

Article

Risk for Seasonal Affective Disorder (SAD) Linked to Circadian Clock Gene Variants

Thanh Dang ¹, William A. Russel ², Tazmilur Saad ^{1,3}, Luvna Dhawka ⁴, Ahmet Ay ^{2,3} and Krista K. Ingram ^{2,*} 

¹ Department of Computer Science, Colgate University, Hamilton, NY 13346, USA; thdang@colgate.edu (T.D.); ssaad@colgate.edu (T.S.)

² Department of Biology, Colgate University, Hamilton, NY 13346, USA; wrussel@colgate.edu (W.A.R.); aay@colgate.edu (A.A.)

³ Department of Mathematics, Colgate University, Hamilton, NY 13346, USA

⁴ Feil Family Brain & Mind Research Institute, Weill Cornell Medicine, New York, NY 10065, USA

* Correspondence: kingram@colgate.edu

Simple Summary: Depression and related mood disorders, such as seasonal affective disorder (SAD), affect nearly one-fifth of the global population and have substantial economic and personal costs. Mutations in circadian clock genes are associated with mood disorders, suggesting that molecular pathways affecting mood are influenced by the circadian clock, but the molecular mechanisms underlying this link are poorly understood. We use machine learning and statistical analyses to determine the circadian gene variants and clinical features most highly associated with symptoms of seasonality and SAD. Our results support previous links between CRY2, PER2, and ZBTB20 genes and identify novel links for CLOCK and PER3 with symptoms of seasonality and SAD. Our findings reinforce the sex-specific influences of circadian genes on seasonality and SAD and highlight the multiple pathways by which clock variants affect downstream mood pathways via direct and indirect mechanisms. These results enhance our understanding of the molecular mechanisms underlying seasonality in mood and may elucidate new therapeutic targets for circadian-related mood disorders, including SAD, depression, and anxiety.



Citation: Dang, T.; Russel, W.A.; Saad, T.; Dhawka, L.; Ay, A.; Ingram, K.K. Risk for Seasonal Affective Disorder (SAD) Linked to Circadian Clock Gene Variants. *Biology* **2023**, *12*, 1532. <https://doi.org/10.3390/biology12121532>

Academic Editors: Gianluigi Ubaldo Mazzoccoli and Marina Maria Bellet

Received: 17 November 2023

Revised: 11 December 2023

Accepted: 13 December 2023

Published: 15 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Molecular pathways affecting mood are associated with circadian clock gene variants and are influenced, in part, by the circadian clock, but the molecular mechanisms underlying this link are poorly understood. We use machine learning and statistical analyses to determine the circadian gene variants and clinical features most highly associated with symptoms of seasonality and seasonal affective disorder (SAD) in a deeply phenotyped population sample. We report sex-specific clock gene effects on seasonality and SAD symptoms; genotypic combinations of CLOCK3111/ZBTB20 and PER2/PER3B were significant genetic risk factors for males, and CRY2/PER3C and CRY2/PER3-VNTR were significant risk factors for females. Anxiety, eveningness, and increasing age were significant clinical risk factors for seasonality and SAD for females. Protective factors for SAD symptoms (in females only) included single gene variants: CRY1-GG and PER3-VNTR-4,5. Clock gene effects were partially or fully mediated by diurnal preference or chronotype, suggesting multiple indirect effects of clock genes on seasonality symptoms. Interestingly, protective effects of CRY1-GG, PER3-VNTR-4,5, and ZBTB20 genotypes on seasonality and depression were not mediated by chronotype, suggesting some clock variants have direct effects on depressive symptoms related to SAD. Our results support previous links between CRY2, PER2, and ZBTB20 genes and identify novel links for CLOCK and PER3 with symptoms of seasonality and SAD. Our findings reinforce the sex-specific nature of circadian clock influences on seasonality and SAD and underscore the multiple pathways by which clock variants affect downstream mood pathways via direct and indirect mechanisms.

Keywords: circadian clock; seasonal affective disorder; SAD; depression; seasonality; machine learning; mood disorders; chronobiology; molecular clockwork

1. Introduction

The global prevalence of depression and related mood disorders, including seasonal affective disorder, incurs significant personal and economic costs [1,2]. Individuals affected by mood disorders are frequently plagued by sleep problems and disruptions in their circadian rhythms and daily cycles in physiological processes and behavior, indicating a connection between circadian physiology and mood pathways [3–5]. In addition, mutations in core regulatory clock genes are associated with sleep–wake patterns and mood disorders, suggesting that molecular pathways regulated by the circadian clock may influence mood [6].

Circadian misalignment or ‘social jetlag’, the degree to which the circadian phase of an individual is advanced or delayed relative to their entrained (or socially influenced, i.e., work demands) sleep–wake cycle, is a likely factor influencing negative affect [7–11]. Multiple candidate gene studies have reported circadian clock polymorphisms associated with both chronotype and mood disorders [4,5,12–17], with some clock variants displaying sex-specific patterns [18]. Genome-wide association studies (GWAS) also link circadian genes with chronotype, depression, and other psychological disorders [19–22]. In addition, studies examining individuals with severe circadian disorders, such as the familial advanced sleep phase (FASP) and delayed sleep phase (DSP) phenotypes, support the high frequency of comorbidity between circadian disorders and negative affect [6,22,23]. Although the results from these studies show robust support for the influence of circadian misalignment on mood, the mechanisms driving these gene–environment interactions are poorly understood.

Seasonal affective disorder (SAD) is an excellent model with which to test for gene–environment interactions. SAD is a sub-type of depression that normally occurs in the fall/winter seasons and is most frequently reported at latitudes that experience shorter photoperiods or hours of daylight (but see [24] for a review of the heterogeneity of SAD symptoms in general populations). The disorder is associated with hypersomnia, social withdrawal, overeating, low energy, and carbohydrate cravings [25]. The diagnosis of SAD requires that patients experience symptoms for a minimum of two years. SAD’s prevalence in the general population can be as high as 4–15%, and SAD is four times more prevalent in women [25]. While depression and anxiety can have a myriad of contributing risk factors [26], seasonal affective disorder is likely to be more tightly linked to the circadian clock. Another important consideration is the fact that global populations harbor diverse genetic backgrounds that may or may not be well-adapted to short-light days. There is ample evidence to suggest that frequencies of circadian gene variants differ with latitude, population history, and environmental factors [27]. Finally, the prevalence of SAD tends to be higher in females. A recent study highlighted the sex-specific association of circadian gene variants with MDD [18], suggesting the possibility that sex-specific molecular pathways may be linked to SAD.

Large-scale approaches to genomic association studies (e.g., GWAS) utilize genomic databases to identify single gene variants associated with a phenotype, including complex mental health disorders like seasonal depression. Still, many GWAS studies fail to identify clinically or biologically relevant associations. Newly described methods, like phenome-wide association studies (PheWAS), attempt to solve this issue by using a genotype-to-phenotype approach utilizing medical databases with more robust phenotyping. However, these studies are costly, and it is still challenging to extrapolate beyond single-gene associations.

In this study, we seek to explore a complementary approach to determine key genetic associations of genotypes with complex behavioral disorders. Taking advantage of PheWAS approaches, we use a combination of machine learning and statistical approaches to identify haplotypes or combinations of multi-gene variants associated with SAD in a population of individuals that are deeply phenotyped. This approach is particularly relevant for the study of complex behavioral disorders, as previous studies of genetic influence on mental health strongly suggest an increased risk due to the additive effects of multiple gene

variants [4,17,28]. It is also becoming increasingly clear that environmentally influenced mental health disorders, like seasonal affective disorder, are strongly influenced by genes–environment interactions. These subtleties are difficult to test using ‘lightly’ phenotyped or non-clinical population samples.

In cross-sectional studies, we and others have found significant associations between circadian gene variants and depression, anxiety, sleep disturbance, sleep architecture, and SAD (reviewed in [29,30]). Our longitudinal study in a predominantly Caucasian population found a prevalence of 12.5% for SAD symptoms, with the risk of SAD being five times higher in females than in males. By targeting previously identified SNPs from depressive mood disorders, we use a deeply phenotyped population sample to identify sex-specific genetic associations of SAD with circadian gene variants, clinical features (depressive symptoms/anxiety symptoms/chronotype/sleep issues, gender), and environmental features (seasons, socioeconomic status) using machine learning and Bayesian approaches.

2. Materials and Methods

The genetic and statistical methods described below were used in our previous study on circadian influences on anxiety with slight modifications in this study. More details can be found in Zafar et al., 2022 [31].

2.1. Participants

Data were collected from a multi-year study of a population of predominantly young Caucasian adults (total sample: $n = 891$; participants with complete genetic data: $n = 273$; 76 males, 197 females; median age: 19, range 17–79). Twenty-nine participants were faculty at the university; removing these older individuals did not alter the results, so they were included in the study. Participants were recruited primarily via an introductory Biology course comprising first- or second-year undergraduates during the years 2013–2019. Sampling occurred in late fall or early spring (November–March). The participation rate for each sampling period ranged from 94 to 99%. All methods adhered to the principles of the Declaration of Helsinki; the Institutional Review Board at Colgate University approved all procedures and consent forms (#FR-F13-07, #ER-F14-12, #F15-13, #ER-F16-19). All participants gave written informed consent.

2.2. Self-Report Surveys

We established a dataset using deep phenotyping via computer-based surveys, including the Seasonal Pattern Assessment Questionnaire (SPAQ ($n = 221$); Rosenthal et al., 1984 [32]), the Beck Depression Inventory (BDI-II ($n = 254$); Beck and Beamesderfer 1974 [33]), Spielberger’s State–Trait Anxiety Scale (State–Trait Anxiety Inventory (STAI; Spielberger et al., 1983 [34]), and the short form of the Patient-Reported Outcomes Measurement Information System Sleep Disturbance (PROMIS®; Yu et al., 2012 [35]). SPAQ is a self-reported measure of seasonality and seasonal affective disorder, ranging from 0 to 24, with scores < 11 indicating non-seasonality. BDI-II is a diagnosis of depression, combining scores on screening tests with clinical interviews. BDI-II scores are treated as quantized values and range from 0 to 60; individuals with scores < 14 are not depressed, those with 14–19 are mildly depressed, those with 20–28 are moderately depressed, and those with > 28 are severely depressed. SAD is a diagnostic measure for seasonal depression, already recorded in binary. The Munich Chronotype Questionnaire (MCTQ; Roenneberg et al., 2007 [36]) is a self-reported measure of sleep–wake chronotype; mid-sleep point on a free day was used to designate the chronotype score, with the highest quartile MSF scores indicating eveningness (F:M ratio = 2.5) and the lowest quartile scores indicating morningness (F:M ratio = 5.8). Social jetlag was measured as MSF–MSW, the mid-sleep on workdays. We also reported diurnal preference using the Morningness–Eveningness Questionnaire (MEQ; Horne and Ostberg, 1976 [37]).

