


# SCIENTIFIC REPORTS



OPEN

## Chronotype Genetic Variant in *PER2* is Associated with Intrinsic Circadian Period in Humans

Anne-Marie Chang<sup>1,2,3,4</sup>, Jeanne F. Duffy<sup>2,3</sup>, Orfeu M. Buxton<sup>1,2,3,5</sup>, Jacqueline M. Lane<sup>4,6</sup>, Daniel Aeschbach<sup>2,3,7</sup>, Clare Anderson<sup>2,3,11</sup>, Andrew C. Bjornes<sup>4,6</sup>, Sean W. Cain<sup>2,3,11</sup>, Daniel A. Cohen<sup>2,3</sup>, Timothy M. Frayling<sup>8</sup>, Joshua J. Gooley<sup>2,3,12</sup>, Samuel E. Jones<sup>8</sup>, Elizabeth B. Klerman<sup>2,3</sup>, Steven W. Lockley<sup>2,3</sup>, Mirjam Munch<sup>2,3,13</sup>, Shantha M. W. Rajaratnam<sup>2,3,11</sup>, Melanie Rueger<sup>2,3</sup>, Martin K. Rutter<sup>9,10</sup>, Nayantara Santhi<sup>2,3,14</sup>, Karine Scheuermaier<sup>2,3,15</sup>, Eliza Van Reen<sup>2,3,16</sup>, Michael N. Weedon<sup>8</sup>, Charles A. Czeisler<sup>2,3</sup>, Frank A. J. L. Scheer<sup>2,3</sup> & Richa Saxena<sup>2,4,6</sup>

The *PERIOD2* (*PER2*) gene is a core molecular component of the circadian clock and plays an important role in the generation and maintenance of daily rhythms. Rs35333999, a missense variant of *PER2* common in European populations, has been shown to associate with later chronotype. Chronotype relates to the timing of biological and behavioral activities, including when we sleep, eat, and exercise, and later chronotype is associated with longer intrinsic circadian period (cycle length), a fundamental property of the circadian system. Thus, we tested whether this *PER2* variant was associated with circadian period and found significant associations with longer intrinsic circadian period as measured under forced desynchrony protocols, the 'gold standard' for intrinsic circadian period assessment. Minor allele (T) carriers exhibited significantly longer circadian periods when determinations were based on either core body temperature or plasma melatonin measurements, as compared to non-carriers (by 12 and 11 min, respectively; accounting for ~7% of inter-individual variance). These findings provide a possible underlying biological mechanism for inter-individual differences in chronotype, and support the central role of *PER2* in the human circadian timing system.

<sup>1</sup>Department of Biobehavioral Health, Pennsylvania State University, University Park, Pennsylvania, 16802, USA. <sup>2</sup>Division of Sleep and Circadian Disorders, Department of Medicine and Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts, 02115, USA. <sup>3</sup>Division of Sleep Medicine, Harvard Medical School, Boston, Massachusetts, 02115, USA. <sup>4</sup>Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, 02142, USA. <sup>5</sup>Department of Social and Behavioral Sciences, Harvard Chan School of Public Health, Boston, Massachusetts, 02115, USA. <sup>6</sup>Department of Anesthesia, Critical Care and Pain Medicine and Center for Genomic Medicine; Massachusetts General Hospital, Boston, Massachusetts, 02114, USA. <sup>7</sup>Department of Sleep and Human Factors Research, Institute of Aerospace Medicine, German Aerospace Center, Cologne, 51147, Germany. <sup>8</sup>Genetics of Complex Traits, University of Exeter Medical School, Exeter, United Kingdom. <sup>9</sup>Division of Endocrinology, Diabetes & Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. <sup>10</sup>Manchester Diabetes Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK. <sup>11</sup>Present address: Monash Institute of Cognitive and Clinical Neurosciences and School of Psychological Sciences, Monash University, Clayton, VIC, Australia. <sup>12</sup>Present address: Program in Neuroscience and Behavioral Disorders, Duke-NUS Medical School, Singapore, Singapore. <sup>13</sup>Present address: Sleep/Wake Research Centre, College of Health, Massey University, Wellington, New Zealand. <sup>14</sup>Present address: Surrey Sleep Research Centre, University of Surrey, Guildford, UK. <sup>15</sup>Present address: Wits Sleep Laboratory, Brain Function Research Group, School of Physiology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. <sup>16</sup>Present address: Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, Providence, RI, USA. Frank A. J. L. Scheer and Richa Saxena contributed equally. Correspondence and requests for materials should be addressed to A.-M.C. (email: [amchang@psu.edu](mailto:amchang@psu.edu)) or F.A.J.L.S. (email: [fscheer@bwh.harvard.edu](mailto:fscheer@bwh.harvard.edu))

| SNP                                  | E/A | Categorical Chronotype<br>(n = 335,789) Beta (SE) | p                             | Extreme Chronotype*<br>(n = 26,056 cases, 80,065<br>controls) OR (95% CI) | p                             |
|--------------------------------------|-----|---|-------------------------------|---|-------------------------------|
| rs35333999                           | T/C | 0.058 (0.008)                                     | <b>9.7 × 10<sup>-14</sup></b> | 1.16 (1.10–1.22)  | <b>1.5 × 10<sup>-9</sup></b>  |
| rs35333999 conditioned on rs80271258 | T/C | 0.063 (0.008)                                     | <b>6.0 × 10<sup>-16</sup></b> | 1.18 (1.12–1.23)  | <b>4.6 × 10<sup>-11</sup></b> |
| rs80271258                           | T/C | 0.055 (0.006)                                     | <b>3.5 × 10<sup>-23</sup></b> | 1.14 (1.10–1.18)  | <b>5.1 × 10<sup>-13</sup></b> |
| rs80271258 conditioned on rs35333999 | T/C | 0.058 (0.006)                                     | <b>2.7 × 10<sup>-25</sup></b> | 1.15 (1.11–1.19)  | <b>2.0 × 10<sup>-14</sup></b> |

**Table 1.** Results from association testing of *PER2* rs35333999 and regional lead SNP rs80271258 with chronotype in unrelated UK Biobank participants of white British ancestry. E/A = effect/alternative allele. Beta and standard error (SE) derive from a genetic regression model showing per allelic effect of rs35333999 and rs80271258 on chronotype. SNPs rs80271258 and rs35333999 are uncorrelated in European populations ( $r^2 < 0.01$ ). Significant results ( $p < 5 \times 10^{-8}$ ) are shown in bold. Results from association testing of rs35333999 and rs80271258 conditioned on the other SNP demonstrate independence of both association signals. \*Evening types coded as cases and morning types as controls.

