenous oscillator in the suprachiasmatic nuclei in the brain that influences the timing of the

Received: 8 July 2022
Accepted: 21 September 2022
Published: 21 October 2022

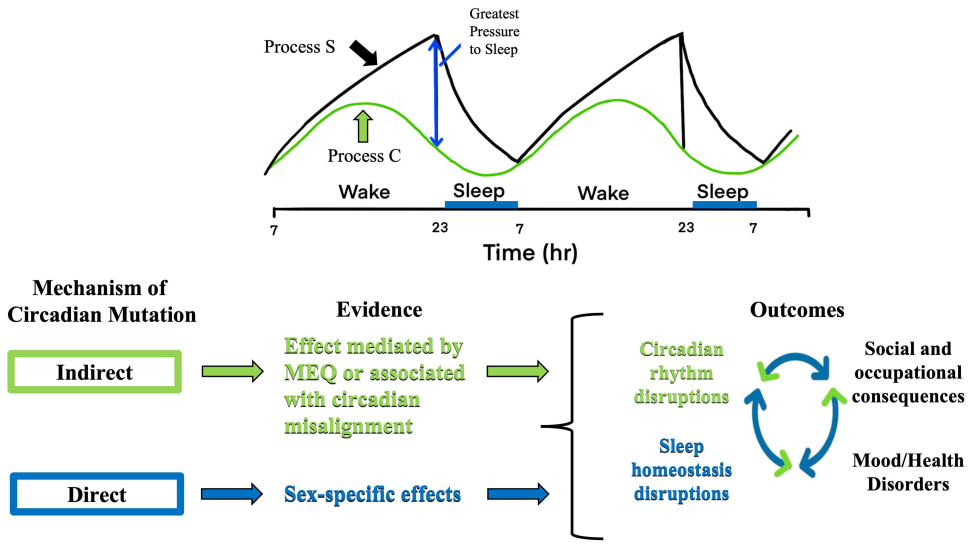

Figure 1 Dual processes controlling the sleep-wake cycle and potential mechanisms by which circadian mutations may influence sleep disturbance and related outcomes. Process S describes the sleep homeostat; the body's drive for sleep builds during the day as our time awake increases and energy is expended. Sleep drive gets stronger the longer we are awake. After a full night of sleep, the drive reaches its lowest point when we have no more need for sleep. Process C describes the circadian rhythm—a self-sustained oscillator that influences the timing of the sleep-wake cycle. These two processes of sleep regulation are continuously interacting but are guided by independent mechanisms.

sleep-wake cycle. These two processes of sleep regulation are continuously interacting but are regulated independently.[3,4] Many of the negative effects of sleep disturbance and sleep deprivation may be attributed to the mismatch between these two processes.

One leading factor in sleep disturbance is circadian misalignment—a mismatch between an individual's endogenous circadian rhythms and their lifestyle. Internal circadian rhythms in the brain operate on a 24-hour cycle and are entrained every day by environmental factors such as temperature and exposure to light. These molecular rhythms interact with the sleep homeostat to regulate the sleep-wake cycle via transcription-translation feedback loops in core clock genes which, in turn, regulate associated transcription factors that promote or reduce sleep tendencies.[5–7] In mammals, core clock genes include the transcription factors CLOCK and BMAL1, which regulate the transcription of *Period* (*Per1, Per2*, and *Per3*) and *Cryptochrome* (*Cry1* and *Cry2*) complexes. In the negative loop, the protein products, PER and CRY dimerize in the cytoplasm and reenter the nucleus to inhibit E-box activation of the CLOCK:BMAL1 dimer.[8] In the alternate loop, *Ror* and *Rev-erbα*, are also regulated by the E-box, and their protein products control the rate of *Bmal1* transcription; slower transcription when REV-ERBα is bound and accelerated transcription when ROR is bound. The interaction of these dual feedback loops stabilizes circadian gene oscillations as well as associated downstream, clock-controlled processes. In humans, the misalignment of the circadian clock with behavioral timing can disrupt downstream regulation of the sleep-wake cycle which can lead to circadian rhythm sleep disorders such as delayed sleep phase disorder (DSPD).[9]

Previous research on the human circadian/sleep homeostat interface has identified circadian genes directly involved in sleep homeostasis.[4,10,11] For example, a variable number of tandem repeat (VNTR) region of PER3 is associated with variance in homeostatic sleep pressure and sleep architecture.[3] Direct changes in the sleep homeostat can also result from altered light sensitivity which can be mediated by the function of ipRGC's. These changes can be treated by directly targeting the ipRGC response via light therapy or melatonin supplements.[12,13] The interactions between sleep drive changes and circadian rhythms have also been well studied in mice and other rodents.[14–16] For example, sleep deprivation in hamsters attenuates phase shifts induced by light, suggesting that an increase in sleep drive reduces the circadian clock response to a zeitgeber, in this case, light.[20] Interestingly, attenuated phase advances were also found after sleep restriction in humans.[17] However, sleep disturbance can also be modulated by behavioral chronotypes via indirect pathways involving phase changes in the circadian clock. For example, Russo et al (2017) found that sleep problems occurring more than once a week or frequent struggling to fall asleep were much higher in evening type adolescents than

morning types.[18] Collectively, these studies underscore the multiple potential links between the circadian clockwork and the sleep homeostat.

Mutations in core clock genes could thus affect sleep quality via direct effects (eg, sex-specific transcriptional regulation of downstream pathways) or via indirect effects (eg, mediation via behavioral and physiological rhythms driven by phase changes in circadian rhythms and circadian misalignment.[19] To better inform therapeutic sleep technologies, we need to determine which circadian clock gene mutations affect sleep quality and to understand the mechanisms by which these mutations influence sleep. To test for direct and indirect circadian-related mechanisms influencing sleep disturbance, we performed deep phenotypic analysis using the PROMIS sleep disturbance survey,[20] a well-established measure of diurnal preference, the Horne-Ostberg Morningness-Eveningness Questionnaire (MEQ)[21] and molecular chronotyping (via in-vivo gene expression analyses based on the phase of clock gene oscillations). We analyzed genotype associations with sleep disturbance using machine learning and statistical approaches to identify the effects of single and multiple circadian mutations and clinical features on sleep disturbance scores.

Methods

Experimental Data Collection

Study participants were recruited from Colgate University and the community residing in Madison County, NY, USA (n = 982; males = 318, females = 664, ages 17–79; median = 19). Participants were predominantly Caucasians of European descent with less than 8% of respondents indicating a socioeconomic status of lower middle class or poor. Participants gave written informed consent, and all procedures followed the principles of the Declaration of Helsinki. The Institutional Review Board at Colgate University authorized all consent forms and procedures (#FR-F13-07, #ER-F14 -12, #F15-13, and #ER-F16-19). The study population and similar methodologies were used in a previous study focused on circadian influences on anxiety.

Self-Report Surveys

Participants completed computer-based surveys, including the short form of the Patient Reported Outcomes Measurement Information System (PROMIS™) Sleep Disturbance[20] and the Horne Östberg Morningness-Eveningness Questionnaire (MEQ) survey.[21] PROMIS scores indicate self-reported assessments of sleep disturbance, ranging from 0 to 40, with scores of 0–25 indicating "no or low sleep disturbance", and 25–40 indicating "high sleep disturbance." Twenty three percent of the sample population (55 (17%) males and 174 (26%) females) scored 25 or above on the PROMIS scale, indicating high sleep disturbance. The MEQ survey measures diurnal preference, with scores of <41 representing evening-type preferences and scores >59 representing morning-type preferences.

