RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART B
Neuropsychiatric Genetics

A Genome-Wide Association Study of Sleep Habits and Insomnia

Enda M. Byrne,[1,2]* Philip R. Gehrman,[3] Sarah E. Medland,[1] Dale R. Nyholt,[1] Andrew C. Heath,[4] Pamela A. F. Madden,[4] Ian B. Hickie,[5] Cornelia M. Van Duijn,[6] Anjali K. Henders,[1] Grant W. Montgomery,[1] Nicholas G. Martin,[1] Naomi R. Wray,[1,2] and The Chronogen Consortium

[1]Queensland Institute of Medical Research, Brisbane, Queensland, Australia

[2]The University of Queensland, Queensland Brain Institute, St. Lucia, Queensland, Australia

[3]Department of Psychiatry & Penn Sleep Center, Behavioral Sleep Medicine Program, University of Pennsylvania, Philadelphia, Pennsylvania

[4]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri

[5]Brain and Mind Research Institute, University of Sydney, Sydney, Australia

[6]Department of Epidemiology, Unit of Genetic Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands

Manuscript Received: 27 June 2012; Manuscript Accepted: 11 April 2013

Several aspects of sleep behavior such as timing, duration and quality have been demonstrated to be heritable. To identify common variants that influence sleep traits in the population, we conducted a genome-wide association study of six sleep phenotypes assessed by questionnaire in a sample of 2,323 individuals from the Australian Twin Registry. Genotyping was performed on the Illumina 317, 370, and 610K arrays and the SNPs in common between platforms were used to impute non-genotyped SNPs. We tested for association with more than 2,000,000 common polymorphisms across the genome. While no SNPs reached the genome-wide significance threshold, we identified a number of associations in plausible candidate genes. Most notably, a group of SNPs in the third intron of the *CACNA1C* gene ranked as most significant in the analysis of sleep latency ($P = 1.3 \times 10^{-6}$). We attempted to replicate this association in an independent sample from the Chronogen Consortium (n = 2,034), but found no evidence of association ($P = 0.73$). We have identified several other suggestive associations that await replication in an independent sample. We did not replicate the results from previous genome-wide analyses of

How to Cite this Article:
Byrne EM, Gehrman PR, Medland SE, Nyholt DR, Heath AC, Madden PAF, Hickie IB, Van Duijn CM, Henders AK, Montgomery GW, Martin NG, Wray NR The Chronogen Consortium. 2013. A Genome-Wide Association Study of Sleep Habits and Insomnia.
Am J Med Genet Part B 162B:439–451

IB Hickie declares the following potential conflicts of interest: Director of Headspace, National Youth Mental Health Foundation Ltd.; Member, Mental Health Expert working group, Department of Health and Ageing, Australian Federal Government. Member of Bupa Australia Medical Advisory Panel; Educational and Research programs/grants that are supported by Pharmaceutical Industry: Servier, Pfizer, Astrazeneca, Eli Lilly.
Consortium contributors listed in the Supplementary Material.

Grant sponsor: Pharmaceutical Industry Servier, Pfizer, Astrazeneca, Eli Lilly (Educational and Research programs); Grant sponsor: EMB NHMRC 613608; Grant sponsor: Australian National Health Medical Research Council; Grant numbers: 241944; 339462; 389927; 389875; 389891; 389892; 389938; 442915; 442981; 496739; 552485; 552498; 613608; Grant sponsor: Australian Research Council; Grant numbers: A7960034; A79906588; A79801419; DP0770096; DP0212016; DP0343921; Grant sponsor: FP-5 GenomEUtwin Project; Grant number: QLG2-CT-2002-01254; Grant sponsor: US National Institutes of Health; Grant numbers: NIH AA07535; AA10248; AA13320; AA13321; AA13326; AA14041; MH66206; Grant sponsor: Genetic Cluster Computer, Netherlands Scientific Organization; Grant number: NWO 480-05-003; Grant sponsor: National Health and Medical Research Council (NHMRC) Fellowship Scheme.

*Correspondence to:
Enda Byrne Ph.D., Queensland Brain Institute, The University of Queensland, Upland Road, St. Lucia, Brisbane, QLD 4069 Australia.
E-mail: e.byrne3@uq.edu.au

Article first published online in Wiley Online Library
(wileyonlinelibrary.com): 31 May 2013
DOI 10.1002/ajmg.b.32168

self-reported sleep phenotypes after correction for multiple testing.

Key words: insomnia; genetics; mood; sleep; circadian

INTRODUCTION

Despite the fact that insomnia is the most common sleep disorder, little is known about the contribution of genetics to its etiology and pathophysiology. Between 6% and 10% of individuals experience insomnia that is chronic in nature, while another 25% report occasional difficulties with sleep [Ohayon, 2002]. Insomnia is associated with a number of negative sequel including fatigue, irritability and impaired concentration and memory. Longitudinal studies have also repeatedly shown that insomnia is a risk factor for the development of new-onset mood, anxiety, and substance-use disorders [Ford and Kamerow, 1989]. Given the prevalence of insomnia and its associated public health impact, advances in our understanding of the genetic underpinnings of the disorder could lead to prevention and treatment efforts that would benefit a substantial proportion of the population.

One of the difficulties in studying the genetics of insomnia is the lack of standardized phenotypes. Human genetic studies have largely relied on self-report, including one or more questions related to sleep patterns of characteristics such as sleep latency, time spent awake during the night, or total sleep time. A number of studies have demonstrated that part of their variability can be attributed to genetic factors. Several groups have conducted classical twin studies, comparing concordance in MZ and DZ twins [Partinen et al., 1983; Heath et al., 1990; McCarren et al., 1994; Heath et al., 1998]. With only a few exceptions, heritability estimates were consistently in the range of 0.25–0.45, regardless of the exact insomnia question or phenotype used, indicating that self-reported insomnia has moderate genetic influences. Studies where individuals were asked to report on sleep patterns of family members also provide support for genetic influences [Abe and Shimakawa, 1966; Hauri and Olmstead, 1980; Bastien and Morin, 2000; Dauvilliers, 2005; Beaulieu-Bonneau et al., 2007]. The search for specific genes that are associated with sleep patterns and insomnia is in its infancy, but initial studies point in a number of directions. Candidate gene studies in animals and humans have found associations between insomnia phenotypes and circadian clock genes such as BMAL/Mop3 [Laposky, 2005], PER3 [Viola, 2007], and CLOCK [Serretti, 2003]. However, many of these analyses had small sample sizes by comparison to those used in genome-wide association studies. A recent analysis of common variants located in genes known to be involved in the circadian clock revealed an association between TIMELESS and symptoms of depression and sleep disturbance [Utge, 2010].

A number of genome wide association studies (GWAS) have been conducted on sleep phenotypes in humans. Gottlieb et al. [2007] studied a subset (n = 749) of the Framingham heart study offspring cohort using both linkage and association analysis. Their survey included assessments of self-report sleep onset time (SSOT) and sleep duration, phenotypes that might have some relevance for insomnia. Linkage analysis failed to find any peaks with LOD >3, but five peaks with LOD >2 were found, including a linkage between usual bedtime and *CSNK2A2*, a gene known to be a component of the circadian molecular clock. In population-based association tests, an association between an intronic SNP in the *PDE4D* gene and sleepiness reached the genome-wide threshold of significance. Usual bedtime was associated with the SNP rs324981, located in the gene *NPSR1*, which is a component of the neuropeptide S receptor. More recently, a genome-wide association study of identified a genome-wide significant SNP in the *ABCC9* gene as influencing sleep duration [Allebrandt, 2011], while a genome-wide scan of insomnia induced by caffeine failed to identify any genome-wide significant signals, but did replicate a previously reported association with a variant in the *ADORA2A* gene [Byrne, 2012]. These investigations are important in establishing the feasibility of finding genetic associations with self-reported sleep phenotypes.

In order to advance our understanding of the genetics of sleep/wake regulation and insomnia there is a need for gene discovery studies that include a wider range of insomnia phenotypes and that have adequate sample sizes to detect what are likely to be small effects. Here we present the results of a GWAS of sleep and circadian phenotypes in a sample of >2,000 Australian twins. We tested >2,000,000 common genetic polymorphisms for association with sleep latency, sleep time, sleep quality, sleep depth, sleep duration and an insomnia factor score. We also tested the top hits from the sleep latency analysis for replication in an independent sample of four cohorts from the Chronogen Consortium.

METHODS

Participants

Between 1980 and 1982 a Health and Lifestyle Questionnaire was administered by mail to 5,867 complete pairs of twins who had been registered with the Australian Twin Registry. Responses were received from a total of 7,616 individuals (2,746 males and 4,780 females) and they had a mean age of 34.5 years (SD = 14.3). Phenotypic and genotypic data collection was approved by the Queensland Institute of Medical Research (QIMR) Ethics Committee and informed consent was obtained from all participants. A total of 2,323 individuals provided both phenotypic and genotype information for the study (601 males and 1,721 females). The mean age for the genotyped sample was 31.4 years (SD = 11.0). A breakdown of the participants by zygosity is given in Table I.

