CLOCK Gene Variants Associate with Sleep Duration in Two Independent Populations

Karla V. Allebrandt, Maris Teder-Laving, Mahmut Akyol, Irene Pichler, Bertram Müller-Myhsko, Peter Pramstaller, Martha Merrow, Thomas Meitinger, Andreas Metspalu, and Till Roenneberg

**Background:** Sleep is an active and complex behavior, yet it has two straightforward properties—timing and duration. Clock genes are associated with dysfunctional timing of sleep, mood, and obesity disorders, which are commonly associated with sleep duration.

**Methods:** Sleep duration was assessed in Central Europe, Estonia, and South Tyrol ($n \approx 77,000$) with the Munich ChronoType Questionnaire. It showed a Gaussian distribution in all investigated populations after averaging over a standard workweek and normalization according to age and gender. A follow-up, two-stage design, linkage disequilibrium-based association study was conducted with subjects from South Tyrol (discovery sample; $n = 283$) and with short (< 7 hours) and long (> 8.5 hours) sleepers from Estonia (confirmation sample; $n = 1011$). One hundred ninety-four single nucleotide polymorphism markers covering 19 candidate clock genes were genotyped in the discovery sample, and two of the best association signals (analyzed by a linear regression model) were investigated in the confirmation sample.

**Results:** Single and multi-marker associations were found within a **CLOCK** gene intronic region (rs12649507 and rs11932595). In a meta-analysis between South Tyrol and Estonia association signals, rs12649507 ($p = .0087$) remained significant. Significance persisted only for the multiple-marker association signal of the rs12649507/rs11932595 haplotype GGAA with long sleep ($p = .0015$).

**Conclusions:** We report an association between variants of the human **CLOCK** gene and sleep duration in two independent populations. This adds another putative function for **CLOCK** besides its possible involvement in circadian timing, depression, obesity, and personality.

**Key Words:** **CLOCK**, sleep duration, clock genes, MCTQ, short sleepers, long sleepers

Sleep is one of the most conspicuous aspects of our existence, drastically changing our state of consciousness for approximately one-third of our lives. The question “why do we sleep?” is one of the hot topics in biology. On the one hand, even small doses of sleep deprivation cause decrements in performance, and complete sleep deprivation is fatal in experimental animals; but on the other hand, pharmacological agents such as modafinil can be used to allay sleep with remarkably small side effects. Hypotheses regarding why we sleep concern neuronal reorganization and repair (1–4). We remain, however, far from fully understanding the function of sleep.

Despite its complexity, sleep has two straightforward characteristics—timing and duration. They are influenced by two physiological mechanisms: a homeostatic (how long have we been awake) and a circadian process (5,6). The latter refers to the circadian clock, an endogenous mechanism that specifies a temporal window—a phase—within the day for a myriad of processes, from gene expression to behavior, including sleep. The timing of sleep is often used to assess entrainment of the circadian clock. This entrainment is specific to individuals, yielding different so-called chronotypes (7).

Sleep timing can be distinguished from sleep duration. There are just as many short and long sleepers among early chronotypes as there are among late ones (7). Sleep duration has been associated with disease and mortality (8); personality traits, such as self-criticism, emotional reliance on another person, or agreeableness (9,10); and learning and memory (11). Lack of sleep has been associated with obesity and diabetes (12). The effect of sleep deprivation on performance shows large differences among self-reported short (<6 hours) versus long sleepers (>9 hours); short sleepers apparently tolerate a higher homeostatic sleep pressure than long sleepers (13), presumably because a genetically regulated mechanism allows them to recover sleep homeostasis faster or to better withstand sleep pressure relative to long sleepers. Other sleep qualities, besides timing, have been associated with so-called clock genes (those involved in generating the circadian rhythm at the cellular level). Sleep structure and homeostasis, for example, correlate with **PER3** gene variants (14), and clock genes (**PER1** and **PER2**) expression correlates with sleep deprivation and recovery phase in mice (15). The temporal regulation of non-rapid eye movement sleep is compromised in **NPAS2** −/− mice (16).