Genotyping

DNA samples were obtained from ten to twenty hair follicles from each participant. Follicles were digested with Proteinase K at 56°C for 24-hours and purified using the Qiagen DNAeasy Micro Kit. Genotyping of single nucleotide polymorphisms (SNPs) was completed using a TaqMan SNP Genotyping assay (Applied Biosystems, Foster City, CA) on an ABI QuantStudio real-time qPCR instrument. A fragment length analysis of the PER3 VNTR length polymorphism repeat region was conducted using GeneScan software on an ABI 3100 sequencer using PCR fluorescent primers. The forward primer fluorescently labeled with 6-FAM was used with the following PCR primers: forward, 5'-CAAAATTTTA TGACACTACCAGAATGGCTGAC-3', and reverse, 5'-AACC TTGTACTTCCACATCAGTGCCTGG-3'.[22] The PCR cycling conditions were 3 min at 94°C, followed by 35 cycles of 45 s at 94°C, 45 s at 58°C, and 45 s at 72°C, with a final step at 72°C for 3 min. Capillary electrophoresis on an ABI 3100 sequencer was used to separate and size bands using ABI ROX standards. The genotype of each participant was identified as PER3 4/4, PER3 4/5, or PER3 5/5.

Molecular Chronotyping via Circadian Gene Expression Analysis

Ten to twenty additional hair follicles were collected at four different time points: 8 a.m., 4 p.m., 5 p.m., and 8 p.m.[23] We optimized the estimation of clock phase from four time points and two genes using Stochastic Ranking Evolution

Strategy (SRES) methods in previous publications (methods in Supplementary File) 24. All follicle samples were stored in RNAlater solution at −80°C prior to analysis. RNA was extracted and purified from hair follicles using the RNeasy Micro purification kit according to the protocol provided by Qiagen. The purified RNA was converted to cDNA using rt-PCR (TaqMan Gold rt-PCR, ABI) and quantified using Nanodrop. Expression levels of clock genes, PER3 and NR1D2, and control genes, GUSB and 18S, were measured in triplicate using rt-qPCR on an ABI QuantStudio instrument (Applied Biosystems). Relative mRNA levels were determined using the standard curve method and converted into z-scores per individual. The trained curve was fitted to the four data points of each subject using the parameter estimation method, Stochastic Ranking Evolutionary Strategy (SRES).[24] For phase shift estimation, the training curve was obtained from known intermediate types (n = 20; individuals not included in the current study but included in previous studies).[23] The phase difference between each subject's curve and the training curve estimates the circadian phase shift of the individual.

Feature Generation and Selection

Similar methodologies to those described below were used in a previous study on circadian influences on anxiety; for more information, see Zafar et al, 2022.[25]

Genotypic and Clinical Features

We used seven genotypic features: CLOCK3111 (rs1801260), CRY1 (rs228716), CRY2 (rs10838524), PER2 (rs10462023), PER3A (rs228697), PER3B (rs17031614), and PER3 VNTR (rs57875989), and four behavioral/clinical features: diurnal preference scores, age (≤22 or >22), gender, and socioeconomic status (poor, lower-middle-class, upper-middle-class, affluent). The seven, single nucleotide polymorphisms (SNPs) are circadian clock polymorphisms that have been reported previously in associations with sleep homeostasis and/or circadian phenotypes (diurnal preference/chronotype). To reduce multicollinearity, we performed one-hot encoding for non-binary features, removing the most frequent variant (as the baseline condition). We created 2-way genotypic combinations using the seven genes and their respective variants, giving us 7C2 * 9 total features. The most frequent class was removed for each group of 9 combinations and treated as a reference category for a combined data set of 174 total features for analysis.

Feature Selection

We used four feature selection methods (each method ten times with ten-fold cross-validation) to find epistatic combinations predictive of mood disorders: InfoGain (IG), ReliefF (ReF), Minimum Redundancy Maximum Relevance (MRMR) and Joint Mutual Information (JMI).[25–28] A subset of robust features was used for in-depth statistical analysis. A feature was considered robust if it appeared in 95% of the runs for a certain feature selection method and in at least three out of the four feature selection methods.

Classifiers

We modeled the relationship between risk factors and sleep disturbance using three classifiers: Random Forests (RF), XGBoost (XGB), and Support Vector Machines (SVM).[29–31] We evaluated the performance of our classifiers using accuracy scores and the area under the receiver operating characteristic (AUROC) curves. We used preliminary testing to determine an appropriate range of hyperparameters for each classifier and employed grid searching to determine the optimal combination of hyperparameters for maximizing accuracy. We used stratified tenfold cross-validation to determine each model's generalizability. We performed the k-nearest neighbors' imputation to fill in missing values for each fold.[32] We repeated the cross-validation procedure ten times to ensure robust results, each time using a different random number generator seed.

SMOTE Analysis

We employed Synthetic Minority Oversampling (SMOTE) technique to increase model accuracy by balancing our unbalanced dataset (with k (nearest neighbor) = 5).[33] We performed our analysis with and without SMOTE and then reported the balanced dataset results.

Statistical Analyses

All statistical analyses were performed using R.[34] We performed logistic regression analysis on each feature individually, keeping one-hot encoded features grouped together for the regression. We performed multivariate logistic regressions using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), employing a sequential replacement method to identify subsets of features with low multicollinearity and strong association with the target variable.[35] We adjust p-values from the regression analyses using the Benjamini–Hochberg method.[36] We used mediation analysis to determine whether MEQ scores were statistically significant mediators between genotypic and clinical factors and mood disorders using the R package mediate.[37] We created heatmaps illustrating patterns of association between diurnal preference and sleep disturbance using Fisher's exact tests. To identify sex-specific differences in the average PROMIS Sleep Disturbance scores for different two-way gene combinations, we used the car library in R to conduct a Type-3 Sum of Squares two-way ANOVA.[38] Tukey's follow-up tests on significant factors were performed using the emmeans library. The normality of data was assessed by Shapiro–Wilk's test in R.

We performed association rule learning analysis using the arules package.[39] We limited our analysis to rules of size at most six, with at least 90% confidence. For each sex, we found rules that code for both categories of the target variable and then sorted them by their respective lift values. We visualized these relationships in sex-specific network plots, using the igraph library.[40]

We used the Algorithm for the Reconstruction of Gene Regulatory Networks (ARACNE) to create gene networks of direct and indirect interactions between genes, clinical features, MEQ, and mood disorders using mutual information.[41] We used the minet package in R to create the network and then plotted it using Rgraphviz.[42] We employed bootstrapping to determine the frequency (ie, confidence level) with which each link in the network appears.

Results

Diurnal Preference (MEQ) and Misalignment are Strongly Associated with Sleep Disturbance

In this young adult population, we found an overall average sleep disturbance score of 20.68 (SD = 5.91), with 25% of individuals reporting sleep disturbance (n = 227). There was a significant difference in sleep disturbance scores between chronotypes ($F(2856)=15.82$, $p = 0.004$) with evening-types reporting higher sleep disturbance scores (Dunn post-hoc analysis, adj.$p<0.001$). In addition, a heat map of significant Fisher exact tests reveals that evening-type individuals (low MEQ scores) have significantly higher high sleep disturbance scores relative to morning-type individuals (high MEQ scores) (Figure 2). Additionally, after further separation according to sex females with higher MEQ scores indicative of morning preference were strongly associated with low sleep disturbance scores (Supplementary Figure 1A). While males with lower MEQ scores indicative of evening preference are associated with higher sleep disturbance scores (Supplementary Figure 1B).