2.3. Genotyping

DNA was extracted from 10 to 20 hair follicles from each participant. Hair follicles were digested for 24 h at 56 °C with Proteinase K, and these samples were further purified using the Qiagen DNAeasy Micro Kit. Using TaqMan SNP Genotyping assays (Applied Biosystems, Foster City, CA, USA), we identified single-nucleotide polymorphisms (SNPs) using real-time qPCR protocols on an ABI 3700HT instrument. Homozygous or heterozygous genotypes for the major and minor alleles were recorded for each participant.

The PER3-VNTR length polymorphism repeat region was analyzed via fragment length analysis on an ABI 3100 sequencer using GeneScan[®] software (1.0). We fluorescently labeled the forward primer with 6-FAM (5'-CAAATTTTATGACTACTACCAGAATGGCTGAC-3'). PCR was performed using the forward and reverse primer (5'-AACC TTGACTTCCACATCAGTGCCTGG-3'). The 25- μ L reactions were run with the following cycling conditions: 3 min at 94 °C, 35 cycles of 45 s at 94 °C, 45 s at 58 °C, and 45 s at 72 °C, with a final 3 min step at 72 °C. PER3 alleles were separated by repeat number/size as PER3 4/4, PER3 4/5, or PER3 5/5 and recorded for each participant.

2.4. Feature Generation, Selection and Classification

Each individual has nine genotypic features—*CLOCK3111* (rs1801260), *CRY1* (rs2287161), *CRY2* (rs10838524), *PER2* (rs10462023), *PER3A* (rs228697), *PER3B* (rs17031614), *PER3C* (rs10462020), *PER3VNTR* (rs57875989), and *ZBTB20* (rs139459337)—which can be homozygous dominant, homozygous recessive, or heterozygous. We also collect five behavioral/clinical features: gender, diurnal preference scores (MEQ and MSF), age (≤ 22 or > 22), and socioeconomic status (poor, lower-middle-class, upper-middle-class, affluent). We thus generated 14 features (we will refer to these as one-way features). We further created pairwise interactions among the nine genotypic features and their respective variants, generating 9C2*9 additional features (we will refer to these as two-way features).

We performed one-hot-encoding for the non-binary categorical features. To reduce multicollinearity and high dimensionality, we further removed the most frequent group for each feature as the reference point, generating a total of 286 features in our expanded dataset. The final dataset for each disorder target is obtained by stratifying the expanded data by the respective disorder metric (SPAQ, SAD, or BDI) and removing encoded features that were not present in any sample. The corresponding dimensions of these final datasets are as follows: 221 subjects and 261 features for SPAQ, 152 subjects and 258 features for SAD, and 254 subjects and 268 features for BDI, respectively.

Datasets are cleaned, fully imputed with k-nearest neighbors' imputation [38], and normalized before the machine learning process. When the dataset is imbalanced, the synthetic minority oversampling technique (SMOTE) is employed with $k = 5$ to ensure equal numbers of samples for positive and negative cases. Specifically, SMOTE creates an artificial sample by randomly selecting a point between a minority class sample and another random neighbor in its k-neighbor feature space [39].

We analyzed the dataset through a ten-fold cross-validation process. In this process, the data are split into 10 subsets. Each subset is used once as a test set, while the remaining ninety percent of the data serve as the training set. Stratified sampling was used to split the training and testing sets to avoid overfitting or underfitting. This cross-validation procedure was repeated ten times. During each iteration, we aggregated the features selected in each fold, applying this across all the methods we used. We repeated the procedure described below ten times and aggregated the features selected at each of the folds across all listed methods. We used five feature selection methods to find features such as epistatic combinations that are possible robust predictors of mood disorders. Two ranking-based feature selection methods, Chi-squared (CHI2) and Relief (ReF), sort features in order of correlation with the disorder [40,41]. Three subset-based feature selection methods—correlation-based feature selection (CFS), maximum relevance minimum redundancy (MRMR), and joint mutual information (JMI)—use information theory-based criteria to identify potential subsets from the feature space [42–44]. Each feature selection method

was run on the expanded datasets using ten-fold cross-validation. We then extracted copies of the dataset which included only the top-20 most frequently selected features.

We applied four classifiers, including logistic regression (LR), random forests (RF), support vector machine (SVM), and XGBoost (XGB), on the features selected for the training subsets [45–47]. Grid search is used to identify the combination of hyperparameters that yields the model with the highest accuracy. This final model is applied to the test data to predict the labels, and the results are evaluated accordingly based on model accuracy.

2.5. Statistical Analyses

We ran all of the following analyses separately, with each of SPAQ, SAD, and BDI being the target variable with R [48]. Each analysis (except for ANOVA and ARACNE) was run three times each for all participants and each gender separately. We also used the top-20 datasets wherever possible to see if results from a more-robust dataset match the overall trends.

We ran univariate logistic regression on features selected in 80% of the runs by at least three out of the five feature selection methods to determine directionality, as they are likely effective predictors on their own. After this univariate analysis, we performed multivariate logistic regression on the top-20 risk factors to see which factors were significant predictors of mood disorders when used in tandem with other features. We recorded their odds ratios, *p*-values and 95% confidence intervals. Multivariate linear regression was used for SPAQ and BDI metrics but not SAD since the target variable is categorical. Due to the high-dimensional configuration, we performed multivariate logistic and linear regression with stepwise selection using Akaike Information Criterion (AIC) to create an optimal model without compromising the model accuracy [49].

We used the *mediation* package in R to determine the mediation effects of diurnal preference (MEQ and MSF) in the relationship between independent variables (clinical/behavioral factors) and the dependent variable (mood disorder). We fit the mediator and outcome models with linear regression and used the *mediate* function with non-parametric bootstrap for variance estimation [50]. The results are summarized numerically with point estimates, confidence intervals, and the *p*-values for the average direct, indirect, and total effects.

For the SPAQ and BDI target variables, the subset of genotypes determined significant via univariate and multivariate regression were analyzed in-depth with type-III sums of squares two-way ANOVA. Gender was the second independent variable as we targeted sex-specific paths [51].

For each gender, using one-way features, we performed association analysis to identify likely relations based on measures of interest. We transformed all the variables into categorical data and applied the a priori algorithm in the *arules* package in R [52] to obtain rules of the form $X \rightarrow Y$. *X* refers to the multi-way genotype and clinical feature combination's increasing risks of acquiring *Y*, the mood disorder. We set the baseline threshold of rule size to, at most, 6, with minimum support of 0.04 and confidence of 0.7, and then we sorted the rules by their respective lift values. Network plots were created using the *igraph* library [53], with node size representing support and edge width representing confidence.

We used the algorithm for the reconstruction of gene regulatory networks (ARACNE) [54] to identify the direct and indirect relationships between one-way features and the disorder target via diurnal preference scores (MEQ and MSF). We employed the *mi.empirical* method for ARACNE to compute distance metrics. We constructed a gene network via the *minet* package as ARACNE analyses each qualified the gene triplet and removes the edge with the smallest metric; then, we plotted it using Rgraphviz v2.46.0 [55,56]. ARACNE calculates the frequency of a link via bootstrapping, and we kept strong links by setting a threshold of 46% for bootstrap support. The dataset and code are available at <https://github.com/thanh dang2712/Mood-Disorder-Project>.

3. Results

3.1. Seasonality Prediction

Seasonality was predicted with 62.5% accuracy, 6% higher than a baseline accuracy of 56.5% (Figure 1a). Sex-stratified classifiers predicted seasonality for males (68.2%) and females (69.4%), with respective ~7% and ~10% increases from baseline (Figure 1b,c).

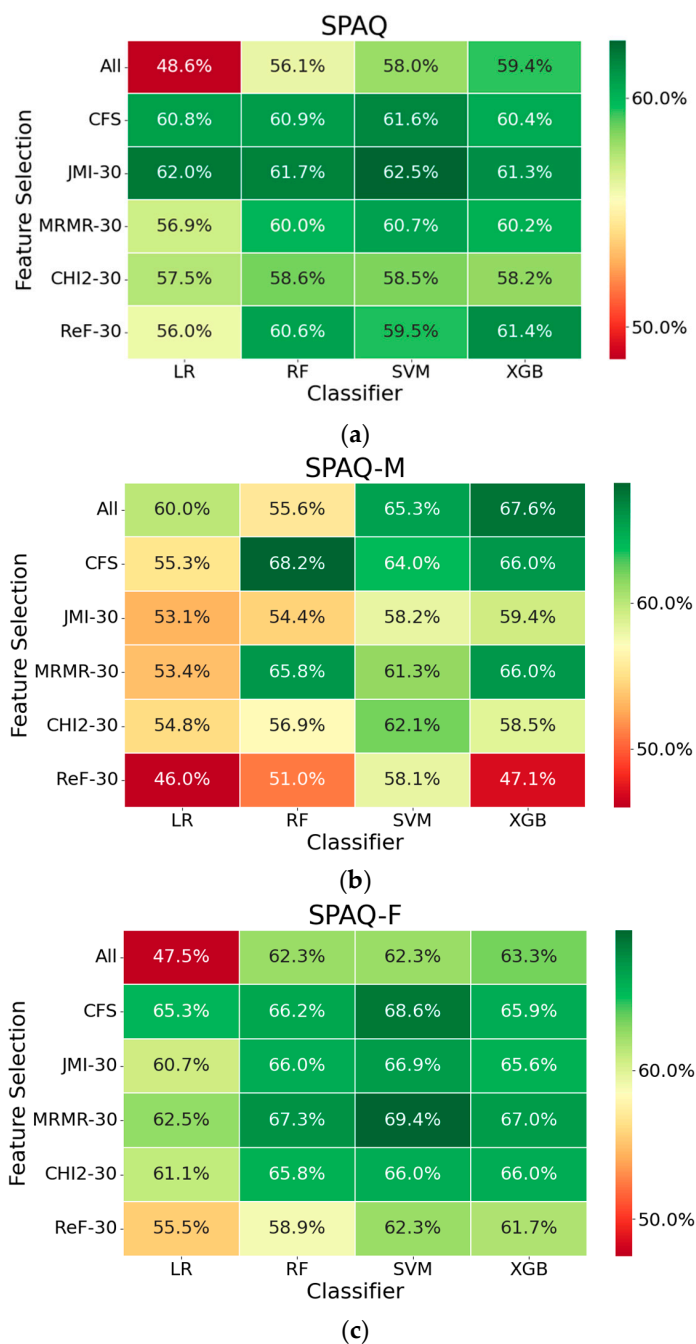


Figure 1. Classifiers exhibit greater accuracy increase from baseline when stratified by sex. (a) Analyses on the overall dataset yielded up to 6% greater accuracy than the baseline accuracy (56.5%) using the support vector machine (SVM) classifier and the joint mutual information (JMI) feature selection. (b) Prediction of male seasonality yielded up to 10% greater accuracy than the baseline accuracy (58%) using the RF classifier and correlation-based feature selection. (c) Prediction of female seasonality yielded up to 7% greater accuracy than the baseline accuracy (62.2%) using the LR and SVM classifiers and mRMR feature selection.