Sleep duration has a genetic component corresponding to 40% heritability, as shown in twin studies (17), but its genetics will surely not be simple (18). Timing and duration of sleep are influenced by the circadian clock, as indicated by functional studies on model organisms (19–22) and genetic association studies (for review see 23,24). Associations between sleep duration and gene variants in general human populations have so far been addressed only once by a small-scale genome-wide study (25).
Although the genetics of sleep duration has been addressed in humans, associations between clock genes and sleep duration have never been investigated systematically. We therefore chose 19 clock genes as candidates and investigated a set of high-density markers in a two-stage design association study, involving populations from South Tyrol (discovery sample) and Estonia (confirmation sample). Single nucleotide polymorphism markers (SNPs) were selected on the basis of a linkage disequilibrium (LD) strategy or on their possible relevance as nonsynonymous or splicing region variants. Sleep duration depends both on age and on gender. We therefore used our Central European and Estonian databases (n ≈ 65,000 and n ≈ 12,000, respectively) to establish algorithms describing the inherent and systematic relationships. These were then used to normalize the phenotype of sleep duration (averaged over a standard week).

Here, we report the association between sequence variants in the human CLOCK gene and sleep duration in two independent European populations. Because we have investigated general populations, it is likely that the influence of the variability of this gene on the distribution of sleep duration is not an isolated case and might apply to other populations.

Methods and Materials

Subjects

All subjects were selected from ethnically homogeneous, Caucasian populations. The local ethic committees approved all subjective assessments and genotyping analyses. The discovery sample was recruited from healthy individuals—participants of Caucasian populations. The local ethic committees approved all subjects (total population size n ≈ 65,000 and n ≈ 12,000, respectively) to establish algorithms describing the inherent and systematic relationships. These were then used to normalize the phenotype of sleep duration (averaged over a standard week).

Here, we report the association between sequence variants in the human CLOCK gene and sleep duration in two independent European populations. Because we have investigated general populations, it is likely that the influence of the variability of this gene on the distribution of sleep duration is not an isolated case and might apply to other populations.

Phenotyping

Average sleep duration was assessed with the short version of the MCTQ (29). Sleep-onset was calculated by adding sleep latency to the time of sleep preparation, and sleep duration was calculated by subtracting sleep onset from sleep end. All values were assessed separately for workdays and for free days (SDw and SDf, respectively). Subjects were also asked how many days/week they worked (Nwd). Average sleep duration (SDav) was calculated by the following equation:

$$SD_{av} = (N_{wd} \times SD_w + [7 - N_{wd}] \times SD_f)/7$$

The SDav, was then corrected for its dependencies on age and gender, yielding the final quantitative phenotype: SDasc. Algorithms for the normalization of age- and gender-dependent changes were based on two databases comprising more than 65,000 subjects from Central Europe (mainly Germans but also Austrians, Dutch, and Swiss) and the Estonian database (n = 12,000; Figures 1A and 1B). The equations of the fits (Table S1 in Supplement 1) were used to normalize sleep duration to a theoretical “neutral” 30 years old.

Marker Selection

We screened 19 clock genes, distributed across 10 chromosomes (Figure 2). Because most of these genes still await validation as clock genes in humans, coding sequence homologies of animal orthologs (e.g., the Drosophila gene jetlag corresponds to the human gene FBXL15) were determined with the help of the BLAST program (http://www.ncbi.nlm.nih.gov/). One hundred ninety-four SNPs (Table S2 in Supplement 1) across the 19 genes were selected from three sources. An LD-based approach with the Caucasian (CEU) HapMap Phase II panel produced 155 tags (criteria were minor allele frequency [MAF] > 20% and r² > .8); the National Center for Biotechnology Information database gave rise to 24 nonsynonymous variants; 5 splice region variants were selected from the Ensembl genome browser. The number of markers selected/gene region (including up and downstream regions) depended mostly on the LD of these markers with surrounding variants (tagging with r² cutoff = .8), assessed with the software Haplovie (30) as predicted by the Tagger-pairwise Tagging algorithm. Redundant tags replaced tag SNPs that did not work with the iPLEX assay (Sequenom, San Diego, California). Nonsynonymous or splicing region variants were not necessarily polymorphisms and were selected independently of the LD-based approach.