Using in-vivo expression analyses to estimate the phase of circadian rhythms (a proxy for misalignment), we found that the relative risk of sleep disturbance was 4.6 times higher if an individual had a delayed phase vs advanced phase, but this risk estimate was not significant (Figure 3; OR = 4.667 (0.5538–39.3242), $p = 0.1566$; RR = 4.6). Individuals with advanced phases and a morning chronotype had significantly lower sleep disturbance scores than individuals, on average (one sample $T = 4.03$; $df = 15$, $p = 0.007$).

Our network analysis using ARACNE depicting the interactions between genotypes, clinical features, and sleep disturbance shows that all features share information with diurnal preference (MEQ), which acts as a mediator (Figure 4).

Two Way-Genotypic Combinations Can Be Predictive or Protective of Sleep Disturbance

RF, SVM, and XGB classifiers predicted sleep disturbance symptoms with an accuracy of 61–71% using all or a subset of genotypic and clinical factors chosen by feature selection methods (Figure 5). The XGB method achieves the highest accuracy (71%) when all features are used. However, if sixty features selected using the JMI method are used, a similar

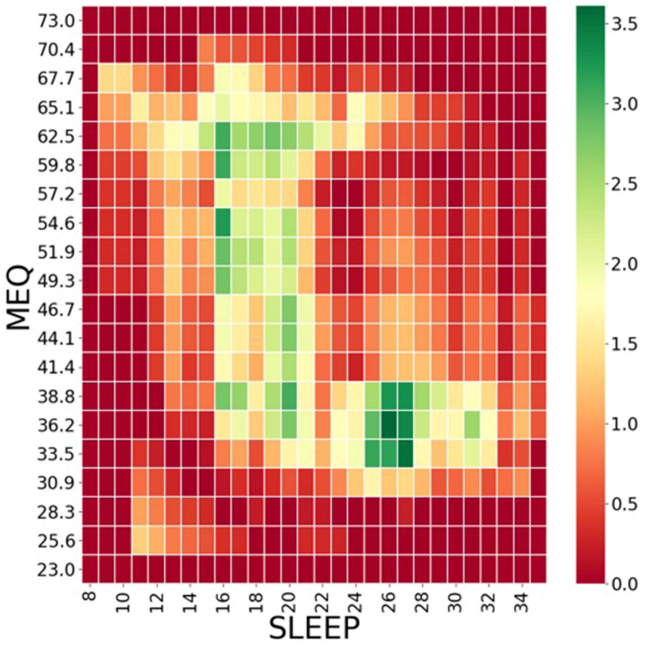

Figure 2 Heat Map of Sleep Disturbance Scores and Diurnal Preference (MEQ). Higher MEQ scores indicative of morning preferences are strongly associated with low sleep disturbance scores and lower MEQ scores indicative of evening preference are associated with higher sleep disturbance scores. The heatmaps were made using Fisher's Exact Test at varying cutoff levels for both dependent and independent variables. Fisher's Exact Test shows significant association between high sleep disturbance scores (25 +) and evening type individuals (MEQ<41). The p-values obtained from the analysis were log-transformed (base 10).

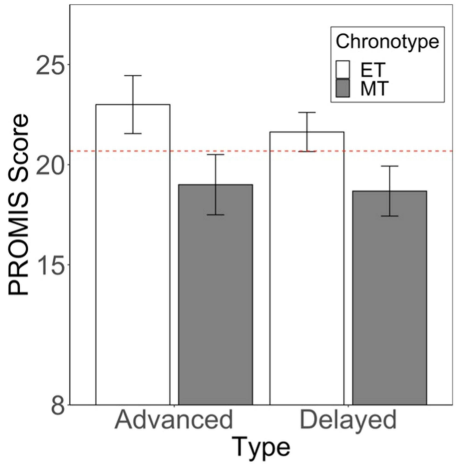

Figure 3 Circadian misalignment predictive of sleep disturbance. Approximately 25% of our sample size reported sleep disturbance with average sleep disturbance score of 20.68 (SD=5.91, SE=0.195). Advanced-morning types have significantly lower sleep disturbance scores than the average (p=0.007). The odds of sleep disturbance tended to be approximately 4.7 times higher if an individual was delayed vs advanced in their chronotype (OR=4.667 (0.5538–39.3242), p=0.157). There was a significant difference in sleep disturbance scores between ET and MT groups for which we had molecular chronotypes (F(1,63)=4.63, p=0.035). Average PROMIS sleep disturbance scores for the population sample is denoted by the red dashed line.

accuracy level (69.9%) can also be obtained using XGB and RF classifiers. These accuracy levels are 19–21% more accurate than random chance in our balanced dataset, which has a baseline accuracy of 50%.

Multivariate logistic regression analysis of the top features revealed that two-genotype combinations predicted a more robust risk of sleep disturbance symptoms when compared to single gene variants (Table 1). In the overall dataset, individuals with the *CLOCK3111*-TC and *CRY2*AG genotype had nearly three times the risk of sleep disturbance. Average sleep disturbance scores in *CLOCK3111*-TC/*CRY2*AG individuals were significantly higher than scores from individuals of other CLOCK/CRY2 genotypes (OR = 2.58 (1.38–4.88), p = 0.041). Average sleep disturbance scores for

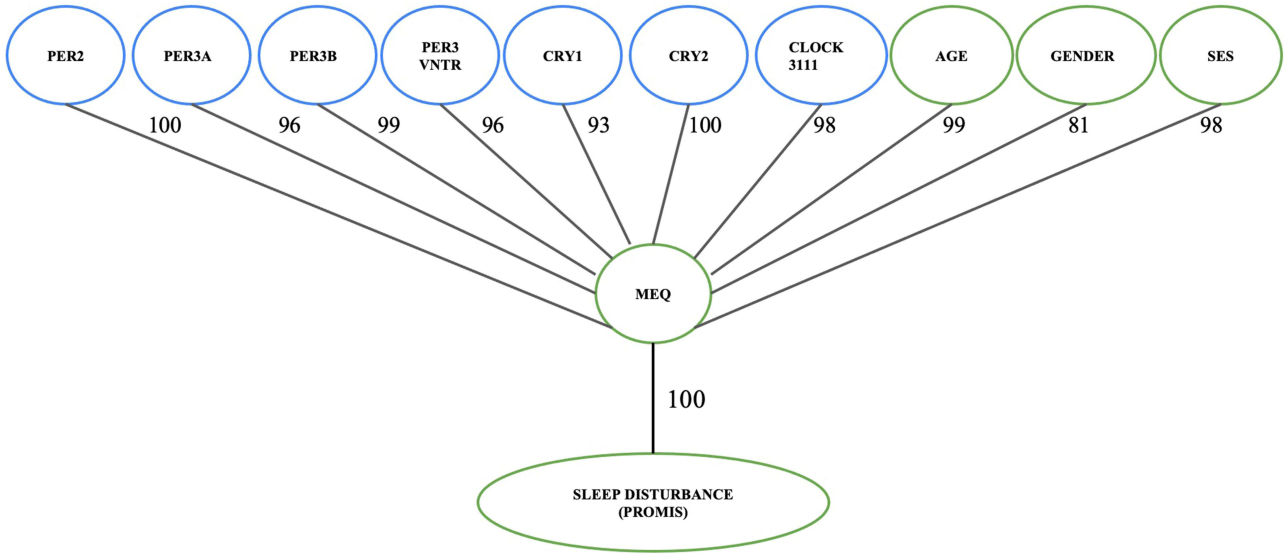

Figure 4 MEQ acts as a mediator between genotypic and clinical characteristics and human sleep disturbance. The network is constructed using the ARACNE method in conjunction with the mi-empirical method and bootstrapping. All links with bootstrap support greater than 50% are shown. MEQ mediates all associations with sleep disturbance.