3.2. Sex-Specific Risk and Protective Factors for Seasonality and SAD Symptoms

In males, the PER2-GG/PER3B-AG combination was associated with a seven-fold increase in odds of seasonality (Table 1; OR 7.3 (0.87–13.7) adj. $p = 0.034$); SPAQ scores were higher in these males than males with other genotypes (Figure 2; Genotype: $F_{1,217} = 4.312$, $p = 0.039$; Gender: $F_{1,217} = 0.005$, $p = 0.946$; Genotype \times Gender: $F_{1,217} = 1.650$, $p = 0.200$). In addition, the CLOCK3111-TC/ZBTB20-AT genotypic combination increased odds of seasonality in males almost four-fold (Table 1; OR 4.0 (0.37–7.58) adj. $p = 0.052$).

Table 1. Significant risk and protective factors for male seasonality, female seasonality, and female seasonal affective disorder ($p < 0.05$). Univariate and multivariate linear and logistic regression were performed for genotypic features in association with male and female seasonality (SPAQ M and F), while univariate and multivariate logistic regression were performed for female seasonal affective disorder (SAD F). Results from each separate analysis are grouped by outcome variable (SPAQ M, SPAQ F, SAD F), and univariate logistic odds ratios are enclosed in asterisks (risk factor: *odds ratio* > 1), multivariate linear estimates appear in brackets (risk factor: [estimate] > 0), and multivariate logistic odds ratios appear as plain numbers (risk factor: odds ratio > 1).

Feature	Male		Female	
	Risk	SPAQ M OR [estimate]	SPAQ F OR [estimate]	SAD F OR [estimate]
CLOCK3111-TC/ZBTB20-AT		3.975		
CRY2-AA/PER3C-TG			[4.279]	
CRY2-AA/VNTR-4,5			*8.919*	
PER2-GG/PER3B-AG		7.262		
Anxiety				1.061
Evening type			*1.783*	
Increasing Age			*6.406*	
Sleep Disturbance		[0.320]	[0.173]	
Protective				
VNTR-4,5				0.258
CLOCK3111-TC/PER3B-GG				0.357
CLOCK3111-TC/PER3C-TT				0.383
CRY1-GG				0.491
CRY1-GG/PER2-AA			[−3.595]	
CRY1-GG/VNTR-5,5			*0.112*	
CRY1-GG/ZBTB20-AA				0.454
PER2-AA/PER3C-TT			*0.203*	
PER3B-GG/VNTR-4,5				0.304
PER3C-TT/VNTR-4,5				0.361
VNTR-4,5/ZBTB20-AA				0.321
Morning type		0.937		

In females, the CRY2-AA/PER3-VNTR-4,5 combination was associated with a nine-fold increase in seasonality risk in our logistic regression (Table 1; OR 8.9 (1.71–164.19) adj. $p = 0.037$). Additionally, linear regression identified CRY2-AA/PER3C-TG as a risk factor for seasonality (Table 1; estimate 4.3 (1.55–7.03) adj. $p = 0.0091$), and SPAQ scores were higher in these females than females with other combinations (Figure 3a; Genotype: $F_{1,217} = 4.084$, $p = 0.044$; Gender: $F_{1,217} = 5.496$, $p = 0.020$; Genotype \times Gender: $F_{1,217} = 0.449$, $p = 0.504$). Older age and evening-type behavior were clinical risk factors for seasonality in females (Table 1; OR 6.3 (1.49–44.04) adj. $p = 0.023$; OR 1.8 (1.32–2.50) adj. $p = 0.023$).

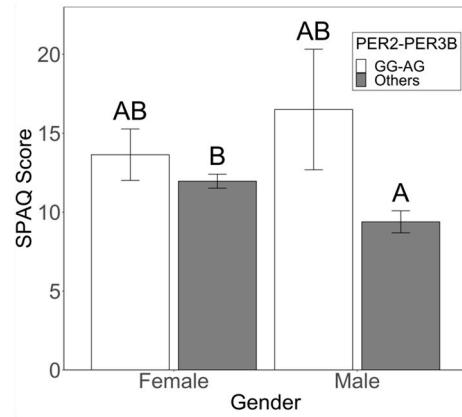


Figure 2. Two-way ANOVA and Tukey post hoc analysis reveal that the PER2-GG/PER3B-AG combination is predictive of seasonality in males. Average SPAQ scores were higher in males with the PER2-GG/PER3B-AG genotypic combination than in males with other genotypic combinations (Genotype: $F_{1,217} = 4.312, p = 0.039$; Gender: $F_{1,217} = 0.005, p = 0.946$; Genotype \times Gender: $F_{1,217} = 1.650, p = 0.200$). Error bars denote SE. A, B denote significant differences for Tukey post hoc analyses.

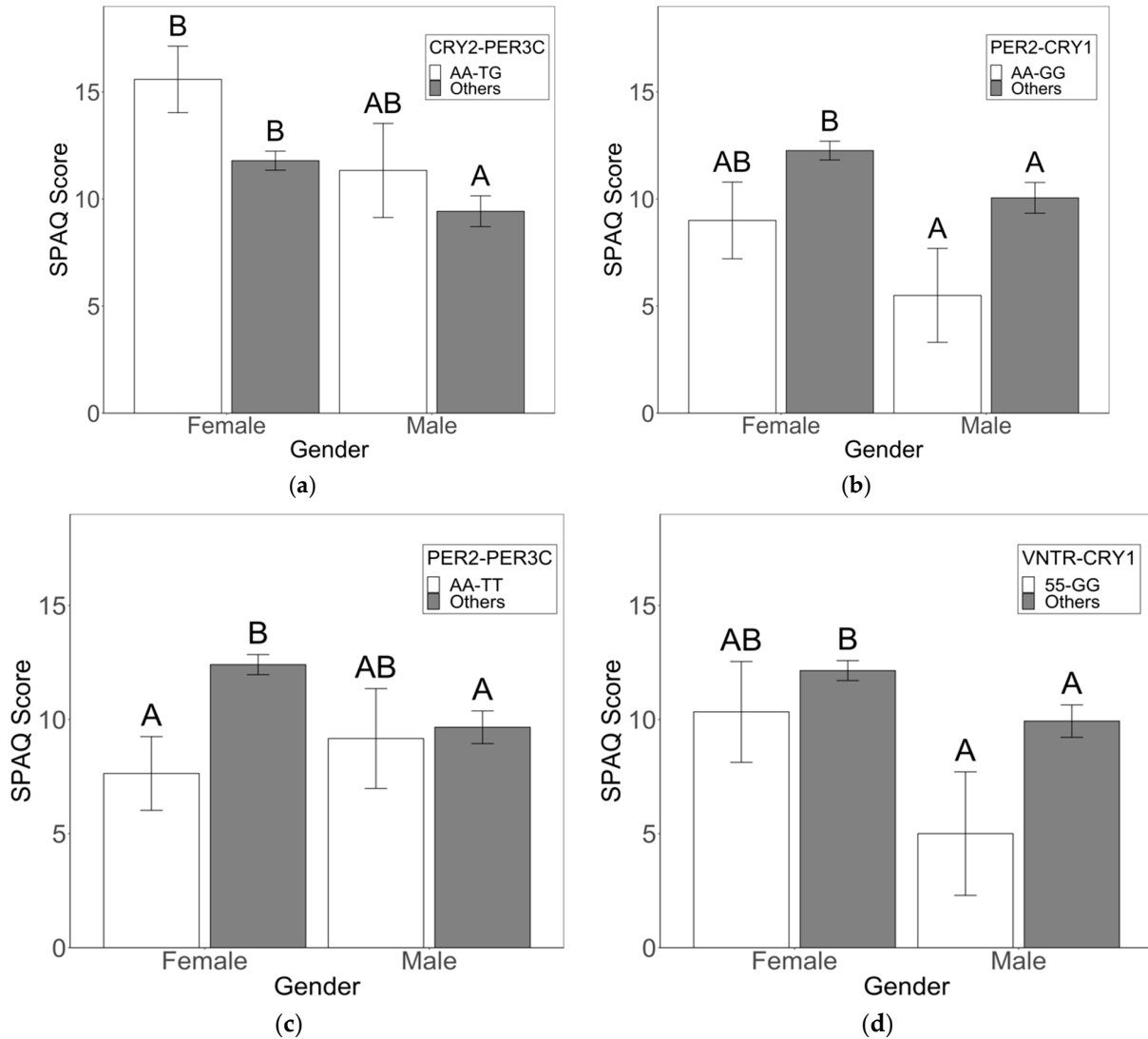


Figure 3. Two-way ANOVA and Tukey post hoc analysis reveal risk and protective combinations for seasonality in females. (a) Significantly higher average SPAQ scores were identified in females with the CRY2-AA/PER3C-TG combination relative to other genotypic combinations (Genotype: $F_{1,217} = 4.084,$

$p = 0.044$; Gender: $F_{1,217} = 5.496$, $p = 0.020$; Genotype \times Gender: $F_{1,217} = 0.449$, $p = 0.504$). (b) Average SPAQ scores were lower for males and females with the PER2-AA/CRY1-GG combination (Genotype: $F_{1,217} = 7.001$, $p = 0.009$; Gender: $F_{1,217} = 3.735$, $p = 0.055$; Genotype \times Gender: $F_{1,217} = 0.192$, $p = 0.662$). (c) Average SPAQ scores were lower for females with the PER2-AA/PER3C-TT combination than for females with other combinations (Genotype: $F_{1,217} = 3.420$; $p = 0.066$, Gender: $F_{1,217} = 0.182$; $p = 0.670$, Genotype \times Gender: $F_{1,217} = 2.257$, $p = 0.134$). (d) Average SPAQ scores were lower for males and females with VNTR-5,5/CRY1-GG combination (Genotype: $F_{1,217} = 3.525$, $p = 0.062$; Gender: $F_{1,217} = 4.417$, $p = 0.037$; Genotype \times Gender: $F_{1,217} = 0.755$, $p = 0.396$) ($N = 221$). A, B denote significant differences for Tukey post hoc analyses.