The *Period* gene was the first circadian gene discovered by Konopka and Benzer in *Drosophila*<sup>1</sup> and the central gene on which the concept of a transcriptional-translational feedback loop was based<sup>2</sup>. Mammalian homologues, including *PER2*, were found to be expressed in the suprachiasmatic nucleus (SCN)<sup>3</sup>, and have been shown to influence circadian period<sup>4,5</sup>. The other two mammalian homologues, *PER1* and *PER3*, have been reported to associate with measures indirectly related to circadian period, i.e., circadian phenotypes and chronotype<sup>6–11</sup>. Together, these studies highlight the central role of the *Period* genes and specifically *PER2* in the generation and maintenance of circadian rhythmicity and timing. However, to date, none of the PERs have been shown to be related to circadian period measured under the ‘gold standard’ forced desynchrony conditions in humans, necessary and designed to uncover the endogenous circadian period independent of the dark/light, sleep/wake, fasting/feeding, and other behavioral influences.

A *PER2* variant, rs35333999 (p.Val903Ile), which is common in people of European ancestry but rare in African and East Asian populations ([https://www.ncbi.nlm.nih.gov/snp/rs35333999#frequency\\_tab](https://www.ncbi.nlm.nih.gov/snp/rs35333999#frequency_tab)), was recently associated with chronotype in a large genome-wide association study (GWAS)<sup>12</sup> showing a more evening-type preference in carriers of the minor allele (T) than in those with the non-T allele (C/C) genotype. Chronotype is genetically influenced, but previous studies using a candidate-gene approach exhibited limited replicability<sup>7,8,13–22</sup>. More recently, GWAS have identified associations of single nucleotide polymorphisms with chronotype using an unbiased approach, as opposed to candidate gene testing, in large cohorts (<https://www.ncbi.nlm.nih.gov/pubmed/30696823>)<sup>12,23</sup>. These studies have identified several genetic associations in genomic regions harboring circadian genes, including genes *PER2*, *PER3*, *Vasoactive intestinal peptide (VIP)*, and *F-Box and Leucine Rich Repeat Protein 3 (FBXL3)*, which is consistent with the hypothesis that common variation influencing circadian genes impacts on phenotypic chronotype. Defining the causal variant and causal gene will require additional work, however. Importantly, *PER2* rs35333999 represents the strongest association signal that peaks within a circadian gene<sup>12</sup>. This *PER2* variant encodes a missense single nucleotide polymorphism (SNP) in exon 19 in the canonical transcript, is conserved across mammalian species, and is predicted to be deleterious to the *PER2* protein (probably damaging, score 0.963)<sup>24</sup>. Furthermore, the SNP lies in the 3’UTR of a non-canonical *PER2* transcript and may also have a regulatory role, as it is predicted to alter several transcription factor binding sites<sup>25</sup>. It is unknown, however, whether this missense variant is causal or simply a marker in linkage disequilibrium with a known or unknown causal variant. Notably, this variant is in strong linkage disequilibrium with regulatory variants rs77939198 (pair-wise  $r^2 = 0.83$ ,  $D' = 0.94$ ) and rs960783 (pair-wise  $r^2 = 0.64$ ,  $D' = 0.94$ ), both of which are annotated with chromatin marks and protein binding sites from the Roadmap Epigenomics and Encyclopedia of DNA Elements (ENCODE) projects<sup>25</sup>. We therefore hypothesized that this polymorphism would affect intrinsic circadian period, in addition to correlated phenotypes chronotype, sleep timing, circadian phase, and circadian phase angle according to genotype in this study.

## Results

We tested the association of *PER2* rs35333999 (p.Val903Ile) with self-reported chronotype (morningness-eveningness preference) in the full UK Biobank dataset of unrelated individuals of European ancestry (n up to 335,789), extending the sample size of our previous genome-wide association analysis based on an interim release of this dataset<sup>12</sup> by up to 235,369 additional individuals. We observed genome-wide significant association of rs35333999 with morningness-eveningness preference; with the rare rs35333999 T allele conferring increased eveningness in analysis of categorical chronotype ( $p = 10^{-14}$ ) and dichotomized extremes ( $p = 10^{-9}$ ; Table 1, Supplemental Fig. 1). This larger analysis confirms the SNP association seen in the previous study<sup>12</sup>, as the statistical evidence for categorical chronotype strengthened from  $p = 10^{-8}$  in the nested smaller sample (n = 100,420) to  $p = 10^{-14}$  in the current sample (n = 335,789), with similar effect estimates. Conditional association analysis adjusting for the primary association signal in this genomic region (represented by SNP rs80271258) confirmed that the association of *PER2* rs35333999 with chronotype is independent of the primary signal and itself robust (Table 1, Supplemental Fig. 2).

To discern underlying mechanisms, we tested the association of *PER2* rs35333999 with circadian phenotypic measures in 196 healthy men and women of European ancestry who participated in highly controlled inpatient physiological protocols. Demographics, circadian measures and results of association testing from these participants are presented in Table 2 and Fig. 1 and results from the entire multiethnic sample (up to n = 252)

| Phenotype              | n   | Mean (SD)     | n  | Mean T Carriers | n   | Mean Non-T Carriers | Beta (SE)    | p            |
|------------------------|-----|---------------|----|-----------------|-----|---------------------|--------------|--------------|
| Age (years)            | 196 | 26.93 (11.16) | 13 | 26.15 (5.11)    | 183 | 26.99 (11.48)       | —            | —            |
| Sex (M/F)              | 196 | 126/70        | 13 | 8/5             | 183 | 118/65              | —            | —            |
| MEQ score*             | 193 | 52.29 (12.14) | 12 | 47.17 (15.27)   | 181 | 52.63 (11.88)       | −3.87 (3.08) | 0.210        |
| CBT period (h)         | 63  | 24.16 (0.20)  | 6  | 24.34 (0.17)    | 57  | 24.14 (0.20)        | 0.20 (0.09)  | <b>0.030</b> |
| MEL period (h)         | 57  | 24.17 (0.19)  | 5  | 24.34 (0.18)    | 52  | 24.15 (0.19)        | 0.19 (0.09)  | <b>0.039</b> |
| CBT phase (clock h)    | 90  | 5:00 (2:07)   | 6  | 4:53 (1:58)     | 84  | 5:01 (2:08)         | 0.19 (0.91)  | 0.836        |
| MEL phase (clock h)    | 102 | 22:25 (1:52)  | 5  | 22:49 (1:32)    | 97  | 22:24 (1:53)        | 0.67 (0.82)  | 0.415        |
| CBT phase angle (h)    | 88  | −3.19 (1.33)  | 6  | −2.88 (1.08)    | 82  | −3.22 (1.35)        | 0.55 (0.60)  | 0.360        |
| MEL phase angle (h)    | 100 | −9.49 (1.21)  | 5  | −9.06 (1.10)    | 95  | −9.51 (1.21)        | 0.57 (0.57)  | 0.304        |
| Bedtime (clock h)      | 152 | 23:57 (1:27)  | 9  | 0:17 (1:44)     | 143 | 23:56 (1:26)        | 0.47 (0.47)  | 0.315        |
| Wake time (clock h)    | 152 | 7:59 (1:27)   | 9  | 8:17 (1:41)     | 143 | 7:58 (1:26)         | 0.42 (0.47)  | 0.380        |
| TIB duration (h)       | 152 | 8:04 (0.13)   | 9  | 7:99 (0.16)     | 143 | 8:05 (0.13)         | −0.06 (0.05) | 0.222        |
| TIB midpoint (clock h) | 152 | 3:58 (1:27)   | 9  | 4:17 (1:43)     | 143 | 3:57 (1:26)         | 0.44 (0.47)  | 0.346        |