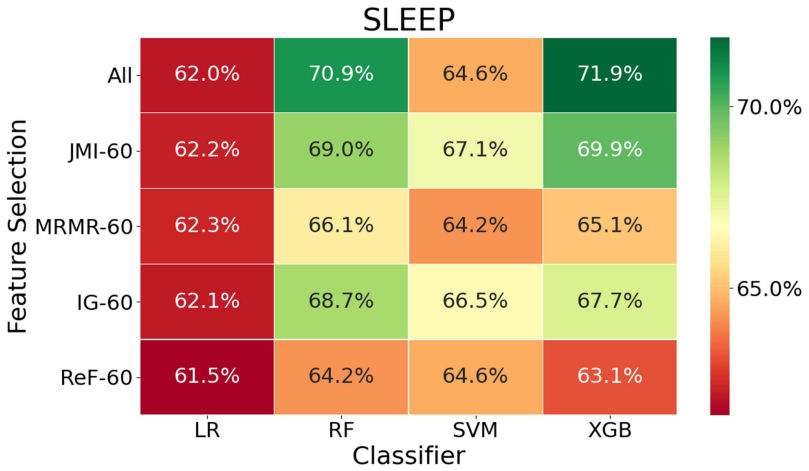

Figure 5 Heat map of prediction accuracy for feature selection and classifier methods. Our analyses yielded up to 19.9% higher prediction accuracy than baseline (50%) on a balanced data set using top features.

individuals with *CLOCK3111*_TC and *CRY2*_AG were higher for both males and females (males: 23.6 ± 1.4; females 24.5 ± 0.88) than for individuals with other genotypes (males: 20.0 ± 0.59; females 21.2 ± 0.35) (Figure 6; genotype F (1463) = 15.94, $p < 0.001$; gender F(1463) = 1.30, $p = 0.254$; genotype by gender F(1463)=0.018, $p < 0.89$). Females with the *CLOCK3111*_TC and *CRY2*_AG genotypes have significantly higher sleep disturbance scores than females with other genotypes ($p = 0.002$) and males with other genotypes ($p < 0.001$). The combination of *CLOCK3111*_TT and *CRY1*_CC also significantly increased the odds of sleep disturbance symptoms (OR = 2.69 (1.29–5.69), $p = 0.054$).

Individuals with the PER3B_AG and CRY1_CC genotypes are eleven times more protected from sleep disturbance. Average sleep disturbance scores in PER3B_AG/CRY1_CC individuals were significantly lower than scores from individuals of other PER3B/CRY1 genotypes (OR = 11.11 (0.00–0.591), $p = 0.091$; Table 1). In the large data set, the highest sleep disturbance score was reported in males with the *PER3B*_AG and *CRY1*_CG genotypes (27.2 ± 2.72) with a significant interaction between the effects of gender and genotype on sleep disturbance (Figure 7; genotype F(1430) = 2.123, $p = 0.146$, genotype F(1430) = 3.03, $p = 0.083$ genotype by gender F(1430) = 5.05, $p = 0.025$).

Table 1 Risk Factors for Sleep Disturbance Symptoms

Variable	Odds Ratio	95% CI	Adj. P
CLOCK3111_TC/CRY2_AG	2.58	1.38–4.88	0.041*
CLOCK3111_TT/CRY1_CC	2.69	1.29–5.69	0.054
PER3B_AG/CRY1_CC	0.09	0.00–0.59	0.091†
SOCIOSTATUS	0.26	0.36–0.94	0.091

Notes: *p-value <0.05; †genotype combination has a significant interaction between gender and genotype (Figure 7; genotype F(1430)=2.123, p=0.146, genotype F(1430)=3.03, p=0.083 genotype by gender F(1430) =5.05, p=0.025). Results from a multivariate logistic regression model based on top-ranked selected features identify two-way gene combinations and one demographic feature, sociostatus, that are strongly associated with sleep disturbance symptoms.

The association rule learning results identified additional genotypes and clinical feature combinations that were predictors of sleep disturbance for males and females. For females, the most frequently appearing genotypes in the top predictors of sleep disturbance were *PER3B*-GG, *CLOCK311*-TC, *CRY2*-AG, *PER2*-AG, with age and MEQ being important clinical features (Figure 8A). *CLOCK311*-TC and *CRY2*-AG were frequently found together without mediation by MEQ. For males, several features were supported by significant lift and frequency values, with no robust co-occurrence of predictive features (Figure 8B).

Discussion

Growing evidence suggests that disruptions in circadian rhythms impact sleep quality and negatively affect both mental and physical health. Previous research on the role of circadian clock gene mutations on sleep disturbance have focused on GWAS or PheWAS studies that have the advantage of large sample sizes and statistical power, but such studies are limited due to weak phenotyping capabilities and the difficulty in analyzing two or more genotype combinations with large datasets. Using a machine learning approach to explore associations between circadian genes and sleep disturbance in a deeply phenotyped population sample, we report three findings: 1) chronotype and circadian misalignment are strong predictors of self-reported sleep disturbance, 2) synergistic interactions in genotype combinations of *Period 3, Clock* and *Cryptochrome* variants (PER3B (rs17031614)/ CRY1 (rs228716) and CLOCK3111 (rs1801260)/CRY2 (rs10838524)) yield strong associations with sleep disturbance, and 3) clock genotypes predictive of sleep disturbance risk tend to have

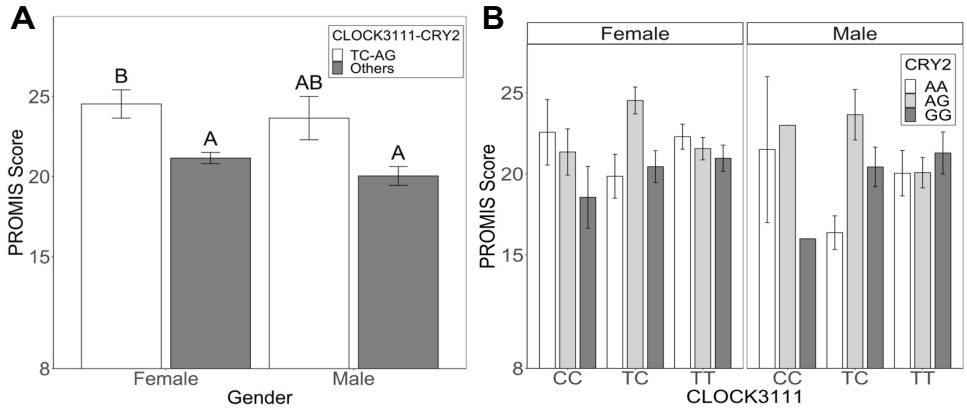

Figure 6 Genotype combination of CLOCK3111_TC and CRY2_AG predictive of sleep disturbance in humans. (**A**) Sleep disturbance scores (±SD), measured using the self-reported Patient Reported Outcomes Measurement Information System (PROMIS™) Sleep Disturbance instrument, for males and females with the *CLOCK3111*_TC and *CRY2*_AG genotype. Average sleep disturbance scores were higher for both males and females with *CLOCK3111*_TC and *CRY2*_AG (males: 23.6 ±1.4; females 24.5±0.88) than for individuals with other genotypes (males: 20.0 ±0.59; females 21.2±0.35; two-way ANOVA: genotype F(1463)=15.94, p<0.001; gender F(1463)=1.30, p=0.254; genotype by gender F(1463)=0.018, p<0.89)). (**B**) Females with the *CLOCK3111*_TC and *CRY2*_AG genotypes have significantly higher sleep disturbance scores than females with other genotypes (p=0.002) and males with other genotypes (p<0.001).