Logistic regression revealed that the CRY1-GG/PER3-VNTR-5,5 combination was associated with a near ten-fold decrease in the odds of seasonality in females (Table 1; OR 0.1 (0.0058–0.72) adj. $p = 0.048$); males and females with this combination had lower SPAQ scores than individuals with other combinations (Figure 3d; Genotype: $F_{1,217} = 3.525$, $p = 0.062$; Gender: $F_{1,217} = 4.417$, $p = 0.037$; Genotype \times Gender: $F_{1,217} = 0.755$, $p = 0.396$). PER2-AA/PER3C-TT reduced the odds of seasonality five-fold (Table 1; OR 0.2 (0.043–0.74) adj. $p = 0.023$); SPAQ scores were lower in females with this genotype combination (Figure 3c; Genotype: $F_{1,217} = 3.420$, $p = 0.066$; Gender: $F_{1,217} = 0.182$, $p = 0.670$; Genotype \times Gender: $F_{1,217} = 2.257$, $p = 0.134$). CRY1-GG/PER2-AA was a significant protective factor identified via multivariate linear regression (Table 1; estimate -3.6 (-6.74 – -0.45) adj. $p = 0.050$); average SPAQ scores were lower for males and females with this combination relative to other combinations (Figure 3b; Genotype: $F_{1,217} = 7.001$, $p = 0.009$; Gender: $F_{1,217} = 3.735$, $p = 0.055$; Genotype \times Gender: $F_{1,217} = 0.192$, $p = 0.662$).

3.3. Association Rule Learning Networks Also Reveal Sex-Specific Rules for Seasonality

Association rule learning networks were used to identify co-occurring features that strongly predicted seasonality (Figure 4). In males and females, the most frequently occurring SNP variants associated with seasonality symptoms were CRY1-CC, PER2-GG, PER3C-TG, and PER3-VNTR-4,5 (Figure 4a,b). CRY2-AA and CRY1-GG were SNP variants exclusive to males and females, respectively. PER3C-TG, PER3-VNTR-4,5, and age were the most frequently occurring factors for males, while PER3C-TG, PER3-VNTR, age, being upper middle class, and eveningness were the most frequently occurring factors for females.

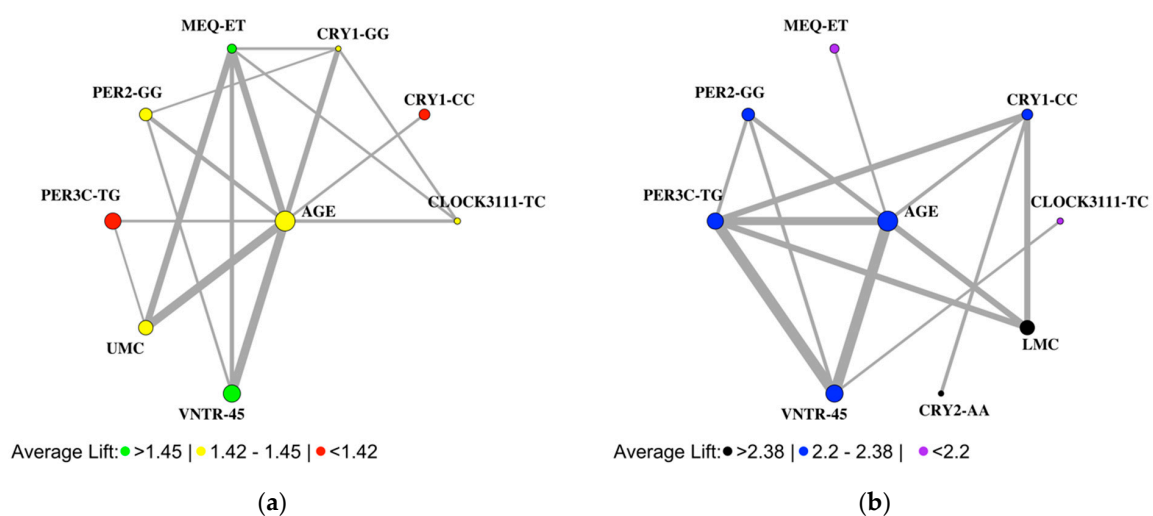


Figure 4. Association rule networks for male and female seasonality. (a) The most frequently occurring SNP variants for males included VNTR-4,5, PER3C-TG, PER2-GG, and CRY1-CC, while age and being lower-middle class (LMC) appeared as important clinical factors. (b) The most frequently occurring SNP variants for females were VNTR-4,5, PER3C-TG, PER2-GG, and CRY1-CC, while age, being upper-middle class (UMC), and eveningness commonly appeared as important clinical factors.

We constructed Venn diagrams to summarize sex-specific gene and clinical feature associations with seasonality symptoms (Figure 5). The CLOCK3111-TC/ZBTB20-AT and PER2-GG/PER3B-AG combinations were significant risk factors exclusive to males, while CRY2-AA/PER3C-TG and CRY2-AA/PER3-VNTR-4,5 were significant risk factors exclusive to females. Risk factors for both sexes included CRY1-CC, PER2-GG, PER3C-TG, and PER3-VNTR-4,5 (Figure 5a). Additionally, anxiety and belonging to the upper middle class were risk factors in males, while belonging to the lower middle class was a risk factor for females (Figure 4b). In both sexes, age and sleep disturbance were clinical risk factors, and diurnal preference was protective (Figure 5b).

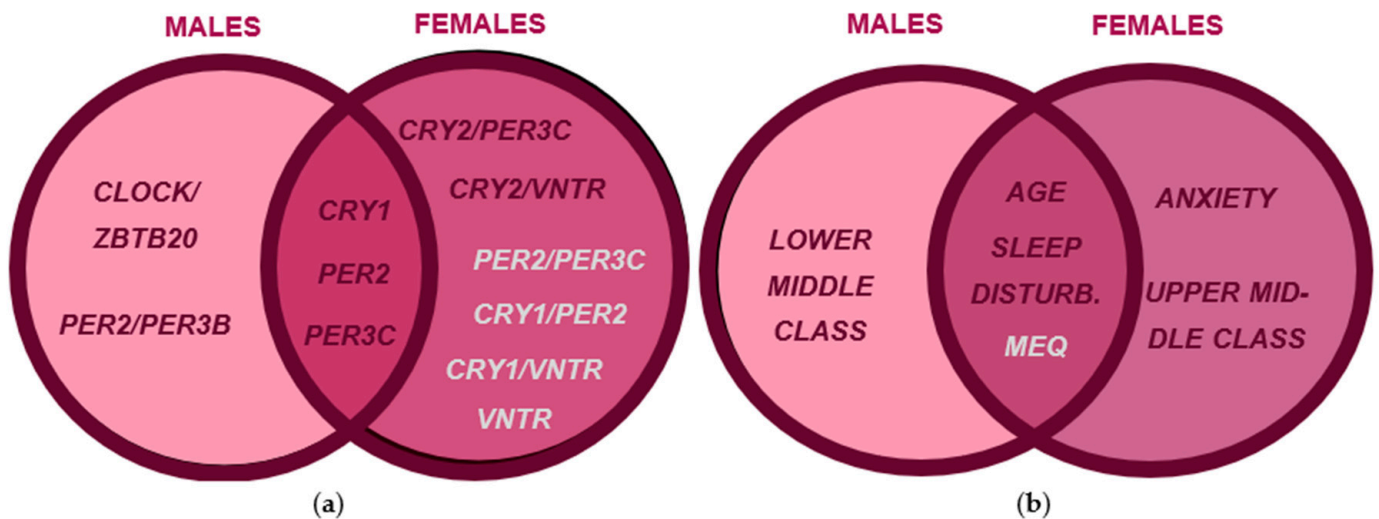


Figure 5. Venn diagram summarizing significant genotypic and clinical features for seasonality in males and females resulting from multivariate linear regression, univariate logistic regression, multivariate logistic regression, and association rule learning network results. Venn diagram comparisons were constructed from male and female seasonality analysis, risk factors were denoted by red text, and protective factors were denoted by silver text. (a) Genotypic factor comparison reveals shared and exclusive risk factors between males and females. (b) Clinical factor comparison reveals shared and exclusive risk factors between males and females.

3.4. Comparisons of Genotypic and Clinical Feature Associations with Seasonality and SAD

CRY2 and PER3C were risk factors exclusive to seasonality symptoms (Figure 6a). PER2 was protective for seasonality; CLOCK3111, PER3B, and ZBTB20 were protective for SAD symptoms; and CRY1 and PER3-VNTR were protective against both outcomes (Figure 6a). Age and anxiety were clinical risk factors for both outcome, while belonging to the upper middle class was specific to seasonality (Figure 6b).

3.5. Mediation Analyses Reveal Robust Mediation of Seasonality Relative to SAD Symptoms

Mediation analysis revealed that MEQ partially mediated the association of PER2-GG/PER3B-GG with seasonality in males (Table 2; $t_{153} = 2.0$, $p = 0.046$). In females, MEQ fully mediated PER2-GG (Table 2; $t_{153} = -2.0$, $p = 0.048$), while MSF partially mediated PER2-GG (Table 2, $t_{153} = 2.13$, $p = 0.034$). Also, MSF was a full mediator for CRY1-CC/PER2-AG (Table 2; $t_{153} = -2.5$, $p = 0.015$), and MEQ was a partial mediator for CRY1-CG/PER3-VNTR-5,5 (Table 2; $t_{153} = -2.5$, $p = 0.014$).

Table 2. MEQ- and MSF-mediated associations of genotypic features with male and female seasonality. Mediation analysis was performed using the Horne–Östburg Morning–Eveningness Questionnaire (MEQ) and mid-sleep point on a free day [37] as potential mediators, and mediation was confirmed for average mediated effect p -values < 0.10 . Mediation was considered full if the average direct effect had a p -value ≥ 0.10 .