Genotyping

The DNA was extracted from ethylenediaminetetraacetic acid-treated blood samples with a salting out procedure (31) and quantified with a NanoDrop 1000 spectrophotometer (Thermo Scientific, Waltham, Massachusetts). Discovery samples were genotyped by polymerase chain reaction (PCR) primer extension assay with a Sequenom iPLEX genotyping platform, following the manufacturer’s instructions. After amplification of target sequences and before the primer extension reaction of the iPLEX protocol, the DNA amplification quality (24 wells/plate) was verified in an agarose gel. Base-extension products were detected in a MassARRAY time-of-flight mass spectrometry system (32). Replication samples were genotyped by using Taqman SNP Genotyping Assays on an Applied Biosystems (Foster City, California) 7900HT Fast Real-Time PCR machine (33). The PCR reactions contained 20 μL (10 ng) of DNA, 10 μL of TaqMan genotyping assays (5 μL forward and 5 μL reverse primers), 2 μL of 2X Taqman Master Mix, and 2 μL of 8X PCR-grade water. The reactions were performed in a 96-well plate. The cycling parameters were: 10 min at 95°C, 40 cycles of 95°C for 15 sec and 60°C for 1 min. The plate was analyzed with the SDS software (33).
(Applied Biosystems) Genotyping Master Mix (2×), and 1 μL of Taqman Pre-Designed SNP Genotyping Assay (20×). Samples were denatured at 95°C for 10 min followed by 40 cycles at 92°C for 15 sec and 60°C for 1 min. Genotyping was performed in two genotyping facilities, the Helmholtz Zentrum München and the Estonian Biocentre.

Statistical Analysis
Quantitative trait association was tested with linear regression model analysis for association of normalized sleep duration with the genotypes with the PLINK (http://pngu.mgh.harvard.edu/purcell/plink/) package version 1.04 (34). Multi-marker analysis was performed with the haplotype-based quantitative trait association test from the PLINK package with a sliding window of two SNPs. A meta-analysis was conducted with the R package rmeta version 2.15 to estimate the magnitude of the combined effect of the results from both samples.

Results

Phenotyping Sleep Duration
Optimal phenotyping is an important prerequisite for achieving unequivocal results in phene-gene association studies. Similar to sleep timing (35), sleep duration depends systematically on both age and gender (Figure 1). Access to large databases on sleep behavior allowed us to establish equations (Table S1 in Supplement 1) that were used for normalizing sleep duration with regard to the influences of age and gender (adults in the Central European sample: women, n = 30,959, r = .87; men, n = 28,236, r = .87; adults in the Estonian sample: women, n = 8035, r = .86; men n = 3676, r = .73). Additional fits for both young women and men (ages 10–20) were possible for the large Central European sample (n = 5111, r = .93 and n = 3904, r = .96, respectively). The relatively small South Tyrolean sample did not allow significant estimations (women: n = 178, r = .14; men: n = 155, r = .13). We applied the correction algorithms established in the Central European database to the South Tyrolean data, on the basis of the high correlation of the age- and gender-dependent changes in Central Europe and Estonia (women: r = .9; slope = 1.3; men: r = .8; slope = 1.1).

Figure 1B shows the distributions of the averaged weekly sleep duration for the Central European, Estonian, and Tyrolean populations before and after normalization. After normalization SDasc distribution and mean were not different between the Estonian and Tyrolean populations (n = 5098, SDasc Ø = 8.07, SDasc Ø = 8.14, respectively; t test: p > .1; analysis of variance p > .05, Figure 1B). Before normalization, they were significantly different (p < .0001).

For the replication study, 507 short (<7 hours) and 504 long (>8.5 hours) sleepers were selected from the distribution of normalized average sleep duration in the Estonian database (Figure 1B). This strategy is powerful for association studies, because genetic differences between the extremes of a quantitative trait distribution should become more evident (36). We have not selected the extremes from the total pool of phenotyped subjects used for normalization (approximately 12,000 subjects, Figure 1), because by the time genotyping started only a subset of the DNA samples were available.