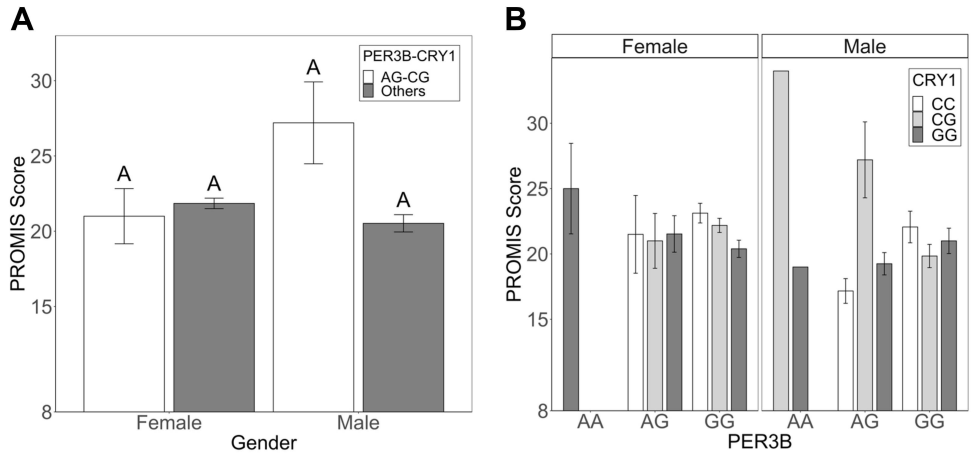

Figure 7 Genotype combination of PER3B_AG and CRY1_CG show significant interaction effect. (**A**) Sleep disturbance scores (±SD), measured using the self-reported Patient Reported Outcomes Measurement Information System (PROMIS™) Sleep Disturbance instrument, for males and females with a combination *PER3B*_AG and *CRY1*_CG genotype. Sleep disturbance scores reveal a significant gender-by-genotype interaction (genotype F(1430)=2.123, p=0.146; gender F(1430)=3.03, p=0.083; genotype by gender F(1430)=5.05, p=0.025). (**B**) The highest average sleep disturbance score was reported in males with the *PER3B*_AG and *CRY1*_CG genotypes (27.2 ±2.72), with a single male of genotype *PER3B*_AA/*CRY1*_CG recording the highest individual sleep disturbance score (34.2).

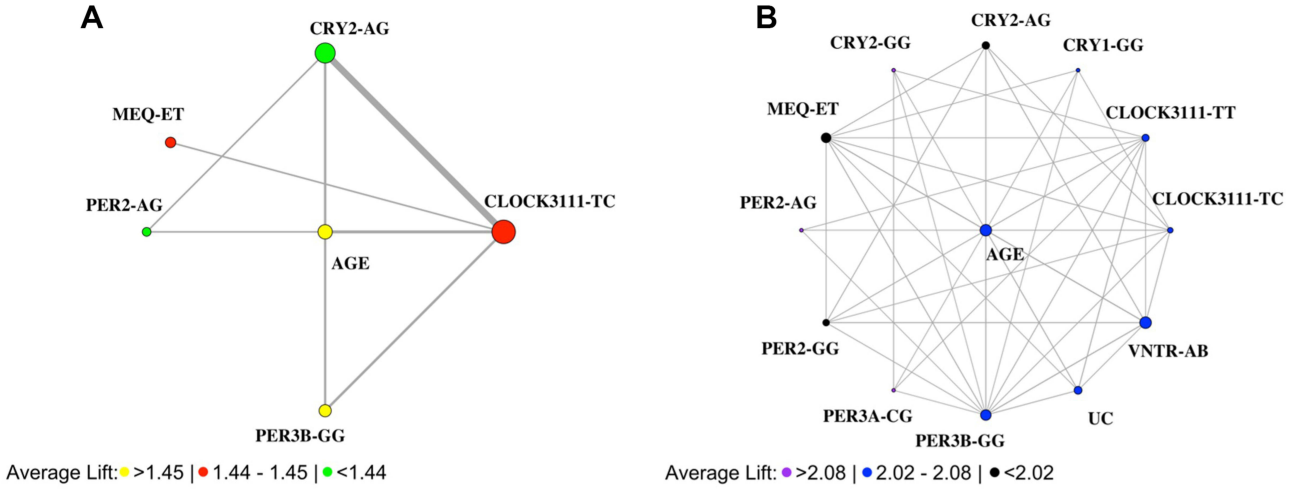

Figure 8 Association rules networks for sleep disturbance. (**A**) In females, *CLOCK311*-TC and *CRY2*-AG co-occurred most frequently with no mediation via MEQ. *PER3B*-GG and age had the highest average lift (>1.45) in the analysis. (**B**)) In males, no robust co-occurrence across genotypic and clinical features was found, although average lift values were higher for males than females.

sex-specific effects. Our results suggest that circadian clock gene variants may have both indirect (circadian clock-mediated) and direct (sex-specific mediated) effects on sleep disturbance.

Association of Sleep Disturbance with Diurnal Preference and Circadian Misalignment

Our results show that high levels of self-reported sleep disturbance are strongly associated with low MEQ scores and evening-type diurnal preferences. A potential explanation for the association between diurnal preference and sleep quality is the influence of circadian misalignment on physiological regulation of sleep drive. Individuals with extreme diurnal preference are more likely to encounter misalignment between their physiological circadian rhythms and their sleep-wake patterns. In the current study, the participants were primarily undergraduates encountering a college lifestyle; evening-type individuals may be inclined to stay up even later due to social and academic pressures but are also compelled to wake up early to attend classes or athletic training. To further test whether increased sleep disturbance and eveningness could result from circadian misalignment, we examined the circadian phase of a subset of participants and

identified participants with a mismatch between their self-reported diurnal preference and their circadian phase. As expected, individuals with delayed phases and evening chronotypes were significantly more sleep disturbed than other individuals. Interestingly, evening-type individuals with advanced phases were significantly more sleep disturbed than individuals with advanced phases and morning chronotypes. Individuals with both an advanced phase and morning chronotype experienced significantly less sleep disturbance than other individuals. Overall, our results confirm that eveningness is strongly associated with sleep disturbance and suggest that circadian misalignment may be an additional contributing factor to sleep disruption.

In addition, our network analysis reveals that all effects of circadian mutations on sleep disturbance tested in this study (as well as clinical factors) are mediated via diurnal preference scores. Connections between diurnal preference or chronotype and circadian-related genes have been reported numerous times in the literature.[43–47] One of the most robust connections involves the circadian clock gene, *Per3*. Several studies have reported that the longer allele of the PER3 VNTR polymorphism (rs57875989) is associated with morning-types.[4,24,46,48–52] Archer et al (2003) also found that found that prevalence of DSPD in individuals with the 5-repeat allele was low.[43] In other studies, associations between PER3 VNTR variants and diurnal preference are less clear. For example, Turco et al (2017) found no significant association between the length polymorphism (rs57875989)[48] and diurnal preference. Overall, there is ample evidence that diurnal preferences in some populations are mediated by circadian gene mutations and these circadian phenotypes can be associated with sleep disturbance.