Phenotypic Measures

As part of the questionnaire, respondents were asked a number of questions about their sleep habits. To assess usual sleep patterns participants were asked the following questions:

"On WEEKDAYS after you go to bed, what time do you usually try to get to sleep?" (Self reported sleep onset time (SSOT))

"On WEEKDAYS, how long in minutes do you think it usually takes you to fall asleep from when you first try to go to sleep?" (sleep latency)

TABLE I. Breakdown of the Sample by Zygosity

MZ female pairs	492
MZ male pairs	102
DZ female pairs	122
DZ male pairs	59
Opposite sex DZ pairs	125
Female siblings of twins	33
Male siblings of twins	5
Female non-twin singletons	335
Male non-twin singletons	150
Total sample size	2,323

"How would you describe the quality of your usual sleep over the last few months?" (Quality)

1 = Very good 2 = Good 3 = Fair 4 = Poor 5 = Very poor

"In particular, how would you describe the depth of your sleep?" (Depth)

1 = Hard to Wake 2 = About average 3 = Easy to wake

"On WEEKDAYS, how long would you usually sleep for?" (sleep duration)

Participants were also asked about how much the quality of their sleep varies and about the frequency with which they wake up in the middle of the night. Further information on the sleep disturbance measures in the questionnaire can be found elsewhere [Heath et al., 1990, 1998; Luciano, 2007]. In total, six variables were analyzed—sleep time, sleep latency, sleep quality, sleep depth, sleep duration, and an insomnia factor score (IFS). A factor analysis applied to an independent dataset that used a similar questionnaire also identified a factor that underscored poor sleep to which sleep latency, waking during the night, and sleep quality loaded strongly [Johns, 1975]. Principal components analysis applied to this dataset previously showed that among the sleep measures assessed in the questionnaire—sleep quality, variability of quality, sleep latency and frequency of night-time waking—appear to load strongly on a general sleep disturbance component that measures general insomnia [Heath et al., 1990]. In the present analysis, principal components analysis was performed on the same variables to derive an overall score for insomnia for each individual. The analysis supported a single factor loading on these variables. Descriptive statistics for each of the traits analyzed are given in Table II. An identical questionnaire (with regard to assessment of sleep habits) was administered to a subsample of 96 individuals who participated in a pilot study several months prior to the main study. This allowed us to test for consistency of responses over time. All of the individual sleep items analyzed in this study showed good reliability ($r^2 > 0.71$, Table II) [Heath et al., 1990]. Previous analysis showed that the variables show strong internal consistency—a further indication of the validity of the subjective reports of sleep disturbance [Heath et al., 1990]. The bivariate correlations between the variables are shown in Supplementary Table I. There were significant age and sex effects, with sex being the most significant by several orders of magnitude. All of the analyses included age and sex as covariates. We also tested for effects of age [Ford and Kamerow, 1989], age × sex, and age × sex [Ford and Kamerow, 1989], but there were no significant effects for any of these polynomial terms. Similarly, we tested for the effect of state of residency to check whether sleep patterns were affected by latitude, but no effects were detected. Due to positive skew in the distribution of sleep latency and the insomnia factor score, both traits were natural log transformed (Table II). In the case of the insomnia factor score, a constant was added to the scores to ensure that all scores were positive prior to transformation. For the SSOT analysis, individuals who said that they usually try to go to sleep between 3 am and 6 pm were removed from the analysis (n = 23), as there was a high likelihood that they were shift workers who were not choosing to sleep at that time by their own preference, or were individuals with delayed sleep phase syndrome. Individuals who reported regular use of sleeping tablets or tranquilizers (n = 15) were removed from the analysis.

As part of the protocol, participants were asked if they had ever had any of a number of conditions. Of potential relevance to sleep phenotypes were the following conditions: high blood pressure, heart attack, stroke, diabetes, thyroid trouble, and cancer. A total of 258 individuals in the genotyped sample reported high-blood pressure, eight reported having diabetes, 95 reported thyroid trouble and 34 reported having had cancer. No participants who were included in this analysis reported having had a heart attack.

Sleep duration has a very low-estimated heritability in our data (9%) but was included in our analyses because it was a phenotype

TABLE II. Descriptive Statistics for Each of the Phenotypes and Test–Retest Correlations

	No of individuals	Heritability (h^2)	Minimum	Maximum	Mean	SD	Reliability
Latency raw scores (minutes)	2,280	0.32	1	180	20.98	22.94	0.72
Latency (natural log)	2,280	0.32	−1.21	5.48	2.67	0.9	—
Quality*	2,315	0.32	1.00	5.00	1.99	0.87	0.78
Depth*	2,314	0.21	1.00	3.00	1.78	0.65	0.82
Sleep time	2,322	0.42	8:00 pm	2:30 am	10:42 pm	52.75 min	0.81
Sleep duration	2,278	0.09	5 hr	12 hr	7.73 hrs	0.91 hrs	0.74
IFS (natural log)	2,267	0.31	−1.91	1.72	0.11	0.67	—

*Higher scores for the quality and depth variables indicate poorer quality and lighter depth of sleep.

included in the only other genome-wide study reported to date [Gottlieb et al., 2007]. To reduce the effect of outliers on the analysis, individuals whose reported sleep duration was less than 5 hr or greater than 12 hr (n = 8) were removed from the analysis. The heritability of the IFS was estimated to be 31%, in line with that found for insomnia phenotypes in other studies. This was also consistent with the estimates of the individual items, indicating there is no increase in heritability when combining information from many sleep phenotypes into a single insomnia phenotype.

Genotyping

Genotype information was collected as part of a number of genotyping projects undertaken by the Genetic Epidemiology group at Queensland Institute of Medical Research. DNA samples were collected in accordance with standard protocols and submitted to different genotype centers using different SNP platforms (Illumina 317K, Illumina Human CNV370-Quadv3, and Illumina Human 610-Quad). Supplementary Figure 1 gives an overview of the number of individuals genotyped on each platform, including those genotyped in more than one genotyping project. SNPs were called using the Illumina BeadStudio software. A standard quality control procedure was used for all of the genotyping projects, prior to imputation. A detailed description of the quality control (QC) steps and procedure for detection of ancestry outliers is given elsewhere [Medland, 2009]. A total of 22 individuals were removed from the present analysis due to being ancestry outliers.

A set of 274,604 SNPs that were common to all of the genotyping chips was used for imputation, which was performed using the program MACH [Li et al., 2009]. The imputation process uses information on the haplotype structure in the human genome from the HapMap project (Release22 Build36) to impute non-genotyped SNPs in the sample. The imputed SNPs were screened further for Mendelian errors, minor allele frequency and missingness. Only SNPs with an imputation quality score (r^2) greater than 0.3 were retained, which resulted in a total number of 2,380,486 SNPs.

Genome-Wide Association Analysis

Association analysis was performed using a score-test in MERLIN [Abecasis et al., 2002; Chen and Abecasis, 2007] with each SNP tested in a single point analysis. This association test is appropriate for use with family data and allows inclusion of MZ and DZ pairs. The test combines information from both a within family test and a between family test to give an overall test of association. The analysis utilized the best-guess genotypes from the imputation analysis. A *P*-value $<5 \times 10^{-8}$ was considered to be genome-wide significant.

Post-GWAS analysis and annotation was carried out using the program WGAViewer [Ge, 2008]. Owing to the highly correlated nature of the results from imputed data because of linkage disequilibrium (LD) between the SNPs, some regions will have many SNPs with similar *P*-values. The clumping algorithm in PLINK [Purcell, 2007] was used to filter results and find the most significant independent signals. SNPs with $r^2 < 0.5$ were considered to be independent signals.

A large Finnish twin study that estimated the heritability of sleep quality found that the most parsimonious model was that the heritability was different for males and females [Paunio, 2009]. We therefore analyzed each trait in males and females separately in addition to the overall. It should be noted the power afforded by combining the results from the separate analyses in males and females affords less power than the overall analysis as the within family information from opposite sex twins is not incorporated into the by sex analyses.

Power

We used a simulation procedure in MERLIN to estimate the power afforded by our sample to detect variants that are associated with traits at the genome-wide significance level. The simulation procedure generates a dataset that has an identical distribution, heritability, marker informativeness, allele frequencies, and missingness patterns and then permits testing for association with an allele that accounts for a specified portion of the phenotypic variance. For the present power analyses, we simulated a SNP with minor allele frequency of 0.25 that explains 1% of the variance for each of three traits—insomnia factor score, sleep latency, and SSOT. These traits have estimated heritabilities of 31%, 32%, and 42% in our sample. We performed 1,000 replicates for each trait and the power was calculated as the proportion of those replicates for which the simulated variant was associated at a genome-wide significance level. We then used the Genetic Power Calculator to calculate the equivalent number of unrelated individuals that would be required to have the same power to detect association with the trait.

Supplementary Table I shows the statistical power of the study for variants explaining different proportions of the phenotypic variance. Our study has >80% power to detect a variant that explains 2% of the phenotypic variance at the genome-wide significance level and approximately 99% power to detect a variant explaining 3%. Further, we estimate that our study had 19.4% power to detect a variant explaining 1% of the phenotypic variation in the insomnia factor score at a genome-wide significant level. We had 16.2% power to detect the same variant in the sleep latency analysis and 19.5% power to detect it in the SSOT analysis. Approximately 1,970 unrelated individuals would be needed to have the same power to detect the same variant. Approximately 4,280 unrelated individuals would be needed to have 80% power to detect a variant explaining 1% of the phenotypic variance at a genome-wide significant level ($P < 5 \times 10^{-8}$).