Figure 1. Sleep duration depends on age and gender. (A) On average, women (open circles) sleep longer than men (filled circles). The age-dependency of not-normalized averaged weekly sleep duration is plotted, separately for the genders, for the large Central European database (cEUR DB) (upper left panel) as well as the Estonian (middle left panel) sample annually and South Tyrolean (bottom left panel) sample on a 5-year basis. For details see text. (B) The distributions of weekly average sleep duration in the different populations. The cEUR DB (upper right panel), Estonia (middle right panel), and South Tyrol (bottom right panel). The distributions of the raw data are drawn as stippled curve, and the distribution of the age and gender corrected data are drawn as bars. For orientation purposes, a vertical line is drawn through sleep duration of 8 hours.
Genotyping Efficiency

Total call rate in the discovery sample was .94. Fifty of the 333 individuals were removed for low genotyping (call rate/individual was between 80% and 89%), because a large number of markers were genotyped. All SNPs included in the analysis had an MAF >.01 and total call rate of .99. Two SNPs were genotyped in the confirmation sample (Estonia; n = 1011) with a call rate of 100% and Hardy–Weinberg-equilibrium (HWE) (p >.05) in both samples. All SNPs with a call rate under 85% were excluded from the analysis (18 of the 194 investigated variants). Nineteen monomorphic variants and 5 variants that were not in HWE (cutoff: p <.01) were excluded from the association analysis. Only two cases of heterozygote underrepresentation were present among investigated SNPs, and those were not in HWE, as expected. All other functional SNPs (except monomorphic ones) had a higher proportion of heterozygotes in relation to the homozygotes. The average allele frequency was .328 ± .110. The MAFs for rs12649507 and rs11932595 in the confirmation sample were .30 and .49, respectively. Marker names, genomic location, gene, function, alleles, HWE, and minor frequencies are listed in Table S2 in Supplement 1.

Association Analysis

The most significant phenotype-genotype associations in the discovery sample (from South Tyrol) were found for SNPs located in the intronic region of the CLOCK gene (rs12649507, p = .0051; rs11932595, p = .0080; Figure 3, Table 1). These alleles were not in LD with each other. The fact that we have removed 50 of the subjects of the discovery sample to achieve a total call rate of 99% did not modify the final results. When including all subjects for the analysis, the two reported associations were still the most significant (rs12649507, p = .0059; rs11932595, p = .0021). From the six markers tagging the CLOCK gene (Figure 2), only five were considered for the association analysis (rs6850524, rs9312662, rs11732723, rs11932595, and rs12649507)—one of them, rs111333911, was eliminated on the basis of a low call rate. The number of investigated SNPs/gene did not bias the discovered associations. Other clock genes investigated in this study had many more variants than the CLOCK gene (Figure 2). The SNPs in four other genes (NPAS2, Timeless, CSNK1D, and Per3) produced p values <.05 for the association with SDasc (Table 1) but were above the threshold of p <.01 that we had set for the second stage of this study. Neither of the aforementioned results (the associations in CLOCK nor those in the four additional genes) remained significant after
Bonferroni correction for multiple comparisons. Assuming that a
two-stage design study would be more meaningful than an
empiric correction, we tried to replicate our findings for the best
SNPs (rs12649507 and rs11932595; \( \alpha \) level \( p < .01 \)) in an
independent sample from our Estonian database. Indeed, signifi-
cant associations were found for rs12649507 and rs11932595
(\( p < .05; \) Table 1). However, only the rs12649507 association
could be replicated, whereas a heterogeneous effect was ob-
served for rs11932595 in comparison with the discovery sample.
Thereby, we conducted a meta-analysis of the results from both
samples validated only the association of rs12649507/
rs11932595 in comparison with the discovery sample. Due to the limited sample size of the
association described here. Due to the limited sample size of the
multiple markers associations.

<table>
<thead>
<tr>
<th>Genes</th>
<th>SNP rs</th>
<th>Location</th>
<th>Discovery Sample</th>
<th>Confirmation Sample</th>
<th>Meta-Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>p/Effect Allele</td>
<td>n</td>
<td>p/Effect Allele</td>
<td>p</td>
</tr>
<tr>
<td>CLOCK</td>
<td>rs12649507</td>
<td>Intron 1</td>
<td>282</td>
<td>.0051/A</td>
<td>.0087</td>
</tr>
<tr>
<td></td>
<td>rs11932595</td>
<td>Intron 11</td>
<td>275</td>
<td>.0080/G</td>
<td></td>
</tr>
<tr>
<td>NPAS2</td>
<td>rs7598826</td>
<td>Intron 1</td>
<td>281</td>
<td>.0255/A</td>
<td></td>
</tr>
<tr>
<td>TIMELESS</td>
<td>rs2291738</td>
<td>Intron 22 - boundary</td>
<td>281</td>
<td>.0207/G</td>
<td></td>
</tr>
<tr>
<td>CSNK1D</td>
<td>rs7209167</td>
<td>Intron 1</td>
<td>277</td>
<td>.0363/G</td>
<td></td>
</tr>
<tr>
<td>PER3</td>
<td>rs10462021</td>
<td>Exon 20</td>
<td>268</td>
<td>.0468/G</td>
<td></td>
</tr>
</tbody>
</table>