Genotypic Feature	SPAQ F	SPAQ M
PER2-GG	Full MEQ, partial MSF Full MSF Partial MEQ	Partial MEQ
CRY1-CC/PER2-AG		
CRY1-CG/VNTR-55		
PER2-GG/PER3B-GG		

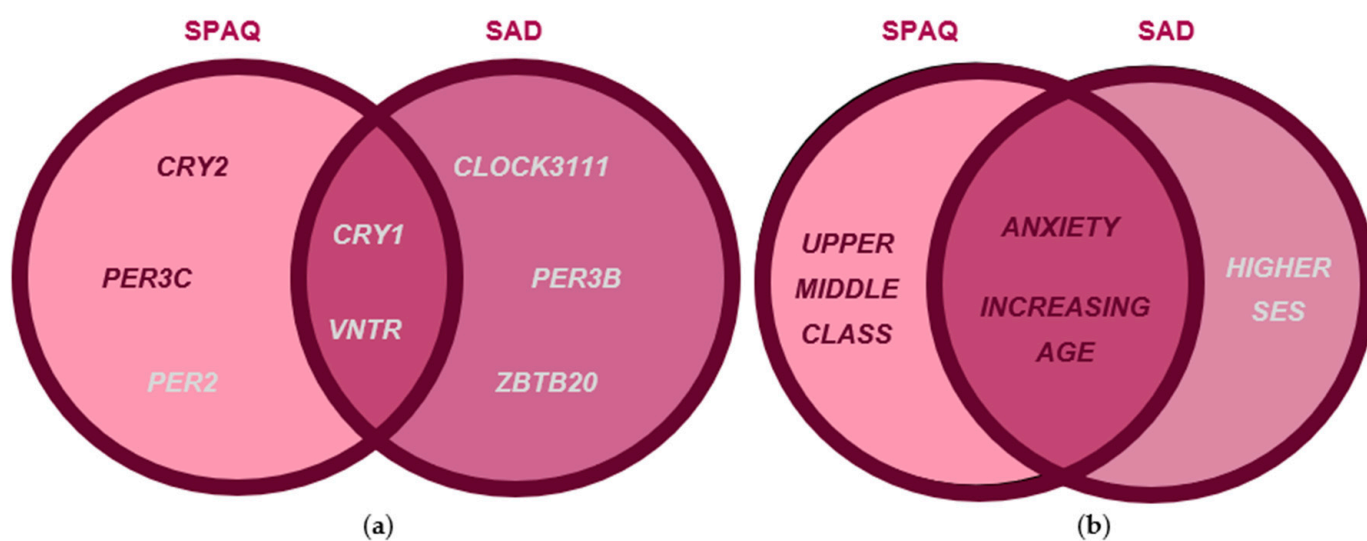


Figure 6. Venn diagram summarizing significant genotypic features for female seasonality and female SAD resulting from multivariate linear regression, univariate logistic regression, multivariate logistic regression, and association rule learning network results. Venn diagram comparisons were constructed from our female seasonality analysis and our SAD regression results, risk factors were denoted by red text, and protective factors were denoted by silver text. (a) Genotypic factor comparison reveals risk factors exclusive to seasonality. (b) Clinical factor comparison reveals shared and exclusive risk factors between seasonality and SAD in females.

3.6. Mutual Information Analyses

ARACNE analysis revealed that CRY1 and ZBTB20 share mutual information directly with seasonality in females, while other genotypes were mediated through diurnal preference; males also show direct links between CRY1 and ZBTB20, as well as CRY2 and PER3-VNTR (Figure 7). When tested independently using depressive symptoms, CRY1, CRY2, PER3-VNTR, and ZBTB20 also had direct associations with depressive symptoms in both males and females (Figure 8).

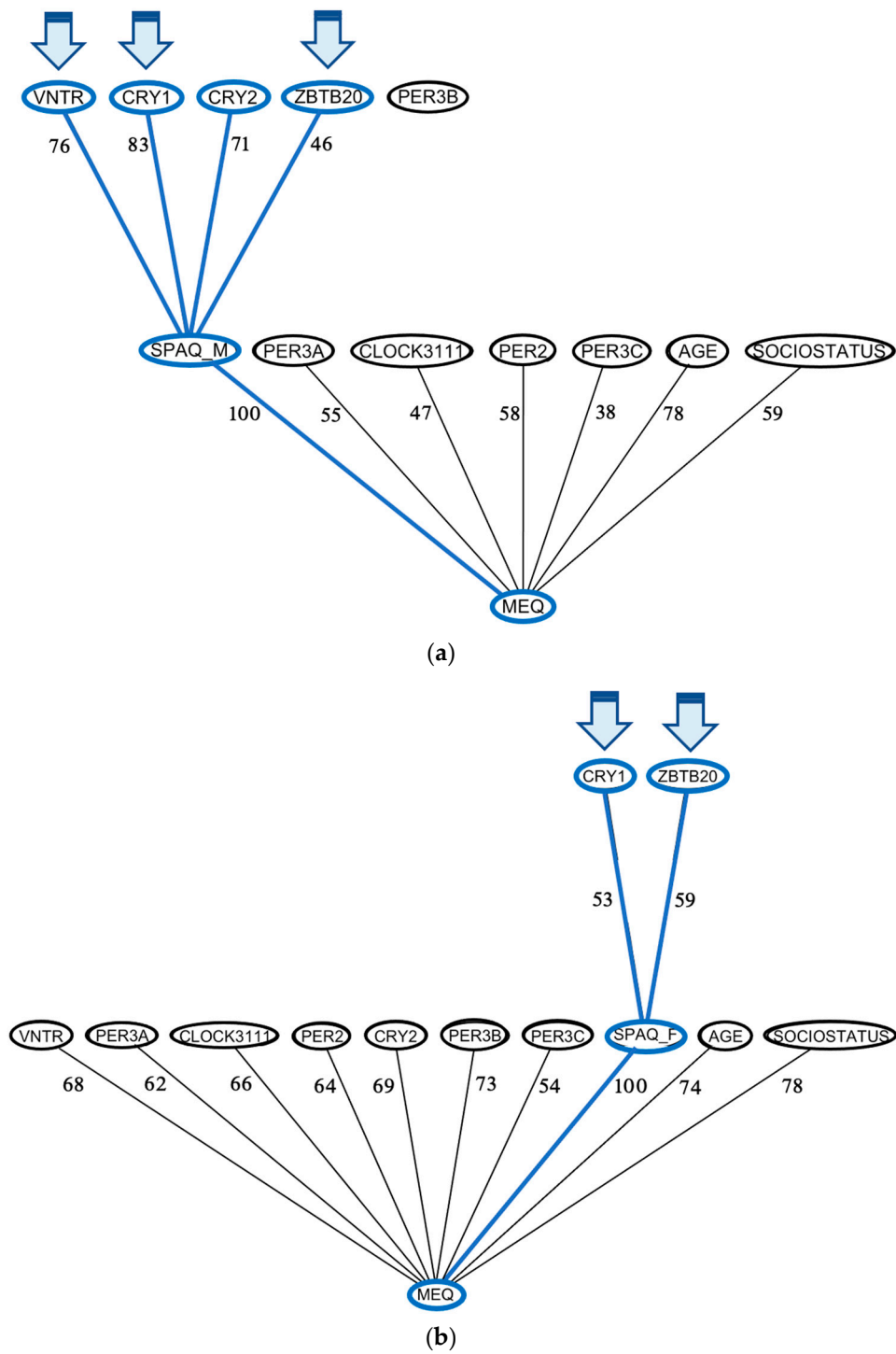


Figure 7. ARACNE networks for seasonality in male and female population subsections. ARACNE networks were constructed for seasonality in (a) males and (b) females using MEQ as a potential mediator variable.

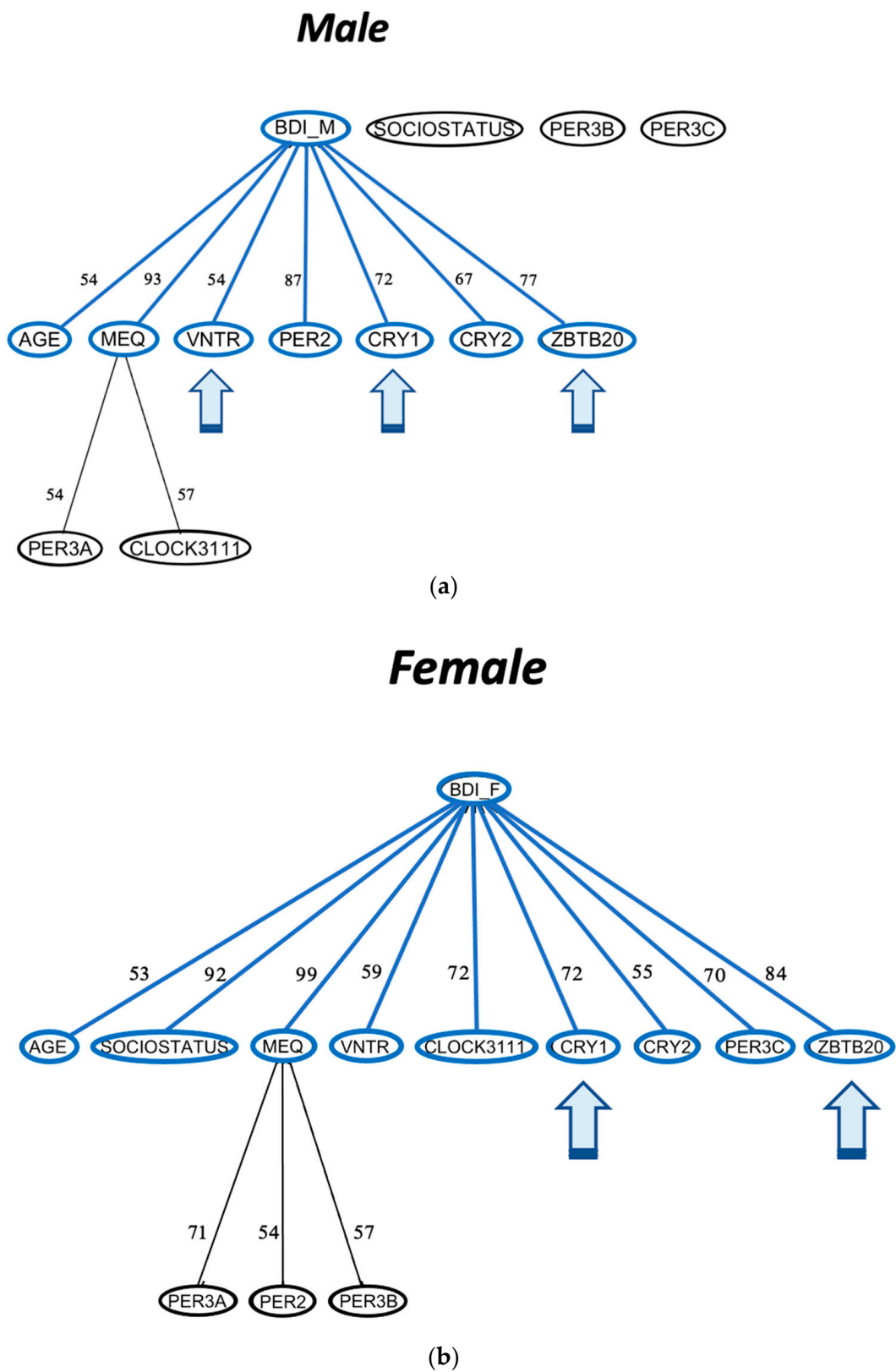


Figure 8. ARACNE networks for depression in (a) male and (b) female population subsections. ARACNE networks were constructed for depression in males and females using MEQ as a potential mediator variable.