Gene-Based Tests

To determine whether there were any genes that harbor an excess of SNPs with small *P*-values, a gene-based test of association was performed [Liu et al., 2010]. A SNP was considered to be part of a gene if it was located within 50 kb of the start or stop site of the gene and so could be allocated to more than one gene. The test uses the *P*-values from the single SNP association analysis and computes an overall gene-based test statistic by aggregating the individual SNP effects in each gene, accounting for the number of SNPs in each gene, and the correlation between them because of LD. The value of this test depends on the unknown true genetic architecture of causal variants, which is likely to differ between genes.

Pathway Analysis

To test whether there was an enrichment of associations in genes that act in the same biological pathway or genes that have strongly related functions, all genes with a *P*-value <0.05 from the gene-based test were included in a pathway analysis in the Ingenuity Pathway analysis software (Ingenuity Systems Release 6.0, Ingenuity Systems, Redwood City, CA). The Ingenuity program collates information from published research articles regarding the structure, function, localization, and interactions of genes, proteins, and biochemical molecules and assigns them to functional and canonical pathways. This permits testing for enrichment of a particular pathway that may be relevant to the trait of interest. Fisher's Exact Test was used initially to test whether a particular pathway was over represented and the Benjamini–Hochberg method was used to correct the *P*-values for multiple testing. A corrected $P < 0.05$ was considered to be significant.

Candidate Loci and Genes

We attempted to replicate the association findings of Gottlieb et al. [2007] for sleep duration and self-report sleep onset time. In addition, using the Ingenuity (Ingenuity Systems, Redwood City, CA) software, we identified 86 genes that have been associated with sleep phenotypes in humans or animal models. We then checked whether SNPs within or near these genes showed evidence of association with the sleep phenotypes or if any of these genes ranked highly in the gene-based test of association.

Replication Sample

For replication of the top hit for sleep latency, the results of a meta-analysis of GWAS performed as a collaborative effort by the Chronogen Consortium were used. This comprised a total of 4,270 subjects with European ancestry and included samples from the Erasmus Rucphen family (ERF), Estonian genome center (EGCUT), the co-operative health research in the Augsburg region (KORA), the KORCULA study in Croatia, the micro-isolates in south Tyrol study (MICROS), the Netherlands study of depression and anxiety (NESDA) and the Orkney complex disease study (ORCADES). However, only four of the studies provided information on SNPs in the LD block that was identified in the discovery analysis. A detailed description of these studies is provided in the supplementary methods. All studies in the replication cohort used the Munich Chronotype Questionnaire [Kantermann et al., 2007] to assess sleep traits. Sleep information only on free days, when a person's sleep pattern was not influenced by professional duties (use of alarm clock was an exclusion criterion), was analyzed. Persons that used medications that may influence sleep were excluded from the analyses. Informed consents were obtained from all study participants and an appropriate local committee approved study protocols. Descriptive statistics for the replication cohorts are given in Table III.

Replication cohorts were genotyped on a variety of platforms (Affymetrix 250K, Illumina 317K, Illumina 370K; Perlegen 600K; Affymetrix 1,000K). Imputations of non-genotyped SNPs in the HapMap CEU v21a or v22 were carried out within each study using MACH [Li et al., 2009] or IMPUTE [Howie et al., 2009]. Quality control was done in each group separately. The overall criteria were to exclude individuals with low call rate, excess heterozygosity, and gender mismatch. Based on sample size and study specific characteristics, different criteria were used.

Individual GWAS was performed using linear regression (under additive model), natural log of sleep latency as the dependent variable, SNP allele dosage as predictor and age and sex as covariates. The association analyses were conducted in ProbABEL [Aulchenko and Struchalin, 2010] or SNPTEST [Wellcome Trust Case Control 2007]. All cohorts with information on the top hit from the sleep latency discovery analysis used a linear mixed model in ProbABEL. The software incorporates the FASTA [Chen and Abecasis, 2007] method and kinship matrix estimated from the genotyped SNPs to correct for relatedness [Amin and van Duijn, 2007] (the ERF and MICROS samples included related individuals). This method also accounts for cryptic population stratification.

A fixed effects meta-analysis was conducted using the inverse variance weighted method as implemented in METAL [Willer et al., 2010]. Genomic control correction was also applied to all cohorts prior to the meta-analysis.

RESULTS

The quantile–quantile (QQ) plots of the observed versus expected −log(p) from the six association analyses are presented in Supplementary Figure 2. There was no evidence for population stratification as demonstrated by the genomic control λ (the median χ^2 association statistic divided by the median expected under the null) being between 0.99 and 1.02 for all of the analyses. No SNPs passed the genome-wide significance threshold ($P < 5 \times 10^{-8}$) and there is no evidence for an enrichment of associations at the tail of the distribution. Manhattan plots for the analyses are given in Supplementary Figure 3. Table IV lists the most significant SNPs that represent independent signals for each trait with $P < 10^{-5}$ in the overall analysis and also their results separately by sex.

A gene that has been previously associated with bipolar disorder *CACNA1C* (calcium channel, voltage-dependent, L type, alpha 1C subunit) on chromosome 12 showed evidence for association with sleep latency and with sleep quality. A set of SNPs in perfect LD located in the 3rd intron (rs7316184, rs7304986, rs7301906, rs16929275, rs16929276, rs16929278, and rs2051990) each with minor allele frequency ~0.014 were the most strongly associated SNPs with sleep latency (Table IV, $P = 1.3 \times 10^{-6}$). The SNPs were genome-wide significant when the analysis was performed on the untransformed residuals ($P = 4.9 \times 10^{-10}$), but this did not remain after transforming the distribution to log normal. These SNPs were not in LD with the validated bipolar variants [Ferreira, 2008] subsequently found to be associated with schizophrenia and recurrent major depression [Green, 2009] (rs1006737 and rs10848635 also in intron 3, r^2 with rs7316184 etc = 0.018 and 0.006) and so represent an independent signal. We attempted to replicate the association with the *CACNA1C* SNPs in an independent sample comprising seven cohorts with a total sample size of 4,260 that had collected information on sleep latency. Of these, four cohorts had the rs7304986 variant genotyped or imputed (sample size = 2,001). Three of the cohorts had the same direction of effect

TABLE III. Annotated List of the Most Highly Associated Independent SNPs ($P < 10^{-5}$) Each Association Analysis