Circadian Genotypes Predictive of Sleep Disturbance

Using feature selection methods, we identified multiple genotype combinations of two or more gene variants that provide robust predictions of sleep disturbance risk; the predictive value of these genotype combinations is higher than any predictive value for single variants (Table 1). Using multivariate regression, we found that participants that were heterozygous for variants in CLOCK (TC) and CRY2 (AG) genes had nearly three times the risk of sleep disturbance relative to individuals of other CLOCK/CRY2 genotypes and double homozygous individuals with CLOCK (TT) and CRY1 (CC) genotypes have a similar risk of sleep disturbance, although this association is only significant at adjusted alpha level of 0.1. Our results suggest that that synergistic effects of multiple mutations in the molecular clockwork may directly or indirectly modulate physiological pathways associated with sleep disturbance.

In the primary circadian feedback loop, the CRY genes work in the retina as light-independent inhibitors of the CLOCK/BMAL heterodimers.[53] These genes feedback to inhibit CLOCK/BMAL activity to eventually suppress their own expression.[54] The CRY1 variant (rs2287161) SNP is located downstream from the CRY1 polyadenylation site, suggesting that other regulatory features are involved in the modulation of CRY1 gene expression. In previous studies, CRY1 has been linked to both depression and disruptions in sleep regulation.[55–57] Wisor et al (2002) found that mice without CRY1 or CRY2 exhibited high non-REM (eg, quiet, non-dreaming, quiescent state) sleep drive, suggesting cryptochrome genes influence the homeostatic regulation of sleep.[58] This condition was also associated with high mRNA levels of PER1 and PER2 which are normally inhibited by CRY1/CRY2 in the CLOCK/BMAL negative feedback loop.[58] In humans, Patke et al (2017) discovered a link between a familial delayed sleep phase disorder (DSPD) and a hereditary dominant, gain of function variant in CRY1. This mutation enhances the function of CRY1, causing reduced expression of specific transcriptional targets and a lengthening of the circadian period, leading to DSPD symptoms such as delayed sleep onset and offset times.[9]

Previous studies of circadian influences on mental health reveal multiple associations of the CRY2 (rs10838524) polymorphism with mood disorders, including associations with dysthymia, winter depression and Type I bipolar disorder.[59–61] The diverse associations between a single CRY2 polymorphism and mood underscores the potential links between circadian clock genes and physiological pathways central to mental health. Because mood disorders are often co-morbid with sleep disorders, it is possible that CRY2 may play a role in mediating pathways affecting sleep disorders as well. Interestingly, a study of bipolar patients by Lavebratt and colleagues (2010) revealed marked differences in CRY2 mRNA levels in control participants relative to bipolar participants following one night of sleep deprivation.[60] These results suggest that dysregulation of CRY2 may be linked to physiological changes caused by sleep

deprivation. Collectively, these findings suggest that cryptochrome genes play a role in both circadian and sleep regulation.

In our study, individuals with a PER3B (AG)/CRY1 (CC) genotype were 11 times more likely to be protected from sleep disturbance, although the association was only significant with an adjusted alpha value of 0.1. Interestingly, the negative association of PER3B (AG) and CRY1 (CC) with sleep disturbance in our study is driven by genotype–gender interactions in males and this effect is completely mediated by diurnal preference. This suggests that circadian features may indirectly influence sleep quality in males with the PER3B-CRY1 (AG-CC/CG) genotype. The PER3 gene has been implicated in sleep architecture and circadian rhythm disruptions in previous studies.[4,10,46–49] Most importantly, researchers have found that links between PER3 polymorphisms, diurnal preference, and mood provide evidence for the "phase-shift hypothesis" which supports that some depressive symptoms in humans can be linked to a misalignment of the circadian phase with the sleep-wake cycle.[47] A seminal study identified two missense variations in PER3 (PER3-P415A/H417R) in humans with familial advanced sleep phase disorder (FASPD) which are also associated with elevated Beck Depression Inventory (BDI) and seasonality scores indicative of seasonal affective disorder (SAD).[62] Additional experiments on these mutations in mice resulted in mice with delayed circadian phases in short photoperiods and longer durations of wheel-running activity in constant light. Further experiments suggested that the products of PER3-P415A /H417R variants effectively reduce the stabilization of PER1 and PER2, leading to the observed changes in phase and lengths of circadian rhythms.[62]

The *Per3* gene has also been linked to sleep phenotypes, independently of circadian rhythms.[3] One interesting finding in our study is the lack of association between the PER3 variable number of tandem repeat region (VNTR; rs57875989) and measures of sleep disturbance. This length polymorphism has been associated with altered responses to sleep loss and gene expression in pathways involved in sleep homeostasis in mice,[60] as well as sleep architecture in humans.[4,10,22,43] For example, Lazar et al (2012) found morning preference, earlier wake time and bedtime, and reduced daytime sleepiness was associated with the longer VNTR allele ($PER3^{5/5}$).[52] This length polymorphism has also been associated with both sleep duration and depressive symptoms, suggesting that PER3 may be involved in the regulation of both sleep homeostasis and mood pathways in humans.[11,43,47,48] In our study, we expected that we might find the PER3 VNTR polymorphism to be predictive of sleep disturbance values, but we did not uncover significant predictions of risk or protection from risk associated with PER3 VNTR genotypes. It is possible that larger sample sizes are required to identify effects of PER3 VNTR on sleep disturbance—or that sleep structure, but not necessarily sleep quality is affected by variants of this gene.

Sex-Specific Circadian Genotype Associations with Sleep Quality

In addition to the significant genotype-by-gender interaction in the PER3B/CRY1 results, our network analyses showed clear differences between males and females in genotype risk factors for sleep disturbance. For females, the most frequently appearing genotypes in the top predictors of sleep disturbance were *PER3B*-GG, *CLOCK3111*-TC, *CRY2*-AG, and *PER2*-AG, with age and MEQ being important clinical features. Our multivariate regression analysis provided strong evidence that individuals with both *CLOCK3111*-TC and *CRY2*-AG genotypes were more likely to report sleep disturbance. For males, a number of features were supported by significant lift and frequency values, with no single predictive features stand out. Average lifts were higher for all of the top rules for males relative to females, but co-occurrence was more widely distributed across the genotypes and clinical characteristics. Previously, Shi et al (2016) revealed that some variants in circadian clock-related genes are associated with major depressive disorder (MDD) in a sex dependent manner.[63] After a sex-stratified analyses they identified two variants CLOCK (rs1801260) and NPAS2 (rs34705978) that were significantly associated with MDD in only one gender. The CLOCK polymorphism was significantly associated with MDD in males ($P = 0.028$) and not females.[63] In our study, the CLOCK genotype TC is significantly associated with sleep disturbance in the population when found in combination with the CRY2 AG genotype. However, our collective analyses reveal a stronger association for this genotypic combination with sleep disturbance in females. Thus, the effects of CLOCK mutations may be mediated through sex-dependent pathways in conjunction with or independent of core clock pathways. Further studies are needed to understand the sex-dependent effects of circadian clock genes on sleep architecture and features, to fully understand this difference.

Limitations

The limitations of the current study include the study of a large number of features in the machine learning analyses using a relatively limited sample size.[64] We did not study, and thus could not remove the confounding variables of environmentally induced sleep disturbance and alcohol and drug use, although we did remove individuals that use melatonin supplements from the population sample. Our population sample includes primarily young adult Caucasians of European descent, which limits the generalization of our findings to more diverse populations. In addition, the average sleep disturbance scores in the study indicate that many young adults in this sample do not experience severe sleep disturbance which may limit the generalization of the findings to older or younger populations. Future studies should assess the precision of our analyses, including our risk factor predictors and top features analyses, using an independent population. Our study is limited outcomes reporting associations between sleep disturbance and particular genotypes or circadian phenotypes; future studies should use this information to further understand causal connections between these variables and how potential therapeutic targets could alleviate sleep distress in humans.