4. Discussion

Although there is considerable evidence in the literature suggesting multiple connections between clock gene variants, circadian disruptions, and seasonal affective disorder, previous GWAS, PheWAS, and gene association studies have only been able to describe the effects of single clock gene mutations on seasonal affective disorder. These traditional techniques are limited in their ability to explore the effects of multiple genotype combinations or to distinguish between the direct or indirect effects of clock genes on SAD symptoms. Using novel machine learning approaches and statistical analyses, we observed effects of both single-genotype and multi-genotype combinations on SAD. Additionally, we were able to discern direct and indirect effects using ARACNE analyses combined with mediation analyses. We report three main findings from these analyses: (1) clock variants predictive of seasonal affective disorder tend to exhibit sex-specific effects; (2) clock variants can have direct transcriptional effects on seasonality and SAD symptoms or indirect effects that are mediated by diurnal preference or chronotype; and (3) clock variants with direct effects on seasonality symptoms also have direct effects on depressive symptoms.

4.1. Circadian Genotypes Exhibit Sex-Specific Effects on Seasonality and SAD Symptoms

Our feature selection results revealed greater prediction accuracy increases from the baseline when the dataset was stratified by sex. In addition, the main findings from our multiple logistic and linear regression analyses revealed sex-specific risk and protective factors. For example, PER2-GG/PER3B-AG and CLOCK3111-TC/ZBTB20-AT were significantly associated with the risk of seasonal symptoms in males, and CRY1-GG/PER2-AA, CRY2-AA/PER3C-TG, CRY2-AA/PER3-VNTR-4,5, PER2-AA/PER3C-TT, and PER3-VNTR-5,5/CRY1-GG had significant associations with the risk of seasonal symptoms in the female population. These findings suggest that these genotypic combinations of clock genes exert sex-specific effects, a pattern previously identified in studies examining major depressive disorder (MDD) [18] and anxiety [31].

Interestingly, Shi et al. (2016) found that the CLOCK variant (rs1801260) was significantly associated with MDD in males only, a pattern similar to our finding of a sex-specific association of a genotype combination (CLOCK3111-TC/ZBTB20-AT) with seasonal symptoms in males. This CLOCK variant may be involved in sex-specific effects through glucocorticoid regulation [18]. Also, since CLOCK3111-TC has been associated with the sleep–wake cycle [57,58], this polymorphism may promote circadian disruption, which is hypothesized to affect mood through serotonin regulation [59]. This mechanistic reasoning is in line with clinical evidence, which has demonstrated that SSRIs can eliminate SAD in affected individuals [60].

4.2. Clock Gene Effects on Seasonality and SAD May Be Direct or Mediated by Diurnal Preference

The relationship between seasonality score and chronotype is well-established [61–63], indicating that circadian gene variants associated with chronotype and circadian phase shifts may indirectly affect seasonal depressive symptoms through a lack of energy, burnout, sleeping problems, and social withdrawal, which may result from physiological misalignment of circadian rhythms with behavioral patterns [64]. In support of this, our current study found that most of the variables in the ARACNE analyses were mediated via diurnal preference (MEQ) and/or chronotype (MSF). For example, CRY2 and PER3C were risk factors exclusive to seasonality symptoms, and variants of these two genes have previously been associated with advanced and delayed sleep phase [65,66], respectively, suggesting that these genes are critical to circadian clock pathways. We also found that PER2, a protective factor exclusive to seasonality, was mediated by diurnal preference and appeared in several genotypic combinations that were also mediated by diurnal preference, further supporting the importance of sufficient PER2 gene function to maintaining circadian clock alignment.

Although the PER3-VNTR and CRY1-associated genotypes were protective for both seasonality and SAD analyses, mediation analysis found that the CRY1/PER3-VNTR

protective combination was partially mediated through diurnal preference, raising the possibility that these genes may also be involved in pathways that indirectly influence mood. Indeed, the PER3-VNTR 4 allele has been associated with anxiety [67,68] and depression [69], in addition to evening chronotype [12,13]. Genotypes with variants of PER3B and CLOCK3111 were also protective for SAD in this population. In the Shi et al. (2016) study, the PER3B A allele (rs17031614) was significantly associated with depression, and the CLOCK3111 C allele was significantly associated with MDD in males [18], suggesting potential sex-specific, direct effects on mood. However, we observed in our ARACNE analyses that genotype combinations with the PER3B and CLOCK3111 genes were mediated by diurnal preference. Moreover, the CLOCK3111 C-allele has previously been associated with evening chronotype [58] and seasonal depression [14], suggesting indirect effects or a combination of direct and indirect effects for this gene.

On the other hand, our ARACNE analysis indicates that genotypes including variants of ZBTB20 and CRY1 may have robust direct effects on mood as these two genes that shared mutual information directly with seasonality and are not mediated by diurnal preference or chronotype. In other studies, CRY1 is robustly associated with depression [4,24]. The ZBTB20 gene is currently the only gene associated with SAD in GWAS studies [70] and directly shares mutual information with seasonality in our study. Together, these findings suggests that CRY1 and ZBTB20 are genes involved directly in mood regulation.

4.3. Clock Variants with Direct Effects on Seasonality Symptoms Also Have Direct Effects on Depression Symptoms

Seasonal affective disorder is classified as a subset of depression where a small portion of individuals exhibiting clinical depression may also exhibit seasonal variability characteristic of SAD. Clock gene associations are commonly consistent between those with SAD [17,71,72] and depression [4,15]. Given this overlap, we utilized patient depression scores (BDI) as the outcome for ARACNE analysis to test the hypothesis that PER3-VNTR, CRY1, and ZBTB20 are directly involved in mood regulation. Here, we observed that the effects of CRY1, PER3-VNTR, and ZBTB20 were also directly associated with depression, underscoring the possibility that these genes directly impact mood regulation via transcriptional (CRY1, PER3-VNTR) or activational (ZBTB20) regulatory mechanisms.

The PER3-VNTR genotype was directly associated with depression in both males and females, although it was only directly associated with seasonality in males only. This SNP (rs57875989) has previously been associated with anxiety and depression [24,68], but has also appeared significantly associated with evening chronotype [12,13], suggesting that PER3-VNTR may have both direct and indirect effects on mood disorders. Liberman et al. (2018) found that the PER3-VNTR 4-repeat allele was associated with both a three-fold increase in odds of depression and significantly greater tendencies toward eveningness in the same dataset [24]. Previously, Archer et al. (2003) identified an association between this polymorphism and extreme diurnal preference, finding that those homozygous for the 5-repeat allele are more likely to be morning-types, while those homozygous for the 4-repeat allele are more frequently evening types and encounter delayed sleep phase disorder (DSPS) [12]. The PER3-VNTR polymorphism occurs in exon 18, encoding a variable number of tandem repeat region (VNTR) of 18 amino acids that repeats four or five times depending upon the variant allele [73]. PER3-VNTR appears to have an important role regulating sleep homeostasis as this gene has been implicated in sleep duration, timing, and structure. These findings have led others to hypothesize that PER3-VNTR may function in a shared molecular pathway involving the regulation of circadian rhythm and sleep structure, although such a pathway has yet to be confirmed by experimental findings. Supporting our observed gender effect, Weiss et al. (2020) found that PER3^{4/4} males had nine times greater odds of being evening-types, while there was no significant effect in females or the combined dataset [73]. Such findings support the notion that PER3-VNTR may influence mood directly through sex-specific pathways illuminated above such as glucocorticoid or serotonin regulation.

Previous research illuminates the potential pathways by which CRY1 and ZBTB20 variants may directly affect seasonal affective disorder. The CRY1 SNP (rs2287161) is located upstream of the CRY1 promoter [4], and it has been predicted that this SNP fosters increases in CRY1 gene expression [24]. CRY1 has a critical role in HPA axis regulation via the repression of glucocorticoid receptors [74], and dysregulation of this pathway may promote the development of depressive symptoms [75]. Therefore, differential expression of CRY1 could affect glucocorticoid receptor repression and HPA axis stability, contributing to depressive symptoms. Altered expression of CRY1 also likely affects the expression of other circadian genes [24], which may then directly alter mood through physiological changes in other downstream mood pathways.

The ZBTB20 (rs139459337) SNP results in reduced expression of the ZBTB20 gene [22], which has been associated with seasonal affective disorder and major depressive disorder [22,76]. ZBTB20 is a zinc finger transcriptional repressor protein that is abundant in the hippocampus and plays an important role in the regulation of clock gene expression. Indeed, the loss of ZBTB20 is associated with impaired output of circadian rhythms [77,78]. As a transcription factor, ZBTB20 is also involved in synaptic transmissions, which could influence glucose homeostasis, glucocorticoid stress pathways, or monoamine pathways associated with depression symptoms. Our observations from the ARACNE analyses suggest direct effects of CRY1 and ZBTB20 on seasonal mood symptoms as well as depressive symptoms, supporting these previous findings of associations of CRY1 and ZBTB20 on MDD [4,24]. However, the precise mechanisms by which ZBTB20 and CRY1 exhibit direct effects on seasonal depressive symptoms are not yet known.

Our findings should be considered in light of the following limitations. First, our study population was relatively small and was composed predominantly of Caucasians of European descent, which limits the generalizability of our findings. Additionally, our analysis was constrained by the small prevalence of seasonal affective disorder in the male population, which prevented us from determining features important to SAD in males. Another limitation is that the current study focuses on the effects of chronotype, and we did not measure other aspects of circadian rhythmicity, like circadian amplitude or distinctness, that may intersect with seasonal affective disorder. Finally, seasonality in undergraduate students is potentially confounded by different sleep–wake schedule requirements in the summer months or with other unmeasured variables that differ by season.

5. Conclusions

In this study, we report the sex-dependent and direct and indirect effects of circadian genes on seasonal affective disorder. For example, several clock-related genes show sex-specific effects that may alter glucocorticoid and serotonin regulatory pathways. Our results also suggest that CRY1 and ZBTB20 are involved in pathways that directly influence mood symptoms involved in both SAD and MDD, while PER2, CRY2, PER3C, and PER3-VNTR variants show indirect effects via mediation by diurnal preference or chronotype. Understanding the physiological pathways directly or indirectly associated with SAD can lead to effective clinical targets or behavioral therapies, respectively, to help reduce the burden associated with SAD and other mood disorders.