**Table 2.** Phenotype measures by *PER2* rs35333999 genotype and results from association testing for in-laboratory participants of European ancestry. Phenotype means (standard deviation; SD) and the number of males/females are listed by rs35333999 genotype: T-allele carriers and non-T carriers. Beta and standard error (SE) derive from a genetic regression model showing per allelic effect of rs35333999 on circadian phenotypes, adjusted for age, sex and PCs. Significant results ( $p < 0.05$ ) are shown in bold. Circadian phenotypes include MEQ; circadian period, circadian phase and phase angle of CBT and melatonin (MEL). Sleep timing measures include bedtime, wake time, time-in-bed duration (TIB), and midpoint of TIB. Sleep phenotypes included in this analysis were assessed using call-in data during a fixed 8-hour schedule maintained the week prior to admission. \*A higher MEQ score reflects more morningness<sup>26</sup>.

are presented in Fig. 2, Supplemental Table 1 and Supplemental Fig. 3. Circadian phenotypes included chronotype ( $n = 193$ ), determined by morningness-eveningness questionnaire (MEQ)<sup>26</sup>; intrinsic period of endogenous circadian rhythms of core body temperature (CBT;  $n = 63$ ) and plasma melatonin ( $n = 57$ ); circadian phase of endogenous CBT ( $n = 90$ ) and melatonin ( $n = 102$ ) rhythms; measures of sleep timing while on a self-selected, fixed, 8-hour sleep schedule one week prior to inpatient studies ( $n = 152$ ); and phase angle between endogenous circadian phase markers and sleep timing for CBT ( $n = 88$ ) and melatonin ( $n = 100$ ).

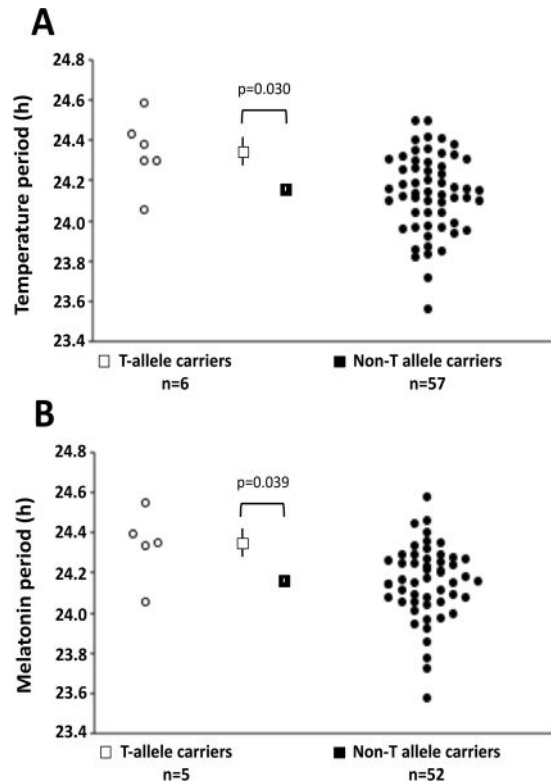
**Circadian period.** The *PER2* polymorphism was associated with intrinsic circadian period of the endogenous rhythms of core body temperature and circulating melatonin concentrations after adjustment for age, sex and population structure using 5 principal components of ancestry (Table 2 and Fig. 1). T-allele carriers exhibited a significantly longer intrinsic circadian period than non-T carriers for CBT (mean  $\pm$  SE;  $24.34 \pm 0.17$  h vs.  $24.14 \pm 0.20$  h;  $p = 0.030$ ) and melatonin ( $24.34 \pm 0.18$  h vs.  $24.15 \pm 0.19$  h;  $p = 0.039$ ). This represents a mean difference of 0.20 hours (12 minutes) for CBT period and a mean difference of 0.19 hours (11 minutes) for melatonin period between genotype groups and represents 7.1% of population variance in temperature period and 6.9% of variance in melatonin period in participants of European ancestry from our sample. Results were similar in the full multiethnic sample (see Fig. 2 and Supplemental Table 1). Sensitivity analysis using the unequal variances t-test to account for unequal sample size (Welch's t-test for rs35333999 T carriers vs CC) shows a consistent direction of effect in those of European ancestry (melatonin period  $p = 0.189$ ; CBT period  $p = 0.097$ ) and reached significance for temperature period in the multiethnic sample (melatonin period  $p = 0.053$ ; CBT period  $p = 0.032$ ).

**Chronotype.** Analysis of rs35333999 with composite MEQ score showed no significant association in the European-only group ( $p = 0.210$ ; Table 2) or in the full multiethnic sample ( $p = 0.120$ ; Supplemental Table 1 and Supplemental Fig. 3). Given the magnitude of the effect observed in the large UK Biobank GWAS study ( $n = 100,420$ )<sup>12</sup>, statistical power to detect the association with chronotype in the current analysis was only 6.2% in participants of European ancestry, and only 6.7% in the multi-ethnic sample.

## Discussion

The T-allele of *PER2* rs35333999 was associated with self-reported evening chronotype in a population-scale sample and with longer intrinsic circadian period as assessed in highly controlled in-laboratory circadian protocols. The consistent direction and magnitude of effect on intrinsic circadian period based on two endogenous circadian phase markers, CBT and circulating plasma melatonin rhythms, suggests a difference in the intrinsic period of the master circadian pacemaker located in the hypothalamic suprachiasmatic nucleus (SCN). The significant association between the minor allele of rs35333999 and longer period of body temperature and plasma melatonin is also consistent with published reports showing an association of evening chronotype with longer circadian period<sup>27,28</sup>.

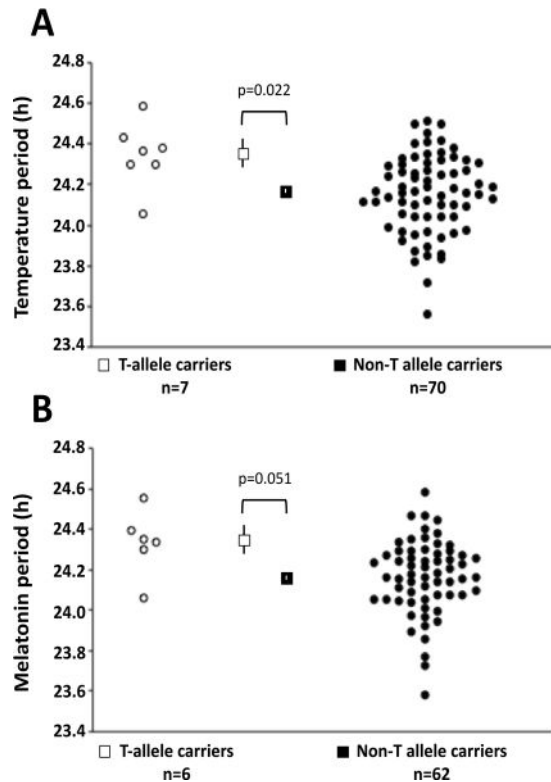
This is the first report of a genetic association with intrinsic circadian period directly measured in humans using forced desynchrony protocols, specifically designed to precisely determine this physiological measure (see Methods section)<sup>29,30</sup>. A previous study reported genetic linkage between another *PER2* marker (D2S395) and Advanced Sleep Phase Disorder in a multi-generational family<sup>31</sup>. This very rare mutation (minor allele frequency



**Figure 1.** Circadian period of core body temperature and melatonin rhythms by *PER2* rs35333999 genotype in in-laboratory study participants of European ancestry. Data for mean (squares) and individual (circles) circadian period of core body temperature (A) and melatonin (B) are shown for T-allele carriers (open symbols) and for non-T allele carriers (filled symbols). Vertical lines denote the standard error (SE) measures of the mean. T-allele carriers had significantly longer temperature and melatonin circadian periods than non-carriers ( $p = 0.030$  and  $0.039$ , respectively).