SNP	Trait	CHR	Position	Location	Closest gene	Distance to gene (BP)	Allele	FREQ	Beta	SE	P	Gen/Imp	P males	Beta Males (SE)	P Females	Beta Females (S.E)
rs7304986	Latency	12	2438105	Intronic	CACNA1C	0	C	0.01	0.49	0.11	1.38E−06	Imp	4.80E−01	0.13 (0.18)	1.13E−06	0.61 (.13)
rs7716813	Latency	5	146651673	Intronic	STK32A	0	T	0.44	0.14	0.03	4.11E−06	Imp	2.84E−03	0.15 (0.05)	8.53E−05	0.13 (0.03)
rs949441	Latency	22	44744884	Intergenic	N/A	N/A	A	0.2	−0.17	0.04	4.66E−06	Gen	2.47E−04	0.04 (0.09)	1.18E−02	−0.15 (0.06)
rs9922235	Latency	16	79096417	Intronic	WWOX	0	G	0.41	−0.13	0.03	5.80E−06	Imp	1.91E−03	0.16 (0.05)	7.33E−04	0.11 (0.03)
rs951896	Latency	10	106428784	Intronic	SORCS3	0	C	0.1	−0.21	0.05	6.26E−06	Imp	1.16E−02	−0.21 (0.08)	1.10E−04	−0.21 (0.05)
rs292410	Latency	5	163043385	Intergenic	LOC391844	84,256	C	0.45	−0.12	0.03	7.37E−06	Gen	3.13E−04	−0.17 (0.05)	2.79E−03	−0.10 (0.03)
rs722258	Sleep time	10	23369421	Intergenic	N/A	N/A	T	0.42	−8.03	1.71	2.85E−06	Gen	6.69E−03	−8.97 (3.31)	3.03E−04	−7.13 (1.97)
rs1478693	Sleep time	16	7164219	Intergenic	N/A	N/A	G	0.27	8.82	1.9	3.61E−06	Imp	1.50E−02	8.525 (3.50)	5.72E−05	8.99 (2.23)
rs1539808	Sleep time	18	5978931	Intronic	L3MBTL4	0	T	0.01	44.44	9.64	4.04E−06	Imp	2.75E−03	51.97 (17.34)	4.95E−04	39.81 (11.43)
rs9804200	Sleep time	10	131673986	Intronic	EBF3	0	C	0.3	8.39	1.86	6.26E−06	Imp	7.03E−03	9.36 (3.47)	3.96E−04	7.65 (2.16)
rs13068101	Sleep time	3	192772295	Intergenic	N/A	N/A	A	0.26	−8.5	1.89	7.13E−06	Gen	2.49E−01	−4.14 (3.59)	4.38E−06	−10.09 (2.19)
rs10734107	Sleep time	10	129353132	Downstream	NPS	2,197	T	0.48	7.4	1.69	1.13E−05	Imp	1.37E−02	8.10 (3.28)	1.86E−04	7.23 (1.93)
rs1986116	Quality	14	77513844	WNCG[a]	AC007686.1	0	T	0.24	0.16	0.03	9.81E−07	Gen	1.60E−01	0.08 (0.06)	2.41E−06	0.18 (0.04)
rs1005956	Quality	12	40341085	Intronic	SLC2A13	0	C	0.41	−0.13	0.03	2.61E−06	Gen	7.00E−02	−0.09 (0.05)	1.74E−05	−0.14 (0.03)
rs17071124	Quality	6	142033726	Intergenic	N/A	N/A	G	0.01	0.75	0.16	2.78E−06	Imp	7.44E−03	0.93 (0.35)	1.15E−04	0.69 (0.18)
rs2302729	Quality	12	2783972	Intronic	CACNA1C	0	T	0.17	0.17	0.04	4.37E−06	Gen	2.30E−01	0.08 (0.06)	4.63E−06	0.2 (0.04)
rs9836672	Quality	3	191912870	Intronic	FGF12	0	T	0.05	0.27	0.06	4.56E−06	Gen	3.26E−01	0.11 (0.11)	2.46E−06	0.33 (0.07)
rs2210430	Quality	9	25501297	Intergenic	TUSC1	175,099	T	0.01	0.65	0.15	8.74E−06	Imp	2.00E−02	0.89 (0.40)	1.77E−04	0.60 (0.16)
rs1949200	Quality	4	67301042	Downstream	RP11-793B9.1	3803	T	0.13	0.18	0.04	9.47E−06	Imp	3.30E−02	0.17 (0.08)	1.24E−04	0.18 (0.05)
rs3110232	Quality	5	38381015	Intronic	EGFLAM	0	G	0.09	0.21	0.05	9.56E−06	Imp	1.21E−03	0.28 (0.09)	8.92E−04	0.18 (0.05)
rs9830368	Quality	3	64396264	Intergenic	N/A	N/A	A	0.2	0.15	0.04	9.66E−06	Gen	3.00E−02	0.14 (0.06)	1.60E−04	0.16 (0.04)
rs679711	Depth	3	109529550	WNCG*	N/A	0	T	0.43	−0.1	0.02	1.18E−06	Imp	1.18E−02	−0.10 (0.04)	8.24E−05	−0.09 (0.02)
rs12913538	Depth	15	62899160	Intergenic	N/A	0	A	0.37	0.1	0.02	6.77E−06	Imp	9.45E−03	0.11 (0.04)	3.75E−04	0.08 (0.02)
rs4780805	Duration	16	19404645	Intergenic	TMC5	−17,216	A	0.15	−0.19	0.04	7.99E−07	Imp	1.42E−01	0.12 (0.08)	1.03E−04	−0.20 (0.05)
rs17737465	Duration	8	98154441	Intronic	PGCP	0	G	0.37	−0.14	0.03	1.50E−06	Imp	7.83E−02	−0.11 (0.06)	4.47E−05	−0.16 (0.04)
rs11640439	Duration	16	57303194	Intronic	PLLP	0	A	0.09	0.24	0.05	2.93E−06	Imp	9.51E−02	0.19 (0.11)	1.21E−05	0.27 (0.06)
rs2042126	Duration	3	175055759	Intronic	NAALADL2	0	G	0.48	−0.14	0.03	2.76E−06	Imp	6.28E−01	−0.03 (0.06)	1.98E−05	−0.16 (0.03)
rs11987678	Duration	8	81228826	Intronic	RP11-941H19.3	0	C	0.05	−0.32	0.07	4.22E−06	Gen	1.40E−01	−0.23 (0.16)	1.28E−03	−0.28 (0.08)
rs10823607	Duration	10	72500763	Coding	ADAMTS14	0	T	0.15	−0.19	0.04	4.57E−06	Imp	5.90E−01	−0.05 (0.09)	2.58E−06	−0.24 (0.05)
rs2278331	Duration	18	71581506	Intergenic	FBXO15	159,082	A		0.14	0.03	4.72E−06	Imp	7.83E−02	0.11 (0.06)	3.19E−05	0.17 (0.04)
rs11174478	I.F.S	12	40354244	Intronic	SLC2A13	0	A	0.4	−0.1	0.02	1.92E−06	Imp	1.10E−01	−0.03 (0.44)	5.42E−06	−0.12 (0.02)
rs2725544	I.F.S	15	48992897	Intergenic	CEP152	26,298	C	0.23	−0.12	0.03	2.28E−06	Gen	6.00E−01	−0.02 (0.04)	−4.00E−02	−0.15 (0.03)
rs12471454	I.F.S	2	200014483	Intergenic	SATB2	−119,740	T	0.41	−0.1	0.02	6.21E−06	Imp	7.30E−02	−0.06 (0.03)	5.90E−05	−0.11 (0.03)

[a]WNCG, within non-coding gene.
Effect sizes for sleep duration are given in decimal hours for sleep time the units are minutes.

TABLE IV. Descriptive Statistics for Cohorts in Chronogen Consortium

Population	Sample	Average sleep latency (SD)	Average [LN (sleep latency + c)] (SD)	Average age (SD)
EGP	Total	15.47 (17.76)	2.31 (1.02)	39.85 (16.08)
	Male	14.38 (15.74)	2.25 (1.01)	38.38 (16.24)
	Female	16.54 (19.70)	2.37 (1.04)	41.03 (15.88)
ERF	Total	17.69 (18.63)	2.56 (0.85)	45.67 (13.00)
	Male	15.53 (14.99)	2.48 (0.80)	47.42 (12.98)
	Female	19.59 (21.16)	2.64 (0.88)	44.12 (12.83)
KORA	Total	10.33 (9.72)	2.08 (0.79)	54.29 (5.49)
	Male	8.71 (7.75)	1.95 (0.77)	54.64 (5.66)
	Female	11.97 (11.16)	2.21 (0.81)	53.96 (5.32)
KORCULA	Total	19.55 (12.15)	2.56 (1.02)	56.41 (12.2)
	Male	17.58 (13.86)	2.52 (0.94)	57.61 (12.97)
	Female	20.69 (18.71)	2.59 (1.07)	55.72 (11.7)
MICROS	Total	13.31 (14.29)	2.22 (0.94)	40.26 (14.52)
	Male	11.94 (12.06)	2.16 (0.89)	41.24 (14.50)
	Female	14.50 (15.88)	2.28 (0.98)	39.44 (14.53)
NESDA	Total	18.52 (18.52)	2.60 (0.90)	41.25 (12.26)
	Male	16.63 (16.95)	2.52 (0.85)	44.15 (12.40)
	Female	19.49 (19.22)	2.64 (0.93)	39.77 (11.94)
ORCADES	Total	15.95 (15.61)	2.36 (1.00)	51.08 (11.08)
	Male	12.67 (11.26)	2.22 (0.89)	51.08 (13.18)
	Female	18.67 (18.2)	2.47 (1.08)	51.26 (13.85)
QIMR	Total	20.98 (22.94)	2.67 (0.90)	31.28 (10.88)
	Male	17.67 (16.26)	2.56 (0.86)	28.17 (8.12)
	Female	22.14 (24.78)	2.71 (0.92)	32.37 (11.51)

as found in the initial GWAS, with the minor allele found to increase sleep latency. The association was not nominally significant however ($P = 0.73$) (Table V). We also performed a meta-analysis of the Australian results and the results from the four cohorts in Chronogen in which results for rs7304986 were available (Table V). The *P*-value for the meta-analysis was 0.01 ($\beta = 0.12$, SE = 0.05).

For the SSOT analysis, there were two SNPs located in or near genes with $P < 10^{-5}$. They are intronic SNPs in the *L3MBTL4* and *EBF3* genes, respectively (Table IV). *L3MBTL4* is a gene on chromosome 18 whose function is not well annotated. *EBF3* is located on chromosome 10 and is known to be expressed in the brain. It is frequently found to be silenced in brain tumors and other forms of cancer and is thought to be a tumor suppressor gene [Zhao, 2006]. No circadian candidate genes harbor SNPs that show strong evidence of association with timing of sleep from our analysis. One SNP–rs10734107–located 2 kb downstream of the *NPS* gene had a *P*-value of 1.1×10^{-5}. *NPS* is an interesting candidate gene for association with sleep timing as it encodes a Neuropeptide S, a molecule that is known to stimulate arousal and that has been associated with anxiety and sleep apnea [Reinscheid and Xu, 2005]. In a previous GWAS of SSOT [Gottlieb et al., 2007], a SNP in the Neuropeptide S Receptor gene (rs324981) was among the most associated variants ($P = 4.5 \times 10^{-5}$). That result did not replicate in our study ($P = 0.133$), but the combined findings of the two genome-wide studies implicate a role for the Neuropeptide S system in sleep/wake regulation.

The *CACNA1C* gene also shows evidence of association with sleep quality (Table IV). The most significant SNP from this region

TABLE V. Results From Replication Analysis of rs7304986 With Sleep Latency

Cohort	SNP	A1	Freq AL1	Imputation quality	Sample size	Beta_SNP_add	Sebeta_SNP_add	*P*
ERF	rs7304986	C	0.023	0.99	746	−0.01	0.07	0.86
KORA	rs7304986	C	0.013	0.99	510	0.07	0.21	0.11
Orkney	rs7304986	C	0.019	0.99	205	0.42	0.34	0.22
NESDA	rs7304986	C	0.029	0.97	540	0.09	0.16	0.57

	SNP	Allele 1	N	Effect size	Standard error	*P*-value	Direction of effect by cohort
Meta-analysis replication sample	rs7304986	C	2001	0.02	0.06	0.73	−+++

—rs2302729 ($P = 4.4 \times 10^{-6}$)—is located in intron 9 of the gene and is not in LD with the variants associated with sleep latency ($r^2 = 0.004$) or with the SNPs associated with bipolar disorder ($r^2 = 0.009$ and 0.034, respectively).