Multiple-Markers Associations

<table>
<thead>
<tr>
<th>Genes</th>
<th>SNP rs</th>
<th>Effect Alleles</th>
<th>Discovery Sample</th>
<th>Confirmation Sample</th>
<th>Meta-Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>p/Effect Allele</td>
<td>n</td>
<td>p/Effect Allele</td>
<td>p</td>
</tr>
<tr>
<td>CLOCK</td>
<td>rs11932595</td>
<td>rs12649507</td>
<td>GG</td>
<td>274</td>
<td>.0003</td>
</tr>
<tr>
<td></td>
<td>rs11932595</td>
<td>rs11932595</td>
<td>AA</td>
<td></td>
<td>.0005</td>
</tr>
<tr>
<td></td>
<td>rs11932595</td>
<td>rs11932595</td>
<td>GA</td>
<td></td>
<td>.4103</td>
</tr>
<tr>
<td></td>
<td>rs11932595</td>
<td>rs11932595</td>
<td>AG</td>
<td></td>
<td>.7995</td>
</tr>
<tr>
<td>CLOCK</td>
<td>rs11732723</td>
<td>rs11932595</td>
<td>GG</td>
<td>275</td>
<td>.0008</td>
</tr>
<tr>
<td></td>
<td>rs11732723</td>
<td>rs11932595</td>
<td>AA</td>
<td></td>
<td>.0862</td>
</tr>
<tr>
<td></td>
<td>rs11732723</td>
<td>rs11932595</td>
<td>GA</td>
<td></td>
<td>.0997</td>
</tr>
<tr>
<td></td>
<td>rs11732723</td>
<td>rs11932595</td>
<td>AG</td>
<td></td>
<td>.6185</td>
</tr>
<tr>
<td>TIMELESS</td>
<td>rs10783783</td>
<td>rs2291739</td>
<td>TC</td>
<td>257</td>
<td>.0008</td>
</tr>
<tr>
<td></td>
<td>rs10783783</td>
<td>rs2291739</td>
<td>TT</td>
<td></td>
<td>.1397</td>
</tr>
<tr>
<td></td>
<td>rs10783783</td>
<td>rs2291739</td>
<td>CT</td>
<td></td>
<td>.3784</td>
</tr>
<tr>
<td></td>
<td>rs10783783</td>
<td>rs2291739</td>
<td>CC</td>
<td></td>
<td>.4096</td>
</tr>
</tbody>
</table>

Gene name, single nucleotide polymorphism marker (SNP) rs code, location, sample size, and nominal \( p \) values/effect alleles are listed for the discovery and replication samples. Dashed lines mean that the respective SNPs have not been genotyped in the confirmation sample.

Table 1. Associations of Single and Multi-Markers with Normalized Average Sleep Duration

Discussion

The distribution of weekly averaged sleep duration depends on age and gender in Central Europe and Estonia (Figures 1A and 1B). The most drastic shortening in sleep duration (by approximately 2 hours) occurs between childhood and the end of puberty. From that age until approximately the age of meno-
pause (at approximately the age of 50), men sleep on average (15 min) less than women, in accordance with earlier studies (37). This difference has been associated with concentration changes of sex hormones (38). Sleep duration decreases steadily in both genders from the end of adolescence (at approximately the age of 20) to the age of menopause (39). From that age on, the gender differences disappear, and sleep duration increases slightly with age (Figures 1A and 1B). The confounding influences of these systematic dependencies were eliminated by our
normalization procedure to optimize phenotyping for the genetic associations described here. Due to the limited sample size of the South Tyrolean database (Figure 1A), normalization for this
discovery sample was based on the Central European database.