<0.1%) that causes a serine to glycine change leading to hypophosphorylation of PER2 protein, was found to be associated with an early circadian timing of sleep, extreme morning chronotype, and a shorter observed period of the sleep-wake and temperature cycles in one individual, though intrinsic period free of influence of self-selected light exposure was not assessed in that study (<https://www.ncbi.nlm.nih.gov/pubmed/10470086>). In the current study, a relatively common variant of *PER2* (5% allele frequency in the 1000 Genomes European population and 3.6% in our population of European ancestry) was associated with intrinsic circadian period assessed under conditions free from the influence of self-selected light exposure. This was achieved by studying participants under dim light conditions (<15 lux), and by distributing that light exposure across all circadian phases outside the range of entrainment of the human circadian pacemaker. This was done in order to minimize the effects of light on circadian period. These conditions are critical for accurate assessment of intrinsic circadian period, as retinal light exposure has a powerful resetting effect on the human circadian pacemaker that would otherwise confound assessment of intrinsic circadian period<sup>32</sup>. Furthermore, because there are aftereffects of prior light exposure on circadian period<sup>33</sup>, we tried to limit differences in prior history by excluding recent shiftwork and transatlantic flights, and had participants maintain a fixed 8-h time in bed for the weeks prior to admission to the laboratory, thereby standardizing the prior light history.

The main strength of our study is the precise assessment of endogenous circadian measures in carefully designed circadian protocols conducted under rigorously controlled laboratory conditions that are considered “gold standard” procedures for determination of circadian period and phase in humans. While chronotype is a measure that is relatively easily obtainable using a questionnaire, and therefore scalable and reflects people’s preferences, it provides less insight into the relative contribution of environmental/behavioral factors (including school, work, and other social demands, exposure to light, caffeine, buildup and discharge of sleep pressure, and social demands) *versus* biology. Furthermore, chronotype cannot distinguish among different biological causes, such as differences in endogenous circadian period, circadian rhythm robustness, circadian light sensitivity, and homeostatic sleep regulation time constants. For example, the recent finding that *PER2* is associated with reduced sensitivity of the circadian pacemaker to light in humans<sup>34</sup>, highlights the relative benefits of biological measures of circadian rhythmicity in studying the potential mechanisms underlying chronotype differences. Circadian period provides a plausible biological mechanism of the link between the *PER2* variant and chronotype, and also provides testable hypotheses for future molecular mechanism studies in animal and *in vitro* models. Intrinsic circadian period is a key factor determining behavioral timing. A longer circadian period, leading to a delay of the circadian timing system—including the circadian drive for wake and sleep—is a plausible pathway to a later



**Figure 2.** Circadian period of CBT and melatonin rhythms by *PER2* rs35333999 genotype in the multiethnic in-laboratory study sample. Data for mean (squares) and individual (circles) circadian period of core body temperature (A) and melatonin (B) are shown for T-allele carriers (open symbols) and for non-T allele carriers (filled symbols). Vertical lines denote the standard error (SE) measures of the mean. T-allele carriers had significantly longer circadian period of temperature ( $p = 0.022$ ) and a trend for longer circadian period of melatonin ( $p = 0.051$ ) than non-carriers.

chronotype. Thus, to assess the most proximal biological mechanism, we focused specifically on intrinsic circadian period. The second reason to focus on intrinsic circadian period as a potential mechanism underlying the link between the *PER2* variant and chronotype is that in rodent mutant models, *PER2* has been shown to play a key role in determining circadian period of the rest activity cycle<sup>4,5,31</sup>. A third reason to focus primarily on circadian period, as opposed to the downstream measures of circadian phase (i.e., timing of circadian markers such as core body temperature and melatonin), sleep timing, circadian phase angle (i.e., relative timing of circadian markers and sleep timing), and chronotype, is that intrinsic circadian period as assessed under forced desynchrony protocols can be determined with great precision<sup>29</sup>, enhancing the statistical power for this outcome measure. This may be why those other circadian measures did not show significant differences, though they trended consistently later in the T-allele carriers than in the non-carriers, consistent with a longer period<sup>27,35</sup>.

While significant associations between rs35333999 and circadian period were identified because of a large effect (Cohen's  $d$  was 1.08 for CBT period and 1.03 for melatonin period), we likely did not see significant associations of the SNP with circadian phase due to the overall smaller effect (Cohen's  $d$  was 0.07 for CBT phase and 0.22 for melatonin phase). One possible explanation is that circadian period is a more proximal phenotype, i.e., one that is influenced directly by the *PER2* SNP, while phase angle is a measure that is influenced by circadian period, but also by exposure to photic and non-photoc *Zeitgebers* (time cues), as well as the accumulation and dissipation of homeostatic sleep pressure<sup>36</sup>. Another potential explanation involves the assessment of sleep timing during the week prior to the inpatient protocols in participants who maintained a self-selected, 8-hour sleep duration schedule (see Methods section). Therefore, we were unable to assess the influence of rs35333999 on unrestricted sleep timing measures or other circadian phenotypes that incorporate them (i.e., phase angle), and may explain why there was no significant association with these phenotypes. Further work assessing this SNP in larger datasets would be necessary to test this. We also did not replicate the association of *PER2* rs35333999 with chronotype in the current study. The lack of a significant effect on chronotype was likely due to insufficient power in this sample ( $n = 246$ ) relative to the UK Biobank nested sample ( $n = 100,420$ )<sup>12</sup> or the larger sample ( $n = 335,789$ ) in the current analysis. Another explanation could be that, similar to phase, chronotype is a less proximal phenotype than circadian period and influenced by other factors. While here we show an association of the *PER2* SNP rs35333999 with human endogenous circadian period, future studies are needed to test whether this SNP is the causal variant. Such verification will require further experimental studies (e.g., *ex vivo* cell culture recordings<sup>37</sup>, or phenotype rescue in animal models)<sup>38</sup>. Furthermore, further studies are needed to test whether this SNP may also relate to homeostatic sleep regulation, which has also associates with chronotype<sup>36</sup>.