Author Contributions: Conceptualization, K.K.I. and A.A.; methodology, K.K.I., A.A., T.D., T.S. and L.D.; software, T.D., T.S. and A.A.; validation, T.D., T.S. and A.A.; formal analysis, K.K.I., L.D., T.D. and A.A.; investigation, L.D. and T.D.; resources, K.K.I.; data curation, T.D.; writing—original draft preparation, W.A.R., T.D. and K.K.I.; writing—review and editing, K.K.I.; visualization, T.D. and W.A.R.; supervision, K.K.I. and A.A.; funding acquisition, K.K.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: All methods adhered to the principles of the Declaration of Helsinki; the Institutional Review Board at Colgate University approved all procedures and consent forms (#FR-F13-07, #ER-F14-12, #F15-13, #ER-F16-19).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available at <https://github.com/thanh dang2712/Mood-Disorder-Project>.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fineberg, N.A.; Haddad, P.M.; Carpenter, L.; Gannon, B.; Sharpe, R.; Young, A.H.; Joyce, E.; Rowe, J.; Wellsted, D.; Nutt, D.J.; et al. The size, burden and cost of disorders of the brain in the UK. *J. Psychopharmacol.* **2013**, *27*, 761–770. [[CrossRef](#)] [[PubMed](#)]
2. Hafner, M.; Stepanek, M.; Taylor, J.; Troxel, W.M.; Van Stolk, C. Why Sleep Matters-The Economic Costs of Insufficient Sleep: A Cross-Country Comparative Analysis. *Rand Health Q.* **2017**, *6*, 11.
3. McCarthy, M.J.; Welsh, D.K. Cellular Circadian Clocks in Mood Disorders. *J. Biol. Rhythm.* **2012**, *27*, 339–352. [[CrossRef](#)] [[PubMed](#)]
4. Soria, V.; Martinez-Amoros, E.; Escaramís, G.; Valero, J.; Pérez-Egea, R.; Garcia, C.; Gutiérrez-Zotes, A.; Puigdemont, D.; Bayés, M.; Crespo, J.M.; 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*, 1279–1289. [[CrossRef](#)]
5. Zhang, L.; Hirano, A.; Hsu, P.-K.; Jones, C.R.; Sakai, N.; Okuro, M.; McMahon, T.; Yamazaki, M.; Xu, Y.; Saigoh, N.; et al. A PERIOD3 variant causes a circadian phenotype and is associated with a seasonal mood trait. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1536–E1544. [[CrossRef](#)]
6. Lee, H.-J.; Rex, K.M.; Nievergelt, C.M.; Kelsoe, J.R.; Kripke, D.F. Delayed sleep phase syndrome is related to seasonal affective disorder. *J. Affect. Disord.* **2011**, *133*, 573–579. [[CrossRef](#)] [[PubMed](#)]
7. Henderson, S.E.M.; Brady, E.M.; Robertson, N. Associations between social jetlag and mental health in young people: A systematic review. *Chronobiol. Int.* **2019**, *36*, 1316–1333. [[CrossRef](#)]
8. Islam, Z.; Hu, H.; Akter, S.; Kuwahara, K.; Kochi, T.; Eguchi, M.; Kurotani, K.; Nanri, A.; Kabe, I.; Mizoue, T. Social jetlag is associated with an increased likelihood of having depressive symptoms among the Japanese working population: The Furukawa Nutrition and Health Study. *Sleep* **2020**, *43*, zsz204. [[CrossRef](#)]
9. Levandovski, R.; Dantas, G.; Fernandes, L.C.; Caumo, W.; Torres, I.; Roenneberg, T.; Hidalgo, M.P.L.; Allebrandt, K.V. Depression Scores Associate with Chronotype and Social Jetlag in a Rural Population. *Chronobiol. Int.* **2011**, *28*, 771–778. [[CrossRef](#)]
10. Murray, J.M.; Sletten, T.L.; Magee, M.; Gordon, C.; Lovato, N.; Bartlett, D.J.; Kennaway, G.J.; Lack, L.C.; Grunstein, R.R.; Lockley, S.W.; et al. Prevalence of Circadian Misalignment and Its Association with Depressive Symptoms in Delayed Sleep Phase Disorder. *Sleep* **2017**, *40*, zsw002.
11. Nguyen, C.; Murray, G.; Anderson, S.; Filipowicz, A.; Ingram, K.K. In vivo molecular chronotyping, circadian misalignment, and high rates of depression in young adults. *J. Affect. Disord.* **2019**, *250*, 425–431. [[CrossRef](#)] [[PubMed](#)]
12. Archer, S.N.; Robilliard, D.L.; Skene, D.J.; Smits, M.; Williams, A.; Arendt, J.; von Schantz, M. A Length Polymorphism in the Circadian Clock Gene Per3 is Linked to Delayed Sleep Phase Syndrome and Extreme Diurnal Preference. *Sleep* **2003**, *26*, 413–415. [[CrossRef](#)] [[PubMed](#)]
13. Hida, A.; Kitamura, S.; Katayose, Y.; Kato, M.; Ono, H.; Kadotani, H.; Uchiyama, M.; Ebisawa, T.; Inoue, Y.; Kamei, 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*, srep06309. [[CrossRef](#)] [[PubMed](#)]
14. Kim, H.I.; Lee, H.J.; Cho, C.H.; Kang, S.G.; Yoon, H.K.; Park, Y.M.; Lee, S.-H.; Moon, J.-H.; Song, H.-M.; Lee, E.; et al. Association of CLOCK, ARNTL, and NPAS2 gene polymorphisms and seasonal variations in mood and behavior. *Chronobiol. Int.* **2015**, *32*, 785–791. [[CrossRef](#)]
15. Lavebratt, C.; Sjöholm, L.K.; Partonen, T.; Schalling, M.; Forsell, Y. PER2 variant is associated with depression vulnerability. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2010**, *153*, 570–581. [[CrossRef](#)]
16. Lavebratt, C.; Sjöholm, L.K.; Soronen, P.; Paunio, T.; Vawter, M.P.; Bunney, W.E.; Adolfsson, R.; Forsell, Y.; Wu, J.C.; Kelsoe, J.R.; et al. CRY2 Is Associated with Depression. *PLoS ONE* **2010**, *5*, e9407. [[CrossRef](#)]
17. Partonen, T.; Treutlein, J.; Alpmann, A.; Frank, J.; Johansson, C.; Depner, M.; Aron, L.; Rietschel, M.; Wellek, S.; Soronen, P.; et al. Three circadian clock genes Per2, Arntl, and Npas2 contribute to winter depression. *Ann. Med.* **2007**, *39*, 229–238. [[CrossRef](#)]
18. Shi, S.-Q.; White, M.J.; Borsetti, H.M.; Pendergast, J.S.; Hida, A.; Ciarleglio, C.M.; de Verteuil, P.A.; Cadar, A.G.; Cala, C.; McMahon, D.G.; et al. Molecular analyses of circadian gene variants reveal sex-dependent links between depression and clocks. *Transl. Psychiatry* **2016**, *6*, e748. [[CrossRef](#)]
19. Courtet, P.; Olie, E. Circadian dimension and severity of depression. *Eur. Neuropsychopharmacol.* **2012**, *22*, S476–S481. [[CrossRef](#)]
20. McCarthy, M.J.; Nievergelt, C.M.; Kelsoe, J.R.; Welsh, D.K. A Survey of Genomic Studies Supports Association of Circadian Clock Genes with Bipolar Disorder Spectrum Illnesses and Lithium Response. *PLoS ONE* **2012**, *7*, e32091. [[CrossRef](#)]
21. Terracciano, A.; Tanaka, T.; Sutin, A.R.; Sanna, S.; Deiana, B.; Lai, S.; Uda, M.; Schlessinger, D.; Abecasis, G.R.; Ferrucci, L.; et al. Genome-Wide Association Scan of Trait Depression. *Biol. Psychiatry* **2010**, *68*, 811–817. [[CrossRef](#)]
22. Mendlewicz, J. Disruption of the circadian timing systems: Molecular mechanisms in mood disorders. *CNS Drugs* **2009**, *23* (Suppl. 2), 15–26. [[CrossRef](#)]
23. Robilliard, R.; Naismith, S.L.; Rogers, N.L.; Ip, T.K.; Hermens, D.F.; Scott, E.M.; Hickie, I.B. Delayed sleep phase in young people with unipolar or bipolar affective disorders. *J. Affect. Disord.* **2013**, *145*, 260–263. [[CrossRef](#)]

24. Liberman, A.R.; Halitjaha, L.; Ay, A.; Ingram, K.K. Modeling strengthens molecular link between circadian polymorphisms and major mood disorders. *J. Biol. Rhythm.* **2018**, *33*, 318–336. [CrossRef]
25. Meesters, Y.; Gordijn, M.C. Seasonal affective disorder, winter type: Current insights and treatment options. *Psychol. Res. Behav. Manag.* **2016**, *9*, 317–327. [CrossRef]
26. Boyle, E.A.; Li, Y.I.; Pritchard, J.K. An expanded view of complex traits: From polygenic to omnigenic. *Cell* **2017**, *169*, 1177–1186. [CrossRef]
27. Ciarleglio, C.M.; Ryckman, K.K.; Servick, S.V.; Hida, A.; Robbins, S.; Wells, N.; Hicks, J.; Larson, S.A.; Wiedermann, J.P.; Johnson, C.H. Genetic differences in human circadian clock genes among worldwide populations. *J. Biol. Rhythm.* **2008**, *23*, 330–340. [CrossRef]
28. Robillard, R.; Hermens, D.F.; Naismith, S.L.; White, D.; Rogers, N.L.; Ip, T.K.; Mullin, S.J.; Alvares, G.A.; Guastella, A.J.; Smith, K.L.; et al. Ambulatory sleep-wake patterns and variability in young people with emerging mental disorders. *J. Psychiatry Neurosci.* **2015**, *40*, 28–37. [CrossRef]
29. Partonen, T. Clock gene variants in mood and anxiety disorders. *J. Neural Transm.* **2012**, *119*, 1133–1145. [CrossRef]
30. Partonen, T. Circadian Clock Proteins in Mood Regulation. *Front. Psychiatry* **2014**, *5*, 195. [CrossRef]
31. Zafar, A.; Overton, R.; Attia, Z.; Ay, A.; Ingram, K. Machine learning and expression analyses reveal circadian clock features predictive of anxiety. *Sci. Rep.* **2022**, *12*, 5508. [CrossRef] [PubMed]
32. Rosenthal, N.; Bradt, G.; Wehr, T. *Seasonal Pattern Assessment Questionnaire (SPAQ)*; National Institute of Mental Health: Bethesda, MD, USA, 1984.
33. Beck, A.T.; Beamesderfer, A. Assessment of depression: The depression inventory. *Mod. Probl. Pharmacopsychiatry* **1974**, *7*, 151–169. [PubMed]
34. Spielberger, C.D.; Gorsuch, R.L.; Lushene, R.; Vagg, P.R.; Jacobs, G.A. *Manual for the State-Trait Anxiety Inventory*; Consulting Psychologists Press: Washington, DC, USA, 1983.
35. Yu, L.; Buysse, D.J.; Germain, A.; Moul, D.E.; Stover, A.; Dodds, N.E.; Johnston, K.L.; Pilkonis, P.A. Development of Short Forms from the PROMIS™ Sleep Disturbance and Sleep-Related Impairment Item Banks. *Behav. Sleep Med.* **2011**, *10*, 6–24. [CrossRef] [PubMed]
36. Roenneberg, T.; Kuehne, T.; Juda, M.; Kantermann, T.; Allebrandt, K.; Gordijn, M.; Mero, M. Epidemiology of the human circadian clock. *Sleep Med. Rev.* **2007**, *11*, 429–438. [CrossRef] [PubMed]
37. Horne, J.A.; Ostberg, O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int. J. Chronobiol.* **1976**, *4*, 97–110.
38. Short, R.; Fukunaga, K. The optimal distance measure for nearest neighbor classification. *IEEE Trans. Inf. Theory* **1981**, *27*, 622–627. [CrossRef]
39. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
40. Huan, L.; Setiono, R. Chi2: Feature selection and discretization of numeric attributes. In Proceedings of the 7th IEEE International Conference on Tools with Artificial Intelligence, Herndon, VA, USA, 5–8 November 1995.
41. Robnik-Sikonja, M.; Kononenko, I. An adaptation of Relief for attribute estimation in regression. In *International Conference on Machine Learning*; Citeseer: Princeton, NJ, USA, 1997; Volume 5, pp. 296–304.
42. Ding, C.; Peng, H. Minimum redundancy feature selection from microarray gene expression data. *J. Bioinform. Comput. Biol.* **2005**, *3*, 185–205. [CrossRef]
43. Hall, M. Correlation-Based Feature Selection for Machine Learning. Ph.D. Thesis, The University of Waikato, Department of Computer Science, Hamilton, New Zealand, 2000; p. 19.
44. Yang, H.H.; Moody, J.E. Feature Selection Based on Joint Mutual Information. In Proceedings of the International ICSC Symposium on Advances in Intelligent Data Analysis, Rochester, NY, USA, 15 October 1999.
45. Boser, B.; Guyon, I.; Vapnik, V. A Training Algorithm for Optimal Margin Classifier. In *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*; Association for Computing Machinery: New York, NY, USA, 1992.
46. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
47. 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, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
48. RStudio Team. *RStudio: Integrated Development for R*; RStudio, PBC: Washington, DC, USA, 2020. Available online: <http://www.rstudio.com/> (accessed on 10 December 2023).
49. Bozdogan, H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika* **1987**, *52*, 345–370. [CrossRef]
50. Imai, K.; Keele, L.; Tingley, D.; Yamamoto, T. Causal Mediation Analysis Using R. In *Advances in Social Science Research Using R*; Springer: Berlin/Heidelberg, Germany, 2009.
51. Fox, J.; Weisberg, S. *An R Companion to Applied Regression*; Sage: New York, NY, USA, 2019.
52. 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.
53. Csardi, G.; Nepusz, T. The Igraph Software Package for Complex Network Research. *InterJournal Complex Syst.* **2005**, *1695*, 1–9.