For the sleep depth analysis, there were only two independent regions associated with $P < 10^{-5}$, neither of which were located within or near annotated genes. The most significantly associated SNP with sleep duration was rs4780805 ($P = 2.66 \times 10^{-6}$). This SNP is located on chromosome 16, 17 kb from the nearest gene *TMC5*.

A SNP–rs11174478 ($P = 1.92 \times 10^{-6}$) in the *SLC2A13* gene is the most strongly associated with the insomnia factor score. This gene is located in the same region of the genome as LRRK2, a gene known to be associated with Parkinson's disease.

No SNPs reached genome-wide significance when males and females were analyzed separately. The results from those analyses are available upon request.

Gene-Based Tests and Pathway Analysis

At least one SNP mapped to 17,695 autosomal genes. A conservative genome-wide threshold for significance was set at 2.83×10^{-6}, which corresponds to a nominally significant *P*-value of 0.05 corrected for 17,695 tests. This threshold does not correct for analyzing multiple (albeit correlated) traits. No genes reached this significance threshold. A list of the five most significant genes for each trait is given in Table VI. The most significant association across the six traits was *ZNF695* with sleep duration ($P = 1.14 \times 10^{-4}$). None of the most strongly associated genes on the list have a known role in circadian rhythms or have previously been identified as candidate genes for sleep phenotypes.

After correction for multiple testing, no biological functions or pathways were found to be enriched in the gene-based test. Supplementary Table III gives the most significant functions and pathways for each of the gene-based analyses.

Candidate Genes

From the Ingenuity Pathway Analysis software, we identified 86 genes that have been associated with circadian rhythms and sleep phenotypes in humans or animal models. In addition, we examined the association statistics for a further nine genes identified in the study of Gottlieb et al. [2007]. The most strongly associated SNP and results from the gene-based test for each of the candidate genes

TABLE VI. The Five Most Significant Genes From the Gene-Based Test for Each of the Traits Analyzed

Gene	Trait	Chr	*P*-value	nSNPs	Start	Stop	Best-SNP	SNP-*P* value
RSPRY1	Duration	16	1.30E−05	69	55777741	55830448	rs11640439	1.50E-06
NIP30	Duration	16	2.00E−05	84	55743878	55777477	rs767505	5.77E-05
CPNE2	Duration	16	4.40E−05	127	55684010	55739377	rs767505	5.77E-05
ARL2BP	Duration	16	9.20E−05	36	55836538	55845046	rs11640439	1.50E-06
GPR68	Duration	14	1.74E−04	76	90768628	90789977	rs2540871	1.93E-06
SIP1	Latency	14	3.12E−04	82	38653238	38675928	rs8011494	6.23E-05
LOC284009	Latency	17	3.48E−04	61	2257024	2265480	rs898751	8.86E-04
TRAPPC6B	Latency	14	4.97E−04	73	38686765	38709385	rs8011494	6.23E-05
SEC23A	Latency	14	5.23E−04	121	38570873	38642188	rs8011494	6.23E-05
C13orf39	Latency	13	6.06E−04	151	102136097	102144855	rs679331	3.75E-04
NGRN	Sleep time	15	4.53E−04	87	88609898	88616447	rs1044813	5.29E-05
TTLL13	Sleep time	15	6.41E−04	84	88593767	88603316	rs1044813	5.29E-05
FLT3	Sleep time	13	6.83E−04	136	27475410	27572729	rs9554235	2.01E-04
CIB1	Sleep time	15	6.84E−04	80	88574480	88578283	rs1044813	5.29E-05
C15orf58	Sleep time	15	7.66E−04	82	88578490	88586316	rs1044813	5.29E-05
ZNF695	Quality	1	1.14E−04	82	245215248	245237978	rs10802457	8.56E-05
SLC2A13	Quality	12	1.22E−04	568	38435089	38785928	rs1005956	2.61E-06
TM4SF20	Quality	2	3.52E−04	107	227935117	227952266	rs11693555	1.28E-03
SLC39A2	Quality	14	4.30E−04	111	20537258	20539870	rs1889774	6.71E-05
METT11D1	Quality	14	4.39E−04	105	20527804	20535034	rs1889774	6.71E-05
CLNS1A	Depth	11	1.29E−04	50	77004846	77026495	rs17135809	9.74E-05
RSF1	Depth	11	1.32E−04	115	77054921	77209528	rs1544274	9.74E-05
PNMA2	Depth	8	1.71E−04	174	26418112	26427400	rs1372882	3.14E-05
LRRN2	Depth	1	4.26E−04	142	202852925	202921220	rs2772232	1.50E-04
LYPLA3	Depth	16	5.12E−04	38	66836747	66852462	rs6499163	7.15E-04
LRRK2	IFS	12	3.64E−04	583	38905079	39049353	rs11564146	1.33E-04
IER5L	IFS	9	5.07E−04	122	130977651	130980361	rs1107329	5.71E-04
CSDA	IFS	12	5.89E−04	114	10742954	10767171	rs797168	1.43E-04
CRAT	IFS	9	6.18E−04	110	130896893	130912904	rs12346996	4.66E-04
DOLPP1	IFS	9	6.53E−04	82	130883226	130892538	rs12346996	4.66E-04

are given in the Supplementary File. With the exception of *NPS* in the sleep latency analysis, none of the candidate genes ranked among the most associated genes for any of the traits. Strikingly, the *NPS* ranked top of the candidate genes for SSOT ($P = 0.001$) and fourth in the latency analysis ($P = 0.03$), indicating that variants within the gene may influence several different sleep phenotypes. This result is not surprising given that the principal components analysis showed that the variables load on one common factor for insomnia. The number of genes with $P < 0.05$ ranged from zero for sleep duration to seven for sleep latency and there was no overall evidence for an enrichment of associations in the candidate genes. A list of SNPs located either in the gene or within 50 kb of the start of stop site with $P < 10^{-3}$ for any of the association analyses are listed in Supplementary Table IV.

Gottlieb et al. identified 34 SNPs that showed evidence of association—in either a population-based or family-based test—or linkage with sleep duration, SSOT or sleepiness. We attempted to replicate those SNPs in our sample initially with the phenotypes for which associations had been reported, and then with the other phenotypes in our study. No measure of sleepiness was available in this study and so it was not possible to try to replicate the top SNPs for that phenotype. Only one SNP replicated with the same phenotype–rs2985334 with SSOT ($P = 0.0062$ $\beta = 5.3$ min), survived multiple testing. Several of the other SNPs replicated with other phenotypes in the sample, but none of these results were significant after accounting for multiple testing. A list of SNPs from Gottlieb et al. with $P < 0.05$ for association with any of the phenotypes is given in Supplementary Table IV.

DISCUSSION

A GWAS of six insomnia-related traits in a sample of over 2,000 Australian twins and their siblings (with power equivalent to 1,970 unrelated individuals) was performed. One previous GWAS of sleep and circadian phenotypes has been reported, but the analysis was limited to ~71,000 SNPs with minor allele frequency >0.1. The present study used >2,000,000 SNPs in the analysis and so surveys a larger fraction of the common variation in the human genome and has a larger sample size. No SNPs reached the genome-wide significance level for any of the traits. The Q–Q plots show that the distribution of the association test statistics closely follows the expected distribution under the null hypothesis of no association. This is not an unexpected finding given the sample size of the study and the effect sizes of variants detected in genome-wide association studies of other complex trait [Visscher and Montgomery, 2009; Lango Allen, 2010; Speliotes, 2010]. This contrasts with the Q–Q plot for a similar study of hair morphology that used the same sample used in the present analysis and found a genome-wide significant hit [Medland, 2009].

The top ranked region for the sleep latency analysis was the 3rd intron of the *CACNA1C* gene. Variants in this intron have previously been found to be associated with a number of psychiatric disorders including bipolar disorder and schizophrenia [Ferreira, 2008; Green, 2009]. While the SNPs identified here are not in LD with the risk alleles and therefore represent an independent signal, there is widespread evidence to suggest a link between sleep disturbance and mood disorders and several studies have reported associations between circadian genes and mood disorders [Kripke et al., 2009; Soria, 2013; Utge, 2013]. It is therefore plausible that variants in genes known to increase risk to mood disorders may also play a role in sleep disturbance. The association did not replicate in an independent sample however ($P = 0.71$, meta-$P = 0.01$), indicating that it may simply be a chance occurrence that these SNPs are associated in our sample.

The lack of replication may be caused by several factors. The replication sample size was 2,001 individuals, which includes some related samples. Under the assumption that the true causal variant at the *CACNA1C* locus has been detected and the effect size has been estimated without error, the estimated proportion of variance in sleep latency explained is 0.5% [Purcell et ai. 2003]. From the Genetic Power Calculator it can be calculated that the replication sample had 89.12% power to detect the same effect with $P < 0.05$. However, because of the "winner's curse effect," the true effect size may have been overestimated in the discovery sample and hence the power to replicate the finding may in fact be less than estimated. In addition, there were some differences between the discovery and replication samples that could have affected the results. Firstly, the questionnaires used to collect latency information were different, with the discovery sample asking about sleep latency on weekdays while the replication sample asked about free days which may have led to slight differences in the phenotypes. Secondly, differences in inclusion criteria between the discovery and replication cohorts and between the individual replication cohorts may have decrease the power to replicate the finding. The NESDA sample removed individuals from the analysis who had major depressive disorder, whereas the Australian questionnaire did not include a diagnostic interview for mood disorders and so could not remove individuals with depression from the analysis. Moreover, the mean age of the discovery cohort (31.28 years) was younger than all of the replication cohorts (Supplementary Table II), which may have affected the power to replicate. This heterogeneity between cohorts is likely to be an issue in many genetic association studies of sleep and insomnia (not just those that rely on self-report information) and so very large sample sizes may be required to have enough power to have enough power to find variants of small effect.