Conclusion

Using a machine learning approach to study the association between circadian clock genes, diurnal preference, and clinical factors with sleep disturbance, our study reveals both direct and indirect, via mediation by circadian phenotypes, effects of circadian clockwork on self-reported sleep quality. Our results support the conclusions that chronotype and circadian misalignment are strong predictors of self-reported sleep disturbance. In addition, we report novel synergistic interactions between clock gene variants that yield strong associations with sleep disturbance and these associations tend to be sex-specific. Our results show evidence for circadian influences on sleep quality, but the role of circadian genes in regulating sleep factors is at a nascent stage. Understanding how clock genes and circadian features influence sleep quality has the potential to reveal clinical targets and other forms of therapy that may help alleviate sleep disturbance.

Data Sharing Statement

All data and coding software are available on Github: https://github.com/aziz-zafar/Mood-Disorder.

Funding

Funding was provided by Colgate University Research Council and Picker Interdisciplinary Science Institute grants to KKI.

Disclosure

The authors declare no competing interests.

References

1. Fang H, Tu S, Sheng J, Shao A. Depression in sleep disturbance: a review on a bidirectional relationship, mechanisms and treatment. *J Cell Mol Med*. 2019;23(4):2324–2332. doi:10.1111/jcmm.14170
2. Utge SJ, Soronen P, Loukola A, et al. Systematic analysis of circadian genes in a population-based sample reveals association of TIMELESS with depression and sleep disturbance. *PLoS One*. 2010;5(2):e9259. doi:10.1371/journal.pone.0009259
3. Borbély AA, Daan S, Wirz-Justice A, Deboer T. The two-process model of sleep regulation: a reappraisal. *J Sleep Res*. 2016;25(2):131–143. doi:10.1111/jsr.12371
4. Dijk DJ, Archer SN. PERIOD3, circadian phenotypes, and sleep homeostasis. *Sleep Med Rev*. 2010;14(3):151–160. doi:10.1016/j.smrv.2009.07.002
5. Evans JA, Pan H, Liu AC, Welsh DK. Cry1-/- circadian rhythmicity depends on SCN intercellular coupling. *J Biol Rhythms*. 2012;27(6):443–452. doi:10.1177/0748730412461246
6. Walker WH, Walton JC, DeVries AC, Nelson RJ. Circadian rhythm disruption and mental health. *Transl Psychiatry*. 2020;10(1). doi:10.1038/S41398-020-0694-0
7. Jagannath A, Taylor L, Wakaf Z, Vasudevan SR, Foster RG. The genetics of circadian rhythms, sleep and health. *Hum Mol Genet*. 2017;26(R2):R128–R138. doi:10.1093/hmg/ddx240
8. Reppert SM, Weaver DR. Coordination of circadian timing in mammals. *Nature*. 2002;418(6901):935–941. doi:10.1038/nature00965
9. Patke A, Murphy PJ, Onat OE, et al. Mutation of the Human Circadian Clock Gene CRY1 in Familial Delayed Sleep Phase Disorder. *Cell*. 2017;169(2):203–215.e13. doi:10.1016/j.cell.2017.03.027
10. Archer SN, Schmidt C, Vandewalle G, Dijk DJ. Phenotyping of PER3 variants reveals widespread effects on circadian preference, sleep regulation, and health. *Sleep Med Rev*. 2018;40:109–126. doi:10.1016/j.smrv.2017.10.008

11. Weiss C, Woods K, Filipowicz A, Ingram KK. Sleep Quality, Sleep Structure, and PER3 Genotype Mediate Chronotype Effects on Depressive Symptoms in Young Adults. *Front Psychol*. 2020;11(August):1–11. doi:10.3389/fpsyg.2020.02028
12. Duffy JF, Abbott SM, Burgess HJ, et al. Circadian rhythm sleep-wake disorders: gaps and opportunities. *Sleep*. 2021;44(5):1–15. doi:10.1093/sleep/zsaa281
13. Chellappa SL. Individual differences in light sensitivity affect sleep and circadian rhythms. *Sleep*. 2021;44(2):1–10. doi:10.1093/sleep/zsaa214
14. Challet E, Caldelas I, Graff C, Pévet P. Regulation of Molecular Oscillations in Extra- SCN Brain and Peripheral Oscillators. *Acta Physiol*. 2003;384:711–719. doi:10.1515/BC.2003.079
15. Mistlberger RE, Landry GJ, Marchant EG. Sleep deprivation can attenuate light-induced phase shifts of circadian rhythms in hamsters. *Neurosci Lett*. 1997;238(1–2):5–8. doi:10.1016/s0304-3940(97)
16. van Diepen HC, Lucassen EA, Yasenkov R, et al. Caffeine increases light responsiveness of the mouse circadian pacemaker. *Eur J Neurosci*. 2014;40(10):3504–3511. doi:10.1111/ejn.12715
17. Bottaro LB. Partial sleep deprivation reduces phase advances to light in humans. *J Biol Rhythms*. 2010;25(6):460–468. doi:10.1177/0748730410385544.Partial
18. Russo PM, Biasi V, Cipolli C, Mallia L, Caponera E. Sleep habits, circadian preference, and school performance in early adolescents. *Sleep Med*. 2017;29:20–22. doi:10.1016/j.sleep.2016.09.019
19. Albrecht U. Molecular mechanisms in mood regulation involving the circadian clock. *Front Neurol*. 2017;8(FEB). doi:10.3389/fneur.2017.00030
20. Yu L, Buysse DJ, Germain A, et al. Development of Short Forms From the PROMISTM Sleep Disturbance and Sleep-Related Impairment Item Banks. *Behav Sleep Med*. 2012;10(1):6–24. doi:10.1080/15402002.2012.636266
21. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int J Chronobiol*. 1976;4(2):97–110.
22. Ebisawa T, Uchiyama M, Kajimura N, et al. Association of structural polymorphisms in the human period3 gene with delayed sleep phase syndrome. *EMBO Rep*. 2001;2(4):342–346. doi:10.1093/embo-reports/kve070
23. Nguyen C, Murray G, Anderson S, Filipowicz A, Ingram KK. In vivo molecular chronotyping, circadian misalignment, and high rates of depression in young adults. *J Affect Disord*. 2019;250:425–431. doi:10.1016/j.jad.2019.03.050
24. Liberman AR, Kwon SB, Vu HT, Filipowicz A, Ay A, Ingram KK. Circadian Clock Model Supports Molecular Link between PER3 and Human Anxiety. *Sci Rep*. 2017;7(1):1–10. doi:10.1038/s41598-017-07957-4
25. Quinlan JR, Quinlan JR. Induction of decision trees. *Mach Learn*. 1986;1:81–106. doi:10.1007/BF00116251
26. Robnik-Sikonja M, Kononenko I An Adaptation of Relief for Attribute Estimation in Regression. In: Proceedings of the Fourteenth International Conference on Machine Learning. ICML '97. Morgan Kaufmann Publishers Inc; 1997:296–304.
27. Ding C, Peng H Minimum redundancy feature selection from microarray gene expression data. In: Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003; 2003: 523–528. doi:10.1109/CSB.2003.1227396.
28. Yang HH, Moody JE Feature Selection Based on Joint Mutual Information. 1999.
29. Boser BE, Guyon I, Vapnik VN A training algorithm for optimal margin classifiers. In: COLT '92; 1992.
30. Chen T, Guestrin C XGBoost: a Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16. Association for Computing Machinery; 2016:785–794. doi:10.1145/2939672.2939785.
31. Statistics LB, Breiman L. Random Forests. *Mach Learn*. 2001;1;5–32.
32. Fukunaga K. The Optimal Distance Measure for Nearest Neighbor Classification. *IEEE Trans Inf Theory*. 1981;27(5):622–627. doi:10.1109/TIT.1981.1056403
33. Chawla N, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic Minority over-Sampling Technique. *J Artif Int Res*. 2002;16(1):321–357.
34. R Core Team. R: a Language and Environment for Statistical Computing; 2020. Available from: https://www.r-project.org/. Accessed October 13, 2022.
35. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. *Psychometrika*. 1987;52(3):345–370. doi:10.1007/BF02294361
36. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Statistical Soc*. 1995;57(1):289–300. doi:10.1111/J.2517-6161.1995.TB02031.X
37. Imai K, Keele L, Tingley D, Yamamoto T. Causal Mediation Analysis Using R. In: Vinod HD, editor. *Advances in Social Science Research Using R*. Springer-Verlag; 2010.
38. Fox J, Weisberg S An {R} Companion to Applied Regression. Sage; 2019. Available from: https://socialsciences.mcmaster.ca/jfox/Books/Companion/. Accessed October 13, 2022.
39. Hahsler M, Chelluboina S, Hornik K, Buchta C. The arules R-Package Ecosystem: analyzing Interesting Patterns from Large Transaction Datasets. *J Mach Learn Res*. 2011;12:1977–1981.
40. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal*. 2006;1;1695.
41. Margolin AA, Nemenman I, Basso K, et al. ARACNE: an Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. *BMC Bioinform*. 2006;7(1):1–15. doi:10.1186/1471-2105-7-S1-S7
42. Meyer PE, Lafitte F, Bontempi G. MINET: an open source R/Bioconductor Package for Mutual Information based Network Inference. *BMC Bioinform*. 2008;9. doi:10.1186/1471-2105-9-9
43. Archer SN, Robilliard DL, Skene DJ, et al. A length polymorphism in the circadian clock gene Per3 is linked to delayed sleep phase syndrome and extreme diurnal preference. *Sleep*. 2003;26(4):413–415. doi:10.1093/sleep/26.4.413
44. Partonen T, Treutlein J, Alpman A, et al. Three circadian clock genes Per2, Arntl, and Npas2 contribute to winter depression. *Ann Med*. 2007;39(3):229–238. doi:10.1080/07853890701278795
45. Zhang L, Hirano A, Hsu PK, et al. A PERIOD3 variant causes a circadian phenotype and is associated with a seasonal mood trait. *Proc Natl Acad Sci U S A*. 2016;113(11):E1536–E1544. doi:10.1073/pnas.1600039113
46. Hida A, Kitamura S, Katayose Y, et al. Screening of clock gene polymorphisms demonstrates association of a PER3 polymorphism with morningness-eveningness preference and circadian rhythm sleep disorder. *Sci Rep*. 2014;4:6309. doi:10.1038/srep06309
47. Liberman AR, Halitjaha L, Ay A, Ingram KK. Modeling Strengthens Molecular Link between Circadian Polymorphisms and Major Mood Disorders. *J Biol Rhythms*. 2018;33(3):318–336. doi:10.1177/0748730418764540
48. Turco M, Biscontin A, Corrias M, et al. Diurnal preference, mood and the response to morning light in relation to polymorphisms in the human clock gene PER3. *Sci Rep*. 2017;7(1):1–10. doi:10.1038/s41598-017-06769-w