Although rs35333999 is associated with self-reported chronotype<sup>12</sup> and intrinsic circadian period as shown here, it is not known if this is the causal variant, or what effect, if any, the allele may have on cell-autonomous circadian rhythms. This missense SNP is predicted to have deleterious effects on PER2 protein and is also predicted to play a role in the circadian molecular mechanism by potentially altering transcription factor binding sites<sup>24</sup>. Future studies are needed to determine the molecular and cellular mechanism by which rs35333999 affects circadian period. Different research approaches for testing the effect of genetic variants have different strengths and limitations. The use of self-reported data collected in humans under ‘free living conditions’, as is the case for the association of the SNP with self-reported chronotype in the UK Biobank, have the benefit that these assessments can be conducted in large scale studies and have the largest generalizability to the human condition. Such populations can be useful for discovery of genetic variants influencing behavioral preferences. The downside is that they cannot assess physiological mechanisms. The benefit of the association of a SNP with an accurately assessed physiological phenotype, as is the case for endogenous circadian period under the forced desynchrony protocol, is that the physiological mechanism can be tested. The molecular mechanism, however, cannot be determined in these studies and requires genetic manipulation studies *in vivo*, or *in vitro*. The systemic integration of the effect of the SNP on the physiology can only be assessed in an *in vivo* system, while the physiological circadian mechanism in humans can only accurately be tested in human circadian protocols. The combination of these approaches will give a complete picture of the molecular and physiological mechanisms as well as the relevance in humans. Thus, ultimately, *in vitro* experiments are needed to test the influence of rs35333999 on molecular mechanisms, including potential alteration of the role of the PER2 protein in the repressor complex of the circadian clock and its translocation to the nucleus<sup>39</sup>.

**Limitations.** Genetic analyses of the in-laboratory population included a small sample size of rs35333999 minor-allele (T) carriers for several circadian measures. This was due to the relatively low frequency of this variant, and the intensive nature of the studies required to obtain accurate and precise assessments of human intrinsic circadian period. Despite the limited sample size, we found significant differences of large magnitude in intrinsic circadian period, both when using core body temperature and plasma melatonin assessments during forced desynchrony protocols. Future studies in larger populations, including other genetic ancestries, are needed to verify and expand this finding. As previously noted, we did not find an association between rs35333999 and chronotype, as assessed by composite MEQ score, in the in-laboratory sample. The MEQ has been widely used as a validated measure of diurnal preference and it has been shown that the MEQ, the reduced MEQ (rMEQ) and the Composite Scale of Morningness (CSM) have high correlative validity<sup>40</sup>; however, these and other scales do not necessarily measure the same thing, especially in a cross-cultural context. For example, in a sample of Korean adults, the global CSM score as well as two of three subscales (“morningness” and “activity planning”) were associated with a non-synonymous *PER2* SNP rs934945<sup>20</sup>. Interestingly, this same *PER2* SNP was tested for association with chronotype using the CSM in a sample of Columbian adults and found that two subscales (“activity planning” and “morning alertness”) were associated but not total CSM score or the “morningness” subscale<sup>22</sup>. While the current work does not replicate or directly overlap those findings (i.e., a different, uncorrelated *PER2* SNP tested and different measure of chronotype), our findings and previous reports, taken together, suggest that circadian regulation of behavior and diurnal preference involves *PER2*, and as we now show, may involve circadian period.

## Conclusion

This is the first study to demonstrate a human genetic variant that associates with intrinsic circadian period. Our results show that the *PER2* rs35333999 T allele is associated with a longer intrinsic circadian period, which provides a likely underlying mechanism for the evening chronotype that was previously associated with this SNP. Remarkably, the single SNP explained 7% of the inter-individual differences in intrinsic circadian period (11–12 min). To put this into perspective, this difference in intrinsic circadian period conveyed by this single SNP is twice as large as the average difference in intrinsic circadian period found between men and women, which has been proposed to underlie sex differences in bedtimes, wake times, and sleep disturbances<sup>41</sup>. Differences in intrinsic circadian period and chronotype may also influence the susceptibility for diabetes and other metabolic disorders (<https://www.ncbi.nlm.nih.gov/pubmed/30705391>)<sup>42,43</sup>. Future studies will be necessary to reveal the molecular mechanism by which the genetic variant influences circadian period and chronotype and to determine its effect on other chronotype-associated physiologic or disease phenotypes.

## Materials and Methods

**UK Biobank sample.** Study participants were from the UK Biobank prospective study (n = 503,325 adults aged 40–69 years), described in detail elsewhere<sup>44</sup>. For the current analysis, individuals of self-reported non-white ethnicity and relatives were excluded to avoid confounding effects. Phenotype and genotype data from 335,789 unrelated individuals of self-reported British white ancestry who gave informed consent for genetics were used in the current analysis. The UK Biobank study was approved by the National Health Service National Research Ethics Service (ref 11/NW/0382), all participants provided written informed consent to participate in the UK Biobank study, and procedures were conducted in accordance with the relevant guidelines and requirements.

**In-lab patient sample.** The population sample consisted of 252 healthy men and women who had previously participated in inpatient physiological research protocols at the Brigham and Women’s Hospital between 2001 and 2011. Pre-study screening criteria for inclusion consisted of general good health, no medical or psychological/psychiatric conditions and no medications other than certain oral contraceptives; no shiftwork in the prior 3 years; and no travel across more than 1 time zone in the prior 3 months. Written informed consent was obtained from participants prior to enrollment in this genetic study and they were compensated for their

participation. All study procedures received approval by the Partners Health Care Human Research Committee and were conducted in accordance to the *Declaration of Helsinki*.

**Phenotypes.** *Chronotype.* In the UK Biobank sample, chronotype was self-reported by a touch-screen questionnaire at the baseline visit<sup>12</sup>. Chronotype was examined both as a categorical trait ( $n = 335,789$ ) and a dichotomous trait, with definite morning responders set to control ( $n = 80,065$ ) and definite evening responders set to case ( $n = 26,056$ ).

In the in-laboratory sample, chronotype was assessed by MEQ<sup>26</sup> from study participants prior to their admission to the laboratory. This 19-question survey was scored using published criteria and summed to obtain a composite score. MEQ composite score was used as a continuous variable in this analysis.

*Sleep timing.* For a subset of the inpatient laboratory studies, participants were required to maintain a stable, self-selected, 8-hour time-in-bed (TIB) schedule for 1–3 weeks prior to admission to the inpatient protocol. Adherence to this schedule was verified by several means including: wrist actigraphy, sleep diaries, and time-stamped daily call-ins at bedtimes and wake times. Bedtime, wake time, TIB duration, and midpoint of TIB, were determined from the average call-in times during the week (7 days) immediately prior to admission. Participants who did not maintain an 8-hour schedule were not included in the analysis of sleep timing.