Despite the large variables set (152 variants distributed over 19 genes) investigated in the discovery sample, the top
two most significant signals were located within the same
gene, the \( \text{CLOCK} \) gene—an essential transcription factor of the
molecular circadian clock. \( \text{CLOCK} \) encodes a basic helix-loop-
helix–PAS (a domain named after three proteins it occurs in: Per, Arnt, and Sim) transcription factor (40), which forms a
complex with the protein BMAL1. This heterodimer is the
activating component of a negative feedback-loop that is
critically involved in circadian rhythm generation at the
cellular/molecular level (41,42). In addition to its role within
the circadian system, functional studies showed \( \text{CLOCK} \) to be
also involved in sleep regulation (21).
related to the circadian mechanism. The CLOCK
CLOCK
CLOCK
tions, indicating that sleep duration has a genetic component
gene variants with sleep duration in two independent popula-
the discovery (South Tyrol) and confirmation (Estonia) samples.

K.V. Allebrandt

region. Between the analyzed
of this gene (CEU HapMap data), mostly located in intronic
tigated in this study cosegregate with 59 other polymorphic sites
CLOCK
CLOCK
CLOCK
a marker for a polymorphic synonymous coding variant of
CLOCK
CLOCK
CLOCK
CLOCK
CLOCK
CLOCK
CLOCK

In this study, we have found the association of the CLOCK
gene variants with sleep duration in two independent populations,
indicating that sleep duration has a genetic component
related to the circadian mechanism. The CLOCK markers inves-
tigated in this study cosegregate with 59 other polymorphic sites
of this gene (CEU HapMap data), mostly located in intronic region. Between the analyzed CLOCK SNPs, one (rs9312662) was
a marker for a polymorphic synonymous coding variant of
CLOCK (rs3736544) and for several intron boundary variants, but
significant association of this marker with sleep duration was not
found in this study. The CLOCK gene nonsynonymous exonic variants described until now are mostly rare mutations and
cannot be responsible for a quantitative phenotype such as sleep
duration. Short (<7 hours) and long sleepers (>8.5 hours) comprise approximately one-third in each of the investigated
populations (Figure 1B). We are thus searching for frequent polymorphisms that could be associated with the relatively
frequent extreme cases of sleep duration. If low frequency alleles
would be contributing to this phenotype, we would rather
expect them to be at coding or splicing region, so most of those
have been selected in addition to the tagSNPs.

The CLOCK variant rs12649507 association signal was unidirec-
tional in both populations and remained significant in a meta-
analysis combining the independent results (Table 1). This
polymorphism is located in intron one of the CLOCK gene, a
region of limited variation (Figure S1 in Supplement 1). It
cosegregates ($r^2 > .80$) with several neighboring variants (Table
S3 in Supplement 1), covering a region of approximately 112 kb
of the chromosome 4, including the CLOCK gene 5’ flanking
region. The association of this SNP (rs12649507) showed the
same tendency in both samples (i.e., sleep duration was de-
creased in subjects carrying the “A” allele). This was not the case
for the SNP rs11932595, because “AA” homozygous for this
polymorphism showed the shortest sleep duration in the South
Tyrolean and the longest in the Estonian population (association
signal that lost significance in the meta-analysis). This heteroge-
neous effect could have at least three reasons. The first reason is
the different levels of LD of this marker with the causal allele.
This SNP rs11932595 is located within a recombination region
(Figure S1 in Supplement 1), subjected to fluctuations on LD
levels in recently split populations (43). Allele frequencies for
rs11932595 were different ($p = 0.0320$) between the discovery
(A/G = .54/ .46) and confirmation samples (A/G = .49/.51). The
discovery sample had relatively more Ad “homozygous, whereas
the confirmation sample had more GG homozygous, being
homozygous frequencies significantly different ($p = 0.0346$).
The second reason is different genetic environment. The extent of
the contribution of other variants, implicated in this polygenic inher-
tance, plays a critical role in shaping the phenotype distribution.
The third reason is phenotypic heterogeneity. The number of
subjects assessed in the discovery sample was not representative for
a quantitative phenotype distribution such as sleep duration. Only
the confirmation sample haplotype association signal held its sig-
nificance in the meta-analysis (Table 1). Indeed, the confirmation
sample should give us a more precise estimate of the associations
found than the discovery sample, because: 1) a sample comprising
subjects with extreme sleep duration should increase the power to
find genetic associations with this phenotype, because the larger
effect observed on the tails of the distribution should reflect the
contribution of several alleles from many quantitative trait loci; and
2) the sample size of the confirmation sample is three times larger
than the discovery sample.