In spite of the lack of replication, *CACNA1C* represents an interesting candidate gene for sleep phenotypes, not only because of its known association with bipolar disorder. An association study of narcolepsy in a Japanese population implicated another SNP in the 3rd intron of *CACNA1C* (rs10774044, $P = 4.2 \times 10^{-4}$) [Shimada, 2013]. The SNP identified in the narcolepsy study is not in LD in the European population with the SNPs identified in the present study ($r^2 = 0.001$), but it is nominally associated with sleep latency in our sample ($P = 0.035, \beta = 0.054, \text{MAF} = 0.048$). An independent region of *CACNA1C* was also suggestively associated with sleep quality in our sample and it is known that hypocretin 1, a neuropeptide that promotes wakefulness, activates the L-type voltage-dependent calcium channels among other signaling pathways in the brain [Selbach and Haas, 2006]. There is therefore evidence from a number of sources implicating *CACNA1C* in sleep/wake regulation, and despite the lack of replication for the SNPs identified here, further studies of the role of this gene in regulating sleep are warranted.

Studies in rats have shown that increased concentrations of neuropeptide S can activate the hypocretin-1 system, and this may explain the effect of neuropeptide S on arousal. NPS is a strong candidate gene for circadian phenotypes due to its established effects on wakefulness. Mice exposed to even small amounts of NPS show increased locomotion and NPS has been shown to decrease paradoxical and slow wave sleep in rats [Reinscheid and Xu, 2005]. While the result did not replicate in our study, the finding of a significant SNP located in the gene encoding the receptor for NPS in a previous study also strongly implicates the biological pathway in which NPS acts in controlling timing of sleep in humans.

The single SNP analysis and the gene-based test both implicate a region on chromosome 12 near the *SLC2A13* and *LRRK2* genes as the most strongly associated with the insomnia factor score. This region has previously been identified as being associated in Parkinson's disease [Satake, 2009]. One of the most common features of Parkinson's is sleep disruption however, none of the genome-wide significant SNPs from the Parkinson's GWAS were nominally significant in our sample.

Our study had several limitations that need to be borne in mind when interpreting the results and that may have affected the power of our study. We estimated the power to detect a variant that explains 2% of the phenotypic variance in any of the traits as ~80%, while the power to detect a variant explaining 1% of the variance was ~18%. However, this may be an overestimate for several reasons.

Firstly, we were unable to systematically examine whether participants undertook shift work. We did remove those individuals who said that their usual bedtime was after 2 am, and who may therefore have worked unusual hours that could have contributed to sleep difficulties. However, there may have been others in the sample that did usually go to sleep in the evening, but who had to do shift work that caused significant circadian disruption.

Secondly, as mentioned above, we were also unable to establish whether any of the participants were suffering from a psychiatric disorder, as questions relating to symptoms were not included. While we did ask about conditions that may have affected sleep such as serious heart disease and stroke, we did not ask about a range of other conditions that may have affected responses to the sleep questionnaire.

Thirdly, even in the absence of any comorbidities, self-report sleep phenotypes are subject to cognitive and perceptual biases that may reduce power to detect associations with genetic variants. Individuals who experience sleep disturbance may be prone to underestimate their usual sleep duration. The gold standard for the assessment of sleep is polysomnography, with actigraphy offering another objective measure that also can be used to measure activity rhythms in humans. These methods provide more objective measures of sleep and circadian phenotypes which may be more amenable to genetic analysis. As an example, certain polysomnographic components have been shown to be >90% heritable [De Gennaro, 2008]. The disadvantage of these methods is that they are expensive and time-consuming and so large genetically informative samples measured for these phenotypes will be difficult to obtain. Studies comparing self-reported sleep information to objective measures have shown a strong correlation between them and those reporting poorer sleep tend to have increased time to fall asleep, less total sleep duration and increased night waking. The self-report items used in the present study showed good test–retest correlations indicating their stability over time and the items were internally consistent which provides another check of the validity of the self-report items. Moreover, the measures used in this study have been validated against laboratory-based EEG measures of sleep [Lewis, 1969]. However, EEG-based measures of sleep remain the most desirable phenotypes for genetic analysis.

In this study, we focused primarily on analyzing both sexes together in order to maximize statistical power. Some studies have identified sex-specific genetic effects, including an increased heritability of sleep quality in females [Paunio, 2009]. For the majority of the associations identified in the overall analysis, there was a nominally significant association in both sexes. Males comprised a smaller percentage of the sample and therefore there was less power to detect associations when analyzing males alone. However, several of the associations showed more evidence of association in one sex when compared to the other. Future replication efforts might be successful by trying to replicate the results in males or females only.

Confirmation of the association in an independent sample is required before an association can be considered "real" rather than simply a chance event. All our associations are at a level expected under the null hypothesis given the extent of multiple testing. However, this study has identified a number of suggestive associations that can be prioritized for replication in other samples. We attempted to replicate the top SNP for sleep latency in an independent consortium of cohorts. However, attempted replication in an even larger sample would be desirable, while the top hits for the other five traits will also be necessary.

The somewhat mixed results from candidate gene studies for sleep/wake regulation also highlight the need for replication in association studies. This study also permitted us to attempt replication of candidate genes and polymorphisms identified in candidate gene association studies for sleep, but none of them were replicated in our study.

In spite of the limitations of the study, we have identified a number of common variants that are suggestively associated with variation in sleep habits in the population, some of which are located in or near candidate genes. These variants should be targeted for replication in other samples. It is likely that larger sample sizes (likely on the order of tens of thousands of individuals) will be required to identify common variants that influence self-report sleep habits in the population. However, any identified variants, genes or biological pathways may have a dramatic impact on our understanding of sleep/wake regulation and will have implications for general medicine, given the link between disturbed sleep and cardiovascular disease [Wolk et al., 2005], psychiatric illness [Gregory et al., 2009], life satisfaction and well-being [Paunio, 2009]. Our results will be useful for replication efforts in independent samples and for future meta-analyses.

ACKNOWLEDGMENTS

We thank the twins and their families for their participation. We also thank Dixie Statham, Ann Eldridge, Marlene Grace, Kerrie McAloney, Lisa Bowdler, Steven Crooks, Peter Visscher and Allan McRae. A portion of the genotyping on which this study was based

(Illumina 370K scans) was carried out at the Center for Inherited Disease Research, Baltimore (CIDR), through an access award to our late colleague Dr. Richard Todd (Psychiatry, Washington University School of Medicine, St Louis, WA). G.W.M, D.R.N and S.E.M. are supp orted by the National Health and Medical Research Council (NHMRC) Fellowship Scheme.

REFERENCES

Abe K, Shimakawa M. 1966. Genetic-constitutional factor and childhood insomnia. Psychiatr Neurol (Basel) 152:363–369.

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. 2002. Merlin—Rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101.

Allebrandt KV, Amin N, Müller-Myhsok B, Esko T, Teder-Laving M, Azevedo RV, Hayward C, van Mill J, Vogelzangs N, Green EW, Melville SA, Lichtner P, Wichmann HE, Oostra BA, Janssens AC, Campbell H, Wilson JF, Hicks AA, Pramstaller PP, Dogas Z, Rudan I, Merrow M, Penninx B, Kyriacou CP, Metspalu A, van Duijn CM, Meitinger T, Roenneberg T. 2013. A K(ATP) channel gene effect on sleep duration: From genome-wide association studies to function in Drosophila. Mol Psychiatry 18:122–132.

Amin N, van Duijn CM, Aulchenko YS. 2007. A genomic background based method for association analysis in related individuals. PLoS ONE 2: e1274.

Aulchenko YS, Struchalin, MV, van Duijn, CM. 2010. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinform 11:134.

Bastien CH, Morin CM. 2000. Familial incidence of insomnia. J Sleep Res 9:49–54.

Beaulieu-Bonneau S, LeBlanc M, Merette C, Dauvilliers Y, Morin CM. 2007. Family history of insomnia in a population-based sample. Sleep 30:1739–1745.

Byrne EM, Johnson J, McRae AF, Nyholt DR, Medland SE, Gehrman PR, Heath AC, Madden PA, Montgomery GW, Chenevix-Trench G, Martin NG. 2012. A genome-wide association study of caffeine-related sleep disturbance: Confirmation of a role for a common variant in the adenosine receptor. Sleep 35:967–975.

Chen WM, Abecasis GR. 2007. Family-based association tests for genome-wide association scans. Am J Hum Genet 81:913–926.

Dauvilliers Y, Morin C, Cervena K, Carlander B, Touchon J, Besset A, Billiard M. 2005. Family studies in insomnia. J Psychosom Res 58: 271–278.