*In-laboratory circadian phenotypes.* Some of the inpatient laboratory protocols assessed circadian phase and/or period using specific study procedures. Specifically, timing of the CBT<sub>min</sub> and the dim light melatonin onset (DLMO) were used as circadian phase markers collected during constant routine or constant posture conditions (see<sup>45</sup> and<sup>46</sup> for detailed description). Intrinsic circadian period of CBT and of melatonin were assessed using forced desynchrony protocols<sup>29,41,47,48</sup> as previously described<sup>29,30</sup>. Forced desynchrony protocols involved the collection of continuous CBT measurement and/or hourly blood sampling during a minimum of two weeks while participants maintained a sleep/wake cycle that was substantially greater than 24 hours (e.g.,  $T = 28$  or  $T = 42.85$  hours). This allowed for the assessment of circadian measures under a  $T$  cycle to which the circadian system cannot entrain while light intensity was maintained at 0 lux during scheduled sleep and <15 lux during scheduled wakefulness. Additionally, circadian phase angle for CBT (difference between timing of CBT<sub>min</sub> and wake time averaged from the week prior to admission) and melatonin (difference between DLMO and wake time) were determined.

**Genotypes.** UKBiobank study: Genotyping was performed by the UK Biobank, and genotyping, quality control, and imputation procedures have been described<sup>49</sup>. In brief, blood was collected from participants, and DNA was extracted from the buffy coat samples. Participant DNA was genotyped on two arrays, UK BiLEVE and UKB Axiom with >95% common content and genotypes for ~800,000 SNPs were imputed to the Haplotype Reference Consortium reference panel<sup>50</sup>. We examined association of directly genotyped SNP rs35333999 and genotyped and well-imputed common variants (minor allele frequency >0.001 and imputation quality score >0.3) within a 400 kb region to capture multiple independent association signals.

In-laboratory Study: DNA was extracted from venous blood samples collected from each in-laboratory study participant. The *PER2* SNP rs35333999 was genotyped using the Sequenom platform that distinguishes allele-specific primer extension products by mass spectrometry (MALDI-TOF; Broad Institute, Cambridge, Massachusetts)<sup>51</sup>. Quality control and Hardy Weinberg Equilibrium were assessed using the genetic software PLINK<sup>52</sup>. An additional 58 African-American and Hispanic ancestry informative markers were genotyped using the same platform to identify the subset of subjects with genetic European ancestry and to correct for population stratification<sup>53,54</sup>.

**Association analysis.** Association testing in both cohorts employed an additive genetic model of *PER2* rs35333999 T with self-reported chronotype or circadian measures using linear or logistic regression analysis in PLINK and adjustment for covariates: age, sex, and five principal components of ancestry. In the UK Biobank, single SNP association analysis to chronotype was performed for SNPs in a 400 kb region encompassing the *PER2* gene and included genotyping array as an additional covariate. Conditional genetic association analysis included further adjustment for SNPs (rs35333999 or regional lead SNP rs80271258) in regression models. P-values <  $5 \times 10^{-8}$  were considered significant. Analyses of in-laboratory participants of European ancestry ( $n = 196$ ) as well as for the entire multiethnic sample ( $n = 252$ ) were performed. In the multi-ethnic sample, an additional race/ethnicity covariate was included. P-values < 0.05 were considered significant for associations between *PER2* rs35333999 and circadian phenotypes. Sensitivity analysis was performed using the Welch's t-test for unequal variances<sup>55</sup> testing T allele carriers (TC and TT) vs CC homozygotes on covariate-adjusted residuals (age, sex, and 5 principal components of ancestry) for both melatonin and core-body temperature period. The additive effect of the TT genotype was not considered in this sensitivity analysis.

## Data Availability

The data that support the findings of this study from the UK BioBank will be made available at <https://sleepgenetics.org> and the underlying genotype and phenotype data are available through application to the UK Biobank. Other phenotype data are available on request, due to privacy or other restrictions, through co-corresponding author Dr. Scheer ([fscheer@bwh.harvard.edu](mailto:fscheer@bwh.harvard.edu)).

## References

- Konopka, R. J. & Benzer, S. Clock mutants of drosophila melanogaster. *Proc. Natl. Acad. Sci. USA* **68**, 2112–2116 (1971).
- Hardin, P. E., Hall, J. C. & Rosbash, M. Feedback of the drosophila period gene product on circadian cycling of its messenger rna levels. *Nature* **343**, 536–540 (1990).