This is the first report that associates variants of the CLOCK
gene with average sleep duration in human populations. An
earlier study showed that a deletion of CLOCK’s exon 19 (due to
a point mutation) shortened both the duration of sleep (by 1–2
hours) and that of rapid eye movement phases (21). Concerning
other aspects of sleep behavior in human populations, a CLOCK
gene polymorphism was associated with insomnia, an apparent
inability to consolidate sleep (T3111C; rs1801260) located in the
3’ flanking region of the CLOCK gene (44). It was also associated
with temporal sleep preferences, so-called morningness and
eveningness (45,46). We have not included this SNP in our
analysis, because its genotyping call rate was too low; it coseg-
regates with another SNP (rs7665846) that was included (Table
S3 in Supplement 1). However, significant associations of
rs7665846 were not found in this study, either for sleep duration
or for sleep timing. Candidate gene approaches have yielded
mixed information concerning CLOCK and PER3, with some
studies reporting associations of SNPs or polymorphisms with
sleep timing in a general population and others failing to confirm
this, as reviewed elsewhere (23). The instrument used for
phenotyping the subjects in these studies was a questionnaire
that sought to identify early and late types as opposed to an
instrument that would indicate the actual timing of sleep.
Without taking sleep duration into account, a long sleeper could be
classified as an early and/or a late sleeper, depending on
whether onset of sleep or wake time is considered. Similar
misinterpretations would apply to short sleepers. Therefore,
associations of those early and late types with CLOCK and PER3 polymorphisms could have been biased by sleep duration.

Our top hit (rs12649507) cosegregates with a CLOCK variant (rs6843722) that has been associated with obesity in candidate genes studies (47,48). Although obesity is known to correlate with sleep duration (49), we have not found an association of rs12649507 with body mass index. This is not surprising, because there were only extremely weak correlations between normalized body mass index (Figure S2 in Supplement 1) and normalized sleep duration in our discovery ($r = 0.029$) and confirmation samples (short SD, $r = 0.010$; long SD, $r = 0.141$). Another SNP (rs6850524) that is associated both with obesity (47) and with bipolar disorder (50) was investigated in the first stage of this study (South Tyrolean sample). Again, we found no association of sleep duration with it or other CLOCK SNPs in the discovery and confirmation samples. Because bipolar disorder is—like in the case of obesity—not independent of sleep duration (patients often have a decreased sleep need), a systematic picture might emerge: CLOCK’s association with different pathologies involves effects on sleep duration. However, this remains speculative, because causal links between these disturbances and sleep regulation are still controversial (51,52).

The function of CLOCK as a transcription factor or as a histone acetyltransferase implies that the genetic variability of this gene could influence its function on gene transcriptional activation. Environmental influences could also regulate CLOCK function as a histone acetyltransferase and in turn affect sleep duration. Whether through epigenetics, gene–gene interaction, or isolated polymorphisms, the association of a CLOCK variant with sleep duration is indicated here. This finding should in turn support further functional studies.

Our work was supported by the FPS program EUCLOCK (TR, KVA, MM, MT, and AM), the Dutch Science Foundation (the NWO), The Hersenen Stichting, and the Rosalind Franklin Fellowships of the University of Groningen. The Estonian group was supported by Targeted Financing from Estonian Government SF0180142, European Union through the European Regional Development Fund in the frame of Centre of Excellence in Genomics, and 7 FP Project ECONEGENE and OPENGENE. In South Tyrol, the study was supported by the Ministry of Health and Department of Educational Assistance, University and Research of the Autonomous Province of Bolzano, and the South Tyrolean Sparkasse Foundation. We thank the participants in the Estonian Genome project and the South Tyrolean study participants for donating their time and DNA.

All authors report no biomedical financial interests or potential conflicts of interest.

Supplementary material cited in this article is available online.


www.sobp.org/journal