De Gennaro L, Marzano C, Fratello F, Moroni F, Pellicciari MC, Ferlazzo F, Costa S, Couyoumdjian A, Curcio G, Sforza E, Malafosse A, Finelli LA, Pasqualetti P, Ferrara M, Bertini M, Rossini PM. The electroencephalographic fingerprint of sleep is genetically determined: A twin study. Ann Neurol 64:455–460. 2008.

Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, Fan J, Kirov G, Perlis RH, Green EK, Smoller JW, Grozeva D, Stone J, Nikolov I, Chambert K, Hamshere ML, Nimgaonkar VL, Moskvina V, Thase ME, Caesar S, Sachs GS, Franklin J, Gordon-Smith K, Ardlie KG, Gabriel SB, Fraser C, Blumenstiel B, Defelice M, Breen G, Gill M, Morris DW, Elkin A, Muir WJ, McGhee KA, Williamson R, MacIntyre DJ, MacLean AW, St CD, Robinson M, Van Beck M, Pereira AC, Kandaswamy R, McQuillin A, Collier DA, Bass NJ, Young AH, Lawrence J, Ferrier IN, Anjorin A, Farmer A, Curtis D, Scolnick EM, McGuffin P, Daly MJ, Corvin AP, Holmans PA, Blackwood DH, Gurling HM, Owen MJ, Purcell SM, Sklar P, Craddock N; Wellcome Trust Case Control Consortium. 2008. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet 40:1056–1058.

Ford DE, Kamerow DB. 1989. Epidemiologic study of sleep disturbances and psychiatric disorders. An opportunity for prevention? JAMA 262:1479–1484.

Ge D, Zhang K, Need AC, Martin O, Fellay J, Urban TJ, Telenti A, Goldstein DB. 2008. WGAViewer: Software for genomic annotation of whole genome association studies. Genome Res 18:640–643.

Gottlieb DJ, O'Connor GT, Wilk JB. 2007. Genome-wide association of sleep and circadian phenotypes. BMC Med Genet 8(Suppl 1):S9.

Green EK, Grozeva D, Jones I, Jones L, Kirov G, Caesar S, Gordon-Smith K, Fraser C, Forty L, Russell E, Hamshere ML, Moskvina V, Nikolov I, Farmer A, McGuffin P; Wellcome Trust Case Control Consortium, Holmans PA, Owen MJ, O'Donovan MC, Craddock N. 2009. The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. Mol Psychiatry.

Gregory AM, Rijsdijk FV, Lau JY, Dahl RE, Eley TC. 2009. The direction of longitudinal associations between sleep problems and depression symptoms: A study of twins aged 8 and 10 years. Sleep 32:189–199.

Hauri P, Olmstead E. 1980. Childhood-onset insomnia. Sleep 3:59–65.

Heath AC, Kendler KS, Eaves LJ, Martin NG. 1990. Evidence for genetic influences on sleep disturbance and sleep pattern in twins. Sleep 13: 318–335.

Heath AC, Eaves LJ, Kirk KM, Martin NG. 1998. Effects of lifestyle, personality, symptoms of anxiety and depression, and genetic predisposition on subjective sleep disturbance and sleep pattern. Twin Res 1: 176–188.

Howie BN, Donnelly P, Marchini J. 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5:e1000529.

Johns MW. 1975. Factor analysis of subjectively reported sleep habits, and the nature of insomnia. Psychol Med 5:83–88.

Kantermann T, Juda M, Merrow M, Roenneberg T. 2007. The human circadian clock's seasonal adjustment is disrupted by daylight saving time. Curr Biol 17:1996–2000.

Kripke DF, Nievergelt CM, Joo E, Shekhtman T, Kelsoe JR. 2009. Circadian polymorphisms associated with affective disorders. J Circadian Rhythms 7:2.

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, Segrè AV, Speliotes EK, Wheeler E, Soranzo N, Park JH, Yang J, Gudbjartsson D, Heard-Costa NL, Randall JC, Qi L, Vernon Smith A, Mägi R, Pastinen T, Liang L, Heid IM, Luan J, Thorleifsson G, Winkler TW, Goddard ME, Sin Lo K, Palmer C, Workalemahu T, Aulchenko YS, Johansson A, Zillikens MC, Feitosa MF, Esko T, Johnson T, Ketkar S, Kraft P, Mangino M, Prokopenko I, Absher D, Albrecht E, Ernst F, Glazer NL, Hayward C, Hottenga JJ, Jacobs KB, Knowles JW, Kutalik Z, Monda KL, Polasek O, Preuss M, Rayner NW, Robertson NR, Steinthorsdottir V, Tyrer JP, Voight BF, Wiklund F, Xu J, Zhao JH, Nyholt DR, Pellikka N, Perola M, Perry JR, Surakka I, Tammesoo ML, Altmaier EL, Amin N, Aspelund T, Bhangale T, Boucher G, Chasman DI, Chen C, Coin L, Cooper MN, Dixon AL, Gibson Q, Grundberg E, Hao K, Juhani Junttila M, Kaplan LM, Kettunen J, König IR, Kwan T, Lawrence RW, Levinson DF, Lorentzon M, McKnight B, Morris AP, Müller M, Suh Ngwa J, Purcell S, Rafelt S, Salem RM, Salvi E, Sanna S, Shi J, Sovio U, Thompson JR, Turchin MC, Vandenput L, Verlaan DJ, Vitart V, White CC, Ziegler A, Almgren P, Balmforth AJ, Campbell H, Citterio L, De Grandi A, Dominiczak A, Duan J, Elliott P, Elosua R, Eriksson JG, Freimer NB, Geus EJ, Glorioso N, Haiqing S, Hartikainen AL, Havulinna AS, Hicks AA, Hui J, Igl W, Illig T, Jula A, Kajantie E, Kilpeläinen TO, Koiranen M, Kolcic I, Koskinen S, Kovacs P, Laitinen J, Liu J, Lokki ML,

Marusic A, Maschio A, Meitinger T, Mulas A, Paré G, Parker AN, Peden JF, Petersmann A, Pichler I, Pietiläinen KH, Pouta A, Ridderstråle M, Rotter JI, Sambrook JG, Sanders AR, Schmidt CO, Sinisalo J, Smit JH, Stringham HM, Bragi Walters G, Widen E, Wild SH, Willemsen G, Zagato L, Zgaga L, Zitting P, Alavere H, Farrall M, McArdle WL, Nelis M, Peters MJ, Ripatti S, van Meurs JB, Aben KK, Ardlie KG, Beckmann JS, Beilby JP, Bergman RN, Bergmann S, Collins FS, Cusi D, den Heijer M, Eiriksdottir G, Gejman PV, Hall AS, Hamsten A, Huikuri HV, Iribarren C, Kähönen M, Kaprio J, Kathiresan S, Kiemeney L, Kocher T, Launer LJ, Lehtimäki T, Melander O, Mosley TH Jr, Musk AW, Nieminen MS, O'Donnell CJ, Ohlsson C, Oostra B, Palmer LJ, Raitakari O, Ridker PM, Rioux JD, Rissanen A, Rivolta C, Schunkert H, Shuldiner AR, Siscovick DS, Stumvoll M, Tönjes A, Tuomilehto J, van Ommen GJ, Viikari J, Heath AC, Martin NG, Montgomery GW, Province MA, Kayser M, Arnold AM, Atwood LD, Boerwinkle E, Chanock SJ, Deloukas P, Gieger C, Grönberg H, Hall P, Hattersley AT, Hengstenberg C, Hoffman W, Lathrop GM, Salomaa V, Schreiber S, Uda M, Waterworth D, Wright AF, Assimes TL, Barroso I, Hofman A, Mohlke KL, Boomsma DI, Caulfield MJ, Cupples LA, Erdmann J, Fox CS, Gudnason V, Gyllensten U, Harris TB, Hayes RB, Jarvelin MR, Mooser V, Munroe PB, Ouwehand WH, Penninx BW, Pramstaller PP, Quertermous T, Rudan I, Samani NJ, Spector TD, Völzke H, Watkins H, Wilson JF, Groop LC, Haritunians T, Hu FB, Kaplan RC, Metspalu A, North KE, Schlessinger D, Wareham NJ, Hunter DJ, O'Connell JR, Strachan DP, Wichmann HE, Borecki IB, van Duijn CM, Schadt EE, Thorsteinsdottir U, Peltonen L, Uitterlinden AG, Visscher PM, Chatterjee N, Loos RJ, Boehnke M, McCarthy MI, Ingelsson E, Lindgren CM, Abecasis GR, Stefansson K, Frayling TM, Hirschhorn JN. 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467:832–838.

Laposky A. 2005. Deletion of the mammalian circadian clock gene BMAL1/Mop3 alters baseline sleep architecture and the response to sleep deprivation. Sleep 28:395–409.

Lewis SA. Subjective estimates of sleep: An EEG evaluation. Br J Psychol 60:203–208. 1969.

Li Y, Willer C, Sanna S, Abecasis G. 2009. Genotype imputation. Annu Rev Genomics Hum Genet 10:387–406.

Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM; AMFS Investigators, Hayward NK, Montgomery GW, Visscher PM, Martin NG, Macgregor S. 2010. A versatile gene-based test for genome-wide association studies. Am J Hum Genet 87:139–145.

Luciano M, Zhu G, Kirk KM, Gordon SD, Heath AC, Montgomery GW, Martin NG. 2007. No thanks, it keeps me awake: The genetics of coffee-attributed sleep disturbance. Sleep 30:1378–1386.