54. Margolin, A.A.; Nemenman, I.; Basso, K.; Wiggins, C.; Stolovitzky, G.; Favera, R.D.; Califano, A. ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. *BMC Bioinform.* **2006**, *7*, S7. [[CrossRef](#)]
55. Hansen, K.D.; Gentry, J.; Long, L.; Gentleman, R.; Falcon, S.; Hahne, F.; Sarkar, D. Rgraphviz: Provides Plotting Capabilities for R Graph Objects, R package version 2.42.0. 2022.
56. Meyer, P.E.; Lafitte, F.; Bontempi, G. minet: A R/Bioconductor Package for Inferring Large Transcriptional Networks Using Mutual Information. *BMC Bioinform.* **2008**, *9*, 461. [[CrossRef](#)] [[PubMed](#)]
57. Benedetti, F.; Dallaspezia, S.; Fulgosi, M.C.; Lorenzi, C.; Serretti, A.; Barbini, B.; Colombo, C.; Smeraldi, E. Actimetric evidence that CLOCK 3111 T/C SNP influences sleep and activity patterns in patients affected by bipolar depression. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2007**, *144*, 631–635. [[CrossRef](#)] [[PubMed](#)]
58. Katzenberg, D.; Young, T.; Finn, L.; Lin, L.; King, D.P.; Takahashi, J.S.; Mignot, E. A CLOCK Polymorphism Associated with Human Diurnal Preference. *Sleep* **1998**, *21*, 569–576. [[CrossRef](#)]
59. Daut, R.A.; Fonken, L.K. Circadian regulation of depression: A role for serotonin. *Front. Neuroendocr.* **2019**, *54*, 100746. [[CrossRef](#)] [[PubMed](#)]
60. Pjrek, E.; Konstantinidis, A.; Assem-Hilger, E.; Praschak-Rieder, N.; Willeit, M.; Kasper, S.; Winkler, D. Therapeutic effects of escitalopram and reboxetine in seasonal affective disorder: A pooled analysis. *J. Psychiatr. Res.* **2009**, *43*, 792–797. [[CrossRef](#)]
61. Natale, V.; Adan, A.; Scapellato, P. Are seasonality of mood and eveningness closely associated? *Psychiatry Res.* **2005**, *136*, 51–60. [[CrossRef](#)]
62. Tonetti, L.; Fabbri, M.; Martoni, M.; Natale, V. Circadian type and mood seasonality in adolescents. *Psychiatry Clin. Neurosci.* **2012**, *66*, 157–159. [[CrossRef](#)]
63. Zhang, L.; Evans, D.S.; Raheja, U.K.; Stephens, S.H.; Stiller, J.W.; Reeves, G.M.; Johnson, M.; Ryan, K.A.; Weizel, N.; Vaswani, D.; et al. Chronotype and seasonality: Morningness is associated with lower seasonal mood and behavior changes in the Old Order Amish. *J. Affect. Disord.* **2015**, *174*, 209–214. [[CrossRef](#)]
64. Merikanto, I.; Suvisaari, J.; Lahti, T.; Partonen, T. Eveningness relates to burnout and seasonal sleep and mood problems among young adults. *Nord. J. Psychiatry* **2016**, *70*, 72–80. [[CrossRef](#)] [[PubMed](#)]
65. Groeger, J.A.; Viola, A.U.; Lo, J.C.; von Schantz, M.; Archer, S.N.; Dijk, D.-J. Early Morning Executive Functioning During Sleep Deprivation Is Compromised by a PERIOD3 Polymorphism. *Sleep* **2008**, *31*, 1159–1167. [[CrossRef](#)] [[PubMed](#)]
66. Hirano, A.; Shi, G.; Jones, C.R.; Lipzen, A.; Pennacchio, L.A.; Xu, Y.; Hallows, W.C.; McMahon, T.; Yamazaki, M.; Ptáček, L.J.; et al. A Cryptochrome 2 mutation yields advanced sleep phase in humans. *Elife* **2016**, *5*, e16695. [[CrossRef](#)] [[PubMed](#)]
67. Liberman, A.R.; Bin Kwon, S.; Vu, H.T.; Filipowicz, A.; Ay, A.; Ingram, K.K. Circadian Clock Model Supports Molecular Link Between PER3 and Human Anxiety. *Sci. Rep.* **2017**, *7*, 9893. [[CrossRef](#)] [[PubMed](#)]
68. Silva, A.C.P.E.; Dos Santos, M.J.; Gitaí, D.L.G.; de Miranda Coelho, J.A.P.; de Andrade, T.G. Depression and anxiety symptoms correlate with diurnal preference, sleep habits, and Per3 VNTR polymorphism (rs57875989) in a non-clinical sample. *J. Affect. Disord.* **2020**, *277*, 260–270. [[CrossRef](#)]
69. Viena, T.; Gobin, C.; Fins, A.; Craddock, T.; Tartar, A.; Tartar, J. A PER3 Polymorphism Interacts with Sleep Duration to Influence Transient Mood States in Women. *J. Circadian Rhythm.* **2016**, *14*, 3. [[CrossRef](#)]
70. Jones, S.E.; Lane, J.M.; Wood, A.R.; van Hees, V.T.; Tyrrell, J.; Beaumont, R.N.; Jeffries, A.R.; Dashti, H.S.; Hillsdon, M.; Ruth, K.S.; et al. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat. Commun.* **2019**, *10*, 343. [[CrossRef](#)]
71. Johansson, C.; Willeit, M.; Smedh, C.; Ekholm, J.; Paunio, T.; Kieseppä, T.; Lichtermann, D.; Praschak-Rieder, N.; Neumeister, A.; Nilsson, L.-G.; et al. Circadian Clock-Related Polymorphisms in Seasonal Affective Disorder and their Relevance to Diurnal Preference. *Neuropsychopharmacology* **2003**, *28*, 734–739. [[CrossRef](#)]
72. Kovanen, L.; Saarikoski, S.T.; Aromaa, A.; Lönnqvist, J.; Partonen, T. ARNTL (BMAL1) and NPAS2 Gene Variants Contribute to Fertility and Seasonality. *PLoS ONE* **2010**, *5*, e10007. [[CrossRef](#)]
73. Weiss, C.; Woods, K.; Filipowicz, A.; Ingram, K.K. Sleep Quality, Sleep Structure, and PER3 Genotype Mediate Chronotype Effects on Depressive Symptoms in Young Adults. *Front. Psychol.* **2020**, *11*, 2028. [[CrossRef](#)] [[PubMed](#)]
74. Lamia, K.A.; Papp, S.J.; Yu, R.T.; Barish, G.D.; Uhlenhaut, N.H.; Jonker, J.W.; Downes, M.; Evans, R.M. Cryptochromes mediate rhythmic repression of the glucocorticoid receptor. *Nature* **2011**, *480*, 552–556. [[CrossRef](#)] [[PubMed](#)]
75. McClung, C.A. How might circadian rhythms control mood? Let me count the ways. *Biol. Psychiatry* **2013**, *74*, 242–249. [[CrossRef](#)] [[PubMed](#)]
76. Ho, K.W.D.; Han, S.; Nielsen, J.V.; Jancic, D.; Hing, B.; Fiedorowicz, J.; Weissman, M.M.; Levinson, D.F.; Potash, J.B. Genome-wide association study of seasonal affective disorder. *Transl. Psychiatry* **2018**, *8*, 190. [[CrossRef](#)]
77. Davies, M.N.; Krause, L.; Bell, J.T.; Gao, F.; Ward, K.J.; Wu, H.; Lu, H.; Liu, Y.; Tsai, P.-C.; Collier, D.A.; et al. Hypermethylation in the ZBTB20 gene is associated with major depressive disorder. *Genome Biol.* **2014**, *15*, R56. [[CrossRef](#)]
78. Qu, Z.; Zhang, H.; Huang, M.; Shi, G.; Liu, Z.; Xie, P.; Li, H.; Wang, W.; Xu, G.; Zhang, Y.; et al. Loss of ZBTB20 impairs circadian output and leads to unimodal behavioral rhythms. *Elife* **2016**, *5*, e17171. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.