49. Archer SN, Oster H. How sleep and wakefulness influence circadian rhythmicity: effects of insufficient and mistimed sleep on the animal and human transcriptome. *J Sleep Res*. 2015;24(5):476–493. doi:10.1111/jsr.12307
50. Fares S, Hermens DF, Naismith SL, White D, Hickie IB, Robillard R. Clinical correlates of chronotypes in young persons with mental disorders. *Chronobiol Int*. 2015;32(9):1183–1191. doi:10.3109/07420528.2015.1078346
51. Levandovski R, Dantas G, Fernandes LC, et al. Depression scores associate with chronotype and social jetlag in a rural population. *Chronobiol Int*. 2011;28(9):771–778. doi:10.3109/07420528.2011.602445
52. Lázár AS, Slak A, Lo JCY, et al. Sleep, diurnal preference, health, and psychological well-being: a prospective single-allelic-variation study. *Chronobiol Int*. 2012;29(2):131–146. doi:10.3109/07420528.2011.641193
53. Vieira J, Jones AR, Danon A, et al. Human Cryptochrome-1 Confers Light Independent Biological Activity in Transgenic Drosophila Correlated with Flavin Radical Stability. *PLoS One*. 2012;7(3):e31867. doi:10.1371/journal.pone.0031867
54. Dardente H, Fortier EE, Martineau V, Cermakian N. Cryptochromes impair phosphorylation of transcriptional activators in the clock: a general mechanism for circadian repression. *Biochem J*. 2007;402:525–536. doi:10.1042/BJ20060827
55. Hua P, Liu W, Chen D, et al. Cry1 and Tef gene polymorphisms are associated with major depressive disorder in the Chinese population. *J Affect Disord*. 2014;157:100–103. doi:10.1016/j.jad.2013.11.019
56. Buoli M, Serati M, Grassi S, et al. The role of clock genes in the etiology of Major Depressive Disorder: special Section on "Translational and Neuroscience Studies in Affective Disorders". Section Editor, Maria Nobile MD, PhD. This Section of JAD focuses on the relevance of translational a. *J Affect Disord*. 2018;234:351–357. doi:10.1016/j.jad.2017.11.015
57. Soria V, Martínez-Amorós È, Escaramís G, et al. Differential association of circadian genes with mood disorders: CRY1 and NPAS2 are associated with unipolar major depression and clock and VIP with bipolar disorder. *Neuropsychopharmacology*. 2010;35(6):1279–1289. doi:10.1038/npp.2009.230
58. Wisor JP, O'Hara BF, Terao A, et al. A role for cryptochromes in sleep regulation. *BMC Neurosci*. 2002;3:20. doi:10.1186/1471-2202-3-20
59. Partonen T. Clock gene variants in mood and anxiety disorders. *J Neural Transm*. 2012;119(10):1133–1145. doi:10.1007/s00702-012-0810-2
60. Lavebratt C, Sjöholm LK. CRY2 is associated with depression. *PLoS One*. 2010;5(2):e9407. doi:10.1371/journal.pone.0009407
61. Kovanen L, Kaunisto M, Donner K, Saarikoski ST, Partonen T. *CRY2* genetic variants associate with dysthymia. *PLoS One*. 2013;8(8):e71450. doi:10.1371/journal.pone.0071450
62. Zhang L, Hirano A, Hsu PK, et al. A PERIOD3 variant causes a circadian phenotype and is associated with a seasonal mood trait. *Proc Natl Acad Sci U S A*. 2016;113(11):E1536. doi:10.1073/pnas.1600039113
63. Shi SQ, White MJ, Borsetti HM, et al. Molecular analyses of circadian gene variants reveal sex-dependent links between depression and clocks. *Transl Psychiatry*. 2016;6(December):2015. doi:10.1038/tp.2016.9
64. Quinlan JR, Quinlan JR. Induction of Decision Trees. *Mach Learn*. 1986;2(1):81–106.