3. Takumi, T. *et al.* A new mammalian period gene predominantly expressed in the suprachiasmatic nucleus. *Genes Cells* **3**, 167–176 (1998).
4. Zheng, B. *et al.* The mper2 gene encodes a functional component of the mammalian circadian clock. *Nature* **400**, 169–173 (1999).
5. Bae, K. *et al.* Differential functions of mper1, mper2, and mper3 in the scn circadian clock. *Neuron* **30**, 525–536 (2001).
6. Ebisawa, T. *et al.* Association of structural polymorphisms in the human period3 gene with delayed sleep phase syndrome. *EMBO Journal* **2**, 342–346 (2001).
7. Archer, S. N. *et al.* A length polymorphism in the circadian clock gene per3 is linked to delayed sleep phase syndrome and extreme diurnal preference. *Sleep* **26**, 413–415 (2003).
8. Carpen, J. D., von schantz, M., Smits, M., Skene, D. J. & Archer, S. N. A silent polymorphism in the per1 gene associates with extreme diurnal preference in humans. *J Hum Genet* **51**, 1122–1125 (2006).
9. Lim, A. S. *et al.* A common polymorphism near per1 and the timing of human behavioral rhythms. *Annals of Neurology* **72**, 324–334 (2012).
10. Lee, K. A. *et al.* Circadian regulation gene polymorphisms are associated with sleep disruption and duration, and circadian phase and rhythm in adults with hiv. *Chronobiol. Int.* **32**, 1278–1293, <https://doi.org/10.3109/07420528.2015.1087021> (2015).
11. Zhang, L. *et al.* A period3 variant causes a circadian phenotype and is associated with a seasonal mood trait. *Proc. Natl. Acad. Sci. USA*, <https://doi.org/10.1073/pnas.1600039113> (2016).
12. Lane, J. M. *et al.* Genome-wide association analysis identifies novel loci for chronotype in 100,420 individuals from the uk biobank. *Nat Commun* **7**, 10889, <https://doi.org/10.1038/ncomms10889> (2016).
13. Katzenberg, D. *et al.* A clock polymorphism associated with human diurnal preference. *Sleep* **21**, 569–576 (1998).
14. Mishima, K., Tozawa, T., Satoh, K., Saitoh, H. & Mishima, Y. The 3111t/c polymorphism of hclock is associated with evening preference and delayed sleep timing in a japanese population sample. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **133**, 101–104 (2005).
15. Archer, S. N. *et al.* Polymorphism in the per3 promoter associates with diurnal preference and delayed sleep phase disorder. *Sleep* **33**, 695–701 (2010).
16. Johansson, C. *et al.* Circadian clock-related polymorphisms in seasonal affective disorder and their relevance to diurnal preference. *Neuropsychopharmacology* **28**, 734–739 (2003).
17. Pedrazzoli, M. *et al.* Clock polymorphisms and circadian rhythms phenotypes in a sample of the brazilian population. *Chronobiol. Int.* **24**, 1–8 (2007).
18. Chang, A. M., Buch, A. M., Bradstreet, D. S., Klements, D. J. & Duffy, J. F. Human diurnal preference and circadian rhythmicity are not associated with the clock 3111c/t gene polymorphism. *J Biol Rhythms* **26**, 276–280 (2011).
19. Barclay, N. L. *et al.* Sleep quality and diurnal preference in a sample of young adults: Associations with 5httlpr, per3, and clock 3111. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **156B**, 681–690, <https://doi.org/10.1002/ajmg.b.31210> (2011).
20. Lee, H. J. *et al.* Per2 variation is associated with diurnal preference in a korean young population. *Behavior Genetics* **41**, 273–277, <https://doi.org/10.1007/s10519-010-9396-3> (2011).
21. Osland, T. M., Bjorvatn, B. R., Steen, V. M. & Pallesen, S. Association study of a variable-number tandem repeat polymorphism in the clock gene period3 and chronotype in norwegian university students. *Chronobiol. Int.* **28**, 764–770, <https://doi.org/10.3109/07420528.2011.607375> (2011).
22. Ojeda, D. A. *et al.* A novel association of two non-synonymous polymorphisms in per2 and per3 genes with specific diurnal preference subscales. *Neuroscience Letters* **553**, 52–56, <https://doi.org/10.1016/j.neulet.2013.08.016> (2013).
23. Hu, Y. *et al.* Gwas of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nat Commun* **7**, 10448, <https://doi.org/10.1038/ncomms10448> (2016).
24. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248–249, <https://doi.org/10.1038/nmeth0410-248> (2010).
25. Ward, L. D. & Kellis, M. Haploreg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res* **44**, D877–881, <https://doi.org/10.1093/nar/gkv1340> (2016).
26. Horne, J. A. & Östberg, O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology* **4**, 97–110 (1976).
27. Duffy, J. F., Rimmer, D. W. & Czeisler, C. A. Association of intrinsic circadian period with morningness-eveningness, usual wake time, and circadian phase. *Behav Neurosci* **115**, 895–899 (2001).
28. Duffy, J. F. & Czeisler, C. A. Age-related change in the relationship between circadian period, circadian phase, and diurnal preference in humans. *Neurosci Lett* **318**, 117–120 (2002).
29. Czeisler, C. A. *et al.* Stability, precision, and near-24-hour period of the human circadian pacemaker. *Science* **284**, 2177–2181 (1999).
30. Brown, E. N. & Czeisler, C. A. The statistical analysis of circadian phase and amplitude in constant-routine core-temperature data. *J Biol Rhythms* **7**, 177–202 (1992).
31. Toh, K. L. *et al.* An hper2 phosphorylation site mutation in familial advanced sleep phase syndrome. *Science* **291**, 1040–1043 (2001).
32. Zeitzer, J. M., Dijk, D. J., Kronauer, R. E., Brown, E. N. & Czeisler, C. A. Sensitivity of the human circadian pacemaker to nocturnal light: Melatonin phase resetting and suppression. *J Physiol* **526**, 695–702 (2000).
33. Scheer, F. A., Wright, K. P. Jr., Kronauer, R. E. & Czeisler, C. A. Plasticity of the intrinsic period of the human circadian timing system. *PLoS ONE* **2**, e721 (2007).
34. Akiyama, T. *et al.* An ancestral haplotype of the human period2 gene associates with reduced sensitivity to light-induced melatonin suppression. *PLoS One* **12**, e0178373, <https://doi.org/10.1371/journal.pone.0178373> (2017).
35. Mongrain, V., Lavoie, S., Selmaoui, B., Paquet, J. & Dumont, M. Phase relationships between sleep-wake cycle and underlying circadian rhythms in morningness-eveningness. *J Biol Rhythms* **19**, 248–257 (2004).
36. Mongrain, V., Carrier, J. & Dumont, M. Circadian and homeostatic sleep regulation in morningness-eveningness. *J Sleep Res* **15**, 162–166 (2006).
37. Azzi, A. *et al.* Network dynamics mediate circadian clock plasticity. *Neuron* **93**, 441–450, <https://doi.org/10.1016/j.neuron.2016.12.022> (2017).
38. Antoch, M. P. *et al.* Functional identification of the mouse circadian clock gene by transgenic bac rescue. *Cell* **89**, 655–667 (1997).
39. Aryal, R. P. *et al.* Macromolecular assemblies of the mammalian circadian clock. *Mol Cell* **67**, 770–782 e776, <https://doi.org/10.1016/j.molcel.2017.07.017> (2017).
40. Di Milia, L., Adan, A., Natale, V. & Randler, C. Reviewing the psychometric properties of contemporary circadian typology measures. *Chronobiol. Int.* **30**, 1261–1271, <https://doi.org/10.3109/07420528.2013.817415> (2013).
41. Duffy, J. F. *et al.* Sex difference in the near-24-hour intrinsic period of the human circadian timing system. *Proceedings of the National Academy of Science USA* **108**, 15602–15608 (2011).
42. Reutrakul, S. *et al.* Chronotype is independently associated with glycemic control in type 2 diabetes. *Diabetes Care* **36**, 2523–2529, <https://doi.org/10.2337/dc12-2697> (2013).
43. Yu, J. H. *et al.* Evening chronotype is associated with metabolic disorders and body composition in middle-aged adults. *J Clin Endocrinol Metab* **100**, 1494–1502, <https://doi.org/10.1210/jc.2014-3754> (2015).
44. Allen, N. E., Sudlow, C., Peakman, T., Collins, R. & Biobank, U. K. Uk biobank data: Come and get it. *Science Translational Medicine* **6**, 224ed224, <https://doi.org/10.1126/scitranslmed.3008601> (2014).



45. Cain, S. W. *et al.* Sex differences in phase angle of entrainment and melatonin amplitude in humans. *J. Biol Rhythms* **25**, 288–296 (2010).
46. Klerman, E. B., Gershengorn, H. B., Duffy, J. F. & Kronauer, R. E. Comparisons of the variability of three markers of the human circadian pacemaker. *J. Biol Rhythms* **17**, 181–193 (2002).
47. Cohen, D. A. *et al.* Uncovering residual effects of chronic sleep loss on human performance. *Science Translational Medicine* **2**, 14ra13 (2010).
48. Buxton, O. M. *et al.* Adverse metabolic consequences in humans of prolonged sleep restriction combined with circadian disruption. *Science Translational Medicine* **4**, 129ra143, <https://doi.org/10.1126/scitransmed.3003200> (2012).
49. Bycroft, C. *et al.* Genome-wide genetic data on ~500,000 uk biobank participants. *bioRxiv*, <https://doi.org/10.1101/166298> (2017).
50. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics* **48**, 1279–1283, <https://doi.org/10.1038/ng.3643> (2016).
51. Bradic, M., Costa, J. & Chelo, I. M. Genotyping with sequenom. *Methods Mol. Biol.* **772**, 193–210, [https://doi.org/10.1007/978-1-61779-228-1\\_11](https://doi.org/10.1007/978-1-61779-228-1_11) (2011).
52. Purcell, S. *et al.* Plink: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* **81**, 559–575, <https://doi.org/10.1086/519795> (2007).
53. Lane, J. M. *et al.* Impact of common diabetes risk variant in mtnr1b on sleep, circadian and melatonin physiology. *Diabetes*, <https://doi.org/10.2337/db15-0999> (2016).
54. Chang, A. M. *et al.* Circadian gene variants influence sleep and the sleep electroencephalogram in humans. *Chronobiol. Int.*, 1–13, <https://doi.org/10.3109/07420528.2016.1167078> (2016).
55. Welch, B. L. The generalisation of student's problems when several different population variances are involved. *Biometrika* **34**, 28–35 (1947).