McCarren M, Goldberg J, Ramakrishnan V, Fabsitz R. 1994. Insomnia in Vietnam era veteran twins: Influence of genes and combat experience. Sleep 17:456–461.

Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G, Gordon SD, Ferreira MA, Wright MJ, Henders AK, Campbell MJ, Duffy DL, Hansell NK, Macgregor S, Slutske WS, Heath AC, Montgomery GW, Martin NG. 2009. Common variants in the trichohyalin gene are associated with straight hair in Europeans. Am J Hum Genet 85:750–755.

Ohayon MM. 2002. Epidemiology of insomnia: What we know and what we still need to learn. Sleep Med Rev 6:97–111.

Partinen M, Kaprio J, Koskenvuo M, Putkonen P, Langinvainio H. 1983. Genetic and environmental determination of human sleep. Sleep 6: 179–185.

Paunio T, Korhonen T, Hublin C, Partinen M, Kivimäki M, Koskenvuo M, Kaprio J. 2009. Longitudinal study on poor sleep and life dissatisfaction in a nationwide cohort of twins. Am J Epidemiol 169:206–213.

Purcell S, 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575.

Purcell S, Cherny SS, Sham PC. 2003. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19(1):149–150.

Reinscheid RK, Xu YL, 2005. Neuropeptide S as a novel arousal promoting peptide transmitter. FEBS J 272:5689–5693.

Satake W, 2009. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. Nat Genet 41:1303–1307.

Selbach O, Haas HL, 2006. Hypocretins: The timing of sleep and waking. Chronobiol Int 23:63–70.

Serretti A, Benedetti F, Mandelli L, Lorenzi C, Pirovano A, Colombo C, Smeraldi E. 2003. Genetic dissection of psychopathological symptoms: Insomnia in mood disorders and CLOCK gene polymorphism. Am J Med Genet B 121B:35–38.

Shimada M, Miyagawa T, Kawashima M, Tanaka S, Honda Y, Honda M, Tokunaga K. 2010. An approach based on a genome-wide association study reveals candidate loci for narcolepsy. Hum Genet 128:433–441.

Soria V, Martínez-Amorós E, Escaramís G, Valero J, Pérez-Egea R, García C, Gutiérrez-Zotes A, Puigdemont D, Bayés M, Crespo JM, Martorell L, Vilella E, Labad A, Vallejo J, Pérez V, Menchón JM, Estivill X, Gratacòs M, Urretavizcaya M. 2010. 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 35:1279–1289.

Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Lango Allen H, Lindgren CM, Luan J, Mägi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segrè AV, Estrada K, Liang L, Nemesh J, Park JH, Gustafsson S, Kilpeläinen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard-Costa NL, Hoesel V, Hottenga JJ, Johansson A, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JR, Peters MJ, Preuss M, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson JO, Lawrence RW, Pellikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MT, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebermann H, Blakemore AI, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, Busonero F, Campbell H, Cappuccio FP, Cavalcanti-Proença C, Chen YD, Chen CM, Chines PS, Clarke R, Coin L, Connell J, Day IN, den Heijer M, Duan J, Ebrahim S, Elliott P, Elosua R, Eiriksdottir G, Erdos MR, Eriksson JG, Facheris MF, Felix SB, Fischer-Posovszky P, Folsom AR, Friedrich N, Freimer NB, Fu M, Gaget S, Gejman PV, Geus EJ, Gieger C, Gjesing AP, Goel A, Goyette P, Grallert H, Grässler J, Greenawalt DM, Groves CJ, Gudnason V, Guiducci C, Hartikainen AL, Hassanali N, Hall AS, Havulinna AS, Hayward C, Heath AC, Hengstenberg C, Hicks AA, Hinney A, Hofman A, Homuth G, Hui J, Igl W, Iribarren C, Isomaa B, Jacobs KB, Jarick I, Jewell E, John U, Jørgensen T, Jousilahti P, Jula A, Kaakinen M, Kajantie E, Kaplan LM, Kathiresan S, Kettunen J, Kinnunen L, Knowles JW, Kolcic I, König IR, Koskinen S, Kovacs P, Kuusisto J, Kraft P, Kvaløy K, Laitinen J, Lantieri O, Lanzani C, Launer LJ, Lecoeur C, Lehtimäki T, Lettre G, Liu J, Lokki ML, Lorentzon M, Luben RN, Ludwig B; MAGIC, Manunta P, Marek D, Marre M, Martin NG, McArdle WL, McCarthy A, McKnight B, Meitinger T, Melander O, Meyre D, Midthjell K, Montgomery GW, Morken MA, Morris AP, Mulic R, Ngwa JS, Nelis M, Neville MJ, Nyholt DR, O'Donnell CJ, O'Rahilly S, Ong KK, Oostra B, Paré G, Parker AN, Perola M, Pichler I, Pietiläinen KH, Platou CG, Polasek O, Pouta A, Rafelt S, Raitakari O, Rayner NW, Ridderstråle M, Rief W, Ruokonen A, Robertson NR, Rzehak P, Salomaa V, Sanders AR, Sandhu MS, Sanna S, Saramies J,

Savolainen MJ, Scherag S, Schipf S, Schreiber S, Schunkert H, Silander K, Sinisalo J, Siscovick DS, Smit JH, Soranzo N, Sovio U, Stephens J, Surakka I, Swift AJ, Tammesoo ML, Tardif JC, Teder-Laving M, Teslovich TM, Thompson JR, Thomson B, Tönjes A, Tuomi T, van Meurs JB, van Ommen GJ, Vatin V, Viikari J, Visvikis-Siest S, Vitart V, Vogel CI, Voight BF, Waite LL, Wallaschofski H, Walters GB, Widen E, Wiegand S, Wild SH, Willemsen G, Witte DR, Witteman JC, Xu J, Zhang Q, Zgaga L, Ziegler A, Zitting P, Beilby JP, Farooqi IS, Hebebrand J, Huikuri HV, James AL, Kähönen M, Levinson DF, Macciardi F, Nieminen MS, Ohlsson C, Palmer LJ, Ridker PM, Stumvoll M, Beckmann JS, Boeing H, Boerwinkle E, Boomsma DI, Caulfield MJ, Chanock SJ, Collins FS, Cupples LA, Smith GD, Erdmann J, Froguel P, Grönberg H, Gyllensten U, Hall P, Hansen T, Harris TB, Hattersley AT, Hayes RB, Heinrich J, Hu FB, Hveem K, Illig T, Jarvelin MR, Kaprio J, Karpe F, Khaw KT, Kiemeney LA, Krude H, Laakso M, Lawlor DA, Metspalu A, Munroe PB, Ouwehand WH, Pedersen O, Penninx BW, Peters A, Pramstaller PP, Quertermous T, Reinehr T, Rissanen A, Rudan I, Samani NJ, Schwarz PE, Shuldiner AR, Spector TD, Tuomilehto J, Uda M, Uitterlinden A, Valle TT, Wabitsch M, Waeber G, Wareham NJ, Watkins H; Procardis Consortium, Wilson JF, Wright AF, Zillikens MC, Chatterjee N, McCarroll SA, Purcell S, Schadt EE, Visscher PM, Assimes TL, Borecki IB, Deloukas P, Fox CS, Groop LC, Haritunians T, Hunter DJ, Kaplan RC, Mohlke KL, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP, van Duijn CM, Wichmann HE, Frayling TM, Thorsteinsdottir U, Abecasis GR, Barroso I, Boehnke M, Stefansson K, North KE, McCarthy MI, Hirschhorn JN, Ingelsson E, Loos RJ. 2010. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 42:937–948.

Utge SJ, Soronen P, Loukola A, Kronholm E, Ollila HM, Pirkola S, Porkka-Heiskanen T, Partonen T, Paunio T. 2010. Systematic analysis of circadian genes in a population-based sample reveals association of TIMELESS with depression and sleep disturbance. PLoS ONE 5:e9259.

Viola AU, Archer SN, James LM, Groeger JA, Lo JC, Skene DJ, von Schantz M, Dijk DJ. 2007. PER3 polymorphism predicts sleep structure and waking performance. Curr Biol 17:613–618.

Visscher PM, Montgomery GW. 2009. Genome-wide association studies and human disease: From trickle to flood. JAMA 302:2028–2029.

Willer CJ, Li Y, Abecasis GR, 2010. METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26:2190–2191.

Wolk R, Gami AS, Garcia-Touchard A, Somers VK. 2005. Sleep and cardiovascular disease. Curr Probl Cardiol 30:625–662.

Wellcome Trust Case Control, C 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678.

Zhao LY, et al. 2006. An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis. Cancer Res 66:9445–9452.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1. Venn diagram showing the breakdown of the sample by Illumina genotyping platform.

Table S1. Correlation between the six variables analysed. All correlations were significant with $P < 0.01$.

Table S2. Power to detect variants with sample size used in present study explaining different proportions of the phenotypic variance in a phenotype with total heritability of 30%. Calculations assume that causal variant is directly genotyped or is tagged by a SNP in complete LD and with the same minor allele frequency.

Table S3. Results from ingenuity pathway analysis of sleep traits showing the number of genes with $P < 0.05$ from gene-based test and top functions and pathways.

Table S4. SNPs in or near candidate genes for sleep/wake regulation with $P < 0.001$.

Table S5. SNPs listed as most strongly associated from Gottlieb et al that show nominal association with sleep traits analysed out of 34 SNPs genotyped or imputed in present sample.