## Acknowledgements

We thank the study participants; the staff of the Center for Clinical Investigation of the Brigham and Women's Hospital, the Chronobiology Core, and the recruitment office of the Division of Sleep and Circadian Disorders, for their assistance with the in-laboratory studies. We also thank the staff of the Partners Healthcare Personalized Medicine Biobank and the Broad Institute for their work on the genetic samples. This research has also been conducted using the UK Biobank Resource under application number 6818. We would like to thank the participants and researchers from the UK Biobank who contributed or collected data. This analysis was supported by National Institutes of Health (NIH) grant R21DK089378 and by a pilot grant from the Harvard Catalyst, The Harvard Clinical and Translational Science Center (UL1 TR001102) and financial contributions from Harvard University and its affiliated academic healthcare centers to F.A.J.L.S. and R.S. Collection of DNA from study participants was supported by NIH grants F32HL078360, R01HL080978, P01AG009975, and R21DK089378. Data collection in the inpatient protocol was supported by the following grants: NIH R01AG06072, R01HL077453, P01AG009975, R21AT002571, R01HL080978, R01NS054277, R01MH45130, R01HL093279, R01HL094654, R01HL077399, and UL1 TR001102; AFOSR FA9550-06-0080; NSBRI HFP01601; and NARSAD Young Investigator Award. A.-M.C. was further supported by T32HL007901, F32HL078360, and K01HL115458; E.B.K. by NIH K24HL105664, P01AG009975, R01GM105018, R01HL114088, and R21HD086392 and NSBRI HFP02802; J.M.L. by F32DK102323; K.S. by T32HL07901 and F32AG031690; F.A.J.L.S. by R01HL140574, R01HL094806, R01DK099512, R01HL118601, R01DK102696, and R01DK105072; and R.S. by R01DK107859, R01DK102696, R01DK105072, R21HL121728 and R01HL113338. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, the NIH, AFOSR, or NSBRI.

## Author Contributions

Study design of genetic association study: A.-M.C., J.F.D., O.M.B., C.A.C., F.A.J.L.S. and R.S. Data collection/analysis/interpretation of UK Biobank sample: J.M.L., T.M.F., S.E.J., M.K.R., M.N.W. and R.S. Data collection of in-laboratory phenotypes: A.-M.C., J.F.D., O.M.B., D.A., C.A., S.W.C., D.C., J.J.G., E.B.K., S.W.L., M.M., S.M.W.R., M.R., N.S., K.S., E.V.R. and C.A.C. Data analysis of in-lab phenotypes: A.-M.C. Data analysis of genotypes: A.B., J.M.L. and R.S. Writing of the first draft: A.-M.C. Editing draft: all co-authors. Obtained funding: A.-M.C., J.F.D., O.M.B., D.A., S.W.C., E.B.K., S.W.L., C.A.C., F.A.J.L.S. and R.S.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-41712-1>.

**Competing Interests:** None of the following relationships reported by authors are related to the present article. O.M.B. received travel support and honoraria from Chevron for consulting. E.B.K. has consulted for Pfizer Inc. and a legal firm, and has received travel reimbursements from the Society for Reproductive Investigation, the Association for Professional Sleep Societies, Wire In-Brain Conference, the Sleep Technology Council, Free Health LLC, and the Employer Health Benefit Congress. S.W.L. has no conflicts of interest related to the current manuscript. He has current consulting contracts with Akili Interactive, Consumer Sleep Solutions, Delos Living LLC, Headwaters Inc., Hints Performance AG, Light Cognitive, Mental Workout, OpTerra Energy Services Inc., Pegasus Capital Advisors LP, and Wyle Integrated Science and Engineering; he has received unrestricted equipment gifts through the Brigham and Women's Hospital from Biological Illuminations LLC; Bionetics Corporation; Philips Lighting; ResMed Inc and F. Lux Software LLC; he has served as a paid expert witness in arbitration and legal cases related to sleep, circadian rhythms, work hours and/or light; he serves as a Program Leader in the Cooperative Research Centre for Alertness, Safety and Productivity. S.M.W.R. reports that he has served as a consultant through his institution to Vanda Pharmaceuticals, Teva Pharmaceuticals, Philips Respironics, EdanSafe, The Australian Workers' Union, National Transport Commission, and Transport Accident Commission, and has through his institution received research grants and/or unrestricted educational grants from Vanda Pharmaceuticals, Takeda Pharmaceuticals North America, Philips Lighting, Philips Respironics, Cephalon, and ResMed Foundation, Rio Tinto, and reimbursements for conference travel expenses

from Vanda Pharmaceuticals. His institution has received equipment donations or other support from Optalert, Compumedics, and Tyco Healthcare. He has also served as an expert witness and/or consultant to shift work organizations. He serves as a Program Leader and a consultant to the Alertness CRC. M.R. reports receiving salary from Valkee Oy, Oulu, Finland as well as Shire Pharmaceuticals in Germany. M.R. also reports providing consulting services for Inteliclinic in Warsaw, Poland. C.A.C. is/was a paid consultant to Bose, Boston Celtics, Boston Red Sox, Columbia River Bar Pilots, Institute of Digital Media and Child Development, Jazz Pharma, Merck, Purdue Pharma, Quest Diagnostics, Samsung, Teva, Vanda Pharmaceuticals, Inc. and V-Watch/PPRS; has received lecture fees from Global Council on Brain Health/AARP, Integritas Communications Group, Maryland Sleep Society, National Sleep Foundation and Zurich Insurance Company, Ltd.; holds equity in Vanda Pharmaceuticals, Inc.; receives research/education support from Cephalon, Mary Ann & Stanley Snider via Combined Jewish Philanthropies, Jazz Pharma, Optum, ResMed, San Francisco Bar Pilots, Schneider, Simmons, Sysco, Koninklijke Philips Electronics, Vanda Pharmaceuticals, Inc.; is/was an expert witness in legal cases, including those involving Bombardier, Columbia River Bar Pilots, Continental Airlines, Fedex, Greyhound, Purdue Pharma, UPS; serves as the incumbent of a professorship endowed by Cephalon; and receives royalties from McGraw Hill, Houghton Mifflin Harcourt, and Philips Respironics for the Actiwatch-2 & Actiwatch Spectrum devices. C.A.C.'s interests were reviewed and are managed by Brigham and Women's Hospital and Partners HealthCare in accordance with their conflict of interest policies. F.A.J.L.S. has received lecture fees from Bayer HealthCare, Sentara HealthCare, Philips, Vanda Pharmaceuticals, and Pfizer.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019